A Search for Antineutrino Disappearance with the SoLid Detector Phase-I Data

Daniel William Galbinski

Imperial College London

Department of Physics

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

A search for antineutrino disappearance to a sterile state with the SoLid detector Phase-I data is presented. Sterile neutrinos are well-motivated by recent anomalies in experimental neutrino data at short baseline, most notably the Reactor Antineutrino Anomaly (RAA), and appear in many extensions to the Standard Model. SoLid (Search for Oscillations with a Lithium-6 Detector) is a very-short-baseline neutrino experiment based at the BR2 research reactor in Mol, Belgium. The detector features a dual-scintillator technology and segmented design that enables the products of inverse beta decay (IBD) interactions to be tagged and reconstructed in three spatial dimensions. The detector granularity also facilitates the effective isolation of the IBD positron from the rest of the visible energy. The analyses conducted for this thesis utilise topological information and machine learning tools to suppress the extremely high rates of background from radioactive and atmospheric sources. A method for extracting the antineutrino energy spectrum as a function of position is applied to the full Phase-I dataset, resulting in $29\,479 \pm 603$ (stat.) IBD candidates with a signal-to-background ratio of 0.26. The spectrum is analysed by a rate-independent fit of the 3+1 neutrino model. No evidence of mixing between active and sterile neutrino states is observed and a portion of the relevant parameter space is excluded at the 90% confidence level (CL), including some of the RAA 95% CL region. Finally, SoLid data is successfully employed to make other measurements, such as antineutrino directionality, that have potential applications in reactor monitoring and nuclear safeguards.

Statement of Originality

I hereby declare that this thesis is my own work. All text within is original and in no instances has been lifted verbatim, in part or in full, from any external source. Due credit is given when presenting work based on previous SoLid studies, or methods and results that were developed with, or performed wholly by, members of the SoLid collaboration other than the author of this thesis. Figures, tables, and other information taken directly or adapted from external published sources are referenced appropriately.

For this thesis I have made extensive use of SoLid software packages (SoLidSim, ROSim, and Saffron2) that I was not directly involved in developing. Chapters 1 and 2 do not contain any original work performed by myself, as they are a literature review and a technical summary of the SoLid experiment respectively. In Chapter 3, I employ an algorithm for event classification developed by the SoLid group at LPC Clermont-Ferrand, albeit using slightly different inputs. The rest of Chapter 3 is entirely my own work.

The data-quality controls described in the first section of Chapter 4 are primarily the result of collaboration between myself and numerous other SoLid members. The antineutrino analysis of Chapter 4 builds on the previous analyses by the LPC Caen (see Ref. [206]) and Imperial College London groups. Specifically, I refer to the design of the preselection cuts, choice of nuBDT input features, and use of the BiPonator variable.

Regarding Chapter 5, the basic background-subtraction technique used for the open dataset had already been devised (as in sections 5.2.1–5.2.4) when I started working on this aspect of the analysis (see Refs. [205–207]). I am solely responsible for improving the procedure and applying it to the entire Phase-I dataset (as in section 5.3). The cross-checks and validations described in chapter 5 are also my own work, with the exception of the log-likelihood fit (section 5.4.3) performed by colleagues at LPC Caen.

I executed the oscillation analysis in its entirety to obtain the results presented in Chapter 6, making use of the SoLO and PySoLO frameworks to which I made substantial contributions. The latter was written exclusively by myself. The choice of statistical tools deployed and evaluation of systematic uncertainties arose from discussions with SoLid collaborators and it follows the work described in Ref. [103]. Finally, the antineutrino directionality study was originally commenced by an MSci student at ICL; I expanded and finalised the method along with my supervisor.

Acknowledgements

I would like to thank the Science and Technology Facilities Council for funding this PhD and granting a much-needed extension due to the impact of the Covid-19 pandemic. Furthermore, a big thanks to the High Energy Physics group at Imperial College London for taking on this budding physicist in the first place back in 2019. I am also obliged to the fellow students in my cohort with whom I shared many laughs during our most unusual of postgraduate experiences.

I am grateful to my supervisor, Antonin Vacheret, for guiding me to the completion of this PhD and patiently providing advice and support throughout. We have had an extremely fruitful collaboration and I hope that we will have the opportunity to work together again in the future. The same goes for all my colleagues at the various member institutes of SoLid who put so much work into this experiment, from which I benefited greatly. We were not able to meet in person as often as we might have liked and it has been a difficult journey, but one which has given me many fond memories.

I am deeply indebted to my parents Marjorie and Dorian, my brother Adam, and all of my extended family for their love over the years. It is not lost on me how extremely lucky I am to have been able to rely on your emotional, social, and (occasional) residential support.

I have also been fortunate enough to live with friends throughout my studies who it seems sustained me with sheer whimsy. To my friends then, the family one gets to choose, frankly none of you were of any real help whatsoever and more likely than not severely hindered the progress of my work. At a push, I suppose I could thank George for the year of isolation we endured together in Glebe Court, tempered by endless cups of tea, fish finger sandwiches, and overly ambitious jigsaw puzzles. To the houses in which I later lived on Salmon Lane, Devonshire Road, and Adelina Grove, you were the best of sanctuaries.

Anna, writing this thesis could have been a tiresome burden, and it was, but it was made so much easier by your love and affection. So, you know, cheers.

Finally, I would like to express special thanks to my friend and tutor Mihai Vasilescu, without whom I could never have reached this point in a million years. I'm only in the field of particle physics because I lost a coin flip in Bristol, but I'm very glad that I did.

Contents

1 Neutrino Physics **17**

1.1 Standard Model Neutrinos . . . 17

1.1.1 History of Discovery . . . 17

1.1.2 Basic Properties . . . 18

1.2 Neutrino Oscillations . . . 19

1.2.1 Solar Neutrino Problem . . . 20

1.2.2 Oscillation Phenomenology . . . 21

1.2.3 Theoretical Implications . . . 24

1.2.4 Experimental Status . . . 25

1.3 Anomalies in Neutrino Data . . . 27

1.3.1 Predictions of Reactor Antineutrino Spectra . . . 27

1.3.2 Reactor Antineutrino Anomaly . . . 31

1.3.3 5 MeV Bump . . . 32

1.3.4 Gallium Anomaly . . . 34

1.3.5 Appearance Anomalies . . . 35

1.4 Sterile Neutrino Hypothesis . . . 36

1.4.1 3+1 Model . . . 37

1.4.2 Detecting Sterile Neutrinos . . . 38

1.4.3 Current Experimental Status . . . 39

1.4.4 Alternative Hypotheses . . . 43

2 SoLid Experiment **46**

2.1 Antineutrino Source . . . 46

2.2 Phase-I Detector . . . 48

2.2.1 Detection Principle . . . 48

2.2.2 Prompt Signal . . . 49

2.2.3 Delayed Signal . . . 51

2.2.4 Photon Collection . . . 52

2.2.5 Full-Scale Detector . . . 54

2.2.6 Deployment at BR2 . . . 55

2.3 Data Acquisition and Reconstruction . . . 57
2.3.1 Front-End Electronics . . . 58
2.3.2 Readout and Triggers . . . 59
2.3.3 Time Clustering . . . 61
2.3.4 Cluster Categorisation . . . 63
2.3.5 CCube Algorithm . . . 65
2.4 Simulations . . . 67
2.4.1 Reactor Antineutrino Generator . . . 67
2.4.2 Detector Model . . . 68
2.5 Calibration . . . 69
2.5.1 MPPC Equalisation . . . 69
2.5.2 In-Situ Calibration . . . 70
2.5.3 Relative Calibration with Muons . . . 74
2.5.4 Energy Linearity . . . 77
2.6 Summary . . . 77

3 Positron & Gamma Estimators 79
3.1 Particle Interactions in the Detector . . . 79
3.1.1 Energy Spectra of IBD Products . . . 79
3.1.2 Electrons & Positrons . . . 80
3.1.3 Gamma Rays . . . 82
3.2 Positron-Gamma Separation . . . 84
3.2.1 Cube-Level Energy Deposits . . . 84
3.2.2 Antineutrino Position . . . 85
3.2.3 Variables for Separation . . . 87
3.2.4 Estimator Definitions . . . 88
3.3 Event Categorisation . . . 92
3.3.1 Cube Categorisation . . . 92
3.3.2 Reconstructed Energy Threshold . . . 93
3.3.3 Gamma Bunch Algorithm . . . 96
3.4 Positron Energy Resolution . . . 97
3.4.1 General Parameterisation . . . 98
3.4.2 Method & Results Per Gamma Category . . . 99

4 Antineutrino Analysis 104
4.1 Data Quality of the Phase-I Dataset . . . 104
4.1.1 Notable Incidents . . . 105
4.1.2 Reactor Schedule . . . 105
4.1.3 Environmental Conditions . . . 106

4.1.4 Muon Bursts . . . 107
4.1.5 DAQ & Saffron2 Parameters . . . 108
4.2 Background Sources . . . 109
4.2.1 Fast Neutron . . . 109
4.2.2 Radioactivity . . . 112
4.2.3 Accidental . . . 114
4.2.4 Composition of Reactor-off Data Before Selection . . . 116
4.3 Preselection Cuts . . . 117
4.3.1 BiPo Suppression . . . 119
4.3.2 Topological Considerations . . . 120
4.3.3 Muon Veto . . . 120
4.3.4 Composition of Reactor-off Data After Preselection . . . 120
4.4 Machine Learning Analysis . . . 121
4.4.1 Boosted Decision Trees . . . 121
4.4.2 Training Variables . . . 123
4.4.3 nuBDT Model . . . 126
4.4.4 Overfitting Checks . . . 134
4.4.5 Efficiency Flatness . . . 135
4.5 Validation with BiPo . . . 136
4.5.1 High-Purity BiPo Region . . . 137
4.5.2 Data-MC Comparison . . . 138

5 Background Subtraction 143
5.1 Parameter Space Regions . . . 143
5.2 Open Dataset Subtraction . . . 144
5.2.1 Accidental . . . 144
5.2.2 BiPo . . . 145
5.2.3 Fast Neutron . . . 147
5.2.4 Reactor Subtraction . . . 150
5.2.5 Optimisation of the nuBDT cut . . . 150
5.2.6 Open Dataset Results . . . 151
5.3 Phase-I Subtraction . . . 153
5.3.1 Composite Pressure Model . . . 154
5.3.2 Time-Evolution of Signal & Background . . . 157
5.3.3 Phase-I Results . . . 160
5.4 Cross-Checks & Validations . . . 161
5.4.1 Scaling with Reactor Power . . . 161
5.4.2 ROff-Minus-ROff Subtraction . . . 162

5.4.3 Individual Reactor Cycles . . . 164
5.4.4 Data-MC Comparisons . . . 165

6 Oscillation Analysis 169
6.1 Obtaining the Predictions . . . 169
6.1.1 Response Matrix . . . 170
6.2 Error Propagation . . . 173
6.2.1 Statistical Uncertainties . . . 174
6.2.2 Systematic Uncertainties . . . 176
6.2.3 Summary . . . 180
6.3 Frequentist Oscillation Fit . . . 181
6.3.1 Hypothesis Testing . . . 182
6.3.2 Definition of the Metric . . . 183
6.3.3 Constructing Exclusion Regions . . . 184
6.3.4 Generating Toy Experiments . . . 187
6.3.5 Building the Test Statistic . . . 188
6.3.6 Null Hypothesis . . . 190
6.3.7 Phase-I Sensitivity & Exclusion Contours . . . 191

7 Prospects & Conclusion 195
7.1 Analysis Improvements . . . 195
7.2 SoLid Phase-II . . . 196
7.3 Uranium-235 Spectrum Measurement . . . 198
7.4 Bayesian Oscillation Fit . . . 199
7.5 Reactor Monitoring Applications . . . 200
7.5.1 rBDT Model . . . 200
7.5.2 Reactor Power Measurement . . . 201
7.5.3 Antineutrino Directionality . . . 204
7.6 Conclusion . . . 207

A Positron Energy Estimators 235

B BiPo Data-MC 238

C Background Corrections 240

D Background Characterisation 244

E Positron Spectra 247

F Single-Event Energy Smearing 249

G Wilk's Theorem vs Pseudo-Experiments **252**

H rBDT Training **254**

List of Figures

1.1 Evidence of oscillation of atmospheric neutrinos from Super-Kamiokande . . 21
1.2 Flavour-integrated neutrino flux at Earth as a function of energy 25
1.3 Two-dimensional projections of allowed regions in mixing-parameter space from NuFit 5.2 global fits 28
1.4 Number of antineutrinos per fission for each fissile isotope 30
1.5 The Reactor Antineutrino Anomaly 31
1.6 Allowed regions in parameter space constructed by a rate-plus-shape analysis of the RAA data and GALLEX and SAGE calibration data 32
1.7 5 MeV bump in LEU and HEU spectra 33
1.8 Gallium anomaly and BEST allowed region 35
1.9 LSND beam excess and MiniBooNE neutrino energy spectrum 36
1.10 Neutrino mass hierarchy with and without a fourth mass eigenstate 38
1.11 PROSPECT and STEREO sterile neutrino searches 40
1.12 Neutrino-4 sterile neutrino search 41
1.13 DANSS and joint RENO-NEOS sterile neutrino searches 42
1.14 Summary of the sterile neutrino searches from VSBL reactor experiments, KATRIN, MicroBooNE, and solar neutrino data 44

2.1 Schematic of the BR2 reactor vessel 47
2.2 Photograph of the BR2 reactor core as viewed from above 48
2.3 Tree-level Feynman diagram and antineutrino spectrum of inverse beta decay 49
2.4 SoLid detector voxel and antineutrino detection principle 50
2.5 EJ-200 scintillator response as a function of initial particle energy 51
2.6 Components responsible for photon transport and collection in SoLid 53
2.7 Exploded view of a detector plane and full module 55
2.8 BR2 containment building blueprint and Geant4 visualisation 56
2.9 Geant4 rendering of the detector at BR2 in relation to the reactor core . . . 57
2.10 Photographs of the SoLid detector at BR2 58
2.11 Hardware components of a detector plane and module 59
2.12 Example waveform that fires the NS trigger 61

2.13 Visualisation of the time-clustering algorithm applied to an ES waveform . . 63
2.14 Integral-over-amplitude separation of ES and NS events 65
2.15 CCube algorithm problem for an energy deposit localised to a single cube . . 66
2.16 Predicted number of fissions per day for a 2018 BR2 reactor cycle 67
2.17 Gain determination and dependence of dark-count and cross-talk on overvoltage 70
2.18 CAD renderings of CROSS with all calibration gaps and source positions . . 71
2.19 Absolute neutron calibration from September 2018 calibration data 72
2.20 Absolute and relative neutron efficiency of each detector plane 73
2.21 Calibrations of the electromagnetic signal using a Kolmogorov-Smirnov test and an analytical fit of the Klein-Nishina formula 74
2.22 Vertical and horizontal projections of a horizontal-muon track 75
2.23 Evolution of the mean PVT light yield over time 76
2.24 Response of the PVT scintillator as a function of initial particle energy . . . 77

3.1 Initial energy spectra of IBD products 81
3.2 Electron stopping power and CSDA range in PVT 82
3.3 Gamma attenuation in PVT as a function of energy 83
3.4 Example Compton scattering distributions 84
3.5 Positron energy deposits per cube in each spatial direction 85
3.6 Significant Migration Probability as a function of positron energy 86
3.7 Correlation between the most energetic true-space and reco-space cubes . . . 87
3.8 Cube-level variables of positron and annihilation-gamma energy deposits . . 89
3.9 Evaluation of the positron-gamma separation of each estimator 91
3.10 Cube multiplicities of the Cross and MAGE estimators 93
3.11 Boundaries defining the classification of energetic cubes 94
3.12 Energy distributions of light-leakage, ghost, and annihilation gamma cubes . 95
3.13 Optimisation of the cube energy threshold 96
3.14 Smearing of the true positron energy in 2D and 1D 98
3.15 GaussExp fits on the reconstructed energy distribution of 3 MeV positrons . 100
3.16 GaussExp fitted parameters as a function of true positron energy 101
3.17 Parameterisation of the energy resolution of IBD events 102
3.18 Mean proportion of deposited energy per gamma category 103

4.1 Thermal power of the BR2 reactor over the Phase-I data-taking campaign . 106
4.2 Time-dependence of temperature and humidity in the detector environment . 107
4.3 Selection of the cut on reactor-pool level 108
4.4 Time-evolution of eight DAQ-level and Saffron2 parameters 110
4.5 Histograms of eight DAQ-level and Saffron2 parameters 111
4.6 A cosmic-ray-induced shower simulated with CRY 112

4.7 Simulated BiPo electron energy spectrum . . . 113
4.8 Comparison between BiPonisher and BiPonator . . . 115
4.9 Composition of reactor-off data before selection . . . 117
4.10 Probability density functions of preselection variables . . . 118
4.11 Time between the NS and nearest preceding muon for reactor-off data . . . 120
4.12 Background composition of reactor-off data after preselection cuts . . . 121
4.13 Relative fractions of each gamma category after preselection . . . 124
4.14 PDFs of available BDT training variables I . . . 128
4.15 PDFs of available BDT training variables II . . . 129
4.16 Correlation matrices of training variables for signal and background events . 132
4.17 Output of the nuBDT model . . . 133
4.18 Fischer scores of nuBDT input features . . . 134
4.19 Overfitting check of the nuBDT model in terms of predicted score . . . 136
4.20 nuBDT selection efficiency on signal events . . . 137
4.21 Schematic of three types of BiPo event . . . 138
4.22 BiPo data-MC comparisons of nuBDT score . . . 140
4.23 BiPo data-MC comparisons of the ratio between distributions after and before cutting on nuBDT score . . . 141
4.24 BiPo data-MC comparison plots in multiple variables . . . 142

5.1 Energy and z-position spectra of events in the open-dataset FPNT sample . 145
5.2 Δt distribution of open-dataset reactor-off data after analysis cuts . . . 146
5.3 Pressure model from open-dataset reactor-off data . . . 148
5.4 Effect of pressure correction on open-dataset fast-neutron rates . . . 149
5.5 Figure of Merit and S/B as a function of excess per day (nuBDT cut) . . . 152
5.6 Stacked backgrounds and reactor-on signal region distributions . . . 152
5.7 Data-MC comparison of open-dataset IBD candidates . . . 153
5.8 Evolution of the PVT light yield estimated by ^{22}Na calibrations . . . 154
5.9 Overview of the Phase-I data-taking campaign . . . 155
5.10 Pressure models derived from reactor-off data of different epochs . . . 156
5.11 Evolution of the BiPo-to-Signal factors over Phase-I . . . 158
5.12 Trending plots of excess and background over time . . . 159
5.13 Statistical significance of the number of IBD events per day . . . 160
5.14 Positron energy spectrum per module of the Phase-I IBD yield . . . 161
5.15 IBD rate as a function of reactor power . . . 162
5.16 ROff-minus-ROff subtraction for P_A and P_B . . . 163
5.17 Mean IBD rates and spectra of individual reactor cycles . . . 164
5.18 Comparison between all individually background-subtracted ROn cycles . . . 165

5.19 Distributions of χ^2/ndf values from data-MC comparisons of all ES variables 166
5.20 Data-MC comparisons of eight control variables 168

6.1 Two-dimensional IBD MC prediction in true and reco space 170
6.2 Performance of the response matrix on IBD MC events 171
6.3 Two-dimensional Flat MC prediction in true and reco space 172
6.4 Performance of the response matrix on Flat MC events 173
6.5 1D projections of the background templates 175
6.6 Measured (data) and simulated (Asimov) statistical uncertainties 176
6.7 Light yield systematic derived from ROSim toys 177
6.8 System matrix systematic from ROSim toys 178
6.9 ES2ES1 data-MC comparison of BiPo events 179
6.10 BiPo data-MC discrepancy systematic 180
6.11 Summary of statistical and systematic uncertainties 181
6.12 Chi-square probability density functions 182
6.13 Absolute covariance matrix of statistical and systematic uncertainties 184
6.14 Example sensitivity contours for an Asimov dataset 186
6.15 Validation of the Cholesky decomposition of the covariance matrix 188
6.16 Example $\Delta\chi^2$ distribution under an alternative 3+1 hypothesis and corresponding critical values as a function of confidence level 189
6.17 Heat map of $\Delta\chi^2_{\text{crit}}$ values at 90% CL in the Δm^2_{41}-$\sin^2(2\theta_{14})$ plane 190
6.18 Distribution of best-fit points of 7000 null-hypothesis pseudo-experiments . . 191
6.19 The χ^2 distribution of 10^5 null-hypothesis toy experiments 192
6.20 SoLid Phase-I sensitivity and exclusion contours at 90% and 95% CL 193
6.21 SoLid Phase-I exclusion contours compared to the first results by STEREO and PROSPECT 193
6.22 SoLid Phase-I IBD spectrum against the HM model with and without oscillations 194

7.1 Characteristics of the S14 series MPPCs 196
7.2 Neutron trigger optimisation of Phase-II 197
7.3 SoLid Phase-I positron energy spectrum against the HM model prediction . . 199
7.4 Fischer scores of rBDT input features 201
7.5 Output of the rBDT model 202
7.6 Figure of Merit and S/B as a function of excess per day (rBDT cut) 202
7.7 Open-dataset IBD rate as a function of reactor power with the rBDT 203
7.8 Position and momentum distributions of the Rand. MC sample 205
7.9 Characterisation of the intrinsic and physical neutron direction biases 206
7.10 Results of the antineutrino direction measurement 206

A.1 Proportions of initial particle energy captured by each estimator 236
A.2 Proportions of deposited particle energy captured by each estimator 237

B.1 BiPo data-MC plots in variables not previously displayed 239

C.1 Effects of the pressure correction on the fast-neutron rates of the Phase-I sub-datasets. 242
C.2 Fits on the Δt distributions for the determination of the BiPo-to-Signal factors of the Phase-I sub-datasets 243

D.1 Stacked histograms of ROn (accidental and BiPo) and ROff (pressure-corrected fast neutron) events with pressure-corrected total ROn signal region distributions overlaid 245
D.2 Normalised histograms of accidental, BiPo, and pressure-corrected fast-neutron events on the same axes 245
D.3 Normalised histograms of accidental, BiPo, and pressure-corrected fast-neutron events on different axes 246

E.1 Positron energy spectrum per module for the independent halves of the Phase-I dataset 248

F.1 GaussExp parameters as a function of true positron energy. 250
F.2 Performance of the single-event sampling method 251

G.1 Wilk's theorem vs toy-generation method 253

H.1 Overfitting check of the rBDT model in terms of predicted score 255

List of Tables

1.1 Types of neutrino experiment and relevant mixing parameters 26
1.2 Quantitative summary of the current-generation VSBL reactor experiments . 43

2.1 Summary of the SoLid trigger strategy 60

3.1 Definition of each energy estimator in terms of spatial and energy-based cuts 90
3.2 Optimisation of the MAGE cuts 92
3.3 Relative fractions of different gamma categories according to the applied cube energy threshold 97
3.4 Fitted parameters that quantify the positron energy resolution 102

4.1 BR2 reactor schedule over the Phase-I data-taking campaign 106
4.2 DAQ-level and Saffron2 data-quality criteria 109
4.3 Preselection criteria of the antineutrino analysis 119
4.4 List of available BDT training variables 127
4.5 XGBoost hyperparameters of the nuBDT model 131
4.6 Overfitting check of the nuBDT model in terms of AUC and LogLoss 135
4.7 Summary of the cuts that define the high-purity BiPo region 138

5.1 Rates of different event types in the open dataset 151
5.2 Delineation of the Phase-I dataset into ROn periods, ROff periods, and pressure epochs 155
5.3 Summary of the antineutrino analysis results 161

C.1 Reactor-off BiPo-to-Signal factors and summed pressure corrections 241
C.2 Reactor-on BiPo-to-Signal factors and summed pressure corrections 241

H.1 Overfitting check of the rBDT model in terms of AUC and LogLoss. 255

Chapter 1

Neutrino Physics

The Standard Model (SM) of particle physics describes our current understanding of fundamental particles and their interactions. Throughout the development of the SM, the neutrino sector has been an important source of experimental anomalies that point to new physics. This chapter provides an overview of the discovery of the neutrino and its basic properties, followed by a description of the neutrino oscillation phenomenon. Next, anomalies in neutrino experiments are used to motivate the search for eV-scale sterile neutrinos, with particular focus on experiments at nuclear reactors. The present experimental landscape in this regard is summarised and proposed solutions to the anomalies are provided.

1.1 Standard Model Neutrinos

1.1.1 History of Discovery

By the late 1920s, experiments had firmly established that the energy spectrum of β-decay electrons is continuous [1–3], refuting numerous suggestions of secondary effects that could cause inhomogeneity in an otherwise monoenergetic spectrum [4]. It had previously been assumed that the electron should be emitted with a fixed energy corresponding to the mass difference between the initial and final states, in analogy with α-decay. The contradiction between experiment and theory was difficult to resolve without abandoning the principle of conservation of energy, until Wolfgang Pauli conjectured that a second spin-$\frac{1}{2}$ particle is emitted in the reaction with which the electron shares the available energy. The new particle would have to be electrically neutral, extremely light, and interact very weakly with matter. It was incorporated into E. Fermi's 1934 theory of β-decay as the *neutrino* [5].

The contemporary evidence validated enough of the Fermi theory predictions (see Sargent curves [6], for example) that the majority of physicists accepted the existence of the neutrino even before a direct detection. In 1953, C. Cowan and F. Reines designed and performed an experiment [7, 8] to directly detect electron antineutrinos via inverse beta de-

cay (IBD). The cross section for such an interaction was calculated to be staggeringly small at $6 \times 10^{-44}\,\mathrm{cm}^2$. Their first detector consisted of a large volume of cadmium-doped liquid scintillator enclosed by photomultiplier tubes (PMTs), situated next to a nuclear fission reactor. Taking advantage of the extremely high reactor-antineutrino flux, the experiment detected 0.41 ± 0.20 signal events per minute, thus tentatively establishing the existence of the neutrino. The result was confirmed in an improved experimental setup [9] a few years later. The temporal coincidence technique pioneered in these experiments remains one of the primary tools for tagging IBD events.

In 1962 and 1964 respectively, experiments at Brookhaven National Laboratory (BNL) and Conseil Européen pour la Recherche Nucléaire (CERN) found evidence of two neutrino generations (or "flavours") [10–12], each with an independently conserved lepton number [13]. This idea was originally motivated by the absence of $\mu \to \gamma e$ observations. BNL measured the charged-current interactions of neutrinos from pion decay in an aluminium spark chamber. A lack of electron shower events strongly favoured the hypothesis that neutrinos produced in β-decay (associated with electrons) are different from those produced in pion decay (associated with muons). The existence of a third generation of leptons was proved by the discovery of the charged τ lepton in 1975 [14–16] and the corresponding τ neutrino was detected in 2001 by the DONUT experiment [17].

1.1.2 Basic Properties

Nowadays, it is understood that Fermi theory (of weak interactions) is a low-energy approximation to the Standard Model. The Fermi interaction underwent several revisions to take into account new discoveries including the Gammow-Teller interaction, [18], parity violation [19–25], and the V−A structure of the weak interaction [26, 27]. Its current form—known as the Glashow-Weinberg-Salam (GWS) model [28–30]—unifies electromagnetism and the weak nuclear force into the *electroweak* interaction, described by an SU(2) × U(1) gauge group.

Neutrinos are spin-$\frac{1}{2}$ particles that are singlets under SU(3) colour transformations*. They also have zero electric charge so interact exclusively via the weak force. Motivated by the results of the Goldhaber experiment [31] which established the negative helicity of neutrinos, SM neutrinos are formulated as eigenstates of the left-handed chirality operator P_L. Helicity is closely related to chirality and they are equal for massless particles, as neutrinos were believed to be. The operator P_L and its right-handed version P_R project a general four-component Dirac spinor ψ into its chiral components:

$$P_{L/R}\psi = \frac{1 \mp \gamma^5}{2}\psi = \psi_{L/R}\,, \tag{1.1}$$

*These two properties make the neutrino a lepton, by definition.

where $\psi_{L/R}$ are two-component Weyl spinors ($\psi = \psi_L + \psi_R$) and $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ is the fifth Dirac gamma matrix. Left-chiral neutrino fields are constructed as weak isospin $\mathrm{SU}(2)_L$ doublets with their respective charged-lepton counterparts:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L , \qquad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L , \qquad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L , \tag{1.2}$$

where the subscript L signifies that the particles are eigenstates of P_L. Right-handed charged-lepton fields (e_R, μ_R, τ_R) exist as $\mathrm{SU}(2)_L$ singlets in the Standard Model, but right-handed neutrino fields are absent. As Lagrangian densities, the weak charged-current (CC) and neutral-current (NC) interactions of neutrinos take the form:

$$-\mathcal{L} = \underbrace{\frac{g}{\sqrt{2}} \sum_{\alpha=e,\mu,\tau} \bar{\nu}_\alpha \gamma^\mu P_L l_\alpha W_\mu + \text{h.c.}}_{\mathcal{L}_{cc}} + \underbrace{\frac{g}{2\cos\theta_\mathrm{W}} \sum_{\alpha=e,\mu,\tau} \bar{\nu}_\alpha \gamma^\mu P_L \nu_\alpha Z_\mu}_{\mathcal{L}_{nc}} \tag{1.3}$$

where g is the weak coupling constant, W and Z are the charged and neutral gauge boson fields respectively, θ_W is the Weinberg angle, and γ^μ are the Dirac gamma matrices. Charged-current interactions mix leptons within their generations, exchanging both electric charge and weak isospin. In other words, under an $\mathrm{SU}(2)_L$ transformation a neutrino state becomes the corresponding charged-lepton state and vice versa. Neutral-current interactions mediated by the Z boson cannot change the particle type.

The assumption of massless neutrinos is encoded in the electroweak framework by the exclusion of the right-handed component, since typical fermion mass terms require both left- and right-handed fields. The general form of a Dirac mass term[†] in a Lagrangian density is:

$$\mathcal{L} = m_\mathrm{D}\overline{\psi}\psi = m_\mathrm{D}(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L) , \tag{1.4}$$

where the spinors ψ, ψ_L, ψ_R are as defined above and m_D is the Dirac mass. Explicit Dirac mass terms violate gauge invariance because they couple left-handed and right-handed spinors which have different gauge symmetries. Instead, SM fermions (and the W and Z bosons) acquire mass through the Higgs mechanism in spontaneous symmetry breaking (SSB) [32–34]. Regardless, right-handed spinor components are still necessary so SM neutrinos are massless by default.

1.2 Neutrino Oscillations

Contrary to the mathematical formulation of the Standard Model, the observation of neutrino oscillations requires the masses of at least two neutrinos to be non-zero and that all

[†]Particles with Dirac mass terms only are known as Dirac particles.

three masses are non-degenerate. A brief history of this discovery is provided below, followed by a derivation of oscillation phenomenology and its implications for theory and experiment.

1.2.1 Solar Neutrino Problem

Electron neutrinos are powerful tools for studying the internal structure of the Sun, as they are not scattered or absorbed during their passage to the exterior of the star. The majority of solar neutrinos are emitted in the proton-proton (p-p) fusion chain, which can be summarised as $4\mathrm{p} \rightarrow {}^{4}_{2}\mathrm{He} + 2e^{+} + 2\nu_e$, though in reality it proceeds along many possible decays. In 1964, J. Bahcall and R. Davis published twin papers respectively on the theoretical and experimental merits of a proposed experiment to measure solar neutrinos [35, 36]. The experiment utilised a tank of liquid tetrachloroethylene to detect neutrinos via the reaction $^{37}\mathrm{Cl} + \nu_e \rightarrow {}^{37}\mathrm{Ar}$. The threshold for this reaction is 0.814 MeV so the detector was only sensitive to the highest energy neutrinos from ^{8}B or ^{7}Be decay in the p-p chain. The detector was operated in the Homestake mine in South Dakota, after which it is named. Calculations of the solar neutrino flux were mainly driven by J. Bahcall and collaborators [37–39] as part of the Standard Solar Model (SSM).

The first Homestake result, published in 1968, set an upper limit on the product of the solar neutrino flux and ^{37}Cl cross section [40]. The limit was lower than the theoretical prediction by a factor of 2.5 [41], yet the uncertainties on the latter were extremely large so no definitive conclusions were drawn. Improvements to the analysis method allowed the upper limit to become an actual measurement of the rate in 1970. A significant deficit in the rate of solar neutrinos with respect to the SSM prediction was measured and persisted over more than a decade of data-taking [42], which became known as the *solar neutrino problem*.

Independent and complementary validations of the Homestake results came from the gallium-based experiments SAGE and GALLEX [43, 44] and the water-Cherenkov KamiokaNDE-II experiment [45, 46]. Detection of solar neutrinos in the gallium experiments relied on the CC reaction $\nu_e + {}^{71}\mathrm{Ga} \rightarrow {}^{71}\mathrm{Ge} + e^{-}$. This process has a much lower energy threshold (0.234 MeV) than the ^{37}Cl reaction, allowing the experiments to detect the lower-energy p-p neutrinos. The 7.5 MeV threshold of KamiokaNDE-II made it sensitive only to ^{8}B neutrinos, but it detected them in real-time and with directional information. The results of all four experiments were incompatible with the SSM. An attempt to reconcile the experimental data with the SSM by varying all parameters of the model within their uncertainties was unsuccessful [47].

The deficit had to arise from some physical effect of neutrinos themselves. The most plausible explanation—after rejection of the neutrino decay hypothesis by the detection of neutrinos from the SN1987A supernova [49]—was believed to be neutrino oscillations. The neutrino oscillation (or "mixing") mechanism was originally proposed by B. Pontecorvo

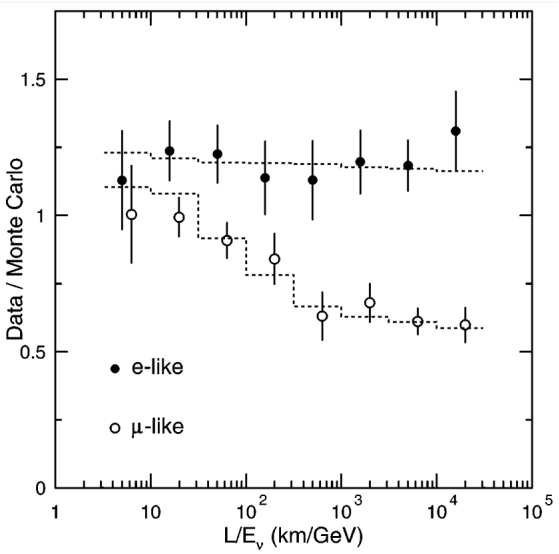


Figure 1.1: Evidence of oscillation of atmospheric neutrinos from Super-Kamiokande published in 1998 [48]. The best-fit of the data (round points) is $\nu_\mu \to \nu_\tau$ oscillation with $\Delta m^2 = 2.2 \times 10^{-3}\,\mathrm{eV}^2$ and $\sin^2(2\theta)$ =1, as illustrated by the dashed lines.

in 1957 [50]. It is a process in which a neutrino ν_α produced in a CC interaction of the corresponding charged lepton l_α, can later interact to produce a charged lepton of a different flavour l_β. Hence, the solar neutrino problem is caused by the *disappearance* of electron neutrinos to other flavours that the existing detectors were incapable of measuring, meaning they only had access to a fraction of the total flux.

In 1998, Super-Kamiokande (the successor of KamiokaNDE-II) presented the first evidence of atmospheric neutrino oscillation [48], as shown in Fig. 1.1. The data can be satisfactorily described by $\nu_\mu \to \nu_\tau$ oscillation with a mass-squared difference $\Delta m^2 = 2.2 \times 10^{-3}\,\mathrm{eV}^2$. This was followed by results from the Sudbury Neutrino Observatory (SNO) collaboration in 2001 [51]. SNO was a heavy-water Cherenkov detector sensitive to CC interactions of electron neutrinos, as well as NC and elastic-scattering interactions of all three neutrino flavours. Evidence of a non-electron-like component in the solar neutrino flux was found. The total flux from all channels was consistent with the prediction of the SSM, thus definitively confirming the neutrino oscillation solution to the solar neutrino problem.

1.2.2 Oscillation Phenomenology

A consequence of neutrino oscillations is that the traditional "flavour eigenstates" of the weak interaction do not obey the Dirac equation, so in some sense are not fundamental SM objects. The flavour fields (ν_e, ν_μ, ν_τ) can be expressed as linear superpositions of three massive neutrino fields (ν_1, ν_2, ν_3) that do obey the Dirac equation:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = V_{\text{PMNS}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} , \tag{1.5}$$

where V_{PMNS} is a 3×3 unitary matrix known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix [52], which transforms between the mass and flavour eigenbases. In general, Greek letters denote lepton flavour (e, μ, τ) and numbers denote mass eigenstate labels (1, 2, 3). The PMNS matrix can be conveniently parameterised as the product of three rotation matrices, with three mixing angles and a phase that violates charge-parity (CP) symmetry:

$$V_{\text{PMNS}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{\delta_{\text{CP}}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} , \tag{1.6}$$

where c_{ij} and s_{ij} are the cosines and sines of the mixing angle θ_{ij}. Each matrix corresponds to the mixing between two mass states. Additional phases that exist in the case of Majorana neutrinos[‡] have no effect on oscillations [53].

The derivation below broadly follows section 14.4 of Ref. [54]. In quantum field theory, the state $|\nu_\alpha\rangle$ is created by the field operator $\nu_\alpha^\dagger$ acting on the vacuum. Thus the flavour eigenstates ($|\nu_e\rangle$, $|\nu_\mu\rangle$, $|\nu_\tau\rangle$) are related to the mass eigenstates ($|\nu_1\rangle$, $|\nu_2\rangle$, $|\nu_3\rangle$) by:

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i}^* \, |\nu_i\rangle \ , \tag{1.7}$$

where the complex conjugate of $U_{\alpha i}$ is taken because the equation refers to states rather than fields [53, 55]. The complex conjugation is not applied for antineutrinos. In a plane-wave model, the mass eigenstate $|\nu_i\rangle$ picks up a quantum mechanical phase related to its energy and momentum as it propagates through time (t) and space ($\boldsymbol{x}$) in a vacuum:

$$|\nu_i(t, \boldsymbol{x})\rangle = e^{-iE_i t} \, |\nu_i(0, \boldsymbol{x})\rangle = e^{-i(E_i t - \boldsymbol{p_i} \cdot \boldsymbol{x})} \, |\nu_i(0, 0)\rangle \ , \tag{1.8}$$

where $E_i = \sqrt{m_i^2 + |\boldsymbol{p_i}|^2}$ is the relativistic energy for mass m_i and 3-momentum $\boldsymbol{p_i}$. Like all quantum mechanical phases, this phase is not directly observable. However, it gives rise to the macro-scale phenomenon of neutrino oscillations because of interference effects between eigenstates that differ in mass. The probability that a flavour eigenstate ν_α is detected as

[‡]See section 1.2.3.

flavour eigenstate ν_β after time t is given by:

$$P_{\nu_\alpha \to \nu_\beta} = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 = |\sum_{i=1}^{3}\sum_{j=1}^{3} U^*_{\alpha i} U_{\beta j} \langle \nu_j | \nu_i(t) \rangle|^2 , \tag{1.9}$$

where the flavour eigenstates have been substituted for the mass eigenstates according to the PMNS matrix. Assuming ultrarelativistic speeds, the energy can be expressed as $E_i \simeq p_i + \frac{m_i^2}{2p_i} \simeq E + \frac{m_i^2}{2E}$ (where E is the average energy of the eigenstates) and the time t can be replaced with the path length $L = |\boldsymbol{x}| \simeq ct$. It can then be derived that the oscillation probability is equal to:

$$P_{\nu_\alpha \to \nu_\beta} = \delta_{\alpha\beta} - 4\sum_{i<j}^{3} \mathrm{Re}[U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j}] \sin^2\left(\frac{L}{4E}\Delta m^2_{ij}\right) + 2\sum_{i<j}^{3} \mathrm{Im}[U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j}] \sin\left(\frac{L}{2E}\Delta m^2_{ij}\right) , \tag{1.10}$$

where $\Delta m^2_{ij} \equiv m^2_i - m^2_j$ is the mass-squared difference (or "mass splitting") between two neutrino states and $\delta_{\alpha\beta}$ is the Kronecker delta function. This equation demonstrates that massless neutrinos—or those with degenerate masses—can not oscillate and that the oscillation probability does not depend on the absolute neutrino mass scale. Note that in the case of the high-energy solar neutrinos measured by Homestake and KamiokaNDE, mixing between different neutrino types is enhanced by the Mikheyev-Smirnov-Wolfenstein (MSW) matter effect.

Assuming only two neutrino flavours, the probability of measuring a neutrino created as flavour α, as a *different* flavour β is equal to:

$$P_{\nu_\alpha \to \nu_\beta} = \sin^2(2\theta)\sin^2\left(1.27\,\Delta m^2[\mathrm{eV}^2]\,\frac{L[\mathrm{m}]}{E[\mathrm{MeV}]}\right) \quad \alpha \neq \beta , \tag{1.11}$$

where θ and Δm^2 are the mixing angle and mass-squared difference between the two flavours respectively. The assumed unitarity of the PMNS matrix means that the two-flavour *survival probability* is then given by:

$$P_{\nu_\alpha \to \nu_\alpha} = 1 - \sin^2(2\theta)\sin^2\left(1.27\,\Delta m^2[\mathrm{eV}^2]\,\frac{L[\mathrm{m}]}{E[\mathrm{MeV}]}\right) . \tag{1.12}$$

In this two-neutrino case, the oscillation probability is independent of the sign of the mass splitting as it only appears in an even function ($\sin^2$).

1.2.3 Theoretical Implications

Given the irrefutable evidence in favour of L/E-dependent oscillation, strongly implying non-zero neutrino mass, its physical origin must explained. Massive neutrinos are themselves evidence of Beyond Standard Model (BSM) physics and open the door to further extensions. For each of the left-handed SM neutrino fields (ν_L) one can postulate a right-handed neutrino field (ν_R) that is a singlet under all SM gauge transformations. Such fields are called *sterile*, as opposed to the *active* neutrinos of the Standard Model.

In analogy with the charged fermions, neutrino masses may then be generated via Yukawa interaction with the Higgs doublet [56, 57]. After SSB, a neutrino field gains Dirac mass $m_{\mathrm{D}} = \lambda v/\sqrt{2}$, where v is the vacuum expectation value of the Higgs field and λ is the coupling constant. In this extension the masses of neutrinos and charged leptons are generated by the same mechanism, yet somehow the former ends up orders of magnitude smaller than the latter (as known from measurements of the neutrino mass scale). To achieve this, the strength of the coupling λ between the neutrino fields and the Higgs field would have to be extremely small. This often forces the introduction of new mechanisms to suppress the neutrino mass scale [58], at which point the "naturalness" of generating neutrino masses in this way becomes questionable.

Another possibility is that SM neutrinos and antineutrinos are objects of the same field with opposite chirality, which can be represented as self-conjugate two-component spinors. These types of particle are known as *Majorana fermions*. Generic Majorana mass terms are of the form:

$$-\mathcal{L} = \frac{m_{\mathrm{M}}}{2}\overline{\psi}\psi = \frac{m_{\mathrm{M}}}{2}(\overline{\psi}_L\psi_R^c + \overline{\psi}_R\psi_L^c)\,, \tag{1.13}$$

where the superscript c denotes charge conjugation and the field ψ obeys the Majorana condition $\psi = \psi^c$. Neutrino oscillation is evidently a process that violates lepton flavour conservation, though global lepton number is conserved in the case of Dirac neutrinos. Majorana mass terms violate lepton number by two units.

The most general models introduce both Dirac and Majorana masses in the form of a mass matrix, which must be diagonalised to obtain the mass eigenstates and corresponding mass terms. Importantly, bare Majorana masses[§] can be arbitrarily large since they are not connected to the Higgs mechanism. If one chooses a large enough Majorana mass scale, say the grand unification scale, the smallness of the active neutrino masses is naturally achieved via some kind of seesaw mechanism [59–61] in which the masses are inversely proportional to the Majorana mass: $m_\nu \propto m_{\mathrm{D}}^2/m_{\mathrm{M}}$. Overall, any BSM theory with massive neutrinos will likely introduce a rich array of possible phenomena (new fields, energy scales, interactions, etc.) that may be experimentally testable [62–65].

[§]A bare mass in this context refers to a mass term that is not generated via the Higgs mechanism.

1.2.4 Experimental Status

There are currently four main neutrino sources employed to study SM neutrino flavour transitions and measure their mixing parameters: solar, atmospheric, reactor, and accelerator. Electron neutrinos from the Sun are predominantly produced by p-p fusion and nuclear reactors release electron antineutrinos from β-decays of fission fragments. The atmospheric neutrino flux is a mixture of electron and muon (anti)neutrinos generated by the decay of muons and pions in cosmic-ray-induced particle cascades. This mechanism also drives (anti)neutrino emission in accelerator experiments, where a proton beam impinges on a fixed target to create mesons that subsequently decay to 1st and 2nd generation neutrinos and charged leptons. The neutrino spectra of these sources span a wide energy range from $\mathcal{O}$(keV) to $\mathcal{O}$(PeV), as illustrated in Fig. 1.2. Accelerator neutrinos are absent from the figure but they typically have MeV–GeV energies, depending on the proton beam. Other types such as geoneutrinos[¶] [66, 67], supernova neutrinos [49], and neutrinos of astrophysical origin [68] have also been detected.

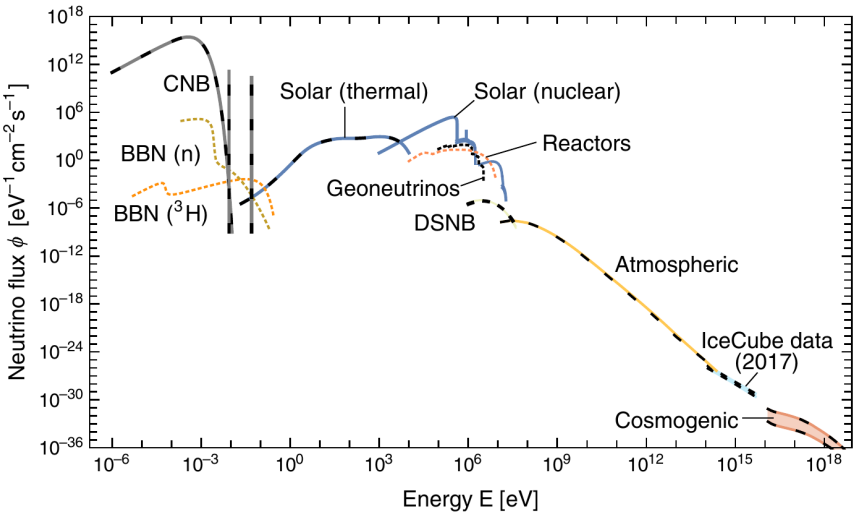


Figure 1.2: Flavour-integrated neutrino flux at Earth as a function of energy, from [69]. Solid and dashed lines denote neutrinos and antineutrinos respectively. Thermal neutrinos (e.g. from plasmon decay) and neutrinos from the cosmic neutrino background (CNB), the diffuse supernova neutrino background (DSNB), and Big Bang nucleosynthesis (BBN) have never been observed.

Neutrino oscillations and the PMNS matrix formalism introduces six new parameters to the Standard Model: three mixing angles, two mass splittings, and one phase. The sensitivity of an experiment to a particular parameter depends on the L/E ratio, which drives the oscillation probability. In some cases, the experimenter has the option of choosing the neutrino energy and/or the baseline (source-detector distance) to maximise sensitivity to a particular parameter. Alternatively, the L/E ratio is determined entirely by the neutrino

[¶]Geoneutrinos are emitted in radioactive decays of naturally occurring isotopes in the Earth's crust and mantle.

Experiments	Dominant (Important)
Solar Experiments (SK, Borexino)	θ_{12} (Δm^2_{21}, θ_{13})
Reactor LBL (KamLAND)	Δm^2_{21} (θ_{12}, θ_{13})
Reactor MBL (Daya Bay, RENO, Double Chooz)	θ_{13}, $\lvert\Delta m^2_{31,32}\rvert$
Atmospheric (SK, IceCube-DC)	($\lvert\Delta m^2_{31,32}\rvert$, θ_{23}, θ_{13}, $\delta_{\rm CP}$)
Accel LBL ν_μ, $\bar{\nu}_\mu$, Disapp (K2K, MINOS, T2K, NOνA)	$\lvert\Delta m^2_{31,32}\rvert$, θ_{23}
Accel LBL ν_e, $\bar{\nu}_e$ App (MINOS, T2K, NOνA)	$\delta_{\rm CP}$ (θ_{13}, θ_{23})

Table 1.1: Types of neutrino experiment and the mixing parameters to which they are sensitive, from the 2022 PDG review [54]. Some previous- and current-generation experiments are given in parentheses. LBL and MBL are acronyms for long-baseline and medium-baseline respectively. Accel, Disapp and App are abbreviations of accelerator, disappearance, and appearance respectively.

creation mechanism, as for solar neutrinos which have a pre-determined (relatively narrow) energy spectrum and baseline. The different types of neutrino experiment and parameters to which they are most sensitive are summarised in Tab. 1.1, adapted from the 2022 Particle Data Group (PDG) review [54]. To date, measurements have been made of all SM neutrino mixing parameters at varying degrees of precision. The latest global fits of neutrino mixing parameters from NuFIT 5.2 [70] are plotted in Fig. 1.3. Notably, there are still two allowed regions of θ_{23} in different quadrants and the constraints on $\delta_{\rm CP}$ are not yet particularly strong.

The 1$^{\rm st}$ and 2$^{\rm nd}$ neutrino eigenstates are almost mass-degenerate with respect to the 3$^{\rm rd}$ generation, i.e. $|\Delta m^2_{31}| \simeq |\Delta m^2_{32}| \gg \Delta m^2_{21}$. Therefore, the mass difference between the first two neutrino generations can often be neglected in experiments and the oscillation probability approximated with only two flavours. The current experimental values are [70]:

$$\Delta m^2_{\rm SOL} \equiv \Delta m^2_{21} \sim 7.4 \times 10^{-5}\,{\rm eV}^2\,, \tag{1.14}$$

$$\Delta m^2_{\rm ATM} \equiv \Delta m^2_{31,32} \sim 2.5 \times 10^{-3}\,{\rm eV}^2\,, \tag{1.15}$$

where the subscripts of $\Delta m^2_{\rm SOL}$ and $\Delta m^2_{\rm ATM}$ define them as the solar and atmospheric mass splittings respectively.

The sign of $\Delta m^2_{31,32}$ has not yet been conclusively established, leading to two possible scenarios for the neutrino mass hierarchy: normal ordering (NO) where $m_1 < m_2 \ll m_3$ and inverted ordering (IO) where $m_3 \ll m_1 < m_2$. Resolving the mass ordering is an important goal of many next-generation precision experiments such as DUNE [71], KM3NeT [72], JUNO [73], Hyper-Kamiokande [74], and IceCube-Gen2 [75]. Furthermore, the absolute

neutrino mass scale can be directly probed by measuring the tritium β-decay spectrum, whose end-point shape depends on $\sum_i |V_{ei}|^2 m_i^2$. The most recent result from the Karlsruhe Tritium Neutrino (KATRIN) experiment is an upper limit of 0.8 eV [76]. Complementary constraints are provided by large-scale cosmological surveys such as those by the PLANCK satellite [77], which can also give an extra handle on the number of neutrino generations.

Whether the neutrino is a Dirac particle or the first fundamental Majorana particle in the SM can be tested in neutrinoless double β-decay [78] experiments (for example, SuperNEMO [79] and KamLAND-Zen [80]). Finally, while CP violation in the quark sector has been understood since the 1960s [81], alone it is insufficient to explain the matter-antimatter asymmetry of the Universe. For this reason, CP-violation in the neutrino sector is of great interest. In a recent analysis, T2K has excluded $\delta_{\mathrm{CP}} = 0, \pi$ at more than 90% confidence level (CL) [82].

1.3 Anomalies in Neutrino Data

New discoveries in particle physics are often preceded by anomalous results that push the experimental effort in a particular direction. Nowhere is this more explicit than in the case of the solar neutrino problem. It is conceivable that the anomalies and tensions in the current experimental landscape will have a similar effect. The main focus of this section is the anomalies in the neutrino disappearance channel, but other types are also discussed. Collectively, they are referred to as the short-baseline (SBL) anomalies. All of the anomalies can be (independently) explained by neutrino mixing with an eV-scale mass-squared difference. The existing mass splittings are clearly unsuitable to explain such mixing so an additional neutrino state is required. For reasons expanded upon in section 1.4, any new light neutrinos must be sterile.

1.3.1 Predictions of Reactor Antineutrino Spectra

Nuclear reactors produce electron antineutrinos ($\bar{\nu}_e$) up to around 10 MeV as β-decay products of the daughters of heavy nuclei that undergo neutron-induced fission. The four primary isotopes burned in nuclear reactors are ^{235}U, ^{238}U, ^{239}Pu, and ^{241}Pu. Each isotope has an associated antineutrino energy spectrum that is an aggregation of the antineutrinos emitted in β-decays of its individual fission fragments. It is estimated that there are around 800 participating isotopes [83]. Research reactors burn high-enriched uranium (HEU) fuel where ^{235}U dominates the antineutrino emission. Commercial reactors burn low-enriched uranium (LEU) comprising mostly ^{235}U and ^{239}Pu, with a contribution from the other isotopes up to around 15%. For a detailed review of reactor neutrinos see [84–86] and references therein. The main advantage of using a nuclear reactor as an antineutrino source is that it produces

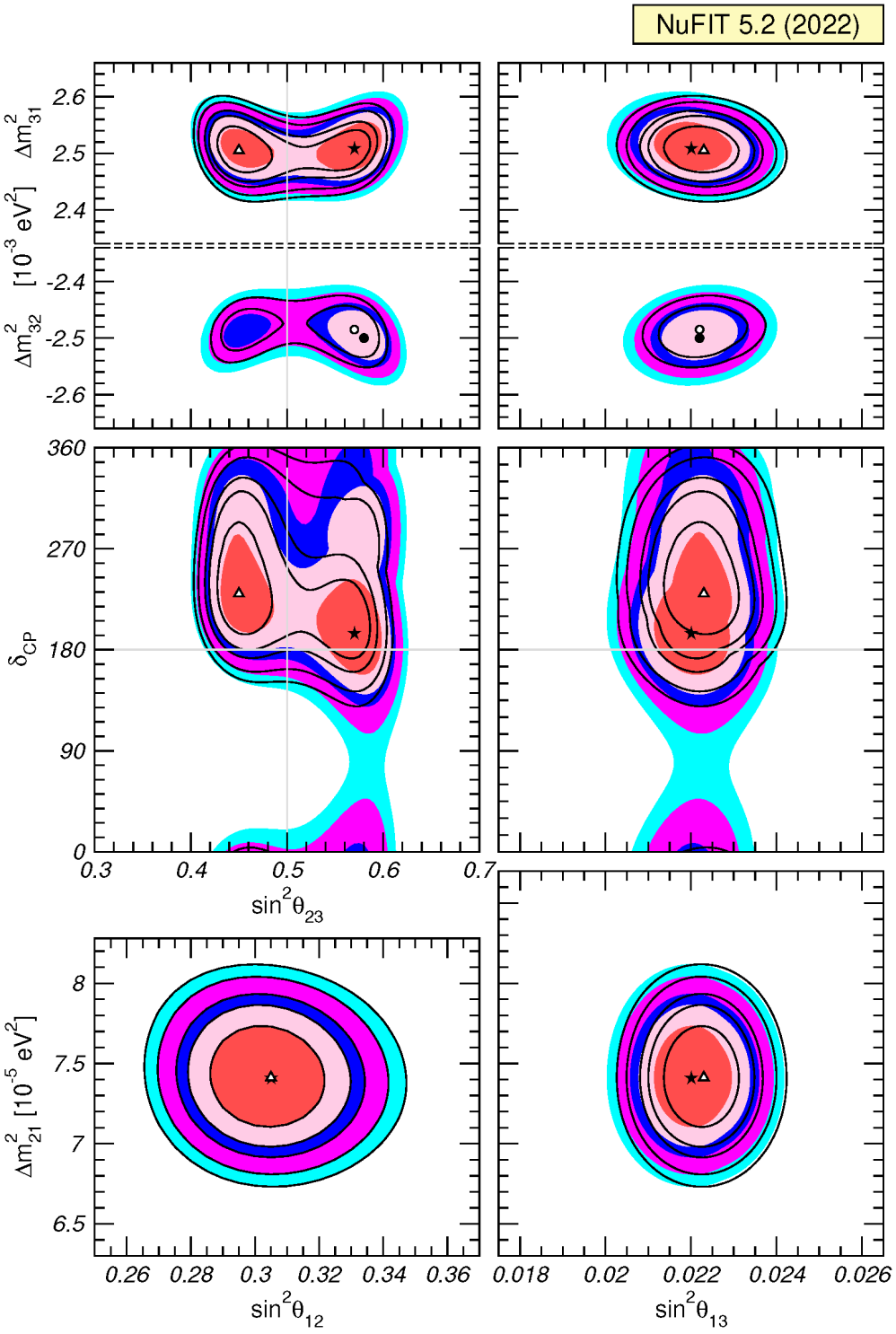


Figure 1.3: Two-dimensional projections of allowed regions in mixing-parameter space from NuFit 5.2 global fits [70]. Filled regions and black lines are computed without and with Super-Kamiokande atmospheric neutrino data respectively, at 68.3%, 90%, 95.5%, 99%, and 99.7% confidence levels. Normal ordering corresponds to Δm^2_{31} and inverted ordering to Δm^2_{32}.

a high-intensity, controllable flux of single-flavour antineutrinos, whose main interaction mechanism—inverse beta decay—is theoretically clean and well-understood. Reactor-based experiments have provided the most precise measurements of θ_{13} to date [87–89].

Labelling the antineutrino energy spectrum of fissile isotope k ($k \in \{$ ^{235}U, ^{238}U, ^{239}Pu, ^{241}Pu$\}$) as $S_k(E_\nu)$, the total number of antineutrinos emitted from a reactor core can be written as:

$$N_{\text{em}} = \int_E \int_t \int_V \underbrace{\frac{P_{\text{th}}(t)}{\sum_k \left[\phi_k(\mathbf{x},t)\,\langle E_{\text{fiss}}\rangle_k\right]}}_{(1)} \underbrace{\sum_k \left[\phi_k(\mathbf{x},t)\,S_k(E_\nu)\right]}_{(2)} \,\mathrm{d}\mathbf{x}^3\,\mathrm{d}t\,\mathrm{d}E\,, \tag{1.16}$$

where part (1) gives the number of fissions in the reactor core at time t and position $\mathbf{x}$ and part (2) expresses the combined antineutrino spectrum. The equation is integrated over the volume of the reactor core, the antineutrino energy range, and the time interval. The parameter $\phi_k(\mathbf{x},t)$ is the fission fraction of isotope k—the number of fissions of that isotope divided by the total number of fissions—which tracks the change in relative proportions of each isotope as the fuel is burnt. Finally, $\langle E_{\text{fiss}}\rangle_k$ is the mean energy released per fission and P_{th} is the thermal power of the reactor.

The antineutrino spectrum for a given isotope $S_k(E_\nu)$ can be computed using two complementary techniques: the conversion method and the summation (*ab initio*) method. In the conversion method, antineutrino spectra are obtained from experimentally measured β-electron spectra. The seminal β-spectra measurements of ^{235}U, ^{239}Pu, and ^{241}Pu were taken in the 1980s at the Institut Laue Langevin (ILL) [90, 91] by irradiating thin foils of each isotope in the thermal neutron flux of the reactor vessel. The β-electron spectrum of ^{238}U measured at FRM II in Garching, Germany, has been available since 2014 [92]. The spectrum of each isotope is summed over all participating β-branches, which cannot be disentangled. Hence, in the original treatments of the data [93, 94] the total electron spectrum was approximated as the sum of thirty "virtual" branches of assumed shapes (from allowed decays only[‖]) with an evenly spaced range of end-point energies:

$$S_k(E_\beta) = \sum_{i=1}^{30} a_i P_i(E_\beta, E_{0i}, Z)\,, \tag{1.17}$$

where Z is the average nuclear charge of parent nuclei in real branches and P_i is the spectral shape of the i^{th} virtual branch. The associated amplitudes a_i were determined as floating parameters in a fit of equation 1.17 to the data. The electron spectrum for a given virtual branch is then converted to an antineutrino spectrum (E_ν) by replacing the electron energy E_β with $E_\nu = E_0 - E_\beta$, where E_0 is the end-point energy of the branch. These converted

[‖]Allowed decays involve no change in parity whereas forbidden decays do.

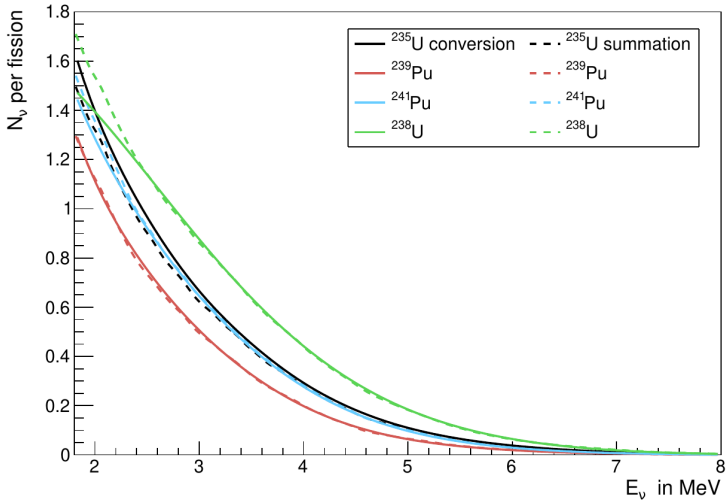


Figure 1.4: Number of antineutrinos per fission for each fissile isotope from a conversion-based method (solid lines) [100, 101] and a summation method (dashed lines) [102]. The displayed energy range begins at the IBD threshold but the spectra extend below this value. Figure from [103].

antineutrino spectra have uncertainties on the absolute rate of about 3%.

In contrast to the top-down approach of the conversion method, the summation method utilises information (fission fractions, β-decay spectra, etc.) about hundreds of individual isotopes from nuclear databases to construct a bottom-up antineutrino spectrum prediction [95–97]. The calculation can also be done on a purely theoretical basis, as when it was first performed [98, 99]. The spectrum per isotope can be expressed as:

$$S_k(E_\beta) = \sum_{f=1} A_f(t)\, S_f(E_\beta)\,, \tag{1.18}$$

where $A_f(t)$ is the activity of fission fragment f and $S_f(E)$ is made by summing the associated β-branch spectra weighted by their branching ratios. The equivalent antineutrino spectrum is again obtained by simply subtracting the electron energy from the spectrum end-point for each β-branch. This method generally has larger uncertainties than the conversion method at around 10% or higher. A comparison between modern conversion [100, 101] and summation [102] spectra is provided in Fig. 1.4.

The conversion method relies on fewer theoretical assumptions than the summation method. However, any uncertainties or biases in the underlying data are propagated to the final prediction. The short irradiation time of the original ILL samples compared to typical reactor cycles also leads to "off-equilibrium" effects, where long-lived fission fragments are under-represented in the measured electron spectra [84, 100]. Accounting for the differences in the temperature (kinetic energy) profile of the neutron flux at different reactors is also essential for an accurate flux prediction. The summation method is frequently impacted by incomplete or inaccurate nuclear databases [104]. Moreover, the Pandemonium effect—

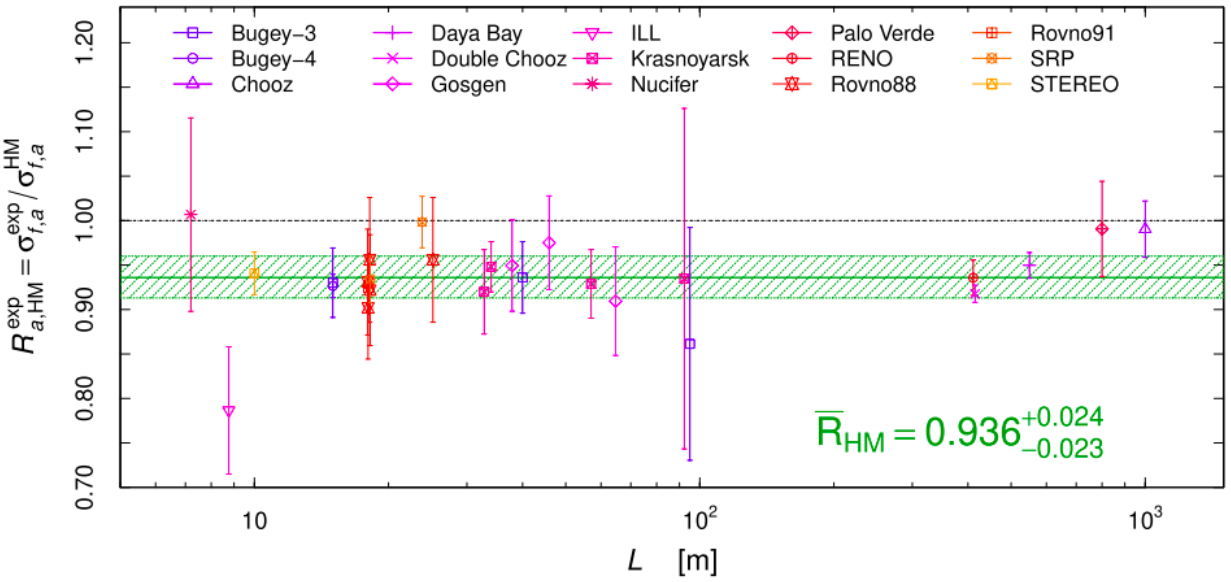


Figure 1.5: The Reactor Antineutrino Anomaly as of 2022 [109]. The mean ratio between data and the HM prediction and associated error are represented by the green line and hatched band. The significance of the RAA is 2.5σ.

detector inefficiencies that cause high-energy de-excitations to be missed when determining a β-spectrum via the associated γ-ray emission—can bias some nuclear database information. Both methods suffer from poor understanding of forbidden decays that make up around 25–30% of the total β-decays in the spectra and significantly enhance the uncertainties [105, 106].

1.3.2 Reactor Antineutrino Anomaly

Both the conversion and summation methods were re-assessed by T. Mueller et al. in 2011, as part of preparation for the Double Chooz experiment and in light of updated nuclear databases [100]. They developed a combined approach to predicting the antineutrino spectrum of ^{235}U, ^{239}Pu, and ^{241}Pu using database information [107, 108] supplemented with ILL data to reduce systematic uncertainties. The model also includes corrections for off-equilibrium and Pandemonium effects, though the latter remained a significant sources of uncertainty in the summation method, as did missing information on exotic β-decaying nuclei. The pure summation method was used for the ^{238}U spectrum. The results of the conversion-based method were independently verified and slightly improved upon by P. Huber in the same year [101]. The Huber-Mueller (HM) model, as it is known, comprises the ^{238}U summation spectrum and the improved conversion spectra of the other three isotopes.

The HM model was found to be compatible with the ILL results in terms of spectral shapes but shifted the overall normalisation of ^{235}U, ^{239}Pu, and ^{241}Pu upwards by 3.1%. A subsequent re-evaluation of existing experimental results from sub-kilometre-baseline reac-

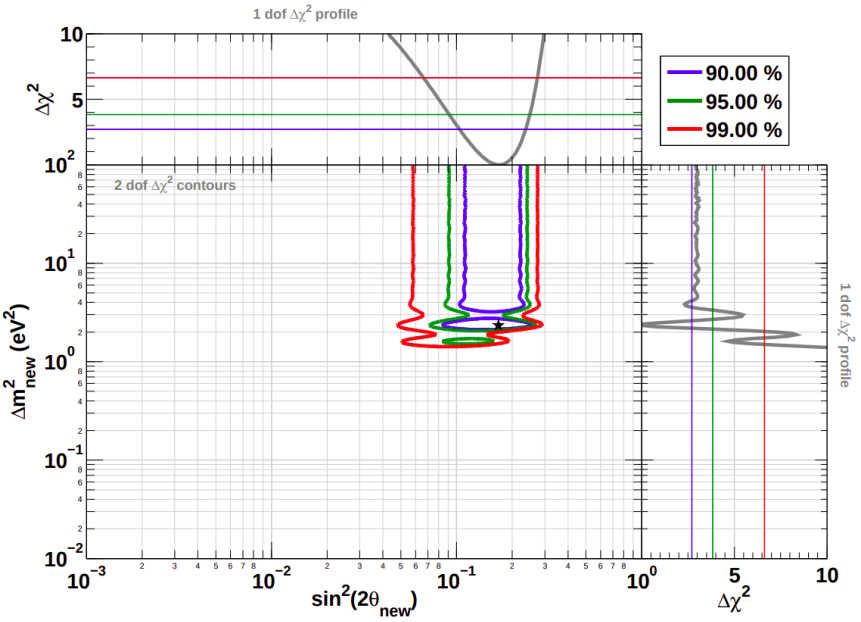


Figure 1.6: Allowed regions in the Δm^2-$\sin^2(2\theta)$ plane constructed by a rate-plus-shape analysis of the RAA data and GALLEX and SAGE calibration data [56]. The best-fit point is indicated by a black star. The eV-scale mass-squared difference implied by the anomaly would require the existence of a BSM neutrino to explain it.

tor experiments revealed a 6% deficit in the mean antineutrino rate compared to the HM prediction. This was dubbed the Reactor Antineutrino Anomaly (RAA) [110]. The RAA is illustrated in Fig. 1.5 which includes experimental results published since the original 2011 paper [109]. The mean deficit of the fifteen experiments is around 6.4% at 2.5σ significance. Various interpretations of the anomaly have been suggested including neutrino disappearance via mixing, errors in the HM model itself, and possible correlated experimental effects.

If the deficit is to be attributed to neutrino oscillations the required mass splitting would be at the eV-scale, suggesting mixing with a BSM neutrino. The 90%, 95%, and 99% CL allowed regions from [56] are shown in Fig. 1.6. The contours are constructed by a rate-plus-shape analysis of reactor neutrino data (including ILL and Bugey-3 spectra) and GALLEX and SAGE calibration data (see section 1.3.4). The data is very well described by the best-fit point $\Delta m^2 = 2.3\,\mathrm{eV}^2$ and $\sin^2(2\theta) = 0.17$. The no-oscillation hypothesis is rejected at the 3.6σ confidence level.

1.3.3 5 MeV Bump

The accuracy of the HM model was further questioned by the discovery of an unexpected feature in the shape of the reactor antineutrino spectrum in 2014, named the 5 MeV bump. Several experiments operating at LEU commercial reactors (Daya Bay, RENO, Double Chooz) reported a 10% deviation in their visible energy spectra, with respect to the HM predic-

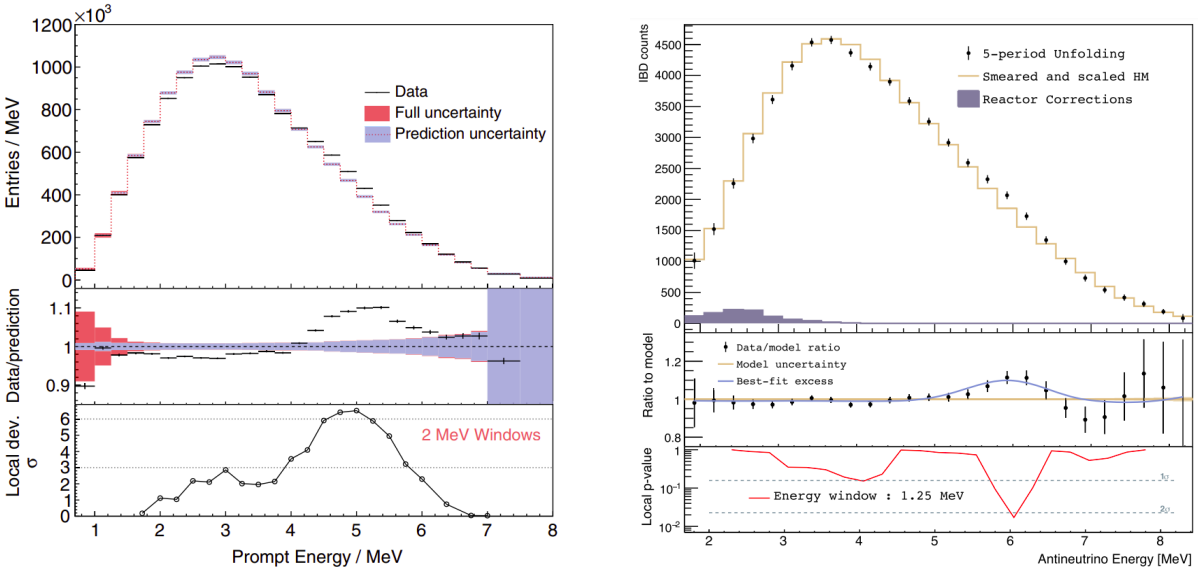


Figure 1.7: Left: Daya Bay LEU prompt-energy spectrum compared to the HM model [118]. The significance of the local deviation in the 4–6 MeV region exceeds 6σ significance. Right: PROSPECT HEU (unfolded) antineutrino energy spectrum compared to the HM model [119]. The local deviation is less significant than the Daya Bay result but still exceeds 2σ.

tion at around 5 MeV [111–113]. Note that the Bugey-3 experiment did not observe any such feature, despite taking measurements at a similar LEU reactor [114]. More recently, the very-short-baseline (VSBL), i.e. sub-ten-metre, experiments PROSPECT and STEREO based at HEU reactors observed similar distortions [115–117].

Fig. 1.7 shows the energy spectrum from Daya Bay (left) as of 2019 [118] and from PROSPECT (right) as of 2023 [119]. The Daya Bay result is plotted as a prompt (visible) energy spectrum, whereas the PROSPECT data has been deconvoluted from detector effects using an unfolding method to extract the antineutrino energy spectrum. The visible energy of IBD events in liquid scintillators is lower than the antineutrino energy by approximately 0.78 MeV because of the interaction threshold and energy of the annihilation gammas. The local p-values in the ratios of the Daya Bay(PROSPECT) spectrum to the HM model exceed 6(2) standard deviations in the 4–6(5.375–6.625) MeV region. The difference in the position of the deviations is largely due to the difference between antineutrino energy and prompt energy.

The source of the distortion is likely not an unidentified background because it is found to be correlated with reactor power. The most probable explanation is deemed to be a mis-modelling of the reactor flux [106, 120, 121]. Strong indications that the bump has a nuclear physics origin, possibly from a particular isotope, comes from measurements by RENO of the antineutrino spectrum as a function of fission fraction [122]. They found non-zero correlation between the size of the bump and the fraction of ^{235}U in the fuel at the 2.9σ level, making it more difficult to attribute the bump to energy-scale uncertainties [114] or

non-standard neutrino interactions [123]. Neither can the bump be reasonably reconciled as the neutrino oscillation pattern of the rate deficit, or otherwise. This discussion is continued in section 1.4.4 in which non-neutrino-mixing solutions of the RAA are also considered.

1.3.4 Gallium Anomaly

The Gallium Experiment (GALLEX) and the Soviet-American Gallium Experiment (SAGE) intended to measure solar neutrinos using gallium-based detectors. The detectors were calibrated using artificial neutrino sources ^{51}Cr (GALLEX AND SAGE) and ^{37}Ar (SAGE only). Both isotopes emit sub-MeV energy neutrinos via electron capture, respectively in four and two monoenergetic lines [124]. The dominant decay branch of ^{51}Cr(^{37}Ar) releases a 747(811) keV electron neutrino. The average event rate of the four calibration runs was lower than expected from the theoretical cross section [125] and detector efficiencies, with a data-MC ratio of 0.84 ± 0.05 [43, 126]. This is called the *gallium anomaly*.

The Baksan Experiment on Sterile Transitions (BEST) experiment was conceived and built for the express purpose of investigating the gallium anomaly in the context of neutrino disappearance. The basic principle of BEST is the same as its predecessors in that it employs a ^{51}Cr neutrino source and liquid ^{71}Ga target to detect electron neutrinos via the $\nu_e + {}^{71}\text{Ga} \to {}^{71}\text{Ge} + e^-$ reaction. As for GALLEX and SAGE, the radioactive ^{71}Ge atoms were extracted after exposure to the neutrino source and placed in a proportional counter device. The number of ^{71}Ge decays is directly related to the number of neutrino interactions. BEST utilises two separate detector volumes, an inner layer and an outer layer, to probe neutrino oscillations on the metre scale. The two layers are simultaneously exposed to the source. BEST performed ten separate exposures lasting around nine days each. The half-life of ^{71}Ge is around eleven days [127] so counting commenced twenty-four hours after each extraction. The neutrinos in the gallium experiments only travel $\mathcal{O}(1)$ m distances from production to detection, meaning they have the shortest baseline of all VSBL experiments.

The 2022 BEST result [128] confirms the gallium anomaly with greater precision and pushes the mean rate of all six measurements to 20% below the expected value. The ^{71}Ge production rates divided by the prediction are $R_{\text{in}} = 0.79 \pm 0.05$ and $R_{\text{out}} = 0.77 \pm 0.05$ for the inner and outer volumes respectively. Like the RAA, the gallium anomaly can be explained by eV-scale neutrino oscillations with an allowed region displayed in Fig. 1.8. The best-fit point of the six gallium measurements corresponds to a relatively large mixing angle of $\sin^2(2\theta) = 0.34$ and a mass splitting of $\Delta m^2 = 1.25\,\text{eV}^2$. The ratio between R_{out} and R_{in} is statistically consistent with unity. Other explanations that have been examined are an overestimation of the neutrino cross section on ^{71}Ga [129–134] and errors or systematic uncertainties on the estimated extraction and counting efficiencies [126, 135, 136]. These explanations are slightly favoured over neutrino disappearance because of the strong tension

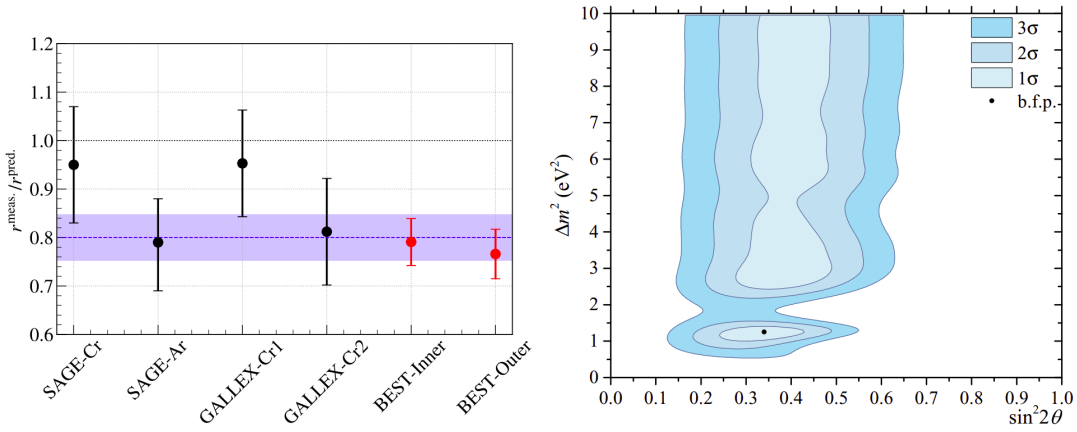


Figure 1.8: Left: Ratio of the observed event rate to prediction (product of theoretical cross section and experimental efficiency) in the gallium experiments GALLEX, SAGE, and BEST [128]. The purple line and band signify the mean value and uncertainty of all six measurements. The most precise values are from BEST (red). Right: Allowed regions of the combined GALLEX, SAGE, and BEST data at three confidence levels from [128]. The black point indicates the best-fit of the data to a two-flavour mixing model. The magnitude of the deficit suggests a larger mixing angle than the RAA.

between the best-fit mixing parameters and solar neutrino data [137–139] and the agreement between the ratios at the two distances.

1.3.5 Appearance Anomalies

The Liquid Scintillator Neutrino Detector (LSND) experiment at Los Alamos searched for evidence of $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillation. The 798 MeV proton beam impinged on a fixed target to produce muon antineutrinos via at-rest μ^+ decay, which travelled a distance of 30 m to the detector. LSND reported an excess of electron-like antineutrino events above the background in the muon antineutrino beam consistent with eV-scale $\bar{\nu}_\mu \to \bar{\nu}_e$ neutrino oscillations [140, 141]. The results were contradicted by the Karlsruhe Rutherford Medium Energy Neutrino (KARMEN) experiment which found no evidence of electron-like neutrino appearance in a similar parameter-space region [142]. The MiniBooNE** experiment at Fermilab [143–145] was devised to confirm or refute the LSND result. MiniBooNE utilises a more energetic neutrino source than LSND, the 8 GeV Fermilab Booster proton beam, thereby avoiding the background events that were relevant in the LSND analysis. The baseline is correspondingly larger at approximately 541 m in order to test the same L/E space.

MiniBooNE recently reported an excess of 638.0 ± 52.1 (stat.) $\pm$ 122.2 (syst.) electron-like events in a beam of primarily muon (anti)neutrinos, from 18.75×10^{20} and 11.27×10^{20} protons on target in neutrino mode and antineutrino mode respectively [146]. The excess

**BooNE stands for Booster Neutrino Experiment.

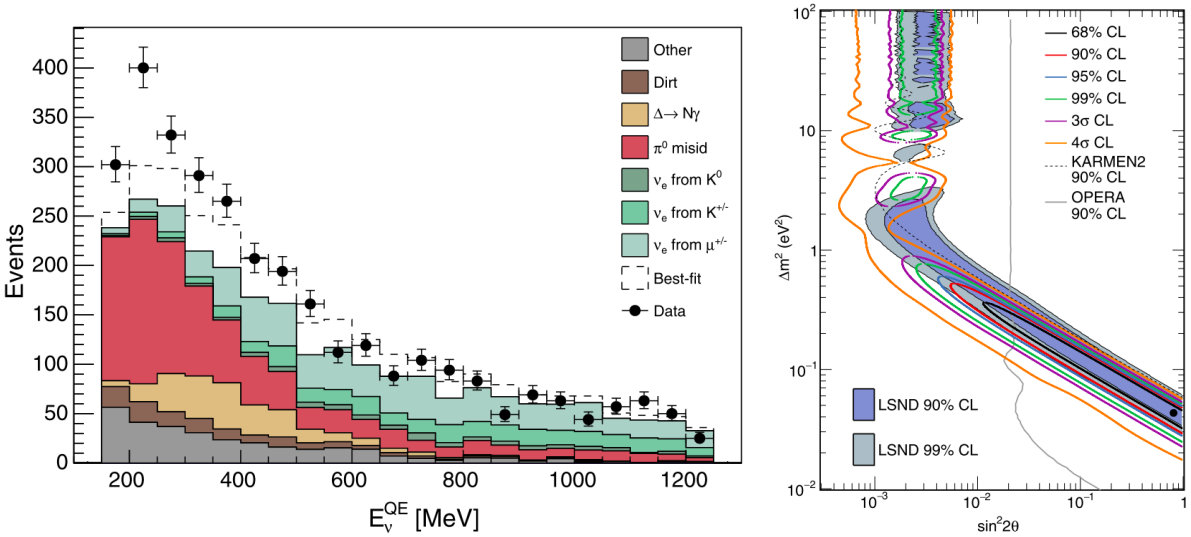


Figure 1.9: Left: The electron neutrino energy spectrum measured by the MiniBooNE detector from 18.75×10^{20} protons on target. A significant excess in the data (black points) can be seen above the predicted backgrounds (filled histograms). The dashed line shows the best fit in a two-neutrino model with oscillation parameters $\sin^2(2\theta) = 0.807$ and $\Delta m^2 = 0.043\,\text{eV}^2$, which is not a particularly good fit to the data. Right: MiniBooNE allowed regions (coloured solid lines) compared to the LSND allowed regions (filled areas). The allowed regions are in tension with the KARMEN2 and OPERA exclusion regions that are also shown. Both figures from [146].

is consistent with the LSND results and referred to as the MiniBooNE low-energy excess (LEE). The MiniBooNE electron neutrino energy spectrum and corresponding limits on oscillation parameters are presented in Fig. 1.9. The data can be fit by a two-neutrino oscillation model with an $\mathcal{O}(1)$eV mass splitting [147]. However, the best-fit oscillation parameters ($\sin^2(2\theta) = 0.807$, $\Delta m^2 = 0.043\,\text{eV}^2$) give a reduced chi-square of 21.7/15.5, thus do not describe the data particularly well.

A variety of SM and other non-SM explanations have also been proposed [148–150], since the LEE is not compatible with muon disappearance searches at [151–153]. MiniBooNE is a mineral-oil Cherenkov detector with poor capability for distinguishing between electron- and photon-generated signals. This limitation is addressed by the liquid-argon time-projection chamber (LArTPC) experiment MicroBooNE [154], which has excellent imaging capabilities for distinguishing between different event types. MicroBooNE is thoroughly scrutinising the MiniBooNE LEE and has so far disfavoured numerous SM backgrounds—such as an underestimated Δ-baryon production rate—as the source of the excess [155–159].

1.4 Sterile Neutrino Hypothesis

The resolution of the anomalies most relevant to this thesis is the sterile neutrino hypothesis. The precise measurements of the Z boson width made at LEP [160] only permit three light

($m_\nu \leq M_Z/2$) active neutrinos. The eV-scale mass splitting mandated by the SBL anomalies, combined with the upper limits on the absolute neutrino mass scale, mean that if the cause of the anomalies is a new neutrino, its mass is indeed less than half the Z boson mass. Therefore, in order to satisfy the LEP constraints it must be sterile, i.e. a singlet under all SM gauge symmetries. Fields of this nature have previously been discussed in the context of generating mass terms for the SM neutrinos by coupling with right-handed chiral fields.

1.4.1 3+1 Model

Sterile neutrinos appear in many BSM theories across a huge range of mass scales, but only eV-scale neutrinos [161] fall within the scope of this work. The 3+1 model is a simple extension of the SM neutrino sector that adds a fourth mass eigenstate ν_4 and corresponding fourth flavour eigenstate ν_s. To accommodate a fourth neutrino in the mixing formalism, the PMNS matrix is expanded to a 4×4 matrix:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \\ \nu_s \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} & U_{\mu 4} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} & U_{\tau 4} \\ U_{s1} & U_{s2} & U_{s3} & U_{s4} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{pmatrix} , \tag{1.19}$$

where the new parameters U_{si} and $U_{\alpha 4}$ describe the mixing between the active states and the sterile state. A parameterisation similar to equation 1.6 gives three new mixing angles (θ_{14}, θ_{24}, θ_{34}), two new CP-violating Dirac phases, and one new CP-violating Majorana phase [162]. A sterile neutrino is accessible to experiments via mixing since it couples to the SM neutrinos through a mass term, assuming that at least one active neutrino comprises a non-zero component of ν_4.

Assuming normal ordering of the active neutrino mass hierarchy, the mixing parameters can be visualised as in Fig. 1.10. In this example, the fourth mass eigenstate is almost entirely aligned with the sterile state and has minimal, but non-zero, active neutrino components. The SM mass eigenstates have correspondingly small contributions from the sterile neutrino to avoid conflict with existing neutrino data that can be fully described within the three-neutrino paradigm. We can make the two-flavour approximation to write the active-sterile oscillation probability as:

$$P_{\nu_\alpha \to \nu_\beta} = |\delta_{\alpha\beta} - \sin^2(2\theta_{\alpha\beta}) \sin^2 \left(1.27 \, \Delta m^2_{\text{SBL}}[\text{eV}^2] \, \frac{L[\text{m}]}{E[\text{MeV}]} \right)| \, , \tag{1.20}$$

where $\sin^2(2\theta_{\alpha\beta}) = 4|U_{\alpha 4}|^2|\delta_{\alpha\beta} - |U_{\alpha 4}|^2|$ and Δm^2_{SBL} is the eV-scale mass splitting required to explain the anomalies.

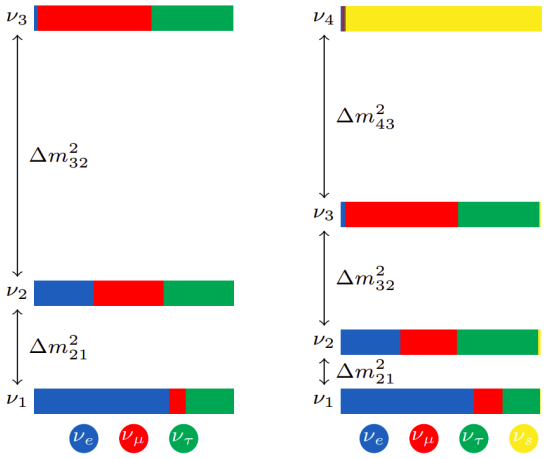


Figure 1.10: Left: Normal ordering of active neutrino masses. Right: Normal ordering of active neutrino masses with additional fourth mass eigenstate composed primarily of the sterile flavour eigenstate. The sterile neutrino is accessible in experiments if the active SM neutrinos have a non-zero ν_4 component. Figure from Ref. [163].

1.4.2 Detecting Sterile Neutrinos

Despite the lack of direct detection channels, there are numerous experimental tests for sterile neutrinos. The 3×3 PMNS matrix is unitary by definition. Evidence to the contrary would strongly suggest that it is a subset of a larger $(3+N)\times(3+N)$ unitary matrix, where N is the number of extra neutrino flavours. The current precision on the matrix elements means the level of non-unitarity must be small, i.e. $|U_{s\alpha}| \ll 1$. Disappearance searches for eV-scale sterile neutrinos can be undertaken using detectors placed at 1–10 m distances to reactor cores, as obligated by the $\mathcal{O}(1)$MeV antineutrino energy. At this L/E ratio, equation 1.20 is valid with $\alpha = e$ and $\sin^2(2\theta_{ee})$ of the form:

$$\sin^2(2\theta_{ee}) = 4|U_{e4}|^2(1-|U_{e4}|^2) = \sin^2(2\theta_{14})\,, \tag{1.21}$$

where $\sin^2(2\theta_{14})$ and the mass splitting $\Delta m^2_{\text{SBL}} = \Delta m^2_{41}$ form the orthogonal axes that define the parameter space of sterile neutrinos in the 3+1 model [164]. Reactor experiments with medium-to-long baselines must also include standard neutrino mixing when calculating the oscillation probability.

Accelerator and atmospheric neutrino experiments can analyse muon (anti)neutrino appearance and disappearance channels with equivalent forms of $\sin^2(2\theta_{e\mu})$ and $\sin^2(2\theta_{\mu\mu})$:

$$\sin^2(2\theta_{e\mu}) = 4|U_{e4}|^2|U_{\mu 4}|^2 = \sin^2(2\theta_{14})\sin^2(\theta_{24})\,, \tag{1.22}$$

$$\sin^2(2\theta_{\mu\mu}) = 4|U_{\mu 4}|^2(1-|U_{\mu 4}|^2) \simeq \sin^2(2\theta_{24})\,, \tag{1.23}$$

where the second equation is derived with a small-angle approximation [162]. Finally, the deflection of the tritium β-spectrum endpoint, which is sensitive to the weighted squared sum of the eigenvalues of *all* mass eigenstates, $\sum_i |V_{ei}|^2 m_i^2$, provides a powerful complementary probe.

1.4.3 Current Experimental Status

Interpreting the SBL anomalies in terms of sterile neutrinos already sets limits on sterile neutrino mixing parameters. The RAA and gallium anomaly were the catalysts for six VSBL reactor experiments searching for eV-scale sterile neutrinos. Four of these experiments (PROSPECT, STEREO, Neutrino-4, and SoLid) are based at HEU research reactors at which ^{235}U is responsible for over 99% of the emitted antineutrinos, whilst the remaining two (DANSS and NEOS) operate at LEU reactors that also burn other isotopes in non-negligible quantities. Almost all of the detectors feature segmented designs to isolate signals at specific distances and improve background rejection.

PROSPECT & STEREO

The PROSPECT experiment [165] uses a segmented detector filled with ^{6}Li-doped liquid scintillator to detector antineutrinos emerging from the High Flux Isotope Reactor (HFIR) at Oak Ridge National Laboratory (ORNL), Tennessee. HFIR is a HEU research reactor that operates at 85 $\mathrm{MW_{th}}$ thermal power. The detector covers a baseline of 6.7–9.2 m from the reactor core. As for all such reactor experiments, antineutrinos are detected by inverse beta decay. The active liquid scintillator volume is separated into 11×14 tubular segments with cross-sectional areas of 0.145 $\mathrm{m^2}$ that are aligned parallel to the reactor vessel wall. Signals are read out at both ends of each segment by PMTs. The experiment site has minimal overburden so the whole detector is nested within passive-shielding layers. An additional layer on the pool wall blocks γ-rays from the reactor.

STEREO is a reactor neutrino experiment with the same physics goals as PROSPECT, that is to search for neutrino oscillations at very-short baseline and measure the ^{235}U antineutrino spectrum. It is based at the ILL research reactor in Grenoble, France, whose nominal power of 58.3 $\mathrm{MW_{th}}$ is slightly lower than HFIR. For reasons of practicality, the STEREO detector is positioned such that its longitudinal axis lies at 17.9° to the line (of length 10.3 m) running from the centre of the detector to the centre of the reactor core, giving STEREO an effective baseline of 9.4–11.2 m. The detector volume is filled with gadolinium-loaded liquid scintillator and split into six identical cells. Collection of the scintillation light is achieved by eight PMTs above each cell. The detector is surrounded by multi-layered shielding, a gamma-catcher, and a water-Cherenkov muon veto to aid background rejection in a limited-overburden environment.

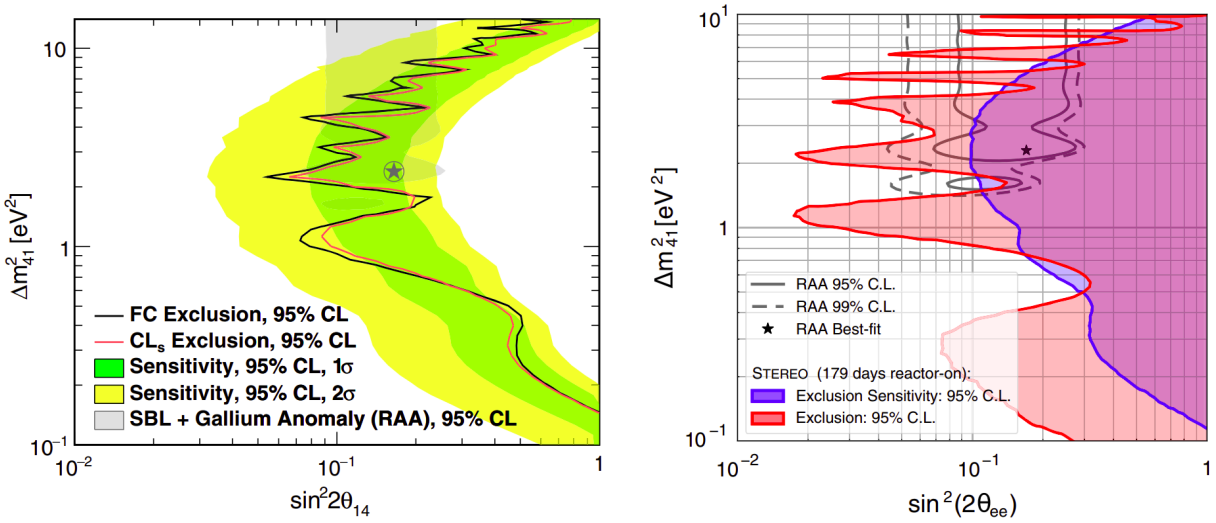


Figure 1.11: Latest searches for sterile neutrino (3+1 model) in the Δm^2_{41}-$\sin^2(2\theta_{14})$ plane by PROSPECT [166] (left) and STEREO [167] (right). Both experiments reject large portions of the RAA allowed region including the best-fit point at the 95% confidence level.

The latest sterile neutrino analyses from STEREO [167] and PROSPECT [166] are highlighted in Fig. 1.11. The RAA-gallium best-fit point is excluded at 2.5σ and over 3σ CL respectively and large portions of the allowed region are excluded. Both experiments achieve signal-to-background ratios (S/B) of $\mathcal{O}(1)$ for a few hundred IBD events per day (see Tab. 1.2). A recent improvement to the PROSPECT analysis [119] significantly increased their average signal-to-background ratio to 4.11, but the oscillation analysis is yet to be renewed. Lastly, the STEREO measurement of the absolute reactor flux confirms the presence of the RAA deficit for a pure ^{235}U source [117].

Neutrino-4

The Neutrino-4 detector is situated at the 90 $\mathrm{MW_{th}}$ SM-3 HEU reactor in Dimitrovgrad, Russia [168]. It is able to cover a baseline of 6.4–11.9 m by sliding horizontally on a rail-mounted platform. Antineutrinos are detected in a volume of gadolinium-loaded liquid scintillator that is divided into 10 × 5 vertical sections capped by PMTs. The entire detector is placed within a 60 t steel- and lead-plated "cabin" lined with 16 cm polyethylene sheets. Neutrino-4 is the only VSBL reactor neutrino experiment to report a significant (2.9σ) sterile neutrino signal [169], as shown in Fig. 1.12. The data is best fit by neutrino oscillations with parameters $\Delta m^2_{41} = 7.3 \pm 1.17\,\mathrm{eV}^2$ and $\sin^2(2\theta_{14}) = 0.36 \pm 0.12$. However, the Neutrino-4 result is in tension with some existing limits and it should be noted that the analysis strategy has been criticised, particularly with regards to possible systematic effects and the statistical approach that was taken [170, 171].

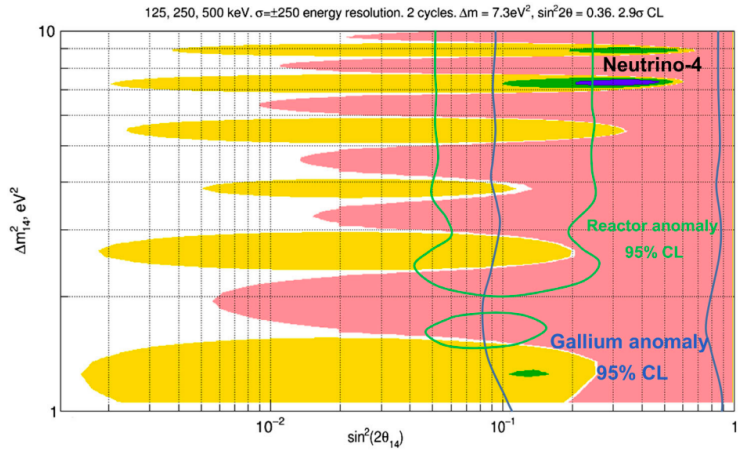


Figure 1.12: Results of the Neutrino-4 sterile-neutrino analysis [169]. Pink areas are excluded at more than 3σ CL. Areas filled in blue, green, and yellow represent 1σ, 2σ, and 3σ allowed regions respectively. An oscillation effect with best-fit parameters $\Delta m^2_{41} = 7.3 \pm 1.17\,\mathrm{eV}^2$ and $\sin^2(2\theta_{14}) = 0.36 \pm 0.12$ is observed at 2.9σ significance, which is in tension with many other measurements.

DANSS & NEOS

DANSS [172] and NEOS [173] are the two VSBL reactor experiments that operate at LEU commercial reactors. DANSS is based at the Kalinin nuclear power station in Russia, placed directly beneath one of the 3.1 GW reactors. This gives it excellent shielding from cosmic backgrounds and consequently it detects up to 5000 antineutrino events per day with a signal-to-background ratio of around 20. A mechanical lifting system can set the detector at one of three possible heights above the floor to measure antineutrinos at different distances from the core, in the 9.7–12.2 m range. The solid-state detector comprises 2500 strips of plastic scintillator sheathed in gadolinium-loaded covers within a shielded housing. Signals are collected by wavelength-shifting fibres and a combination of PMTs and silicon photomultipliers (SiPMs).

The NEOS detector is installed at Reactor Unit 5 of the Hanbit Nuclear Power Plant in Younggwang, South Korea. The reactor burns LEU fuel at 2.8 GW thermal power. The detector is made of a single homogeneous volume of gadolinium-loaded liquid scintillator. As such, it is only able to measure antineutrinos at a single distance of 23.7 m. An array of PMTs are deployed for photon collection and the detector is placed within passive shielding for background suppression. The latter is further enhanced by offline pulse shape discrimination (PSD) techniques, giving a signal-to-background ratio of around 22 for approximately 2000 antineutrino events per day.

Fig. 1.13 depicts the exclusion contours from DANSS [174] and a combined RENO-NEOS analysis [175]. Both experiments reject huge swathes of parameter space and the allowed RAA-gallium region at the 90% confidence level. The RAA best-fit point is excluded by

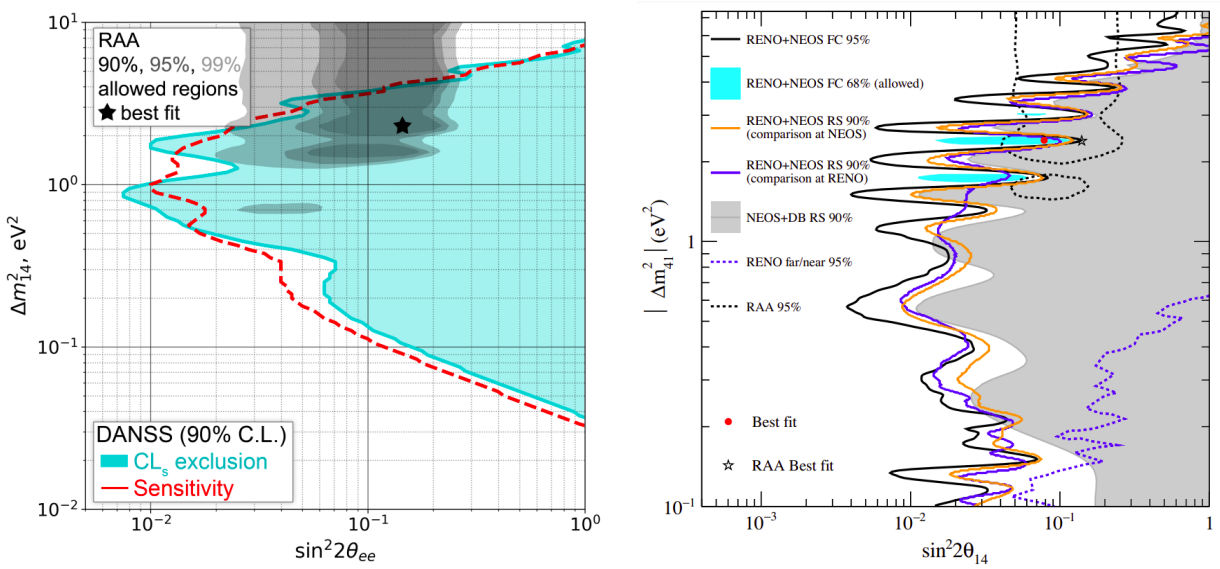


Figure 1.13: DANSS sterile neutrino search [174] (left) and results of a joint RENO-NEOS analysis [175] (right). The former finds no evidence of active-sterile neutrino disappearance and excludes Δm^2_{41} values down to 10^{-2}; the latter grants a one-sigma allowed region with a best-fit point located within the RAA contour, whilst simultaneously rejecting a region comparable to DANSS at 95% CL.

DANSS at greater than 5σ CL. The combined NEOS analysis with the RENO experiment—based near the same reactor—exhibits a 68% CL allowed region with a best-fit point at $\Delta m^2_{41} = 2.41\,\text{eV}^2$ and $\sin^2(2\theta_{14}) = 0.08$ [175].

Summary

The important parameters of the VSBL detectors are summarised in Tab. 1.2, adapted from Ref. [176]. The experiments at LEU reactors benefit from higher power and overburden than those at HEU reactors, thus they can detect far more IBD events per day with less background. On the other hand, the fissile cores of research reactors tend to be far more spatially compact which reduces the smearing of the oscillation pattern. They are also able to measure the pure ^{235}U spectrum with minimal fuel evolution during a cycle. The energy resolution and baseline depend on the detector design and position respectively. Better energy resolution improves the sensitivity to high-frequency oscillations. The experimental sensitivity is extended to lower $\sin^2(2\theta)$ values by parameters such as reactor power, signal-to-background ratio, and antineutrino efficiency. The Δm^2 range is determined by the L/E scale of the experiment, thus why these experiments must operate in such close proximity to the reactor core.

A summary of the VSBL sterile neutrino searches adapted from [128] is shown in Fig. 1.14, which also includes constraints from KATRIN, MicroBooNE, and solar neutrino experiments. Sterile neutrino searches from direct neutrino-mass measurements probe a

Attribute	PROSPECT	STEREO	Neutrino-4	DANSS	NEOS	SoLid
Baseline [m]	6.7–9.2	9.4–11.2	6.4–11.9	9.7–12.2	23.7	6.3–8.9
Overburden [m w.e.]	< 1	15	3.5	50	20	8
Reactor Power [MW_{th}]	85	58.3	100	3100	2800	65
Reactor Fuel	HEU	HEU	HEU	LEU	LEU	HEU
Core Diameter [cm]	51	40	42	320	310	50
Core Height [cm]	44	80	35	370	380	80
Energy Res. @ 1 MeV	4.5%	8%	16%	34%	5%	16%
Signal-to-Background	1.36	0.9	0.54	50	22	0.26
IBD events [d^{-1}]	750	400	200	5000	2000	105

Table 1.2: Quantitative summary of the current-generation VSBL reactor experiments. The sixth experiment, SoLid, is described in the next chapter.

slightly different mass-splitting region to reactor experiments. Most of the Neutrino-4 allowed region is in tension with the STEREO and PROSPECT exclusion contours, but it coincides very well with the allowed BEST region. The RAA and gallium regions are derived by considering both the rate and shape in the case of the former and only the rate in the latter. As such, the allowed regions extend unbounded towards high Δm^2 regions with roughly constant coverage in $\sin^2(2\theta)$. This is because the unresolvable high-frequency oscillations at large mass splitting result in a rate deficit that is independent of Δm^2. Conversely, the sterile neutrino disappearance searches by the dedicated VSBL reactor experiments are only sensitive to the spectrum shape so their exclusion regions are bounded in the absence of a signal.

A global fit of the reactor antineutrino data as of late 2019 can be found in Ref. [177]. DANSS and PROSPECT will probe a larger parameter-space area and fully scrutinise the Neutrino-4 allowed region after planned upgrades [178, 179]. The projected KATRIN sensitivity with the full dataset can investigate both the Neutrino-4 signal and the RAA region at high mass scales [180] that are less thoroughly explored. The final VSBL reactor experiment, SoLid, is the subject of this thesis and is described exhaustively in the next chapter.

1.4.4 Alternative Hypotheses

Analyses by Daya Bay and RENO on the correlation between fuel evolution and antineutrino flux have shed light on the isotopic origin of the anomalies. As mentioned in section 1.3.3, RENO have measured a 2.9σ non-zero correlation between fuel fraction and the proportion of 5 MeV bump events in the spectrum [122]. Simultaneously, they reported a non-zero

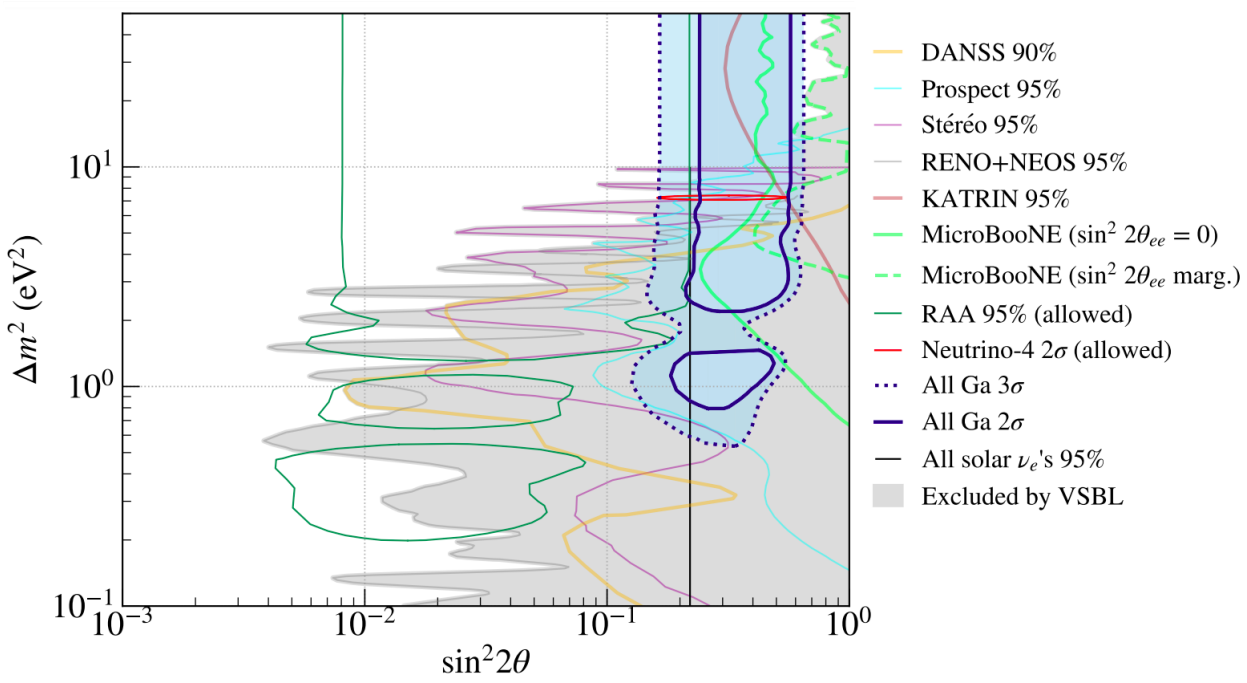


Figure 1.14: Summary of the sterile neutrino searches from various VSBL reactor experiments, KATRIN, MicroBooNE, and solar neutrino data. Figure adapted from Ref. [128]. The Neutrino-4 and BEST results are compatible with each other but in tension with most of the other experiments. The majority of the RAA-only allowed region (green) near the best-fit point is rejected at more than 90% confidence level.

dependence of the IBD yield on the ^{235}U fuel fraction at the 6.6σ level, consistent with the (scaled) HM model. The flux deficit is found to be almost solely caused by ^{235}U over ^{239}Pu, disfavouring the sterile neutrino hypothesis of the RAA which compels a deviation independent of the fissile fractions. Similar measurements from Daya Bay [181, 182] also attribute the rate deficit preferentially to ^{235}U, but suggest that altering either the ^{235}U prediction alone or the predictions of all four isotopes equally can solve the RAA.

Furthermore, comparisons between HEU and LEU spectra are able to probe the origin of the 5 MeV bump [121, 183]. Daya Bay and PROSPECT have combined their data to improve the extraction of the individual contributions of ^{235}U and ^{239}Pu from the Daya Bay spectrum [184, 185]. They observe shape discrepancies in the data compared to the HM model for both isotopes, though only ^{235}U is statistically significant at this stage. The results are consistent with all fissionable isotopes having an equal contribution to the bump [119].

The sophistication of reactor flux calculations has evolved immensely since the Huber-Mueller model. The summation model was updated in 2019 [102] (following [97]) to utilise the latest database information and nuclear decay measurements taken using Total Absorption Gamma-ray Spectroscopy (TAGS), which does not suffer from the Pandemonium effect. The impact of the latter is substantial. The resulting spectrum, known as the Estienne-Fallot

(EF) model, is in better agreement with the Daya Bay antineutrino flux and its evolution as a function of fuel fraction [182]. However, it fails to properly describe the spectral distortion. Recent investigations into the summation method in terms of a phenomenological model [186] and the first complete uncertainty evaluation of all known effects [187] are able to describe the experimental antineutrino spectral shapes and rates, but cannot reproduce the ILL β-spectra, suggesting biases in the reference data in both rate (for ^{235}U) and shape (for all isotopes).

The Hayen-Kostensalo-Severijns-Suhonen (HKSS) conversion method uses a nuclear shell model to demonstrate how a more accurate consideration of forbidden transitions can partially resolve the 5 MeV shape anomaly [188], though the size of the rate deficit is then slightly increased compared to the HM model. Moreover, it has been understood for a number of years that the effect of forbidden decays can induce biases in the spectral shapes and rates [189, 190] and the associated uncertainties—not taken into account in the original HM error budget—can decrease the significance of the RAA [105]. The underlying data on which the conversion method depends has also been reassessed. The authors of Ref. [191] measured the ratio between the β-spectra of ^{235}U and ^{239}Pu at the Kurchatov Institute (KI) research reactor, identifying a roughly constant shift across the entire energy range with respect to the ILL data. The KI ^{235}U normalisation is lower than the ILL spectrum by a factor of 1.054 ± 0.002, bringing the converted flux prediction closer to experimental reactor data and the EF model.

The four prevailing reactor models, HM, KI, HKSS, and EF, are rigorously compared in Ref. [109] in the context of the RAA. They also combine the KI data with the improved HKSS conversion method to form the HKSS-KI model. The work considers the overall normalisation of these models and their flux prediction as a function of fuel fraction, in order to test them against Daya Bay and RENO data. The statistical treatment of the spectra indicates that the significance of the rate deficits predicted by HM, KI, HKSS, EF, and HKSS-KI are 2.8σ, 1.4σ, 3.0σ, 0.8σ, and 1.8σ respectively. In other words, the RAA is heavily dependent on the model. Its solution and that of the 5 MeV bump may be fully identified in the future given additional Pandemonium-free data and more reliable nuclear databases, corrected cumulative β-spectra, or all of the above.

Chapter 2

SoLid Experiment

SoLid (Search for Oscillations with a Lithium-6 Detector) is a very-short-baseline neutrino experiment at SCK CEN (the Belgian Nuclear Research Centre), operating next to the Belgian Reactor 2 (BR2) reactor in Mol [192, 193]. The purpose of the experiment is to test the hypothesis that the anomalies described in the previous chapter can be attributed to active-sterile neutrino oscillations, by measuring the antineutrino energy spectrum as a function of baseline. Furthermore, SoLid aims to measure the pure uranium-235 antineutrino energy spectrum in order to evaluate the contribution of this isotope to the 5 MeV bump. The experiment differs from similarly motivated experiments in terms of detector technology, reactor site, and background discrimination. Therefore, SoLid is able to contribute uniquely complementary physics results. This chapter begins with an overview of the BR2 reactor as an antineutrino source, followed by a description of the Phase-I detector and its deployment at SCK CEN. Lastly, the reconstruction and calibration tools are detailed.

2.1 Antineutrino Source

BR2 is a powerful open-pool materials testing reactor that burns highly enriched uranium for nuclear fuel research and the irradiation of materials for medical and industrial purposes. A technical diagram and photograph of the reactor core are provided in Fig. 2.1 and Fig. 2.2 respectively. The aluminium vessel housing the reactor core is fully submerged in pressurised water for cooling and containment of the high neutron flux. Seventy-nine hollow channels of varying radii are set at inclined angles through a beryllium matrix, such that when the fuel elements are inserted the effective volume for nuclear fission is only 0.5 m in diameter and 0.8 m in height. A compact core is useful for an oscillation measurement because it reduces the positional uncertainty on the antineutrino creation vertex, which would otherwise smear potential oscillation effects. The channels can also accommodate control rods and material for irradiation. The beryllium matrix acts as both a moderator and reflector of neutrons.

The reactor is currently set up to burn highly enriched uranium (93.5% ^{235}U) at 40–

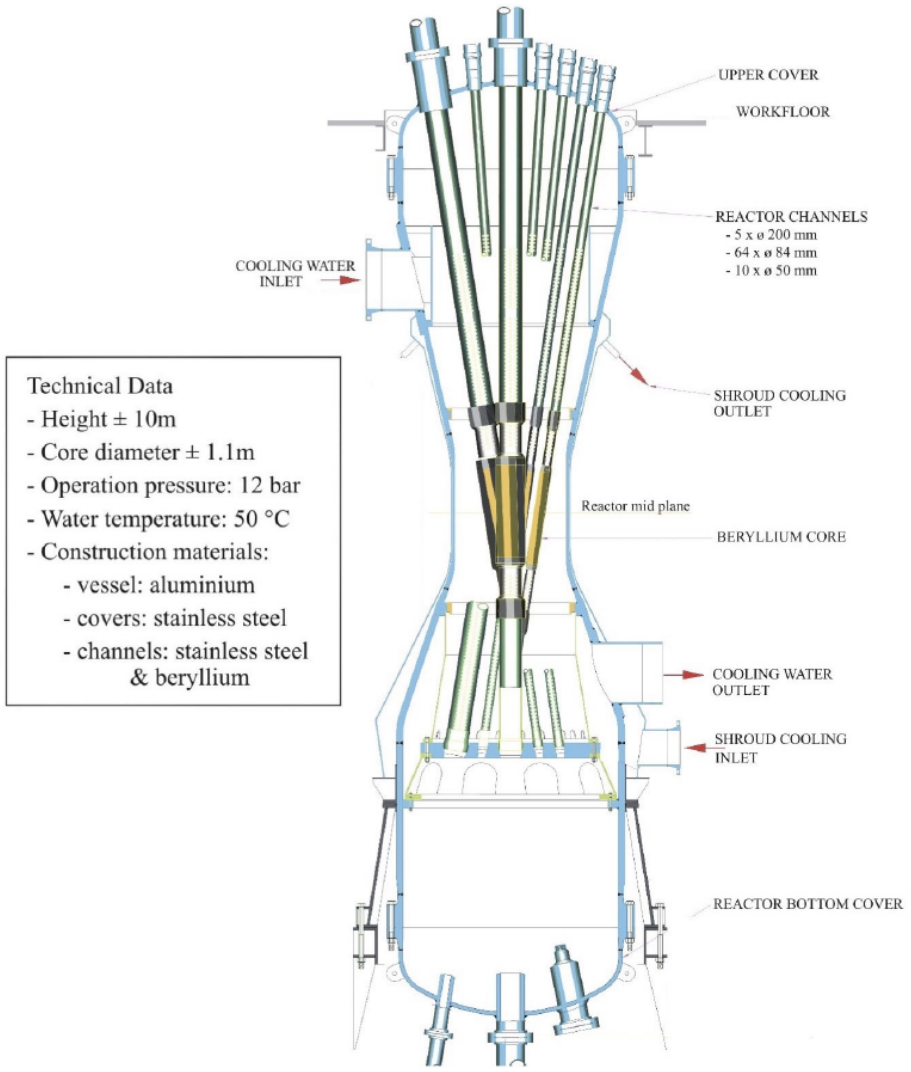


Figure 2.1: Schematic of the BR2 reactor vessel from [193] showing a subset of the inclined channels, the inlets and outlets for water, and the beryllium core. The twisted-channel design means that the effective fissile volume is only 0.5 m in diameter and 0.8 m in height.

100 $\mathrm{MW_{th}}$ thermal power, with a neutron flux on the order of $10^{15}\,\mathrm{cm^{-2}\,s^{-1}}$. The reactor runs at power in cycles of approximately 25–30 days, with periods in between of similar length for maintenance and other tests. These reactor-off periods are used by SoLid for precise background measurements or to perform detector calibration with sources. The SoLid collaboration is granted access to the BR2 Integrated Data Acquisition System for Survey and Experiments (BIDASSE) database which contains power measurements and other environmental data. The antineutrino flux of BR2—computed using equation 1.16—is approximately 2×10^{19} antineutrinos per second at maximum power.

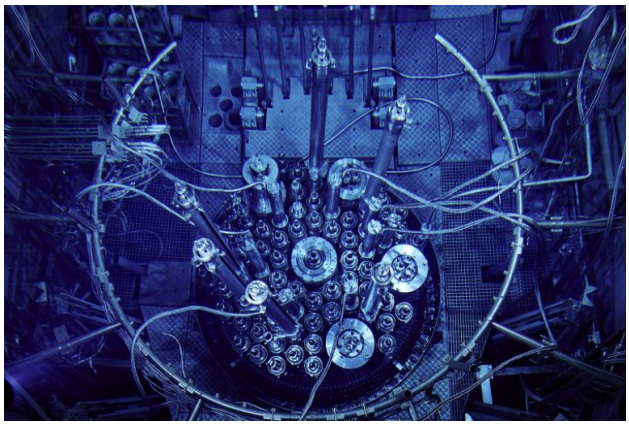


Figure 2.2: Photograph of the BR2 reactor core submerged in the water pool, as viewed from above. Figure courtesy of Ref. [194].

2.2 Phase-I Detector

2.2.1 Detection Principle

The SoLid detector measures antineutrinos via the inverse beta decay (IBD) reaction:

$$\bar{\nu}_e + p \rightarrow n + e^+ \quad \text{for } E_{\bar{\nu}_e} \geq 1.806 \text{ MeV}\,, \tag{2.1}$$

in which a proton (p) and electron antineutrino ($\bar{\nu}_e$) are converted to a positron (e^+) and a neutron (n). The 1.806 MeV threshold is governed by the relative masses of the constituent particles [195]. Inverse beta decay has the largest cross section of any antineutrino interaction with matter at the MeV scale. The antineutrino energy spectrum that is visible to a detector is the product of the true reactor antineutrino flux and the cross section as a function of energy. The tree-level IBD Feynman diagram, in which the process occurs through W boson exchange, is drawn in Fig. 2.3 alongside the true and measurable antineutrino spectra.

The SoLid detector is designed to produce observable signals from both final state particles, separated by a characteristic time delay as per the Cowan-Reines coincidence technique [9]. The available energy of the antineutrino upon destruction is shared between the positron and neutron. However, the mass difference between them is such that the neutron receives less than 1% of the available energy in around 90% of cases, thus the antineutrino energy spectrum information is propagated almost entirely with the positron. It is therefore essential that the energy of the positron be accurately reconstructed. Moreover, the positron spectrum should be measured as a function of position in the detector since oscillation phenomena also depend on the antineutrino path length. The neutron signal is used for triggering on IBD candidates so the detector must maximise neutron detection efficiency and purity as far as possible.

The basic detection unit is a composite object formed of a 5 cm polyvinyl toluene (PVT)

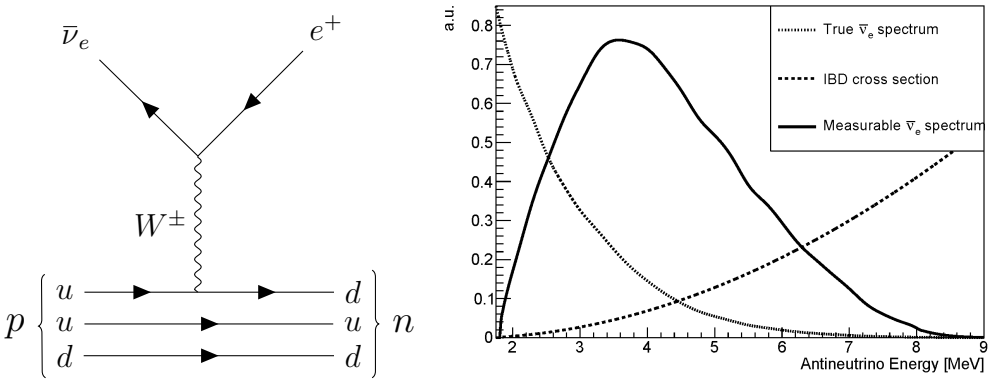


Figure 2.3: Left: Tree-level Feynman diagram of inverse beta decay. Right: Geant4 simulation of the measurable reactor antineutrino energy spectrum (solid line) given by the convolution of the true spectrum (dotted line) and the IBD cross section (dashed line). Note that the apparent "wiggles" in the spectra are statistical fluctuations in the simulation.

cube capped on two sides by $0.25 \times 50 \times 50\,\mathrm{mm}^3$ ^{6}LiF:ZnS(Ag) screens, optically isolated within a 270 µm thick reflective polymer wrapping. A diagram illustrating the main components is provided in Fig. 2.4. PVT is an organic plastic scintillator that acts as proton target for the antineutrino, primary scintillation material for the positron, and moderator for the neutron. Antineutrinos interact within the PVT volume of the detector where the number density of hydrogen atoms is $5.17 \times 10^{16}\,\mathrm{m}^{-3}$. The ^{6}LiF:ZnS(Ag) screens provide a surface for capture of thermalised neutrons on ^{6}Li and scintillation of the resulting daughter nuclei. The screens are placed on the negative z-axis and positive x-axis faces of the cube, with respect to the coordinate system defined in Fig. 2.4. The opacity of the polymer wrapping (DuPont Tyvek, type 1082D [196]) is over 95% which mitigates the amount of light shared between adjacent cubes.

2.2.2 Prompt Signal

The brief passage of the positron over $\mathcal{O}(20)$ mm distances causes fluorescence in the PVT material, emitted promptly after the IBD reaction. The particular PVT scintillator used by SoLid (Eljen EJ-200 [198]) emits approximately 10 000 photons per one MeV electron with a light output of 64% anthracene. The emission spectrum is in the 400–500 nm range and peaks at 425 nm, as per Fig. 2.5, a wavelength to which the material is highly transparent with an attenuation length of 380 cm. A narrow pulse shape is created by the short rise and decay times of 0.9 ns and 2.1 ns respectively.

A major benefit of using a solid plastic scintillator over a liquid scintillator is that the dependence of the light yield on initial electron energy is highly linear over the majority of relevant energy range, as plotted by the manufacturers in Fig. 2.5. The energy calibration is simplified by the fact that non-linearities introduced by quenching effects only become

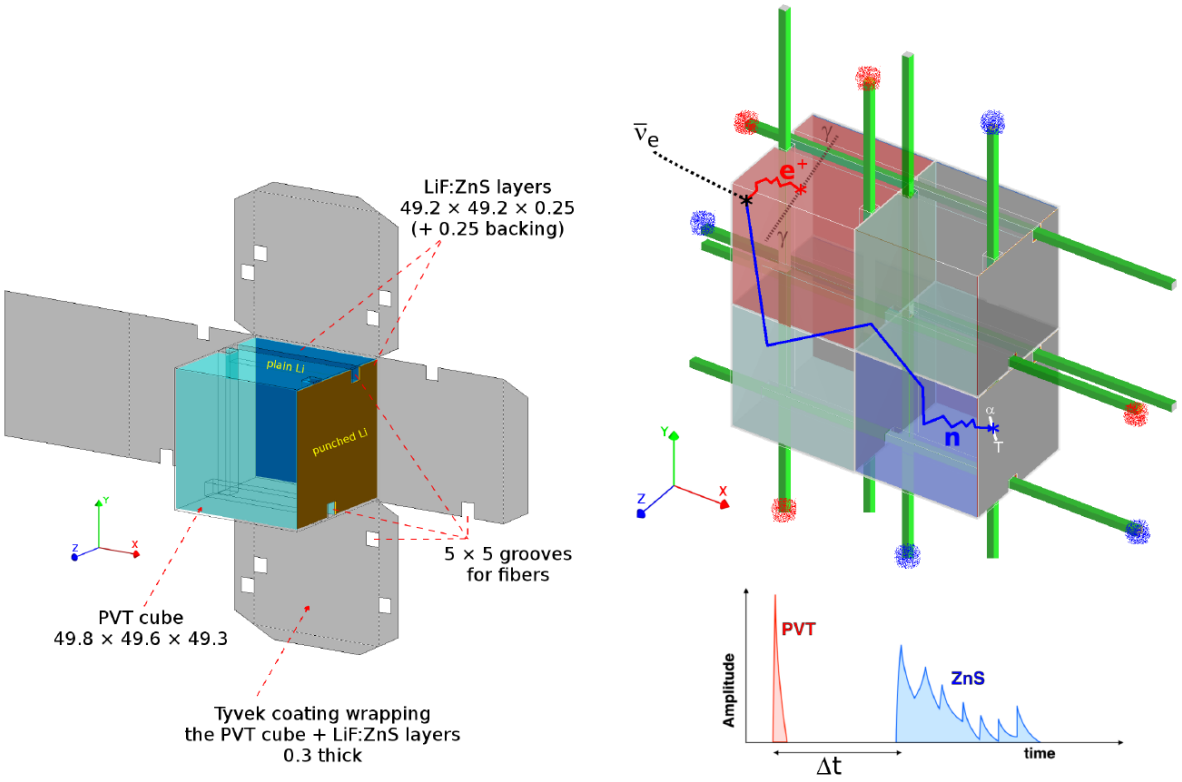


Figure 2.4: SoLid detector voxel and antineutrino detection principle [197]. All dimensions are stated in millimetres. Left: PVT cube with ^{6}LiF:ZnS(Ag) screens in place, in front of (open) Tyvek net. Right: Possible paths of the final state particles through the detector, alongside characteristic shapes of signals produced by each scintillator. The sharp PVT signal is easily distinguishable from the more extended ZnS signal which consists of numerous local peaks over a slowly decaying pulse.

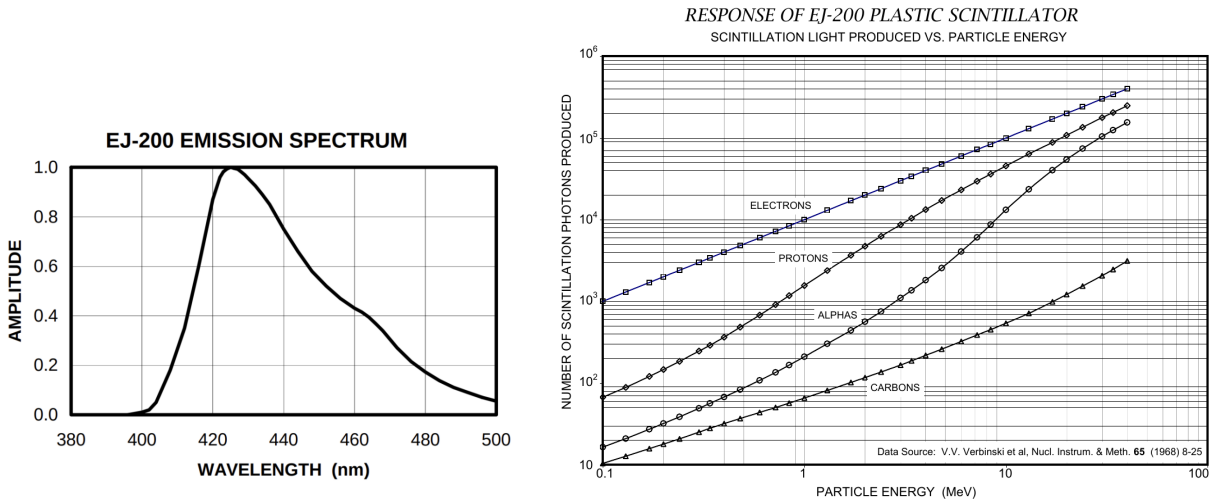


Figure 2.5: Eljen EJ-200 emission spectrum (left) and scintillator response as a function of initial particle energy over three orders of magnitude (right). The number of photons produced by the scintillator is highly linear with energy, particularly for electrons. Both plots provided by the manufacturer [198].

relevant below $\mathcal{O}(100)$ keV. On the other hand, unlike standard liquid scintillators the EJ-200 does not exhibit neutron-gamma PSD properties, which is an extremely useful feature for distinguishing between proton recoil and electron or photon signals.

The second component of the positron-induced signal comes from electron-positron annihilation to two gamma rays ($e^+e^- \rightarrow \gamma\gamma$). If this occurs at rest, the gamma rays are emitted in opposite directions each with an energy of 511 keV. The annihilation gammas interact predominantly via Compton scattering and leave relatively low energy deposits in the detector. The overall prompt signal, hereafter denoted as the Electromagnetic Signal (ES), is the sum of both positron and annihilation gamma deposits but the detector segmentation means that they can be effectively separated. The gamma deposits are a key topological feature of IBD events that is used to differentiate signal from background in offline analysis.

2.2.3 Delayed Signal

The IBD neutron is moderated to thermal energies (~ 0.025 eV) by elastic scatterings on nuclei in the detector. The cross section for neutron capture on ^{6}Li at these energies is around 940 b, over 3000 times higher than the cross section on hydrogen [199], and results in the following reaction:

$$^6\mathrm{Li} + \mathrm{n} \rightarrow {}^3_2\mathrm{He} + \alpha\,, \tag{2.2}$$

in which the tritium ($^3_2$He) and alpha (α) nuclei share a total energy of 4.78 MeV. These highly ionising nuclei cause fluorescence in the ZnS(Ag) at a peak emission wavelength of

450 nm. This is denoted as the Nuclear Signal (NS). The NS is delayed with respect to the ES because of the time taken for the neutron to thermalise prior to capture. Monte Carlo (MC) simulations indicate that the mean time for neutron capture to occur in the SoLid detector is 62 µs, a quantity that depends on the properties of the moderation material and the configuration and capture efficiency of the neutron detection screens.

The ^{6}LiF:ZnS(Ag) neutron detection screens produced by Scintacor [200] are a mixture of lithium fluoride, ^{6}LiF, and silver-activated zinc sulfide, ZnS(Ag), powder set in epoxy resin (250 µm). They are backed by reflective polyester substrates (also 250 µm) to improve light collection. The ^{6}Li enrichment of the lithium fluoride is 95%. The probability of capturing a thermal neutron in one crossing of a screen is 24%. However, the mean free path of a thermal neutron in PVT is around 7 cm, so the capture probability is significantly enhanced by multiple crossings. The decay time of the ZnS(Ag) scintillator, $\mathcal{O}(10)$ µs, is significantly longer than that of the PVT, resulting in an NS waveform that is extended in time compared to the ES. In addition, differences in shape between the ES and NS (the latter consisting of numerous local pulses emitted over the duration of its decay) permit the use of PSD techniques to distinguish between the two. Typical ES and NS waveform shapes are drawn in the lower right part of Fig. 2.4.

2.2.4 Photon Collection

The scintillation photons are collected by double-clad wavelength-shifting (WS) fibres, St Gobain BCF-91A [201], with a cross-sectional area of $3 \times 3\,\text{mm}^2$. Four fibres are threaded through $5 \times 5\,\text{mm}^2$ slots milled into each cube perpendicular to the z-axis. The fibres shift the absorbed photon spectrum from blue to green with an emission peak close to 495 nm. Square holes are punched through the neutron screens and Tyvek wrapping as necessary to accommodate the fibres. The holes in the Tyvek mean a fraction of the scintillation light can escape to neighbouring cubes, causing an effect known as light-leakage (LL) where energy deposited in a single cube is spread across multiple cubes. The fibres are capped at one end by reflective Mylar foil and by a SiPM at the other. The photosensitive area of the SiPM is equal to the cross-sectional area of the WS fibre and coupling between the two is facilitated using optical gel. The SiPMs are housed in 3D-printed sockets to keep the fibres in place and protect the delicate electronics. A schematic of the light collection components is given in Fig. 2.6.

The Hamamatsu S12572-050P Multi-Pixel Photon Counter (MPPC), i.e. SiPM, is an array of 3600 pixels on a silicon substrate. Each pixel is an avalanche photodiode (APD) operated in Geiger mode—under reverse bias greater than its breakdown voltage*—in conjunction with a quenching resistor. A photon incident on the depletion zone of the APD

*The breakdown voltage is the minimum voltage required for the charge carriers to acquire sufficient energy to liberate secondary charge carrier pairs and create a self-sustaining pixel avalanche.

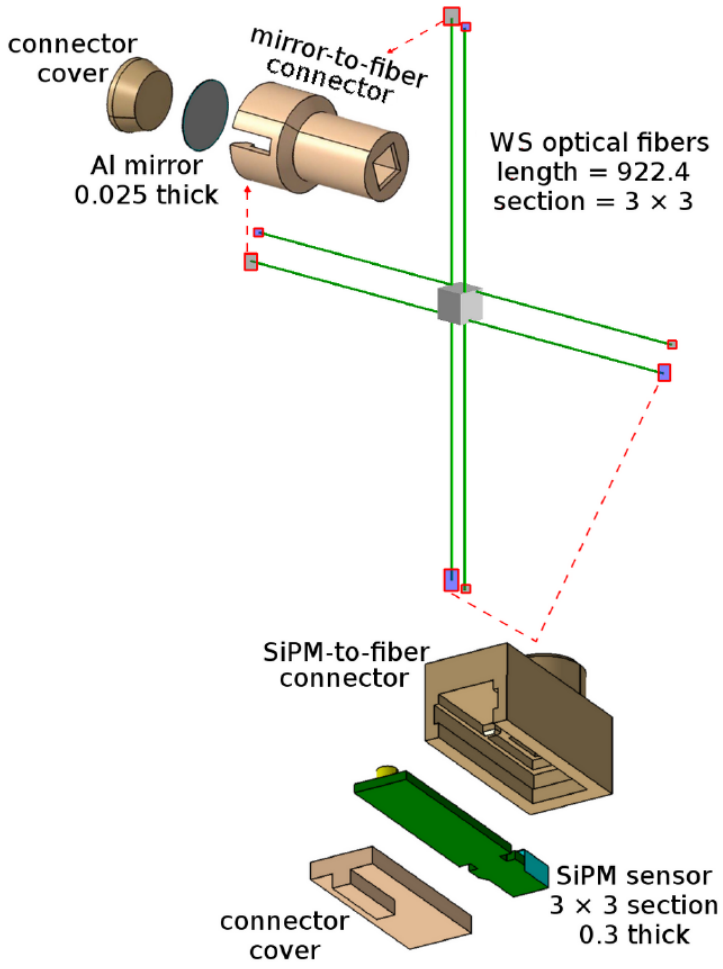


Figure 2.6: Exploded view of the detector components responsible for photon transport and collection in the SoLid detector, including WS fibres, MPPCs and 3D-printed connectors. All dimensions are stated in millimetres. Figure adapted from [193]. The positions of the Mylar foils (mirrors) and MPPCs alternate between planes to mitigate the effect of signal attenuation in the fibres.

p-n junction may liberate an electron-hole pair which is then accelerated by the applied electric field. If the breakdown voltage is surpassed by the applied voltage, the resulting pixel avalanche (PA)[†] generates a growing current which can be detected. The diode is reset by quenching the current with a resistor, yielding a quickly rising and slowly decaying pulse. The difference between the applied voltage and the breakdown voltage is known as the overvoltage (OV).

In this manner, an APD in Geiger mode is capable of counting single photons, but loses all sensitivity to the energy of the incident photons and their multiplicity in the case of simultaneous arrival (pile-up). An MPPC runs many pixels in parallel to recover the proportionality between the measured signal and the number of incident photons, though this relation is lost for extremely high photon rates. The total output of an MPPC at a particular instance is the superposition of signals from all pixels. The peak photon detection efficiency (PDE)—the probability that an incident photon is detected—of the S12572-050P sensor is 35% at 450 nm (25 °C and 2.6 V overvoltage) and it has a fast 250 ps single-photon timing resolution.

2.2.5 Full-Scale Detector

The full-size detector is designed to be highly modular, divided into multiple sub-systems that are capable of operating independently. The smallest such system is a detector plane comprising an array of 16×16 detection cubes and 64 readout channels (WS fibre + MPPC). The $0.8 \times 0.8\,\mathrm{m}^2$ arrays are encased in white high-density polyethylene (HDPE) sealing bars—roughly 46–47 mm thick—to reflect neutrons back into the active detector volume. The HDPE-bound array of cubes is supported by an aluminium frame that also shields the MPPCs and associated cables from external electromagnetic interference. Tyvek sheets are affixed to either side of the frames to negate the sharing of photons between planes. The front-end electronics are contained within an aluminium box attached to the side of the frame. The sixty-four MPPCs per plane are connected to the electronics enclosure via twisted-pair ribbon cables running to an interface plate. Planes are bunched into groups of ten to form independent modules and the full detector is built from five modules. The fiducial mass of the detector, comprising the 12800 PVT cubes, is approximately 1.6 t. Fig. 2.7 depicts the structural components that comprise the individual planes and modules.

[†]The amplitude of an MPPC signal is described in units of PA, i.e. the number of pixel avalanches it comprises (see Tab. 2.1).

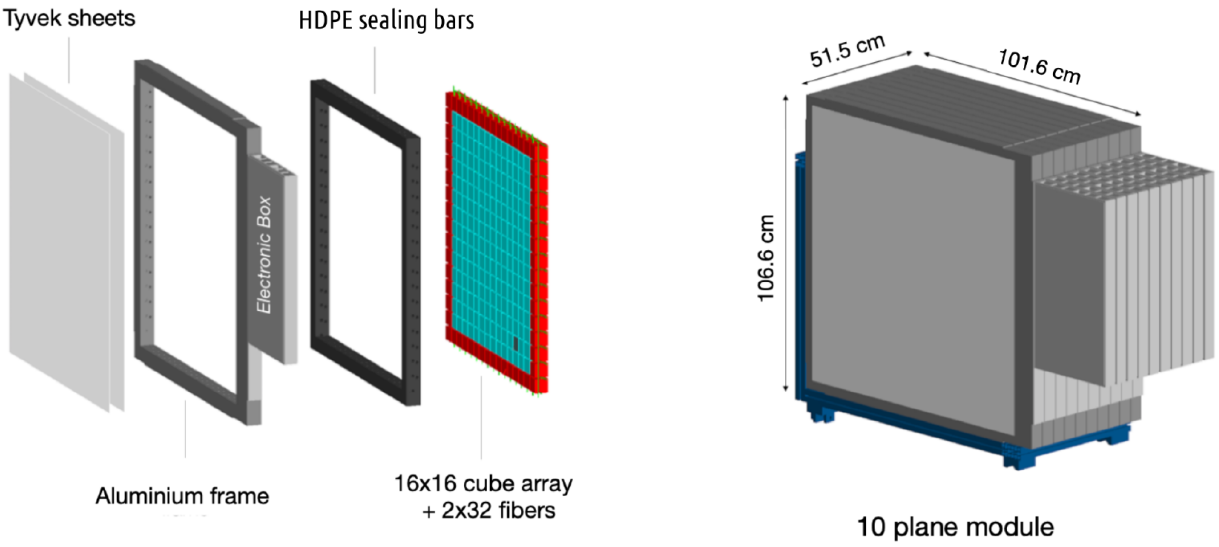


Figure 2.7: Left: Exploded view of a single detector plane. Right: A detector module composed of ten individual planes with electronics enclosures attached. The blue parts of the drawing depict the trolley upon which the module is installed. Diagrams from [193].

2.2.6 Deployment at BR2

The Phase-I detector was successfully commissioned at BR2 by spring 2018 and the period of stable data-taking spans April 2018 to July 2020. It is deployed on level 3 of the BR2 containment building inside a $2.4 \times 2.6 \times 3.8\,\mathrm{m}^3$ thermally insulated, lightproof shipping container. The detector is situated very close to the reactor pool, covering a baseline of 6.300–8.938 m from the centre of the reactor core, as displayed in Figs. 2.8 and 2.9. The planes are oriented perpendicular to the z-axis, which runs from the centre of the core through the lengthwise detector axis.

The expected antineutrino yield in the active detector volume can be obtained by modifying equation 1.16 to take into account the antineutrino cross section and detector acceptance. The total number of antineutrinos interacting in voxel i is given by:

$$N_i = \int_E \int_t \int_V \underbrace{\frac{P_{\mathrm{th}}(t)}{\sum_k \left[\phi_k(\mathbf{x},t)\,\langle E_{\mathrm{fiss}}\rangle_k\right]}}_{(1)} \underbrace{\sum_k \left[\phi_k(\mathbf{x},t)\,S_k(E)\right]}_{(2)} \underbrace{\frac{A_i}{4\pi L_i(\mathbf{x})^2}}_{(3)} \underbrace{n_i\,\sigma_{\mathrm{IBD}}(E)}_{(4)}\ \mathrm{d}\mathbf{x}^3\,\mathrm{d}t\,\mathrm{d}E\,, \tag{2.3}$$

where n_i, A_i and $L_i(\mathbf{x})$ are the number of proton targets, cross-sectional area and distance between fission and IBD vertices respectively of voxel i. $\sigma_{\mathrm{IBD}}(E)$ is the energy-dependent IBD cross section. All other terms are the same as in equation 1.16. The factor (3) represents the reduction in flux due to the geometric detector acceptance and part (4) accounts for the antineutrino interaction probability as it passes through the voxel. Summing over all voxels results in an expected antineutrino yield of around 1200 events per day at nominal power

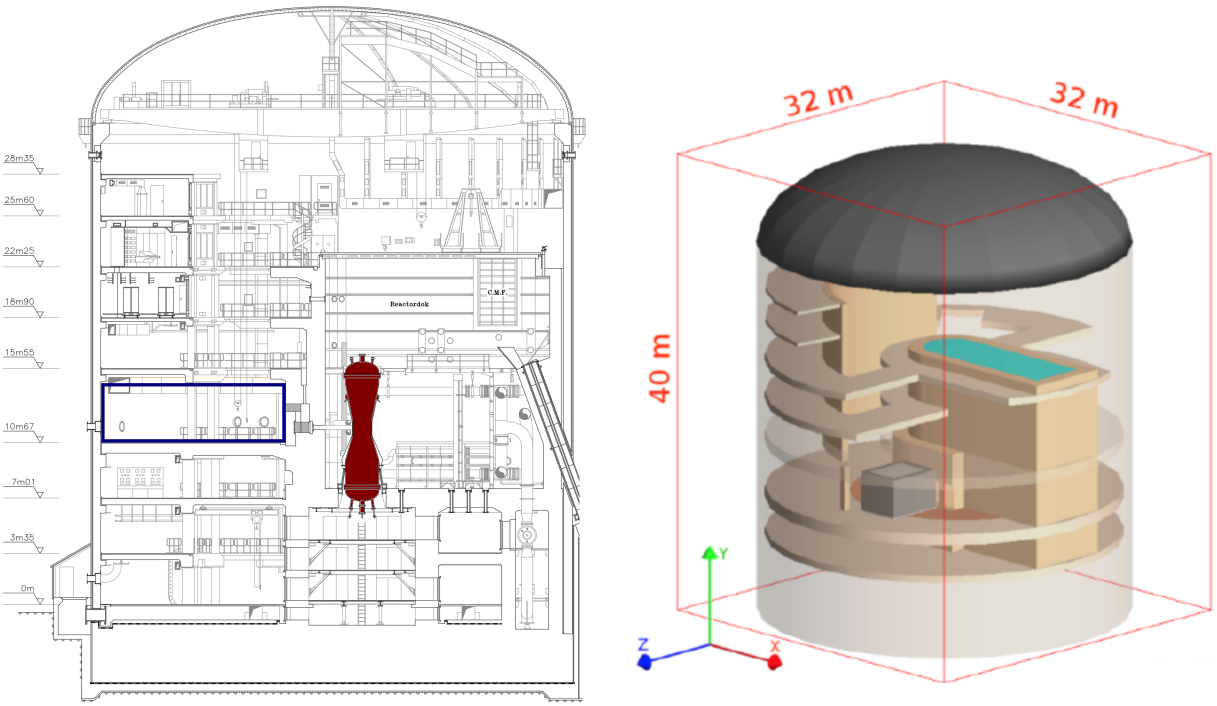


Figure 2.8: Left: Blueprint of the BR2 containment building, adapted from [202]. The reactor vessel is depicted in maroon and the level 3 floor in navy. Right: Visualisation of the BR2 containment building in Geant4 with the SoLid container in place [193].

of 65 MW_{th}.

The container is cooled by a chiller system circulating a 1:4 mixture of propylene glycol and water. The default chiller setting is 5 °C, but the heat generated by the electronics results in an ambient container temperature of around 11 °C. This reduces the MPPC dark-count rates by a factor of three compared to room temperature operation. A dedicated robot, known as CROSS, is installed above the modules for in-situ calibration of the detector using radioactive sources (see section 2.5). The modules slide along horizontal rails to facilitate the placing of radioactive sources in the gaps between them.

The experiment is located above ground with limited overburden from the building itself, meaning there is a significant background rate from particles of atmospheric origin. To mitigate this background, a 50 cm thick passive shield consisting of overlapping HDPE slabs is fitted atop the container, increasing the total overburden to 8 m w.e. In addition, a 50 cm water wall built of separately filled crates surrounds the container to protect it from reactor-induced neutron and gamma background. Finally, a large proportion of the external container area is covered by 2 cm cadmium sheets to block neutrons that emerge from the HDPE shielding. The passive shielding reduces the total atmospheric neutron flux by an order of magnitude. Additional layers of HDPE (9 cm thick) are placed at either end of the detector for shielding and neutron reflection. In terms of reactor-induced background, SoLid benefits from being the only active experiment on level 3 so all beam ports, which

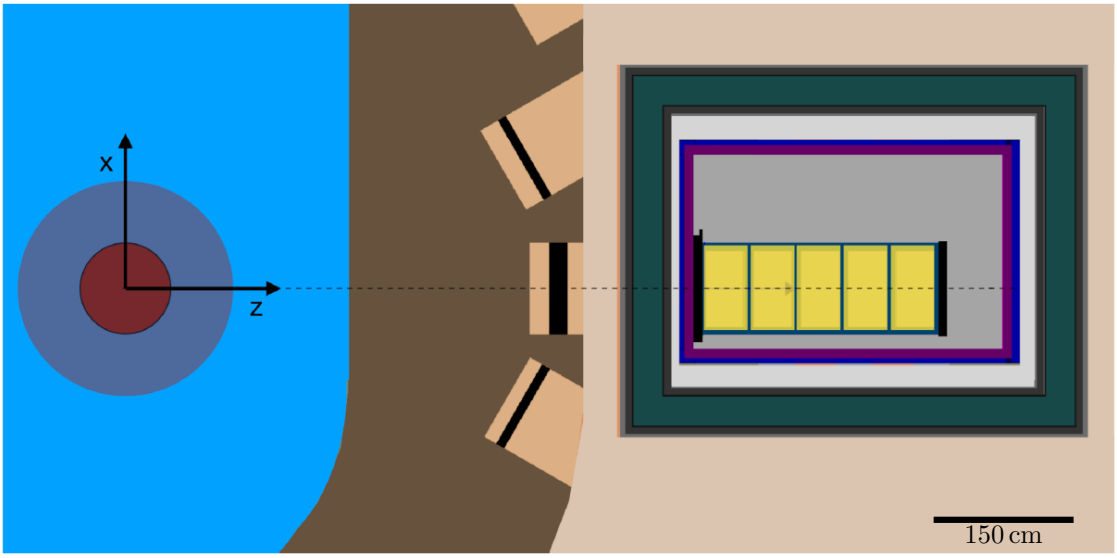


Figure 2.9: Geant4 rendering of the detector at BR2 in relation to the reactor core, adapted from [193]. The five-module detector (yellow) can be seen within the container surrounded by the insulation (purple) and water wall (green). The longitudinal axis of the detector is aligned with the centre of the reactor core (red).

would otherwise contribute to the neutron background, are closed and shielded.

2.3 Data Acquisition and Reconstruction

The conversion of MPPC signals to usable data is a multi-stage process involving communication between hardware, firmware, and software. The data acquisition (DAQ), or *readout*, system combines all of these elements to parse the electronic waveforms, digitising and filtering them before saving the data to disk. In short, the analogue signals from the MPPCs are digitised by sampling the raw waveforms at a particular rate. A so-called "trigger" decision is then made on each digital waveform; the waveform data is only added to the output stream if it satisfies the conditions of at least one trigger. The data from all channels are combined in a single binary file.

The output of the DAQ must be reconstructed before it can be meaningfully used in a physics analysis. Event building and reconstruction is done in a C++ software package called SoLid Analysis Framework (Saffron2). Saffron2 is able to read the binary files, parse the raw waveforms, and group them according to the type of physics event from which they originated. It achieves this using a number of dedicated algorithms. Finally, the event reconstruction is concluded by converting the measured waveforms to localised energy deposits within the detector. Extensive detail is provided in the rest of this section.

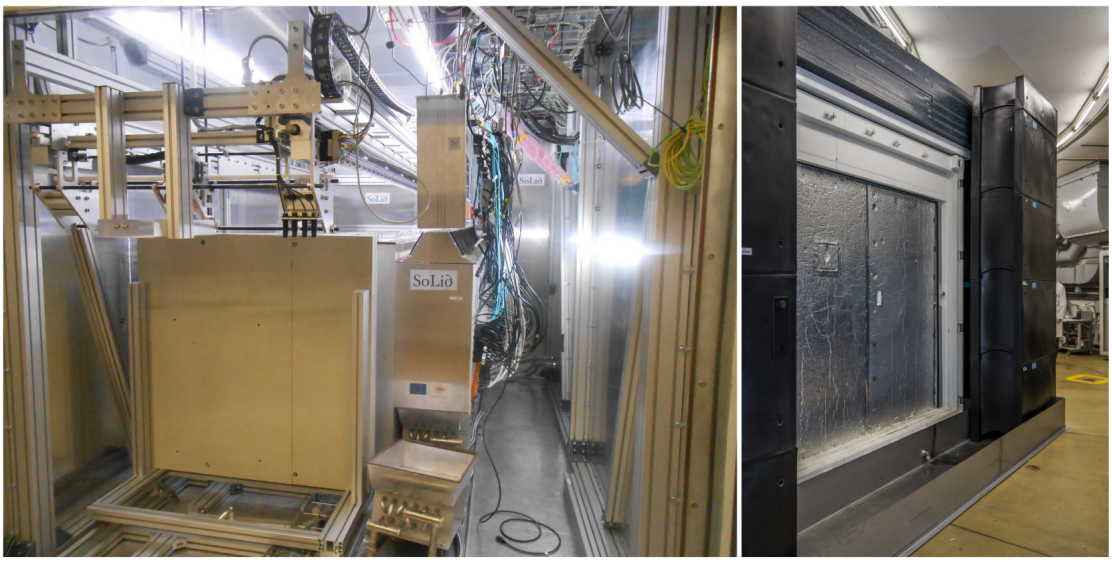


Figure 2.10: Photographs of four SoLid detector modules installed within the shipping container (left) and back of the detector with thermal insulation, water bricks and HDPE slabs in view (right). Images courtesy of the SoLid collaboration.

2.3.1 Front-End Electronics

The electronics that control a given plane consist of two 32-channel analogue boards, one 64-channel digital board, and other boards for power distribution and monitoring of environmental sensors. The analogue boards supply the power and bias voltage to the MPPCs. The bias voltage is applied via a 70 V "high voltage" supply, common to the 32 board channels, plus a second 0–4 V "trim voltage" supply which is tuned per channel in order to equalise the MPPC responses. The MPPC pulses are passed to the digital board after being amplified and extended to $\mathcal{O}(200)$ ns, in order to make them comparable in duration to the digital board sampling rate. The digital board uses eight 8-channel 14-bit LTM9007 analogue-to-digital converter (ADC) devices that sample the raw waveforms at a rate of 40 MHz.

Plane-level trigger and readout logic is programmed onto an Xilinx Artix-7 (XC7A200) field-programmable gate array (FPGA) [203]. The board is designed such that a plane can be run according to either an internal or external clock. Each module has its own DC-DC converter, heat exchanger, clock board, and MinnowBoard computer for power supply, cooling, synchronisation of the planes, and programming of the digital electronics respectively. Finally, a master clock board synchronises the modules to allow them to operate as one detector. The electronics and hardware items are rendered as CAD images in Fig. 2.11, alongside a comparative photograph.

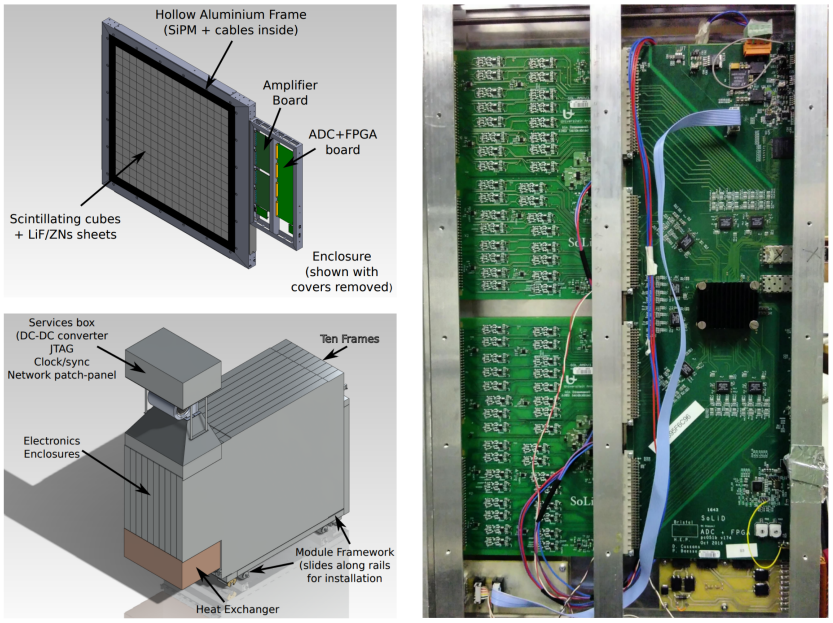


Figure 2.11: Clockwise from top left: CAD representation of a detector plane showing attached (open) electronics enclosure [204]; photograph of an (open) electronics enclosure with visible analogue and digital boards; CAD representation of all hardware associated with a module [204].

2.3.2 Readout and Triggers

The FPGAs control the stream of ADC samples and perform online triggering on digitised waveforms, acting as an intermediary between the ADCs and the readout software that runs on a local server at SCK CEN. The use of triggers for real-time data selection substantially reduces the amount of data saved to disk because data is only recorded once a trigger condition has been fulfilled. Zero-suppression (ZS) is used to further reduce the overall data rate by around a factor of 50, by removing pedestal[‡] and single-PA contributions which contain little useful information, but the ZS threshold can be relaxed if a trigger is fired.

To simplify the trigger approach, decisions are made on a *block* basis rather than on individual samples; a block is defined as 256 consecutive samples (6.4 µs). Trigger decisions are made separately for each channel, but once a trigger is fired the data of the entire plane is read out in order to include all correlated signals. The trigger signal can also be propagated to neighbouring planes. The FPGAs contain per-channel history buffers that enable a number of preceding and succeeding blocks to be read out, thus extending the readout region in time as well as space. The data from all channels is concatenated in a plane-level buffer read out by the DAQ software and written to a binary file for offline event building and analysis.

Normal data-taking (physics mode) is separated into 480 s runs. At the beginning of a run, the readout software resets the boards by re-configuring the firmware registers and

[‡]The pedestal of a SiPM/MPPC is the size of the signal measured in the absence of photons.

Trigger name	Condition	Space-time region read out	ZS thresh.	Trigger rate (Hz)	Data rate (MB/s)
Random	Periodic	All planes, 2 blocks	None	1.2	3.9
Threshold	50 PA, x-y coincidence	Triggered plane, 2 blocks	1.5 PA	2100	2
Neutron	$\geq$ 17 local maxima above 0.5 PA within 1 block	Triggered plane $\pm$3 planes, [-80, 30] blocks	0.5 PA	80	15

Table 2.1: Summary of the parameters that define the SoLid trigger strategy. All three triggers are active during standard physics data-taking. Information from Ref. [204]. Note that the unit "PA" (number of pixel avalanches) describes the amplitude of an MPPC signal.

clearing the buffers. Whilst this occurs, the boards are in so-called "dead-time" as they do not receive any data from the MPPCs. Overflow of the FPGA buffers can also cause dead-time to accumulate during a run. Channel dead-time occurs when a history buffer reaches maximum capacity, at which point further data flow from that channel is suspended until the board is reset at the beginning of the next run. If the main data buffer overflows, plane dead-time is incurred until the buffer is emptied and data flow can resume. Highly energetic muons are the most likely cause of buffer overflow.

There are three types of trigger implemented in the FPGA firmware that are in use simultaneously during normal physics mode, as follows:

- The **random trigger** reads the output of every detector channel for a period of 512 samples (2 blocks, $\sim 36\,\mu s$) at a rate of 1.2 Hz. This allows us to continuously monitor MPPC-related parameters and detector dead-time. Two non-zero-suppressed blocks per channel are read out: the triggered block and the block preceding it.

- The **threshold trigger** is intended for background measurements by detecting high-energy particles passing through the detector volume, such as muons, that generate signals greater than 50 PA (approximately 2 MeV). The deposited energy must be measured by at least one fibre in the horizontal and vertical directions. Two ZS blocks per channel are read out for the entire triggered plane: the triggered block and the block following it.

- The **neutron trigger** is designed to identify the NS signal waveform, indirectly acting as a trigger for antineutrino events. Upon firing, the trigger signal is transmitted to the nearest three planes on either side of the triggered plane (seven in total), which all read out a time window of [-80, 30] blocks ([-512, 192] µs) either side of the triggered

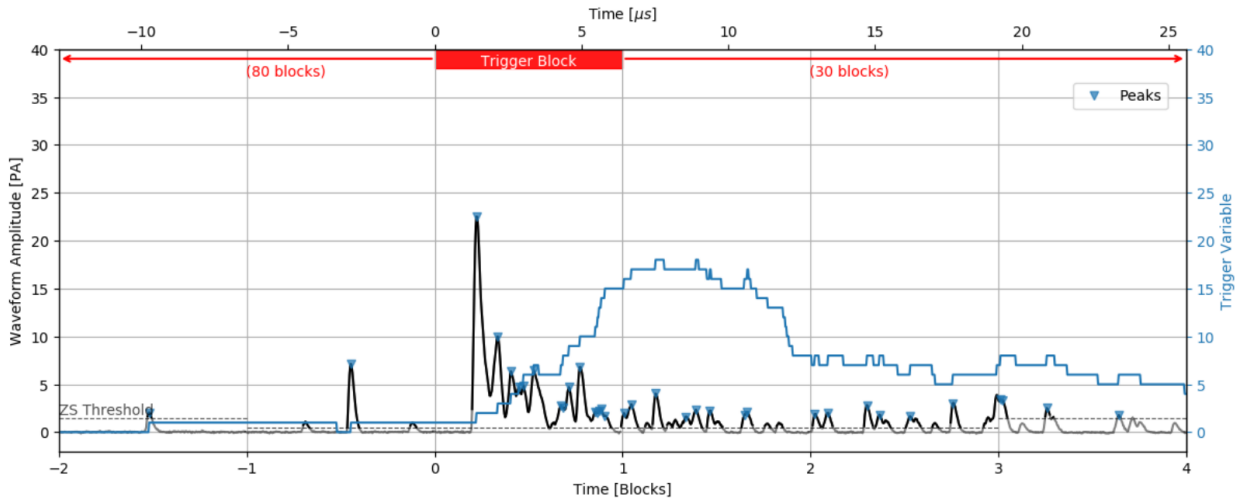


Figure 2.12: Example waveform that fires the NS trigger, adapted from [205]. The trigger conditions are satisfied in the block indicated by the red rectangle, causing the read out of the preceding 80 blocks and subsequent 30 blocks. The zero-suppression threshold is represented as a black dotted line.

block. The size of the readout region is driven by insights from simulation and is sufficiently large in space and time to include all components of the IBD signal, as well as contributions from background events prior to the ES and after the NS. The ZS threshold is lowered to 0.5 PA for four consecutive blocks (including the triggered block) to capture the distinctive low amplitude peaks of the NS waveform.

The neutron trigger algorithm is parameterised by three numbers: the minimum amplitude threshold for a local waveform to be counted as a "peak" (T_{thr}), the number of peaks within a rolling time window (N_{peaks}), and the width of the rolling window in units of ADC samples. The rolling window is fixed at 256 samples, or one block, in width but during commissioning a scan over the remaining parameters—and overvoltage OV—was performed using an AmBe neutron source. The chosen parameters ($N_{\text{peaks}} = 17$, $T_{\text{thr}} = 0.5$ PA, OV = 1.8 V) were designed for maximum trigger efficiency given the data storage constraints and correspond to 75% efficiency and 20% purity for reactor-on data-taking periods. A summary of all three trigger parameters is provided in Tab. 2.1 and an example waveform that fires the NS trigger can be seen in Fig. 2.12. Alternative trigger setups are adopted during calibration runs with sources.

2.3.3 Time Clustering

The starting point of the event reconstruction is a time-ordered list of waveforms, typically spread across numerous detector channels and planes. Time-correlated waveforms are grouped to form "clusters" according to a dedicated algorithm. Each waveform is considered

in turn, resulting in one of three possible outcomes:

1. The waveform is below the minimum amplitude threshold (2.5 PA) or is shorter than two samples and is discarded.

2. The waveform is added to an existing cluster if it meets the necessary conditions.

3. The waveform cannot be added to an existing cluster so becomes the first waveform of a new cluster.

The default minimum amplitude threshold avoids the majority of thermal-noise contributions. Any new cluster is immediately extended by scanning over all other waveforms whose start time falls within ± 7 samples of the initial waveform in that cluster. A waveform within this time window is added to the cluster if it is located in the same plane as the initial waveform and surpasses the amplitude and sample-length thresholds. The necessary conditions for adding a waveform to an existing cluster are:

- The waveform channel is already included in the cluster.

- The waveform time falls within the allowed adding window.

- The integral of the waveform divided by the number of samples is not greater than the previous waveform (i.e. it is not more likely to be the start of a subsequent cluster than the end of the current one).

The adding window is defined as the time difference between the start of the first waveform in the cluster and the end of the current last waveform in the cluster (the cluster length), plus a so-called *inertia window*. The size of the inertia window varies dynamically depending on the current cluster length and number of associated channels, such that the total adding window is long enough to capture all the peaks in the NS tail. A new cluster is discarded if it does not contain waveforms from at least two channels, or added to an existing cluster if the following conditions are met:

- The start times of the new and existing clusters are within 7 samples of each other.

- The new and existing clusters share at least one fibre and the starting times are within 21 samples of each other.

- The times of the waveforms of maximum amplitude of the new and existing clusters are within 7 samples, and the cluster start times are within 21 samples of each other.

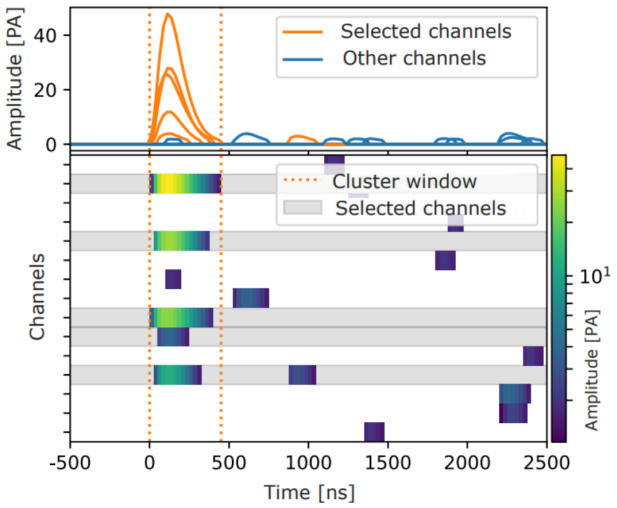


Figure 2.13: Visualisation of the time-clustering algorithm applied to an ES waveform [206]. The dotted lines denote the boundaries of the cluster window within which over-threshold waveforms are added to the cluster. The algorithm captures almost all of the initial high-amplitude signals and rejects most of those at low amplitude.

Thus a cluster may span $\mathcal{O}(100)$ µs in length and multiple planes. The parameters have been tuned to ensure that the algorithm isolates correlated signals from the same physics event with minimal overlap between events. An illustration of the time-clustering algorithm for an electromagnetic signal is provided in Fig. 2.13. Finally, each cluster is assigned a list of cubes based on the channels that were activated by the algorithm; a cube is added to the list if it is intersected by a WS fibre that belongs to a channel in the cluster. Clusters for which no cubes can be assigned (i.e. channels were only activated in a single direction) are discarded.

2.3.4 Cluster Categorisation

The clusters are tagged as one of three distinct types broadly aligned with the physical origins of the clusters: ES, NS, and muon. There is a separate algorithm within Saffron2 for each category.

Muon Clusters

Muons are easily identified because they generate high-energy scintillation in the PVT that is measured by a large number of channels, leaving a distinctive track. They can also decay in the detector to produce Michel electrons which leave similar tracks. A cluster is tagged by the muon algorithm if it contains more than 11 channels over 200 ADC counts in amplitude in both the horizontal and vertical directions. Two linear fits are performed on the x-z and y-z projections of the fibre signals to obtain a coherent track. Muon clusters are further

categorised depending on the results of the fits:

- **Type 0 (T0):** No fit convergence for either projection. Typically muons that clip the detector edge, depositing energy in a small number of cubes.
- **Type 1 (T1):** No fit convergence for either one of the projections, or the track is contained entirely within a single plane so the incoming direction is ambiguous.
- **Type 2 (T2):** Fit convergence for both projections. The track crosses multiple planes so the incoming direction is unambiguous.

ES and NS Clusters

The remaining clusters after muon tagging are then separated into the ES or NS class according to a pulse-shape integral over amplitude (IoA) discrimination technique. Iterating over the clusters, the NS algorithm immediately rejects a cluster if it has already been tagged as a muon, or if it is shorter than 1000 samples (25 µs) in length to account for the long decay time of ZnS scintillation signals. If multiple cubes can be identified in the cluster, the one containing the signal with the highest number of peaks over threshold is selected. The other cubes are re-assessed by the time clustering algorithm, which returns to the NS algorithm any additional clusters that are identified. This is relevant for neutron source runs where multiple NS signals can be generated in quick succession.

For the chosen cube, the mean waveform amplitude per channel and the mean channel integral (over a short 30 sample window) are used to define a boundary between NS and non-NS clusters, as follows:

$$\overline{I} > 10.5 \times \overline{A} - 800\,, \tag{2.4}$$

where $\overline{I}$ is the mean integral and $\overline{A}$ the mean amplitude of the waveform. The corresponding separation between NS and ES clusters is displayed in Fig. 2.14. Furthermore, a selection for avoiding large horizontal-vertical asymmetries within a cube is applied, as detailed in chapter 2.7.2.1 of Ref. [207]. The ES algorithm simply checks that a given cluster is not muon, not NS, and that it is shorter than 75 samples (1.875 µs), in which case the cluster is tagged as ES. A cluster that does not fit into any of the categories is discarded. The selection criteria enhance the purity of NS waveforms to over 99% from the original 20% NS trigger purity.

To begin forming IBD events, NS-ES coincidences are constructed by associating all ES clusters within ±600 µs of an NS cluster to that NS. The time difference between the triggered NS block and the triggered ES block is expressed as $\Delta t \equiv t_{\mathrm{NS}} - t_{\mathrm{ES}}$, so IBD events only exist in the positive Δt region. Note that this means multiple ES clusters can be assigned to a given NS cluster. The ES clusters that become part of NS-ES coincidences

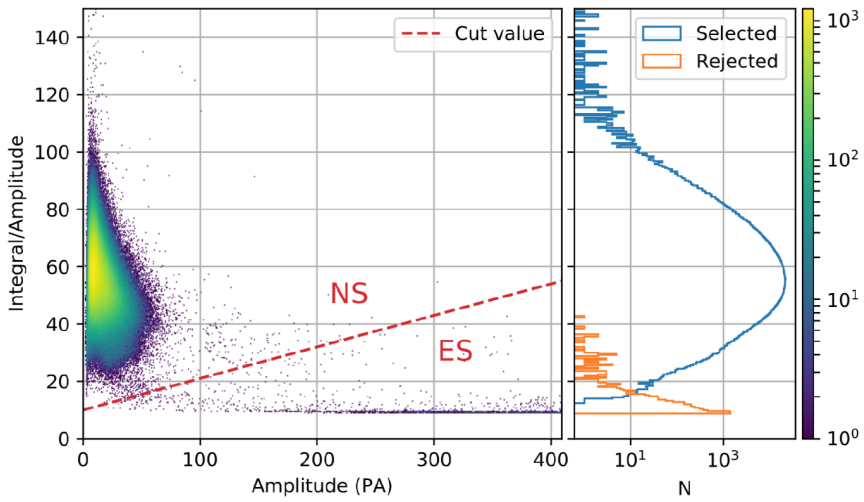


Figure 2.14: Integral-over-amplitude separation of non-muon events by the NS algorithm [193]. The red dashed line denotes the boundary between NS and ES events, which gives a purity of neutron waveforms of over 99%.

are required to have a total energy greater than 1 MeV, which drastically reduces both background rate and file size. The NS-ES coincidences are the primary objects that pass to the next stage of the antineutrino analysis.

2.3.5 CCube Algorithm

The MPPC signals of ES clusters must be transformed back into the original energy deposits per cube in order for a physics measurement to be made. The optical isolation of the planes and absence of longitudinal light-leakage mean that this can be done independently per plane. Hence, transformation of the signals is a two-dimensional linear inversion problem which can be posed mathematically as:

$$\boldsymbol{AE} \equiv \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,256} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,256} \\ \vdots & \vdots & \ddots & \vdots \\ a_{64,1} & a_{64,2} & \cdots & a_{64,256} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ \cdots \\ E_{256} \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_{64} \end{pmatrix} \equiv \boldsymbol{p}\,, \tag{2.5}$$

where $\boldsymbol{A}$ is the 64×256 matrix known as the System Matrix (SMat), $\boldsymbol{E}$ is the 256-element column vector of unknown cube energies in the plane, and $\boldsymbol{p}$ is the 64-element column vector of MPPC signals. The problem is also shown diagrammatically in Fig. 2.15. The element $a_{i,j}$ of the SMat acts as the projector of the energy deposit of cube j to the signal in MPPC channel i. The SMat elements encode all the information related to the conversion of an energy deposit to a digitised signal, including generation of scintillation light, photon attenuation, fibre-MPPC coupling, light-leakage, and PA-to-MeV conversion. In other words, the

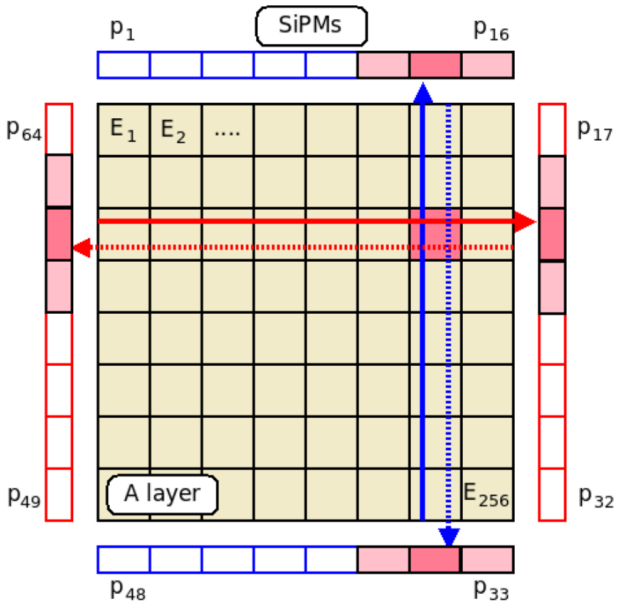


Figure 2.15: CCube algorithm problem for an energy deposit localised to a single cube. In such cases, signals in at least one fibre in each direction allow the deposit to be identified unambiguously. Figure courtesy of H. Chanal.

system matrix is the vital reconstruction object that parameterises the detector response; its derivation is described in section 2.5.3.

If multiple cubes are involved in the cluster then the set of $\boldsymbol{E}$ matrix elements that satisfy equation 2.5 is not unique, so an optimisation algorithm is required to find the most likely distribution of cube deposits for a given list of MPPC signals. SoLid employs a simplified version of the Orthogonal Matching Pursuit (OMP) algorithm [208], in conjunction with a Maximum-Likelihood Expectation-Minimisation (MLEM) procedure [209]. The approach taken, dubbed the CCube[§] algorithm is to pass the output of the OMP algorithm as the initialiser (Bayesian prior) of the MLEM algorithm, thereby greatly improving the overall accuracy of the reconstruction.

The CCube algorithm optimises two parameters: the ghost rate and cube reconstruction efficiency. The latter is defined as the total number of reconstructed cubes that correspond to a true energy deposit, divided by the total number of true deposits. The ghost rate is defined as number of reconstructed cubes that do not correspond to a true energy deposit, divided by the total number of reconstructed cubes. Low-energy cubes are less likely to be reconstructed and are more likely to be ghosts, which becomes important when studying the similarly low-energy annihilation gamma deposits.

[§]Named after Clermont-Ferrand, the location of the institute where it was developed.

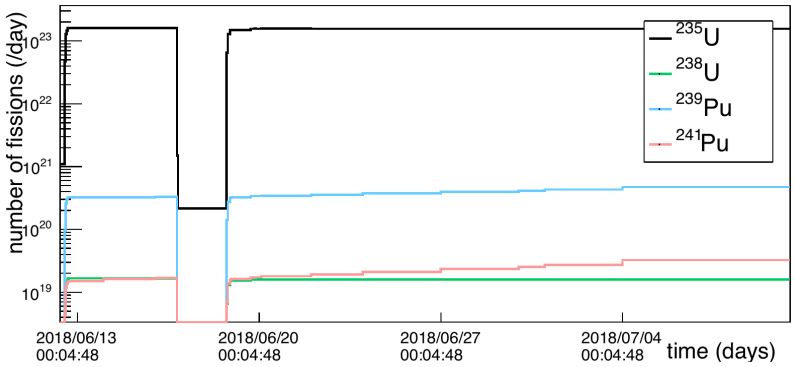


Figure 2.16: Predicted number of fissions per day for a 2018 BR2 reactor cycle. The fission reactions are overwhelmingly dominated by ^{235}U. Figure from [103].

2.4 Simulations

The two main simulations that are crucial to the success of the SoLid experiment are the BR2 reactor antineutrino generator and the Geant4-based detector simulation. The former is embedded in a C++ package known as the SoLid Oscillations (SoLO) code which acts as the interface between the two. As the name suggests, SoLO also contains tools for an oscillation analysis. An overview of the important details of each simulation is provided in this section.

2.4.1 Reactor Antineutrino Generator

Simulation of fission events in the BR2 reactor core and subsequent antineutrino emission relies on multiple software packages. A geometrically accurate representation of the reactor core in the Monte Carlo N-Particle eXtended (MCNPX) package has been created by SCK CEN staff, which is linked with MCNPX-CINDER90 code to compute the fission rates in the core as a function of time. The MCNP Utility for Reactor Evolution (MURE) package is used to track the burn-up of nuclear material over time and correct for off-equilibrium effects. The simulation allows for different configurations of control rods, fuel elements, and materials for irradiation according to a fuel-loading map. It is then coupled with a converted β-spectrum for each fissile isotope—the HM model, for example—to produce an antineutrino spectrum prediction. Summation method predictions are obtained entirely with MURE using the beta decay branches of each isotope from nuclear databases. Fig. 2.16 shows the energy spectrum and number of fissions per day for a randomly chosen 2018 BR2 reactor cycle. The contribution to the antineutrino spectrum from ^{238}U and the plutonium isotopes (mainly created in the decay chain of ^{238}U) is very small, as ^{235}U antineutrinos make up over 99% of the total flux.

Equation 2.3 is implemented in SoLO to generate a distribution of IBD vertices in the detector and nearby proton-rich materials, each with an an associated antineutrino energy.

Random numbers are generated to select three parameters: a fission vertex position in the core (based on the fuel-loading map), an antineutrino energy (according to the computed spectrum), and an interaction vertex position (weighted by the inverse of the distance squared from the fission vertex). The default reactor spectrum used in the simulation is the Huber-Mueller model. Individual reactor cycles are simulated with a factor 100 increase in events compared to the true prediction in order to suppress statistical fluctuations in the output. It has been established that the only significant difference between the simulations of different cycles is the predicted number of events; the antineutrino spectrum shapes differ negligibly [103].

2.4.2 Detector Model

The SoLid Simulation (SoLidSim) package takes the distribution of IBD vertices from the previous step and simulates the transport and interactions of the resulting IBD daughter particles in the detector in Geant4 [210]. The simulation incorporates all materials that could affect the propagation of particles in and around the detector. Structures outside the detector volume such as the HDPE shielding, CROSS robot, shipping container, and BR2 containment building are all included. The level of fine detail in the simulation is maximal nearest the detector and drops off with increasing distance. The simulation of the detector itself comprises the bulk PVT volume and aluminium frames, plus smaller components such as the ^{6}LiF:ZnS(Ag) screens, wavelength-shifting fibres, etc. which are all meticulously replicated. SoLidSim also has the capacity to run without an input vertex map to generate background samples that do not rely on the reactor simulation, for example.

The second part of the detector simulation is implemented within Saffron2. This is the Readout Simulation (ROSim) that transforms the simulated energy deposits into simulated electronic waveforms. ROSim must generate a scintillation response (i.e. a number of photons) according to the measured light yield of the given voxel and modulate the signal as it is passes through the optical components of the detector. The number of photons incident on the MPPC is fluctuated according to Poisson statistics. Moreover, it is dependent on Birk's law of quenching, photon scattering in the PVT, attenuation in the WS fibres, and the coupling between the fibres and the MPPCs. In practise, these readout effects are modelled using the same system matrices that are used to calibrate and reconstruct the ES energy deposits, operated in reverse. Utilising the SMats in this manner is highly convenient as the effects are automatically parameterised during the computation of the matrix.

The final step is to model the electronic response to the received signals. The incident photons are stretched over a certain time distribution depending on whether they originated from scintillation in PVT or ZnS, plus the time travelled in the fibre en route to the MPPC. Pixel avalanches are then created from the incident photons which generate waveforms

in turn. The waveforms are amplified by 31.5 ADC/PA (see section 2.5.1), smeared, and sampled at 40 MHz frequency. Other effects such as photon detection efficiency, pedestal noise, cross-talk, dark counts, and overflowing buffers are also taken into account. Free parameters in the simulation are either sourced from the manufacturer data sheets or tuned in dedicated data-MC studies. The trigger strategy in ROSim is identical to the triggers applied for data. The final output is a set of predicted waveforms for the particular input model applied that can be treated in the exact same way as data and passed to Saffron2 for reconstruction.

2.5 Calibration

The total number of cubes and readout channels in the SoLid Phase-I detector is 12800 and 3200 respectively, corresponding to tens of thousands of parameters to be characterised. The specific details of the numerous detector calibration procedures developed by the collaboration lie beyond the scope of this thesis. Nonetheless, it is important for the overall picture that the main aspects are covered, especially the ES energy calibration and the derivation of the system matrices using cosmic muons.

2.5.1 MPPC Equalisation

The natural variation in breakdown voltages of the MPPCs means that the applied bias voltage necessary to achieve a given overvoltage is different for each MPPC. During detector commissioning it was observed that the gains of the MPPCs varied significantly even under the same overvoltage, defining gain as the difference in ADC samples between the MPPC pedestal and the 1 PA peak, as displayed in the left-hand side of Fig. 2.17. The bias voltages are set such that all channels apply equal gain to a received signal, so waveforms of equal amplitude from different channels are measured as having the same number of pixel avalanches.

The gain of every channel is set to 31.5 ADC/PA in physics mode with an average OV of 1.8 V, as a compromise between three parameters: PDE (32%), dark-count rate (110 kHz), and cross-talk probability (20%). The variation of dark-count rate and cross-talk as a function of overvoltage measured during commissioning is shown in the right-hand side of Fig. 2.17. The overvoltage settings required for a constant gain across the detector may be found by iterating through all channels over a reasonable range of trim voltages, a process that was repeated every few months to account for the drift in breakdown voltage over time. Similarly, the MPPC pedestals were periodically re-measured during data-taking.

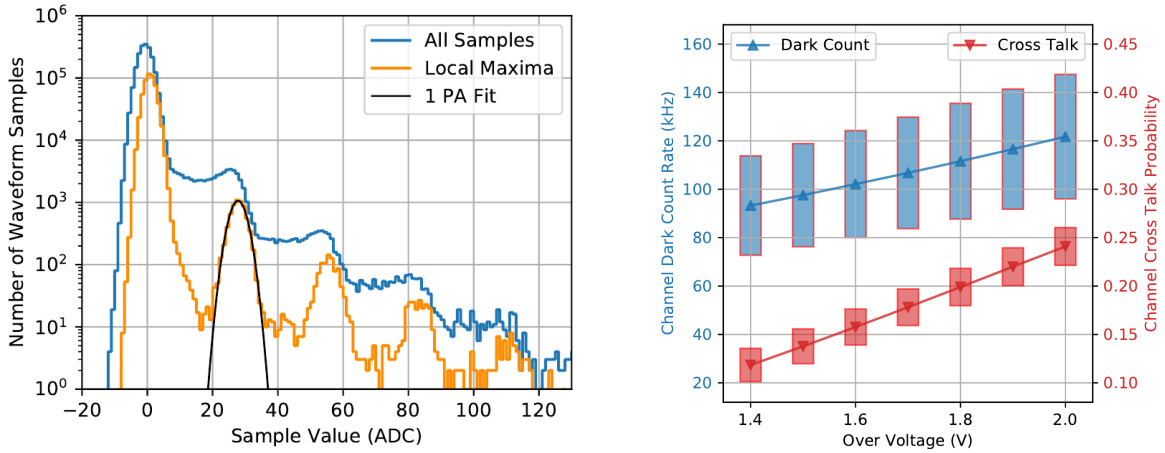


Figure 2.17: Left: Determination of an MPPC gain, defined as the mean of the Gaussian fit on the 1 PA peak after subtracting pedestal contributions. Right: Dark-count rate and cross-talk probability as a function of overvoltage for all detection channels. The rectangles indicate the standard deviation over all the channels. Figures from [204].

2.5.2 In-Situ Calibration

The remote-controlled calibration robot, CROSS, is mounted to the container ceiling above the detector from where it controls the module positions along the rails. It can open a gap between two specified modules and lower a source holder to one of nine fixed positions in the x-y plane, as illustrated in Fig. 2.18. Including the gaps at the front of the first plane and behind the last plane, there are a total of 54 possible source placements covering the entire detector volume. When not in use, the holder can be accessed through a small window in the back of the container for the insertion or withdrawal of sources.

A standard calibration campaign involves moving the holder—containing a radioactive source—through all 54 CROSS points and taking data at each for a few minutes or up to an hour depending on the source activity. Calibration campaigns were undertaken every few months over the two years of Phase-I data-taking, but the most extensive campaign was in September 2018 when the detector was exposed to a wide range of gamma and neutron sources over a week-long period. The September 2018 campaign serves as the reference for all other calibrations.

Neutron Sources

The neutron detection efficiency is directly related to the antineutrino detection efficiency because of the implemented trigger scheme. Hence, precise determination of the neutron detection efficiency for each cube and its evolution over time is crucial for an antineutrino rate analysis. Whilst the overall efficiency is not relevant for a shape-only oscillation analysis, the relative difference in efficiency as a function of distance could cause systematic effects if not properly accounted for. The neutron efficiency (ϵ_{det}) can be expressed as the product of

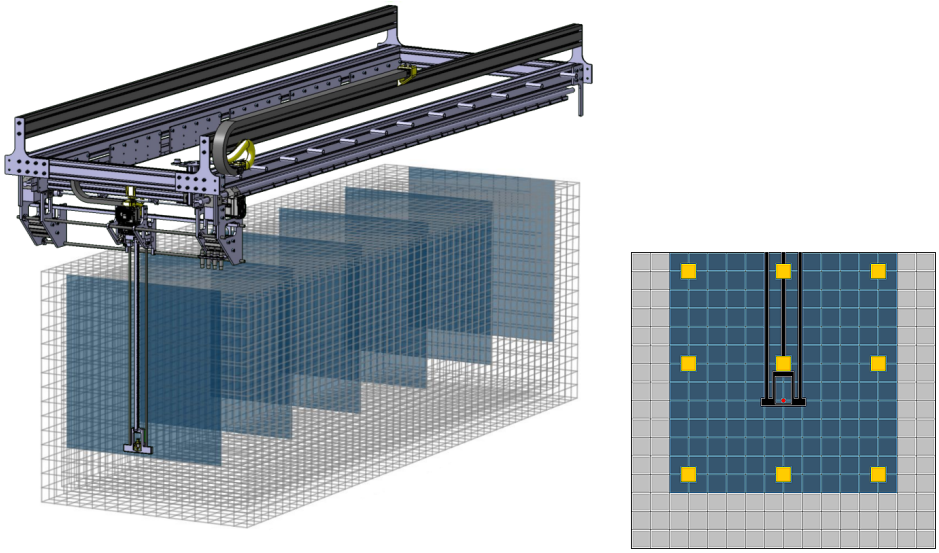


Figure 2.18: Left: CAD rendering of CROSS with the six detector gaps shown in blue [193]. Right: Possible source positions (yellow) from which the detector can be irradiated for a given gap. The red square defines the origin of the CROSS coordinate system. The entire detector can be exposed to varying degrees of radiation in a given time.

the neutron capture efficiency (ϵ_{cap}) and the neutron reconstruction efficiency (ϵ_{reco}). The former quantifies the likelihood that a given neutron is captured upon entering the detector volume and is estimated from MC simulations of neutron interactions in the detector. The latter is the probability that an NS signal is detected and successfully reconstructed, after capture on ^{6}Li has occurred. As such, it is independent of the 4-momentum and position of the initial neutron. Radioactive source data is used to determine ϵ_{reco} since it is affected by both the neutron trigger efficiency and offline particle identification.

Two different neutron sources were used during the September 2018 campaign, ^{252}Cf and AmBe, during which the detector collected over 10^8 reconstructed NS waveforms from each source. The mean reconstruction efficiencies extracted from these datasets are 71% and 77% for AmBe and ^{252}Cf respectively. The origin of the discrepancy between the two is not fully understood so the values are averaged and the difference incorporated as a systematic uncertainty, giving a final value for ϵ_{reco} of $73.9^{+4.0}_{-3.3}$%. Combined with the average neutron capture efficiency from simulation of 72%, the mean IBD neutron detection efficiency is around 52%. The distributions of ϵ_{reco} per cube are provided in Fig. 2.19 which highlights the shift between the results from the two sources.

The neutron reconstruction efficiency as a function of time was tracked by performing periodic calibrations with an AmBe source. The efficiencies per plane are plotted in Fig. 2.20. The absolute efficiencies over the different planes are equal to within 5% of the mean, except for the first and last planes. This trend is the same for all periods but the efficiency curves exhibit a 1.6% per year decrease due to the decline in ZnS(Ag) light yield. The relative

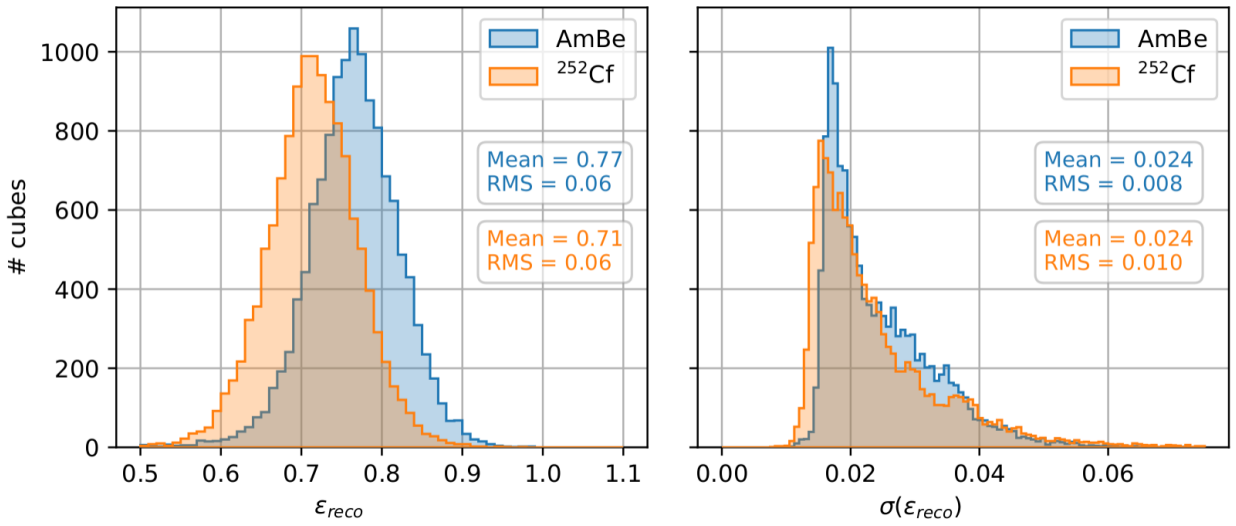


Figure 2.19: Neutron capture efficiency per cube (left) and corresponding uncertainties (right) determined from the September 2018 calibration campaign with AmBe and ^{252}Cf. The disagreement between the two sources is accounted for as a systematic uncertainty in the final result. Figures from Ref. [193].

efficiencies are stable to within one percent over the entire Phase-I period.

Gamma Sources

Numerous sources are available for calibration of the electromagnetic signal, including ^{207}Bi, AmBe, and ^{252}Cf. Before the start of data-taking in 2018 it was decided that ^{22}Na would be used as the primary source for in-situ ES calibration. The other sources are mainly used as cross-checks at higher energy and for validating the linearity of the PVT scintillator response. The objective of the calibration is to determine the light yield of every PVT cube to be able to convert the digitised signals from PA to MeV. In the main decay chain of ^{22}Na, beta decay causes the emission of two 511 keV annihilation gammas followed by a 1275 keV gamma from the de-excitation of ^{22}Ne. The ES calibration is based on the detection of these photons.

The collaboration has developed two distinct methods for ^{22}Na calibration. The first is a numerical fit of the Compton edge[¶] performed with a Kolmogorov-Smirnov (KS) test [207]. The source activity and detector response are accurately modelled in Geant4 at each CROSS position in order to build up a template of the Compton edge profile for a given cube, which is then convoluted with a Gaussian to simulate the detector energy resolution. The values of light yield and Gaussian width are then varied in the simulation and the pair of parameters that minimise the KS goodness-of-fit are selected. The precision of the numerical fit is approximately 6% and the obtained energy resolution is roughly 16% at 1 MeV. The second

[¶]The Compton edge is the energy at which a gamma rays scatters through 180° and deposits the maximum possible amount of energy in the detector.

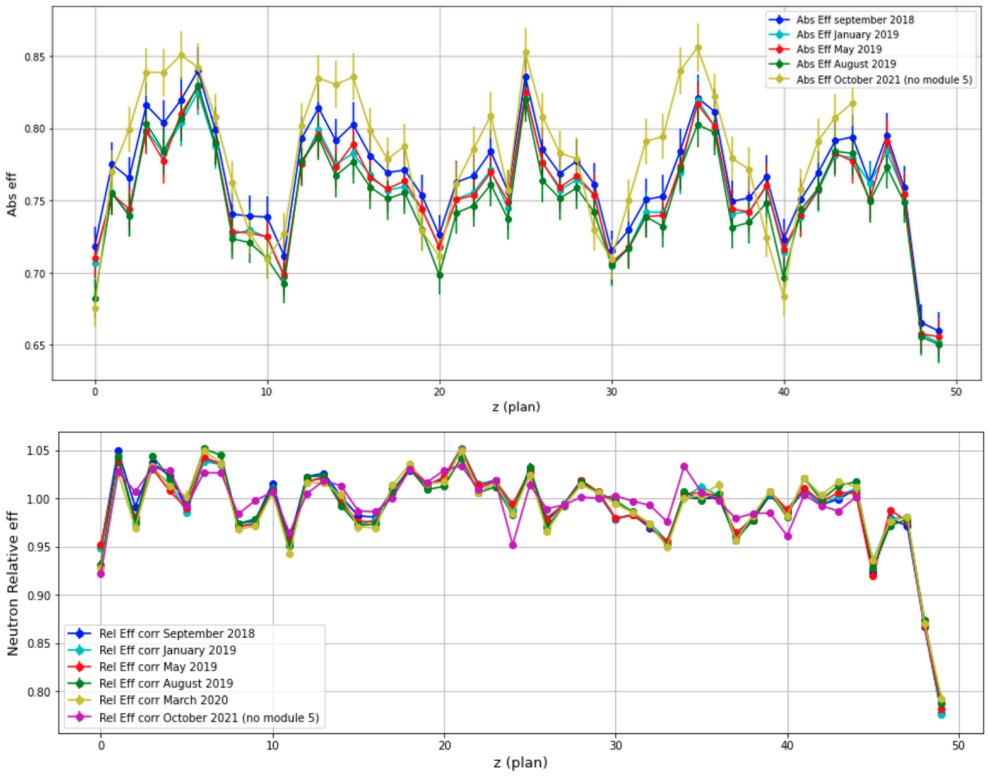


Figure 2.20: Absolute (top) and relative (bottom) neutron reconstruction efficiency of each detector plane across four and five Phase-I calibration campaigns respectively [211]. The lines corresponding to the October 2021 campaign are from the Phase-II data-taking period, which is not relevant to this work. Lines between points are a guide to the eye. The relative efficiency is stable to within 1% but the absolute efficiency decreases by around 1.6% per year because of the degradation of ZnS(Ag) light yield.

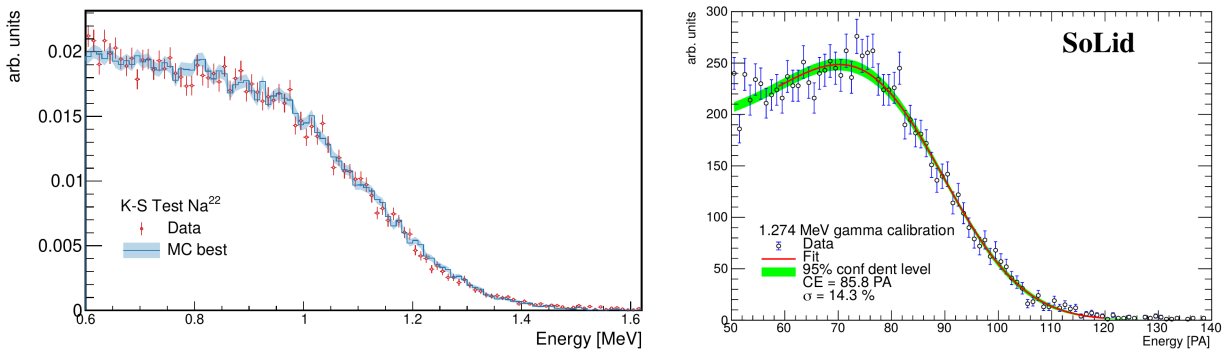


Figure 2.21: Left: Kolmogorov-Smirnov test on Compton edge of ^{22}Na [207]; the MC histogram is obtained by minimising the KS goodness-of-fit metric against the data. Right: Analytical fit of the ^{22}Na Compton edge using the Klein-Nishina formula [212]. The PVT light yields computed by these methods are compatible to within 2%.

method is an analytical fit of the Compton edge profile with the Klein-Nishina formula [212], which describes the cross section of photons scattering on an unbound electron. The two methods agree on the light yield of each cube to within 2%. Examples fits can be found in Fig. 2.21. The September 2018 value averaged over all cubes is 95.5 PA/MeV before accounting for MPPC cross-talk.

A total of eight ^{22}Na calibration runs were taken during Phase-I. Later analysis of a dedicated optical simulation in Geant4 revealed that non-linearities in the photon collection of low-energy deposits can cause percent-level distortions in the calculation of the energy scale using ^{22}Na. The numerical fit method can also be applied to the Compton edge profile of an americium–beryllium (AmBe) source. AmBe releases a 4.44 MeV gamma ray which makes the calibration less prone to the low-energy biases expected from ^{22}Na. Furthermore, a method was developed in early 2023 that exploits the sizeable cross section for electron-positron pair production of the 4.44 MeV gamma. The electron-positron pair effectively acts as a monoenergetic 3.42 MeV source which can be distinguished from the Compton scattered "background" events using similar signal-background discrimination techniques utilised for IBD events (see chapter 4). The uncertainty on the light yield from the pair production method is around 1.5%.

2.5.3 Relative Calibration with Muons

The low overburden of the detector means that the incident rate of cosmic muons is high enough that they can be utilised as a calibration tool. A horizontal muon is defined as a Type 2 muon that leaves a track greater than six cubes in length, enters a cube within 0.5 cm of the border, lies within a polar angle of 40°, and exits the detector. An example is provided in Fig. 2.22. Large samples of horizontal muons are used to derive the SMat elements and perform a relative calibration by homogenising the detector response. They can also provide

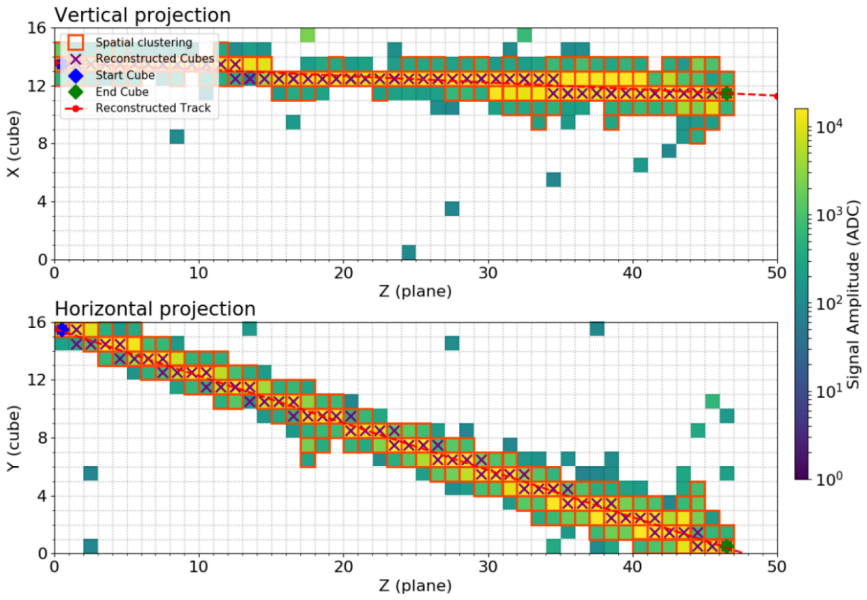


Figure 2.22: Vertical (x-z) and horizontal (y-z) projections of a track left by a horizontal muon passing through the detector. Horizontal muons are used to perform a relative calibration of the detector response and track the evolution of the PVT light yield over time. Figure courtesy of H. Chanal.

an additional handle on the PVT light yield by comparing the $\frac{dE}{dx}$ distributions per cube with an equivalent Geant4 simulation. However, the absolute light yield measurement has a relatively large uncertainty due to the complexity of the energy spectrum of horizontal muons and their energy loss mechanisms in matter.

The relative calibration uses the $\frac{dE}{dx}$ measurement per cube to homogenise the detector response. The optical decoupling of the planes means that each plane can be calibrated independently, in analogy with the CCube reconstruction. If the muon track passes through a single voxel in a given plane, any additional deposits must be caused by light-leakage effects, which can then be evaluated. Simulations indicate that light-leakage only affects neighbouring cubes, so next-to-neighbouring cubes and beyond are safely neglected in this treatment.

The SMat is constructed in two stages. It is initially filled with values based on the light-leakage, for which only the signal amplitudes are required. The initial cube crossed by the muon and the four adjacent cubes in the same plane give five possible affected cubes and twelve involved WS fibres. The total energy deposited in the plane—the sum of the twelve signals—is used to compute the linear energy loss $\frac{dE}{dx}$. In the second stage, the $\frac{dE}{dx}$ distributions of the cubes are fit with a convoluted Landau-Gaussian function. The most probable value of a fit quantifies the response of an individual cube. The SMat elements are then adjusted such that the response of each cube is aligned with the mean detector-wide response. Roughly ten days of data-taking are required to accrue enough horizontal muon

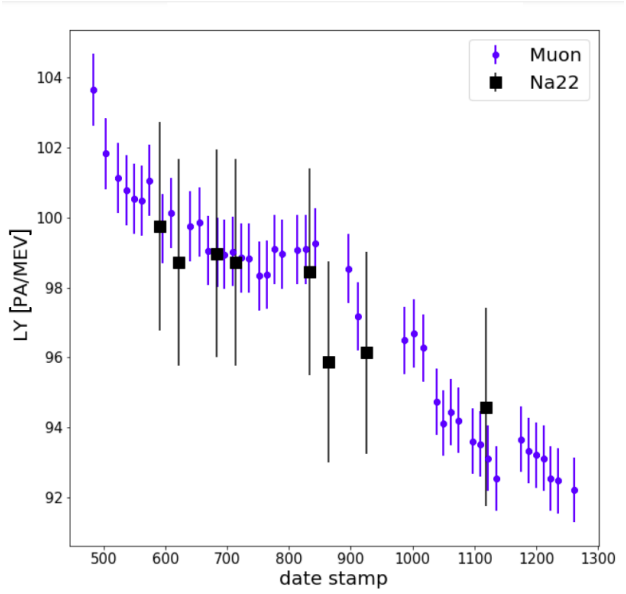


Figure 2.23: Evolution of the mean PVT light yield over time. The time stamp runs over the entire range of Phase-I data-taking from May 2018 to July 2020. The 44 blue points represent the light yield determined using horizontal muons and the 8 black points are the values from ^{22}Na calibration campaigns. The values are scaled such that they are equal at the time of the September 2018 source calibration (the ^{22}Na point earliest in time). Figure courtesy of H. Chanal.

events for a precise derivation of the SMat elements.

If the FPGA buffer of a given channel reaches maximum capacity (overflows) it is disabled for the reminder of the run. This has a number of adverse effects including poor energy reconstruction and fewer NS-ES coincidences. The CCube algorithm deals with this by setting the light yield of an overflown channel to zero in the system matrix, as if the channel had been switched off for the whole duration of the run. Then, a muon track is only included in the calibration procedure if the cube through which it passes has four active (i.e. not overflown) channels, with the exception of the handful of cubes that contain permanently broken fibres.

An advantage of muon calibration over source runs is that the evolution of the detector light yield can be continuously tracked and the only limiting factor to the calibration repetition is the number of horizontal muons reconstructed per day. This is especially important as the only full-detector AmBe scan was performed in September 2018. Moreover, since muons are readily available in the physics data, had this method been developed earlier there would have been less need to pause normal data-taking to perform dedicated calibration runs. The total exposure time of the Phase-I dataset allows for the computation of 44 separate system matrices. The corresponding light yield measurements are compared to the ^{22}Na results in Fig. 2.23, where the values have been scaled such that they are equal at the time of the September 2018 source calibration. The two methods agree very well on the change in light yield over time.

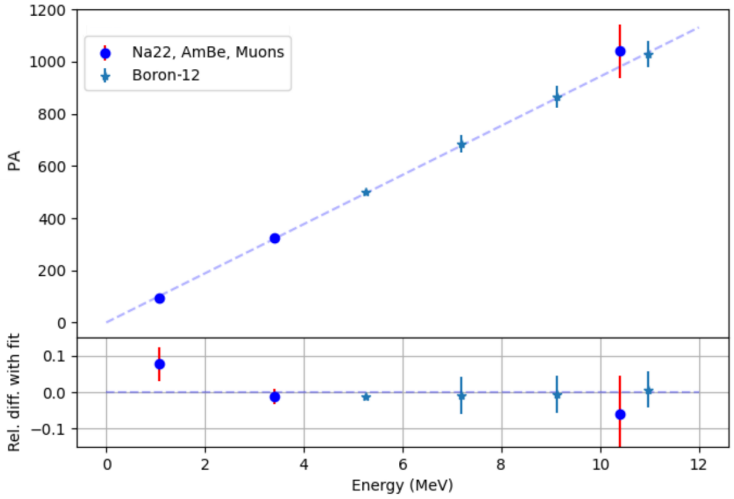


Figure 2.24: Response of the PVT scintillator as a function of initial particle energy, averaged over all detector cubes, measured by a combination of radioactive sources, boron-12 events, and cosmic muons. The linear fit describes the behaviour of the points very well and gives a precise value of the PA-MeV conversion factor. Figure courtesy of H. Chanal.

2.5.4 Energy Linearity

The linearity of the Eljen EJ-200 scintillator response has been verified using the available combination of radioactive sources and cosmic muons, supplemented by the energy spectrum of muon-induced boron-12 events. More detail may be found in Ref. [211]. The scintillator response as a function of energy is plotted in Fig. 2.24. To ensure a valid comparison the data points are taken either from, or were measured in close proximity to, the September 2018 calibration campaign. The gradient of the fit yields a final energy scale of 94.3 $\pm$1.7 PA/MeV.

2.6 Summary

The SoLid detector is designed to measure reactor antineutrinos by generating signals from the products of the IBD reaction. Effective identification of IBD events is achieved by employing the Cowan-Reines technique of introducing a characteristic time delay between positron- and neutron-induced signals. The solid plastic scintillator construction is extremely stable and able to operate over extended data-taking periods. An efficient triggering and readout system has been developed in line with the antineutrino efficiency requirements of the experiment. In addition, the energy response of the PVT has been experimentally validated and found to be linear to a high degree of accuracy.

The relative calibration from horizontal muons allow the detector response to be fully characterised and equalised at the level of individual cubes and channels. The choice of scintillators and overall design of the photon transport results in acceptable light yield for achieving the physics goals of the experiment. The collaboration is thus able to proceed

with the intended antineutrino disappearance and uranium-235 spectrum measurements. The next chapter demonstrates that the high degree of detector segmentation allows the IBD positron to be isolated from the annihilation gamma deposits, giving SoLid a handle on the antineutrino energy with minimal smearing from the annihilation gammas.

Chapter 3

Positron & Gamma Estimators

Accurate estimation of the antineutrino energy is crucial for an oscillation analysis. The energy of the IBD positron is highly correlated with the energy of its parent antineutrino, but the prompt positron signal (ES) also contains contributions from the annihilation gammas. The fraction of gamma energy deposited varies significantly event to event, effectively smearing the measured positron energy and thus degrading the energy resolution. Nonetheless, the three-dimensional imaging capabilities of the SoLid detector means deposits from the positron and annihilation gammas can be disentangled to a great extent.

Firstly, the interaction mechanisms of IBD particles in the detector are discussed. A set of candidates for the positron energy estimator based on some simple cube-level variables are then introduced in section 3.2.4. The goal is to minimise the contamination from annihilation gammas for a given fraction of captured positron energy. Flexibility in the choice of ES energy estimator is made possible by the high degree of detector segmentation and optical isolation of each voxel. A method of event classification based on successful isolation of the positron is also presented. Finally, a study of the positron energy resolution follows naturally and it will be demonstrated that the resolution is correlated with the level of annihilation gamma contamination.

3.1 Particle Interactions in the Detector

3.1.1 Energy Spectra of IBD Products

The energy of an antineutrino that undergoes inverse beta decay is shared between the neutron, the positron, and the IBD energy threshold. The positron then annihilates with an atomic electron to produce two gamma rays. The total visible energy in the prompt signal (E_{vis}) can be expressed as:

$$E_{\text{vis}} = E_{\gamma_1} + E_{\gamma_2} + E_{e^+} \,, \tag{3.1}$$

$$E_{e^+} = E_\nu - E_n - 1.806\,\text{MeV} \,, \tag{3.2}$$

where E_ν is the antineutrino energy, E_{e^+} and E_n are the positron and neutron kinetic energies respectively, and E_{γ_1} and E_{γ_2} are the energies of the gamma rays. In this work, the gamma rays from IBD events are referred to either as "annihilation gammas", "gamma rays", or simply "gammas".

The initial energy spectra of the antineutrino and IBD daughter particles are plotted in Fig. 3.1, after having subtracted the IBD threshold from the antineutrino energy to simplify comparisons with the positron spectrum. The ratio between the threshold-subtracted antineutrino spectrum and positron spectrum is very close to unity over most of the energy range, but diverges significantly at high energy. Accordingly, the share of antineutrino carried by the neutron is extremely small, typically in the 1–150 keV range. The gamma spectrum is strongly peaked at 511 keV—the electron rest mass—which is the energy of each gamma ray if the electron-positron annihilation occurs at rest. The spectrum also exhibits a tail that contains around $\sim 6\%$ of all events, extending up to around 7 MeV, comprising those events for which the annihilation occurs in flight. If the positron annihilation occurs at rest, the annihilation gammas are emitted in anti-parallel directions.

3.1.2 Electrons & Positrons

Energy loss of charged particles through an absorber occurs via two main mechanisms. The first is Coulomb interactions in the material, either with the orbital electrons (causing excitation and ionisation of the atoms) or with the nuclear electric fields. This *collisional* loss scales as $1/E$ for a non-relativistic particle of energy E. The second is a *radiative* effect from the emission of Bremsstrahlung radiation caused by particle deceleration and is proportional to E and Z^2, where Z is the atomic number of the surrounding nuclei. The stopping power of a particle in a given material is defined as the differential energy loss ($\mathrm{d}E$) per unit path length ($\mathrm{d}x$). The total stopping power is the sum of collisional and radiative effects:

$$S(E) = \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{col}} + \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{rad}} \,, \tag{3.3}$$

where the first term can be calculated from the Bethe formula [213, 214], plus corrections for the polarisation of the medium by the particle [215, 216], and the second from theoretical Bremsstrahlung cross sections [217].

The Electron Stopping Power And Ranges (ESTAR) program run by the National In-

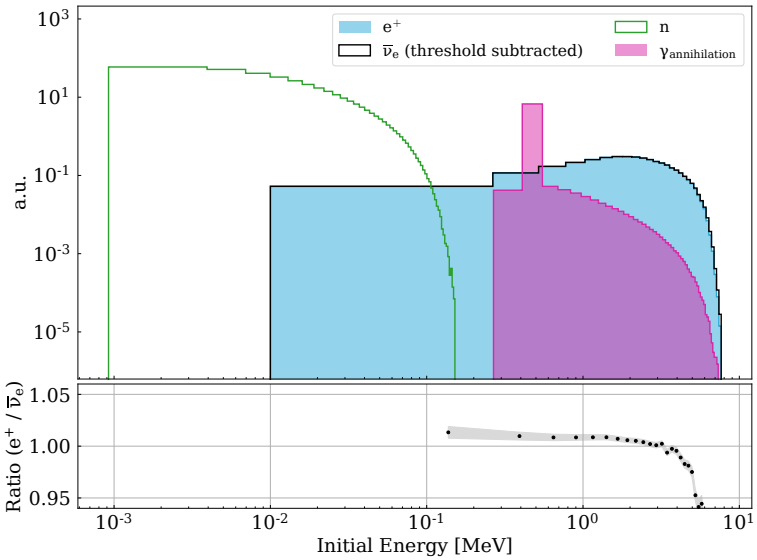


Figure 3.1: Initial energy spectra of the antineutrino and IBD products in SoLidSim. The ratio between the positron and threshold-subtracted antineutrino energy spectra is plotted in the lower pad. The grey band indicates the statistical uncertainty originating from the finite number of simulated events. The positron retains over 99% of the initial antineutrino energy.

stitute of Standards and Technology (NIST) can be used to compute the stopping power of electrons through a medium given the mass density, atomic structure, and mean excitation energy of each element [218]. The last of these is an experimentally determined quantity that appears in the Bethe formula. Integrating the inverse of the stopping power over energy gives an approximation for the average path length of a particle in a material by assuming a constant stopping power, known as the Continuous-Slowing-Down Approximation (CSDA).

Positrons and electrons lose energy through equivalent collisional and radiative processes. They have the same stopping power in a given medium because of their identical mass and charge magnitude. The density-scaled stopping power and CSDA range for electrons in PVT as a function of energy is depicted in Fig. 3.2. The stopping power is dominated by collisional losses over this energy range. Radiative effects only become important above 10 MeV, meaning they are insignificant for IBD positrons. The uncertainties in the calculated stopping power are generally at the 1–10% level depending on the electron energy, atomic number of the absorber, and type of energy loss. The CSDA distance for a 1(10) MeV electron is around 20(50) mm. The typical energy of IBD positron and the dimensions of the voxels in the SoLid detector imply that most of the kinetic energy of the positron is deposited in a single cube.

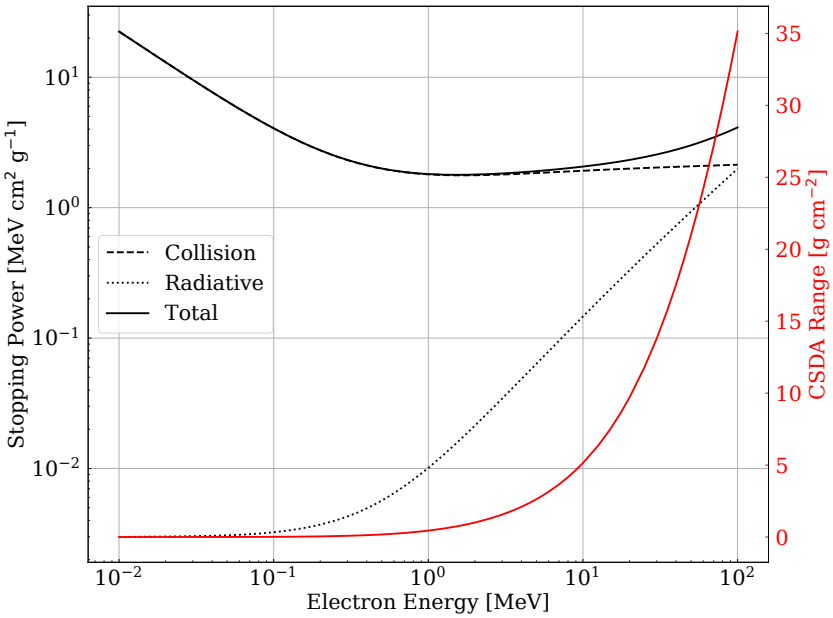


Figure 3.2: Electron stopping power (black) and CSDA range (red) in PVT as a function of energy, computed using the ESTAR program [218]. The total stopping power is the sum of collisional and radiative effects, but the former is dominant over the majority of the energy range shown here.

3.1.3 Gamma Rays

The three main gamma ray interaction mechanisms with matter are as follows:

- **Photoelectric absorption** where the photon incident on a material results in the ejection of an atomic electron. The electron energy is equal to the difference between the initial photon energy and the binding energy of the electron in its shell.
- **Compton scattering** where the photon loses energy by elastic scattering on an electron. The amount of energy transferred depends on the scattering angle.
- **Pair production** where the photon interacts with the electric field of a nucleus to produce an electron-positron pair. This can only occur if the photon energy exceeds twice the electron rest mass.

The total gamma-ray interaction cross section is the sum of the three individual cross sections but Compton scattering dominates at the energy scale of IBD gamma rays, as evident in Fig. 3.3. Photoelectric absorption and pair production are important below $\mathcal{O}(50)$ keV and above $\mathcal{O}(10)$ MeV respectively. The final energy of a photon that scatters on an electron through an angle θ, with respect to its incident direction, is given (in natural units) by:

$$E_f = \frac{E_i}{1 + (E_i/m_e)\,(1 - \cos(\theta))}\,, \tag{3.4}$$

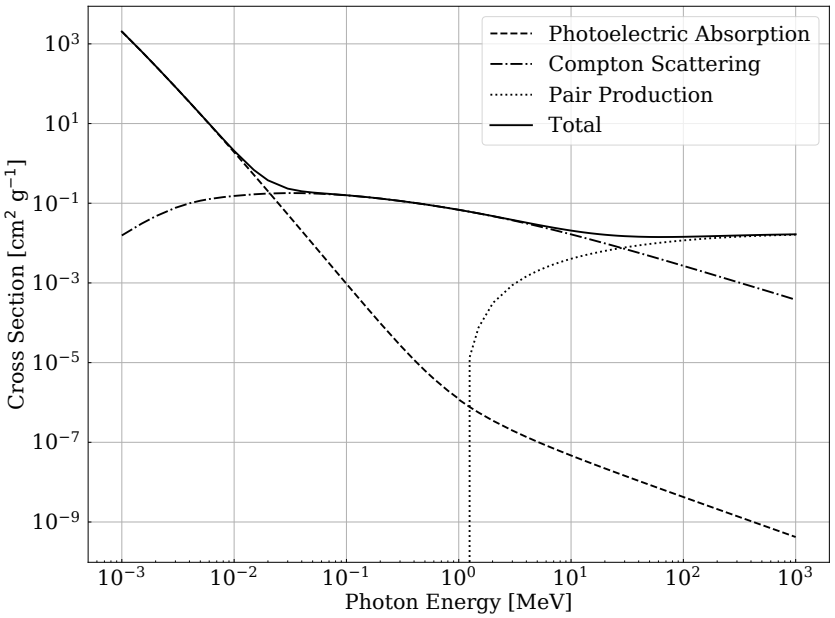


Figure 3.3: Gamma attenuation (proportional to cross section) in PVT as a function of energy [219]. Pair production includes interactions in both nuclear and electron fields. Compton scattering is dominant in the reactor antineutrino range.

where m_e is the electron rest mass and E_i and E_f are the initial and final photon energies respectively. The maximum energy transfer in a single scattering occurs when the scattering angle is π radians, approximately 341 keV for a 511 keV photon.

The Klein-Nishina formula describes the differential cross section for photons scattering on a single free electron:

$$\frac{d\sigma}{d\Omega} \propto \left(\frac{E_f}{E_i}\right)^2 \left[\frac{E_f}{E_i} + \frac{E_i}{E_f} - \sin^2(\theta)\right], \tag{3.5}$$

where all terms have been defined previously. Equations 3.4 and 3.5 are drawn for a selection of scattering angles and initial photon energies in Fig. 3.4. For example, a 1 MeV photon that scatters through π radians transfers 80% of its energy to an electron in the detector, but the cross section for such an interaction is much lower than for scattering at a smaller angle which corresponds to a smaller energy transfer. Combined with the relatively sparse population of IBD gammas above 511 keV, this suggests that annihilation gamma deposits are typically at $\mathcal{O}(100)$ keV, much smaller than the typical positron deposit.

The mean free path of a photon in a medium can be computed from the mass attenuation coefficient of that medium, $\mu_\rho(E)$, using the following equation:

$$\bar{l}(E) = \frac{1}{\rho \times \mu_\rho(E)}, \tag{3.6}$$

where $\bar{l}(E)$ is the mean free path as a function of energy and ρ is the mass density of

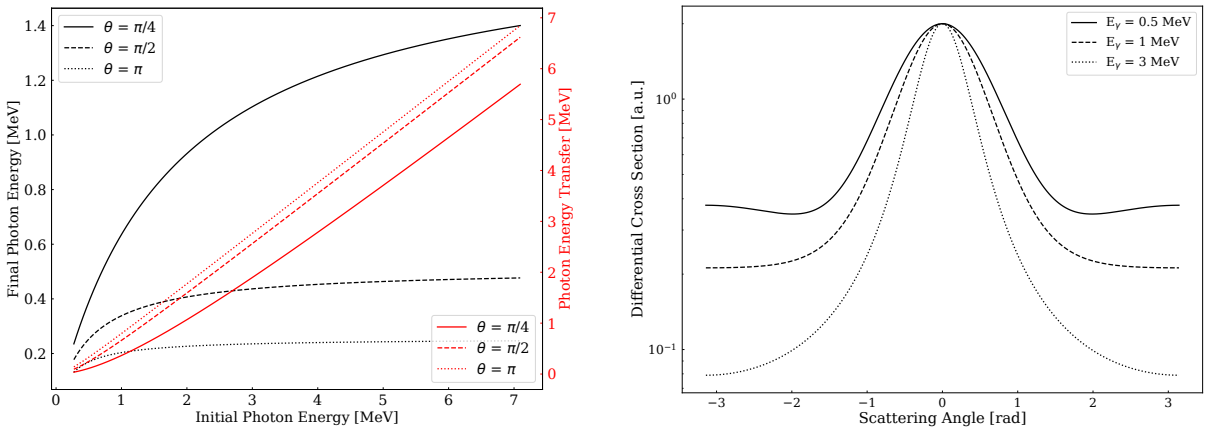


Figure 3.4: Example Compton scattering distributions. Left: Energy transfer as a function of initial photon energy for three scattering angles. Right: Differential cross section as a function of scattering angle for three photon energies. The scattering angle peaks at zero radians where no energy is transferred.

the medium. The density of the EJ-200 scintillator, as provided by the manufacturer, is $1.023\,\mathrm{g\,cm^{-3}}$. The mass attenuation coefficient of PVT for a 0.5 MeV gamma is $0.094\,43\,\mathrm{cm^2\,g^{-1}}$ [220], yielding a mean free path in the detector of approximately 10.8 cm. Hence, energy deposits from gamma rays in the detector are generally displaced from the IBD vertex by at least one cube.

3.2 Positron-Gamma Separation

The main purpose of this section is to compare the energy deposits from IBD positrons and annihilation gammas in order to construct variables that differentiate them. Therefore, the findings presented here are extracted primarily from the (Geant4-based) SoLidSim IBD simulation where complete information on particle IDs, position, 4-momenta, and interactions is accessible. This is to avoid the ambiguities introduced by the CCube algorithm. However, events that do not survive the reconstruction chain are discarded to ensure that the results are more closely aligned with the expectation for experimental data. SoLidSim events are denoted as "true" or as existing in "true space"; Saffron2 events (i.e. after readout simulation and reconstruction) are denoted as "reco" or as existing in "reco space". The one-to-one linking between true and reco events is preserved via the *cycleNum* Saffron2 variable that records the place of a reconstructed event in the ordered list of Geant4 events.

3.2.1 Cube-Level Energy Deposits

The CCube algorithm transforms the fibre signals of the ES cluster to a list of cubes and associated energies, so when considering true-space energy deposits from individual particles

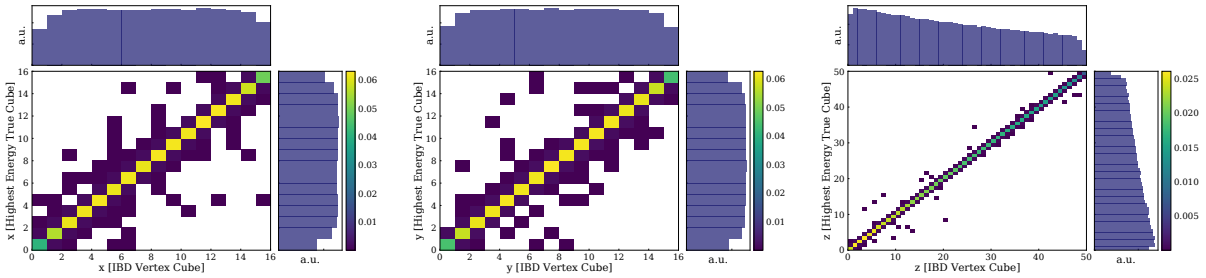


Figure 3.5: Positron energy deposits per cube in each spatial direction, showing the correlation between the IBD vertex cube and the true-space cube containing the largest positron deposit, which are the same in 85% of cases.

they must be combined to form *cube-level* deposits. A cube-level deposit of a given particle is defined as the sum of all contributions from that particle and its progeny in a particular cube. For instance, if a gamma ray were to undergo Compton scattering followed by pair production in the same cube, the cube-level gamma deposit is the sum of the scintillation light from the daughter particles and the initial deposit from the scatter. Annihilation gammas are counted as separate progenitors from the positron, but deposits from both gammas are combined.

Therefore, for each true-space event one can construct a list of cubes and associated energies per particle, analogous to the list of CCube-reconstructed cubes. A proxy for a Saffron2 ES cluster without reconstruction effects can be obtained by merging the per-particle lists. Note that the settings of the neutron trigger mean only three planes either side of the triggered plane survive the readout simulation, but more distant deposits can still be studied in the Geant4 simulation.

3.2.2 Antineutrino Position

In addition to the antineutrino energy, the position of its interaction vertex in the detector must be adequately estimated. The spatial resolution of the detector (5 cm) limits the precision of this estimator to a particular cube. A cube is uniquely defined by its 3D coordinate (x, y, z), where x and y are integers in the range 0–15 and z is an integer in the range 0–49, following the geometry of the detector and the coordinate system defined in the previous chapter. The correlation between the cube containing the IBD vertex and the cube containing the largest cube-level positron deposit is illustrated in Fig. 3.5, in each spatial direction. For events in which these cubes do not coincide they are most likely in directly neighbouring cubes, though there are sporadic instances where they are displaced by multiple cubes.

The Significant Migration Probability (SMP) is defined as the probability that a positron deposits the largest fraction of its energy in any cube other than the one containing the IBD

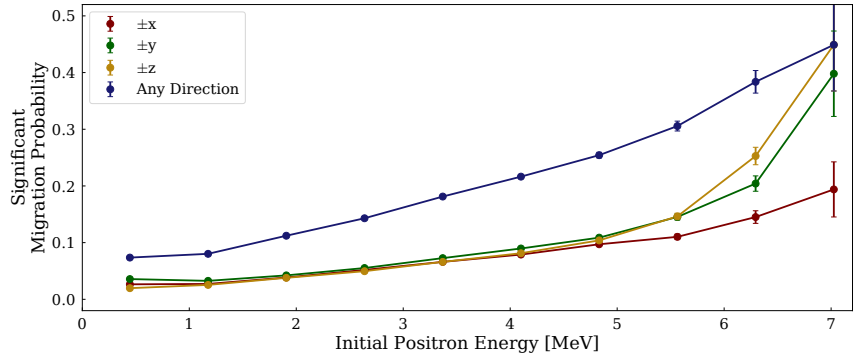


Figure 3.6: Significant Migration Probability as a function of positron energy. Lines between the points are a guide to the eye. Error bars represent the statistical uncertainty on the points. The probability that a positron leaves its largest deposit in a different cube to where it was created increases with initial energy, in agreement with the CSDA range.

vertex. Note that the SMP differs from the general probability that the positron track extends outside its original cube and is a function of initial positron energy and direction. Despite the CSDA range for IBD positrons being smaller than the size of a cube, it is found that the direction- and energy-integrated SMP of a given positron is 15% and the energy-integrated SMP in the z-direction (i.e. migration between detector planes) is 5%. Fig. 3.6 shows the SMP distributions as a function of energy. The SMP increases with energy—as expected from the CSDA curve—and surpasses 0.4 for the most energetic positrons.

Sensitivity to neutrino oscillations is driven in part by the distribution of antineutrino events in the longitudinal direction, i.e. along the z-axis which lies perpendicular to the detector planes. Unlike in the x- and y- directions, the plane assignment of a fibre signal is unambiguous. This means that the main source of smearing along the z-axis is the physical migration of the positron. The effect of non-zero SMP on the sensitivity to neutrino oscillations is small because (a) it affects only a small fraction of the events, (b) the distributions can be binned far more coarsely than the size of the effect, and (c) it is symmetric with respect to negative and positive direction travel. The final reason implies that migration of events between bins in one direction is compensated by migration in the opposite direction, which in any case only affects planes at the bin edges.

As the positron generally dominates the deposited energy in a given event, the cube containing the largest positron deposit differs negligibly from the overall most energetic cube in the merged ES proxy. If the correspondence of the latter with its reconstructed equivalent is sufficiently high, then it is a good candidate for the antineutrino position estimator. The correlation between the most energetic cubes for true and reco events is depicted in Fig. 3.7; the cubes match for over 96% of events. As such, the coordinates of the most energetic cube (MEC) in reco space are taken as the estimator of the IBD vertex position.

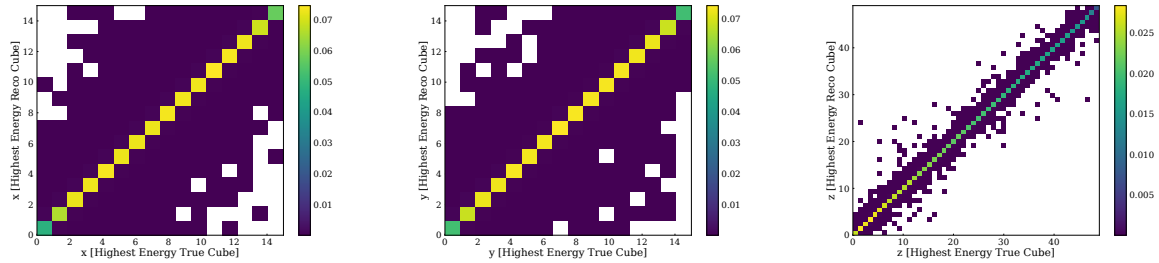


Figure 3.7: Correlation between most energetic true-space (Geant4) and reco-space (Saffron2) cubes in each spatial direction. The probability that these do not coincide is less than 4%.

3.2.3 Variables for Separation

Useful comparisons between positron and gamma cube-level deposits are limited to the variables that are available for both true and reco events. A cube-level deposit can be described by four parameters—(E_{cube}), x, y, z—and its relation in these parameters to the most energetic cube. The MEC defines the local origin $(0, 0, 0)$ of an ES cluster. Three further variables are defined, Δr_{MEC}, E_{cluster} and N_{cluster}, respectively the absolute centre-to-centre distance of a cube from the MEC, total ES energy deposited, and number of energetic cubes:

$$\Delta r_{\text{MEC}} \equiv \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}\,, \qquad (3.7)$$

$$E_{\text{cluster}} \equiv \sum_{i}^{N_{\text{cluster}}} E^{i}_{\text{cube}}\,, \qquad (3.8)$$

where $\Delta x \equiv \Delta x_{\text{MEC}}$, $\Delta y \equiv \Delta y_{\text{MEC}}$, and $\Delta z \equiv \Delta z_{\text{MEC}}$ refer to the displacements from the MEC in the x-, y-, and z-directions respectively. The label i denotes the i^{th} element in the list of cubes. The binned distributions of these variables for each particle type are displayed in Fig. 3.8. Exactly 232 000 simulated IBD events were included in the computation of the histograms, yielding 375 202 and 2 650 408 cube-level deposits from the positron and annihilation gammas respectively, so the statistical uncertainty on the number of entries in each bin can generally be neglected.

Figs. 3.8a and 3.8b show the N_{cluster} and Δr_{MEC} distributions per particle. The binning at low Δr_{MEC} is aligned with the finite set of values that Δr_{MEC} can take. The majority of events, around 70% of the total, contain a single cube-level positron deposit. Of the positrons that leave deposits in cubes other than the MEC, only a small percentage do so at distances larger than one cube. As a result, the total number of cubes in which the positron can deposit energy is automatically constrained by the number of cubes that are directly adjacent to the MEC (3 for corner cubes, 4 for edge cubes, 6 for all others). The relatively high multiplicity of gamma deposits in Fig. 3.8a is partly due to the fact that

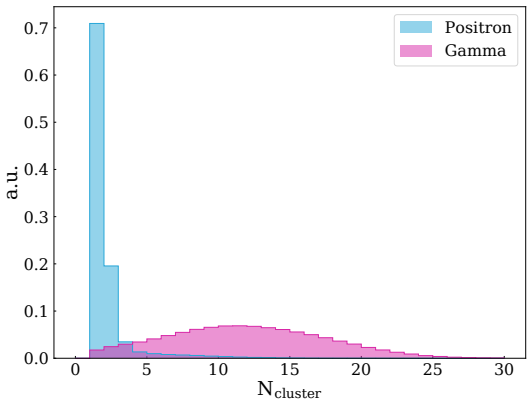
there are two oppositely orientated gamma rays for every one positron and partly owing to the energy loss occurring via Compton scattering rather than Coulomb interactions. In other words, the mean free path of the gamma is not significantly altered after a scattering, unlike the positron whose rate of energy loss increases as it is continuously slowed during passage through the medium.

The distributions of cube energies are given in Fig. 3.8c. The same information is displayed in Fig. 3.8d, but the cube energies have been divided by the total energy of the prompt deposit (E_{cluster}). This has the effect of shifting the main bulge of the positron histogram towards higher energy. The positron distribution broadly follows the shape of the positron energy spectrum of Fig. 3.1 with two notable differences. Firstly, the sharing of positron energy across multiple cubes causes the cube-level distribution to drop to zero at around 5 MeV instead of 7 MeV. In addition, it exhibits a sharp peak at very low energy that is correlated with the highly displaced positron deposits in Fig. 3.8b. Despite the maximum energy deposit from a single Compton scattering of a 511 keV photon being around 341 keV, the gamma distribution extends above 500 keV because of multiple scatterings within a cube, plus a small component from more energetic gammas created by annihilate-in-flight positrons. Overall, it is evident that these variables can provide separation between positron and gamma deposits in the detector. Ergo, there should exist an energy estimator that optimises this separation.

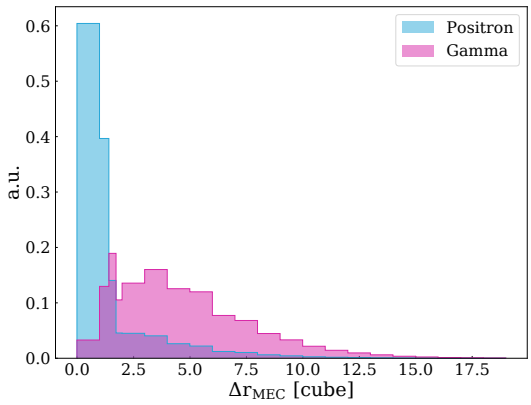
3.2.4 Estimator Definitions

The benchmark energy estimator is the total deposited energy of the prompt signal E_{cluster}. It is effectively the sum of all fibre signals and referred to simply as the *Cluster*. By construction, this estimator captures all of the available ES energy in an event. A first-order separation between positron and gamma deposits can be obtained purely from spatial cuts. Inspection of the Δr_{MEC} distributions in Fig. 3.8b indicate that the vast majority of positron deposits are located within two cubes of the MEC. As the annihilation gamma emission is mostly isotropic, three "envelopes" around the MEC are considered, defined by cubes within radius Δr_{MEC} equal to $\sqrt{3}$, $\sqrt{2}$, and $\sqrt{1}$. These are labelled the *Crown*, *Coronet* and *Cross* respectively. Taking envelopes of successively smaller radii excludes increasingly larger proportions of the deposited gamma energy. The natural limit is the MEC itself which often contains some fraction of the annihilation gamma energy. Note that the MEC is clearly an inadequate energy estimator for events in which the positron deposits energy in multiple cubes.

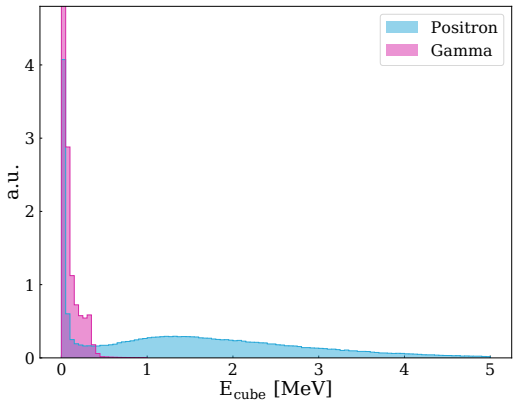
Further rejection of gamma energy is achieved with an energy-based cut to construct a final estimator called the Minimal Annihilation Gamma Estimator (MAGE). The MAGE is essentially the Cross estimator with an additional cut of $E_{\text{cube}}/E_{\text{cluster}} \geq 0.20$, i.e. it



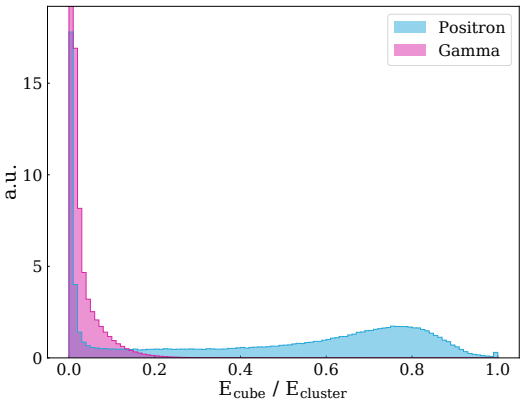
(a) Number of cubes containing energy deposits.



(b) Distance of energy deposits from the event origin.



(c) Energy distribution of cube-level deposits.



(d) Ratio of cube energy to total deposited energy.

Figure 3.8: Cube-level variables of positron and annihilation-gamma energy deposits. The evident differences in shape between positron and gamma distributions should allow them to be effectively separated.

Estimator	**Conditions for cube inclusion**	**Illustration**
Cluster	None	n/a
Crown	$\Delta x \leq 1$, $\Delta y \leq 1$, $\Delta z \leq 1$, $\Delta r \leq \sqrt{3}$	
Coronet	$\Delta x \leq 1$, $\Delta y \leq 1$, $\Delta z \leq 1$, $\Delta r \leq \sqrt{2}$	
Cross	$\Delta x \leq 1$, $\Delta y \leq 1$, $\Delta z \leq 1$, $\Delta r \leq 1$	
MEC	$\Delta r = 0$	
MAGE	$E_{\text{cube}}/E_{\text{cluster}} \geq 0.2$ and $\Delta r \leq 1$	n/a

Table 3.1: Definition of each energy estimator in terms of spatial and energy-based cuts. Note that $\Delta r \equiv \Delta r_{\text{MEC}}$ and similarly for Δx, Δy, and Δz, where MEC stands for the most energetic cube. Only the MAGE is defined based on energy information.

comprises any cube within the Cross that contains an energy deposit larger than 20% of the total deposited energy. The MAGE is an attempt to combine the low gamma-contamination of the MEC with any relatively high-energy positron deposits in the rest of the Cross. The six energy estimators are summarised in Tab. 3.1. The name of each estimator refer to the collection of cubes that the estimator comprises. If the name is written as a subscript it refers to the energy variable formed from the sum of the corresponding cube energies.

It is instructive to quantify the estimators according to the mean proportions of positron and gamma energy that they capture, as in Fig. 3.9, with respect to both the initial particle energy and the deposited particle energy. The latter quantity for annihilation gammas is known as the "gamma contamination" of the estimator. The histograms from which the mean values are obtained can be found in appendix A. The results for the Cluster establish that on average around 90(70)% of the initial positron(gamma) energy is deposited in the active volumes of the detector. These values represent the best- and worse-case scenarios in terms of maximising and minimising the positron and gamma energies respectively.

The amount of positron energy captured is relatively stable across most of the estimators—decreasing by around 3% from the Cluster to the MAGE—but there is a significant drop in positron energy at the MEC. This is in line with the previous expectations regarding the positron migration and multiplicity of cube deposits. The downward trend in the gamma contamination of the Crown, Coronet and Cross confirms that the amount of contamination reduces as fewer displaced cubes are incorporated into the estimator. Gamma deposits in the Cross cubes (other than the MEC) are largely rejected by the $E_{\text{cube}}/E_{\text{cluster}}$ cut of the MAGE.

In summary, the MAGE construction enables the rejection of around 86% of the de-

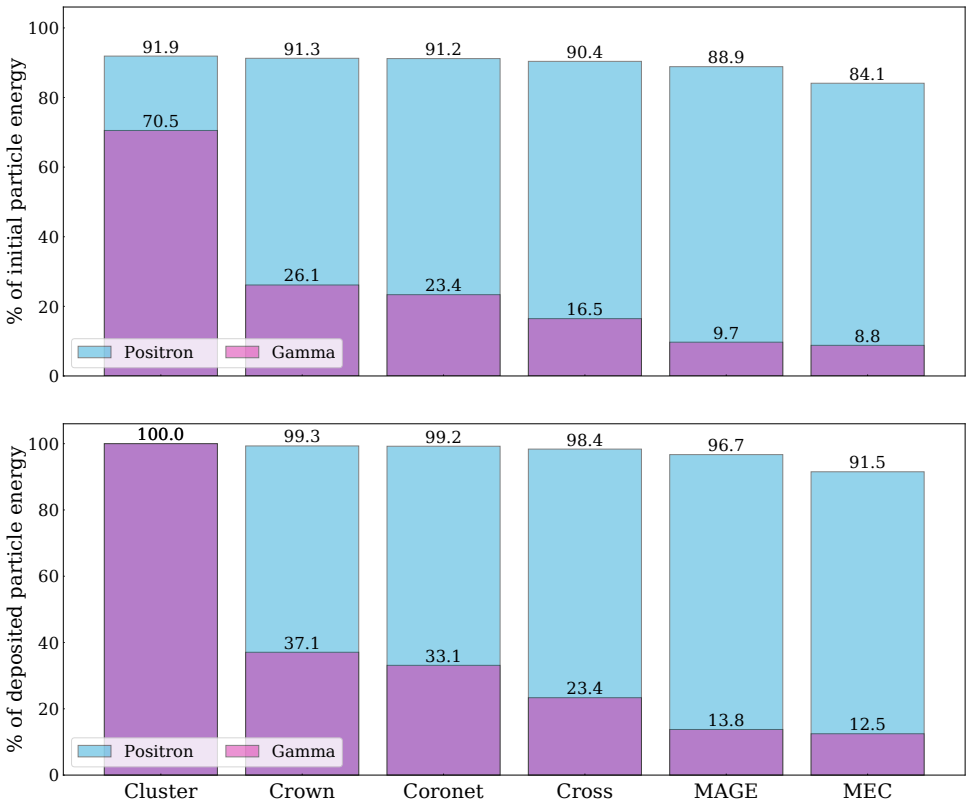


Figure 3.9: Evaluation of the positron-gamma separation of each estimator as the mean proportion of initial (top) and deposited (bottom) energy per particle captured by the estimators. The level of gamma contamination decreases from left to right, ending in the MEC which contains an average irreducible component of 12.5% deposited gamma energy.

Δr_{MEC} \ $E_{\mathrm{cube}}/E_{\mathrm{cluster}}$	0.10	0.15	0.20	0.25
$\sqrt{1}$ (Cross)	97.9 (17.1)	97.4 (14.9)	96.7 (13.9)	95.7 (13.4)
$\sqrt{2}$ (Coronet)	98.7 (20.5)	98.2 (16.3)	97.3 (14.5)	96.1 (13.7)
$\sqrt{3}$ (Crown)	98.8 (21.8)	98.2 (16.8)	97.4 (14.7)	96.2 (13.8)

Table 3.2: Optimisation of the MAGE cuts. Each entry gives the mean percentage of deposited positron (gamma) energy captured, as in Fig. 3.9. The entries are relatively insensitive to changes within the stated ranges.

posited annihilation gamma energy, whilst retaining around 97% of the deposited positron energy. It is very difficult to find further improvement due the irreducible mean gamma energy of 12.5% in the MEC. The threshold of 0.20 was initially chosen by eye but optimisation of the MAGE cuts is possible by scanning over a reasonable range of Δr_{MEC} and $E_{\mathrm{cube}}/E_{\mathrm{cluster}}$ cut values. The results of such an optimisation are summarised in Tab. 3.2. The differences are marginal and it is not obvious whether maximised positron energy or minimised gamma energy should be prioritised. For this reason, it is deemed acceptable for the $E_{\mathrm{cube}}/E_{\mathrm{cluster}}$ cut to remain unchanged. It will also be shown in the next section that there is very little correlation between E_{MAGE} and gamma energy for reconstructed events, indicating that the selection provides sufficient separation between the two.

The cube multiplicities (N_{cluster}) of the Cross and MAGE are displayed in Fig. 3.10. The selection criteria of the MAGE allows for the inclusion of up to five cubes, each sharing a fifth of the available energy, although in practise this does not occur. The Cross may comprise up to seven cubes. The MAGE more accurately reproduces the expected distribution of positron deposits of Fig. 3.8a, comprising a single cube in most cases with a small fraction of events requiring 2–3 cubes to take into account large energy deposits next to the MEC.

3.3 Event Categorisation

3.3.1 Cube Categorisation

The MAGE construction lends itself naturally to the categorisation of energetic cubes at the reconstruction level. The MAGE cuts are designed to isolate the positron energy, so any cubes that satisfy the criteria are designated as "positron-like". The cuts are then inverted to form a "gamma-like" class of cubes comprising low-energy cubes external to the Cross. This categorisation is not yet exhaustive so any remaining cubes are tagged as "leftover". The three categories are defined by regions in Δr_{MEC}-($E_{\mathrm{cube}}/E_{\mathrm{cluster}}$) space, as in Fig. 3.11. They can be summarised as follows:

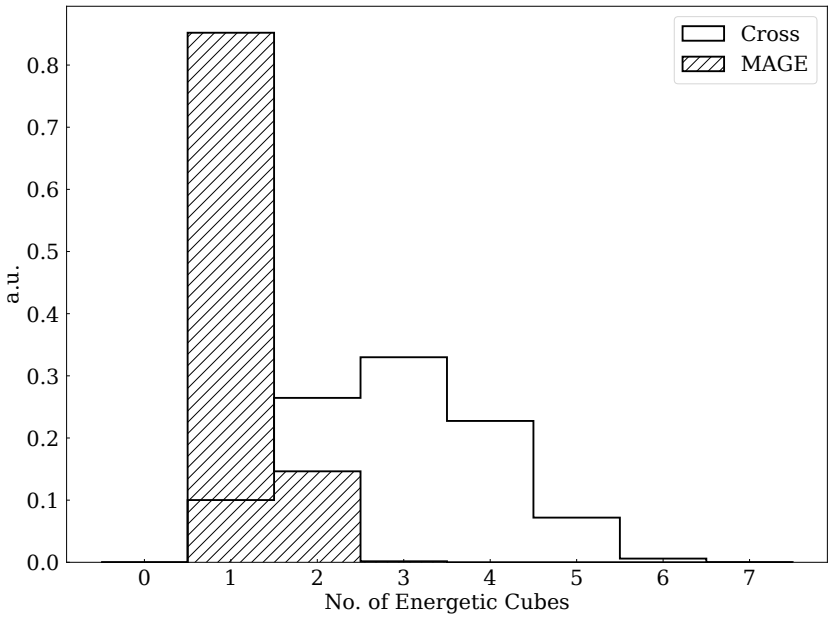


Figure 3.10: Cube multiplicities (N_{cluster}) of the Cross and MAGE. The distributions are both scaled to unity. The energy-based cut of the MAGE makes cube selection far more restrictive compared to the Cross and the distribution better resembles Fig. 3.8a.

- **Positron–like:** $E_{\text{cube}}/E_{\text{cluster}} \geq 0.2$ and $\Delta r_{\text{MEC}} \leq 1$.
- **Gamma–like:** $E_{\text{cube}}/E_{\text{cluster}} < 0.2$ and $\Delta r_{\text{MEC}} > 1$.
- **Leftover:** Not positron–like and not gamma-like.

The reason for introducing the third category, rather than say extending the gamma-like definition, is that whilst IBD events exhibit very few high-energy deposits outside the Cross this is not necessarily the case for certain background events. Hence, variables composed from leftover cubes can be utilised for background discrimination. Furthermore, it is extremely difficult to distinguish between gamma deposits within the Cross and reconstruction artefacts, as demonstrated in the next section.

3.3.2 Reconstructed Energy Threshold

There exist additional types of energy deposits in reconstructed events that are not present at the Geant4 level, including light-leakage, CCube ghost cubes, and dark-count waveforms that are selected by the time-clustering algorithm. These are all examples of "fake" cubes that do not correspond to energy deposits from physical particles. A sample of light-leakage deposits can be obtained by searching for true-space events in which only a single cube measured positron scintillation, then any additional cubes at the reconstruction level are assumed to have occurred from light-leakage. Ghosts are defined as reconstructed cubes

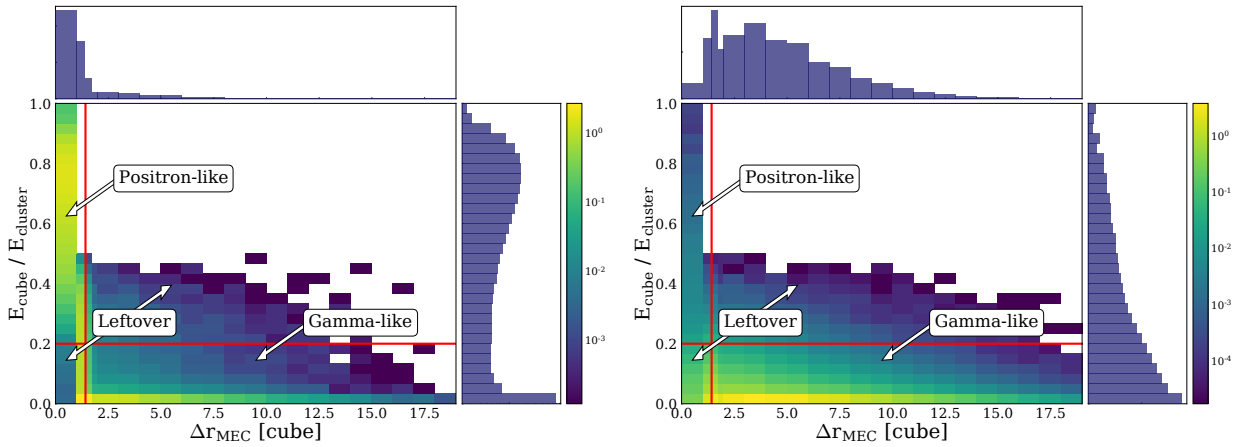


Figure 3.11: Boundaries defining the classification of energetic cubes in Δr_{MEC}-($E_{\mathrm{cube}}/E_{\mathrm{cluster}}$) space for true positron (left) and true gamma (right) cube-level deposits, delineated by red lines. Cubes classes neither as positron-like or gamma-like are dubbed "leftover".

that lack an equivalent truth-level deposit, excluding those that are already defined as light-leakage cubes. The energy spectra of fake cubes and cube-level annihilation gamma deposits are plotted in Fig. 3.12. The annihilation gamma deposits are smeared with a 16% Gaussian sampling—corresponding to the energy resolution estimated from the ^{22}Na calibration—as a simplified detector response. The majority of fakes cubes are found at energies below 100 keV, significantly lower than the total energy of most ES events but commensurate with annihilation gamma deposits. Consequently, fake cubes are classified as gamma-like if they lie outside the Cross, or as leftover otherwise.

Fake cubes could cause a bias between data and MC if not accurately simulated. Moreover, ghosts can mimic annihilation gamma deposits which are a crucial feature of IBD events. A potential effect of ghosts would be to make background events appear more "IBD-like" by artificially enhancing the number of gamma-like cubes detected. To mitigate this effect, it is prudent to introduce a minimum energy threshold that removes the majority of fake deposits and adjust the cube categories accordingly. A scan is performed over fifty threshold values in the range 0–100 keV; at each point the reco-level cubes below threshold are discarded and three metrics are calculated. The most energetic cube of each event is excluded because the 1 MeV cut on ES cluster energy in Saffron2 renders it immune to the threshold.

The three metrics are purity, energy efficiency, and number efficiency. In this context, purity is the mean fraction of reconstructed cubes for which an equivalent cube-level Geant4 deposit exists. This is equivalent to the not-ghost rate. Similarly, the number efficiency is the mean fraction of true cube-level deposits for which an equivalent reco cube exists. Lastly, the energy efficiency is the mean fraction of reconstructed energy of cubes that survive the reconstruction and can be matched with a cube-level Geant4 deposit, i.e. it quantifies the

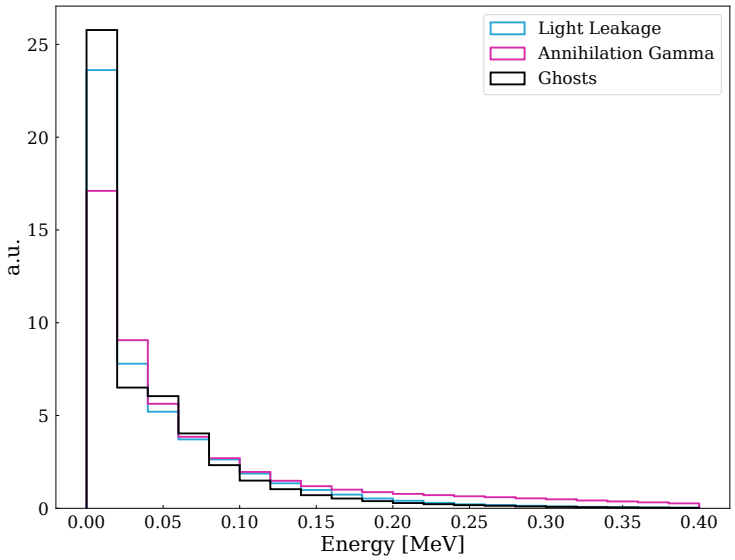


Figure 3.12: Energy distributions of light-leakage cubes, ghost cubes, and annihilation gammas. The level of overlap between the distributions is extremely high, meaning annihilation gamma deposits are difficult to distinguish from fake cubes using energy alone.

retention level of energy from physical particles. Inclusion of the MECs would severely skew the energy efficiency towards higher energy because of the 96% match between the highest energy true-space and reco-space cubes.

Three figures of merit (FoM) are defined:

- **FoM1:** Purity × Number Efficiency
- **FoM2:** Purity × Energy Efficiency
- **FoM3:** FoM1 × FoM2

The metrics and FoMs as a function of cube threshold are drawn in Fig. 3.13. As expected, the purity increases with threshold as a greater number of fake cubes are rejected, while the two efficiency metrics decrease as low-energy positron deposits, gamma-like, and leftovers cubes are removed. The efficiency curves indicates that a large proportion of low-energy cubes are not reconstructed. The three figures of merit are optimised at 63 keV, 92 keV and 74 keV respectively. The 74 keV threshold has been selected for the main antineutrino analysis because of its proximity to an optimal value identified in a parallel analysis by colleagues at LPC Clermont-Ferrand and the fact that it is calculated using all three metrics. A cube threshold of 74(63, 92) keV removes 78.1(74.0, 83.4)% of light-leakage energy, 82.4(77.8, 87.4)% of ghost energy, and 67.8(63.5, 73.1)% of annihilation gamma energy.

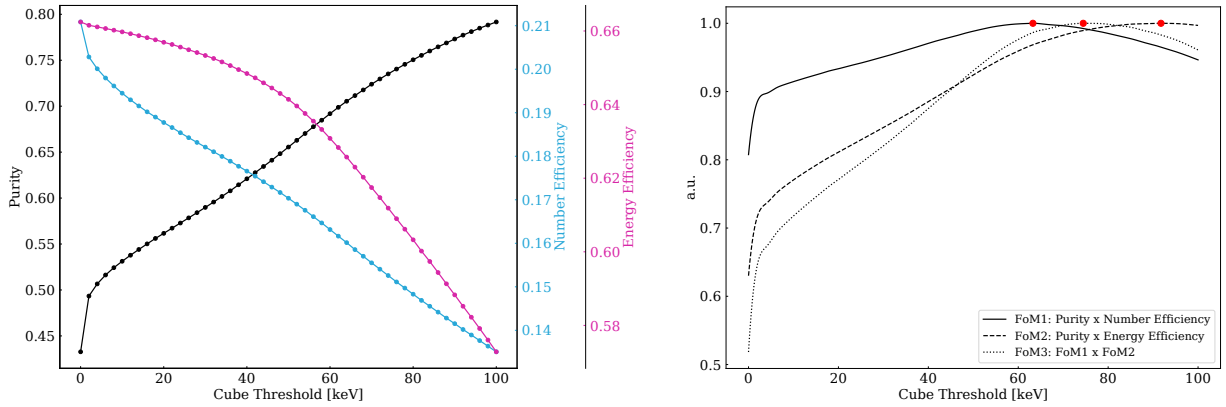


Figure 3.13: Optimisation of the minimum cube energy threshold by three figures of merit defined from three metrics. Left: Purity and efficiency as a function of cube threshold. Right: Figures of merit as a function of cube threshold. The maxima of the curves are indicated by red dots.

3.3.3 Gamma Bunch Algorithm

The cube categorisation can be readily extended to a full event categorisation based on the number of gamma-like cubes identified. This is advantageous for background discrimination since IBD events are distinguished by their annihilation gamma deposits. More specifically, one can use the fact that the annihilation gammas are (mostly) emitted in opposite directions. An algorithm has been developed by the SoLid group at LPC Clermont-Ferrand to identify gamma bunches that are detached from the most energetic cube. The procedure has been applied to the gamma-like cubes defined in this work, as follows:

- Define a set of vectors $V = \{\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_n}\}$ that originate at the centre of the MEC and end at the centres of each gamma-like cube. The algorithm terminates here if there are no gamma-like cubes in the event.
- Compute the dot product of all elements of V with the vector $\boldsymbol{v_i}$ that points to the highest energy gamma-like cube.
- Associate all gamma-like cubes that have dot products greater or equal to zero as a first bunch; all other gamma-like cubes (if any) as a second bunch.
- The energy of each bunch is given by the sum of all corresponding gamma cube energies.

Events are tagged as 2-gamma (2γ), 1-gamma (1γ) or 0-gamma (0γ) depending on the number of identified bunches. The most IBD-like 2-gamma events—in which both annihilation gammas deposit energy above the threshold—are expected to have the lowest background rates. The 1-gamma category is sub-divided into events comprising exactly one gamma-like cube (1γ(=1c)) and events comprising greater than one gamma-like cube

Threshold [keV] \ Gamma Category	2γ	$1\gamma(>1c)$	$1\gamma(=1c)$	0γ
63	0.221	0.197	0.338	0.244
74	0.197	0.164	0.366	0.273
92	0.165	0.119	0.398	0.318

Table 3.3: Relative fractions of different gamma categories according to the applied cube energy threshold. Lowering the threshold results in a higher proportion of 2γ and $1\gamma(>1c)$ events.

($1\gamma(>1c)$). The reasons for this are primarily background discrimination and data-MC comparisons, which are covered in the next chapter. The proportions of IBD events that fall into each category are outlined in Tab. 3.3. Decreasing the threshold results in a larger average number of gamma cubes per event, thereby increasing the proportions of 2γ and $1\gamma(>1c)$ events at the expense of the other categories.

If a significant fraction of positron energy is incorrectly labelled as gamma-like, or vice versa, one might expect to find some correlation between the corresponding energy variables. The density and correlation plots of the three variables (positron energy E_{MAGE}, energy of first gamma, energy of second gamma) are shown in appendix A. The Pearson correlation coefficients between each pair of variables are all below 0.1, indicating extremely weak (linear) correlation, thus helping to validate the cube categorisation and gamma-bunching algorithm.

3.4 Positron Energy Resolution

The distribution of reconstructed energy in a detector or calorimeter for a monoenergetic source can often be modelled by some continuous bounded function, typically a Gaussian. The broadening of the initial energy is caused by energy smearing and loss processes. A measure of the "width" of the curve as a function of energy (the standard deviation in the Gaussian case) is used to quantify the resolution, which is normally quoted as the value at 1 MeV. However, note that this can be a misleading quantifier of the overall energy resolution for non-Gaussian behaviour, especially when comparing the results to standard calorimeters. A poor energy resolution deteriorates the sensitivity of the experiment to neutrino oscillations because it makes it more difficult to resolve oscillations in the energy spectrum, particularly at high values of Δm^2_{41}. Accordingly, the choice of estimator and gamma category for an oscillation analysis is partially guided by the energy resolution.

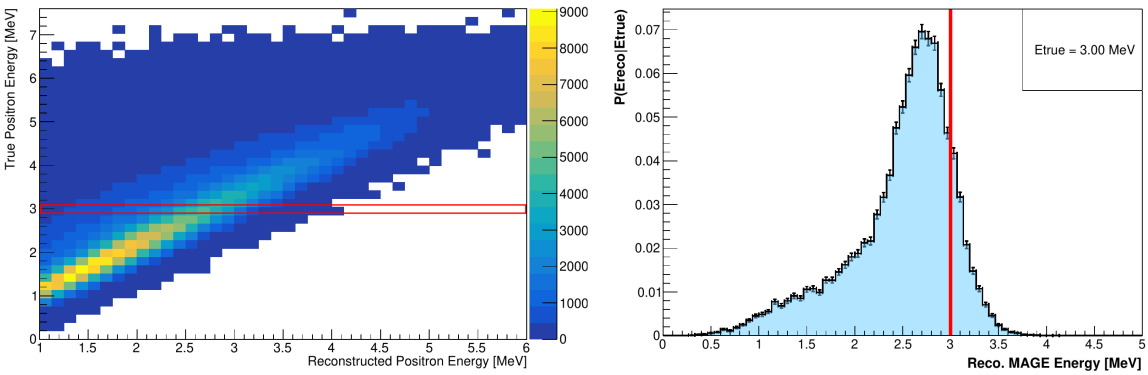


Figure 3.14: Left: Two-dimensional distribution of true-reco positron energy. The red box indicates a 10 keV window around 3 MeV. Right: One-dimensional slice of reconstructed energy for a true positron energy of 3 MeV (within ±5 keV), as marked by the red vertical line. The spread of reconstructed energy is a measure of the energy resolution and the shift between the red line and the most probable reconstructed energy is indicative of energy loss processes.

3.4.1 General Parameterisation

In general, the energy resolution is a combination of a stochastic term (statistical fluctuations on the number of scintillating photons and induced pixel avalanches), an electronic noise term (from the readout electronics, e.g. dark counts or signal amplification) and a constant term (arising from energy-independent effects such as detector geometry and mechanical defects). Expressed as the quadratic sum of these three terms, the energy resolution of the SoLid detector can be parameterised as:

$$\frac{\sigma}{E} = \sqrt{a^2 + \left[\frac{b}{\sqrt{E}}\right]^2 + \left[\frac{c}{E}\right]^2}, \tag{3.9}$$

where a is the constant term, b is the stochastic resolution, c is the noise contribution, and E is the true energy. The presence of annihilation gammas in IBD events adds a further smearing effect that is broadly independent of energy, but acts as a second stochastic term because of the random nature of Compton scattering.

The two-dimensional distribution of true-reco positron energy in Fig. 3.14 exhibits a strong linear correlation, with a small tail extending to low reconstructed energies that is present over the entire true-space energy range. The one-dimensional projection shows the tail more clearly and also exposes a shift of about 250 keV between the true positron energy (3 MeV in this case) and the most probable value (MPV) of the reconstructed energy. The shift can be attributed to positron energy that is not deposited in the active detector volume*. The presence of this tail means a Gaussian is inadequate for describing the energy resolution of IBD events in the SoLid detector. Instead, the energy smearing is modelled with

*As proved by the mean amount of positron energy captured by the Cluster estimator.

a GaussExp function [221]. This is a Gaussian that transitions smoothly to an exponentially decaying tail at some number of standard deviation below the mean. It is similar to the more common Crystal Ball function [222] with one fewer parameter:

$$f(x,\mu,\sigma,\alpha) = \frac{1}{N}\begin{cases} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} & \text{for } \frac{x-\mu}{\sigma} \geq -\alpha \\ e^{\frac{k^2}{2}+k(\frac{x-\mu}{\sigma})} & \text{for } \frac{x-\mu}{\sigma} < -\alpha \end{cases} \tag{3.10}$$

where μ and σ are the mean and standard deviation of the Gaussian part of the equation and N is a normalisation factor. The parameter α determines the transition point between the Gaussian and the exponential. The GaussExp function is preferred over the Crystal Ball because its lower number of parameters makes it less susceptible to instabilities when fitting. For the remainder of this chapter, the true-space positron energy is denoted as E_{true} and the reconstructed energy variable (i.e. the MAGE) is denoted as E_{reco}. The standard deviation of the GaussExp assumes the role of the Gaussian standard deviation in quantifying the energy resolution.

The GaussExp function is able to parameterise the distributions of reconstructed positron energy reasonably well; example fits for each gamma category are given in Fig. 3.15. However, there is usually a small component of the low-energy tail that is not accounted for by the fit. It is conceivable that a more complex function could better model these second-order effects.

3.4.2 Method & Results Per Gamma Category

Reactor antineutrino are not monoenergetic. Even so, a high-statistics sample of IBD MC events can be used to estimate the energy resolution via the following procedure:

1. Select all events within a (non-overlapping) rolling window running over the entire true-space energy range.
2. Bin the reconstructed energy distributions for all events that fall within the windows.
3. Determine the values and uncertainties of the parameters μ, σ, and α by fitting the distributions with a GaussExp function.
4. Fit the standard deviation divided by true energy (σ/E_{true}) as a function of E_{true} with equation 3.9.

The width of the rolling window is limited by the number of available MC events, but in most cases 10 keV is found to provide enough statistics whilst being sufficiently small compared to the expected energy resolution. As a quality control measure the fit parameters are constrained to a lie within a range of physically reasonable values, which in this case simply means that α, σ, and μ must be greater than zero.

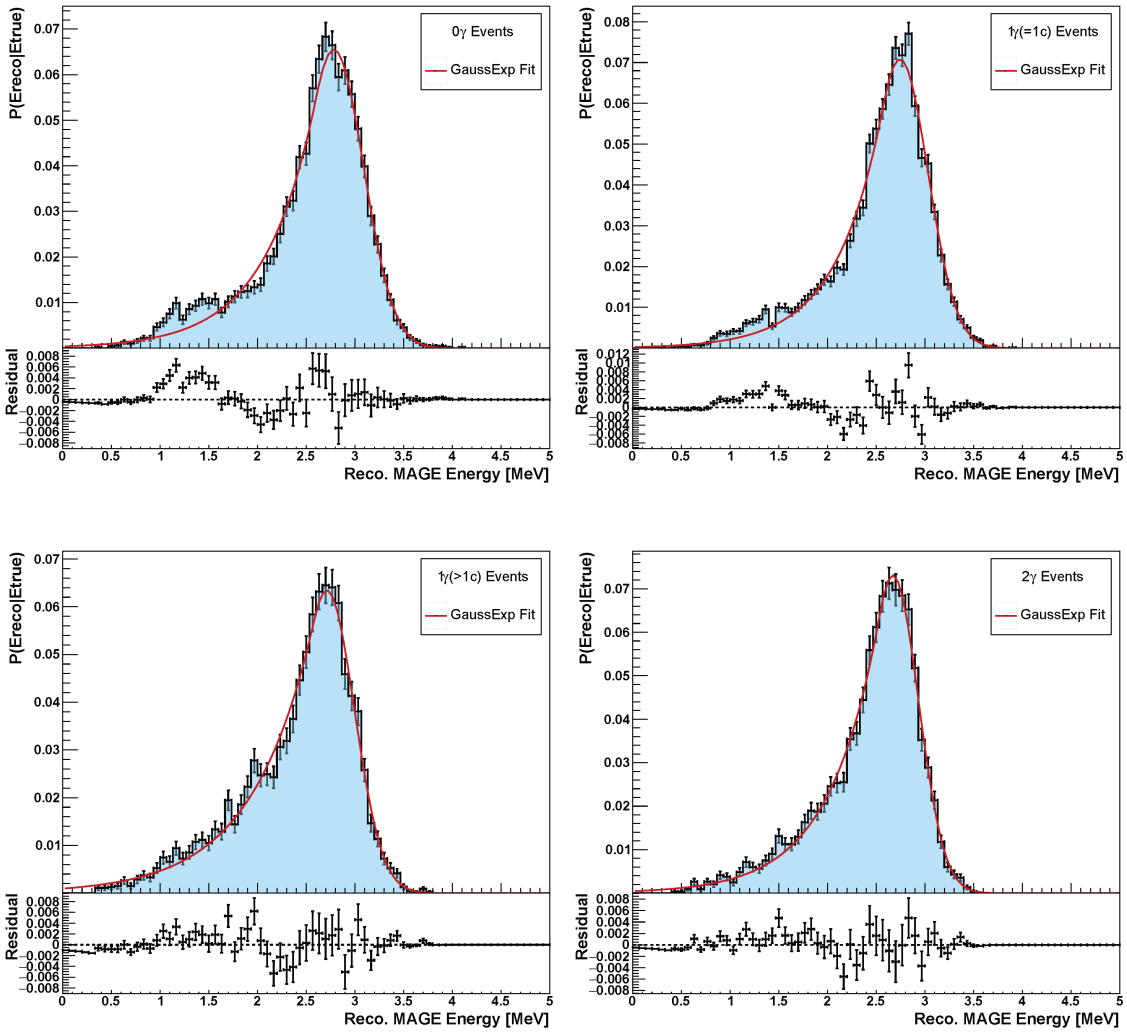


Figure 3.15: GaussExp fits on the reconstructed energy distributions of 3 MeV positrons, for each gamma category. Ratio plots show the fits divided by the histograms entries. Error bars represent the statistical uncertainty on the points. The fits model the Gaussian part of the distributions reasonably well, but significant deviations can be seen in the exponential tails.

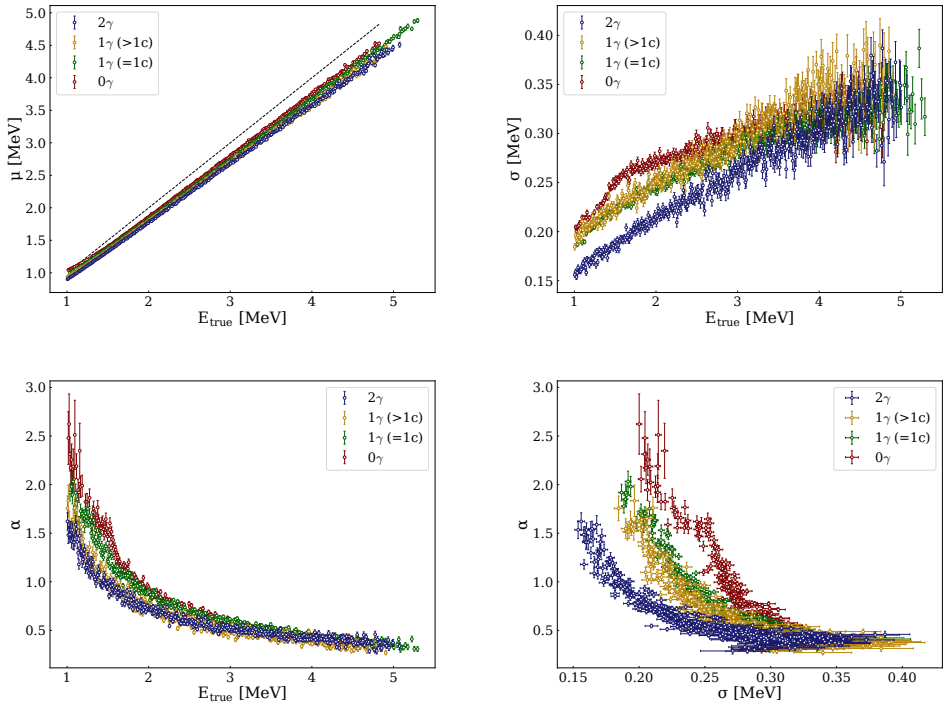


Figure 3.16: Parameters of the GaussExp fits as a function of true positron energy and correlation between α and σ (lower right). Each point and associated error bar represents a GaussExp fit on a rolling window. The correlation between α and σ means more Gaussian-like distributions have larger standard deviations.

The three GaussExp parameters are plotted as a function of E_{true} in Fig. 3.16. The shift between the MPV and the true positron energy is explicit in the behaviour of the $\mu(E)$ curves. The magnitude of the shift from the line of unit gradient increases with positron energy, suggesting that energy is primarily lost in gaps and non-active material between voxels as the positron passes through. The α parameter is plotted against σ in the lower-right pad of Fig. 3.16. The correlation between the two is strongly negative, meaning a narrower Gaussian peak corresponds to an exponential tail closer to the MPV, which is not desirable. Thus there is a trade-off between the width of the Gaussian and the size of the exponential tail. The $\alpha(E)$ curve shows that pure Gaussian behaviour is approached at low energy where the exponential tails begin up to two standard deviations from the mean. It is apparent from the $\sigma(E)$ curves that 2γ events have superior energy resolution. There is only a slight difference between the two 1γ sub-categories. The strange features in the distributions of 0γ points, particularly around 1.5 MeV, are caused by distortions in the spectra which are largest for 0γ events.

Equation 3.9 is used to fit the σ/E curve to parameterise the energy resolution as a function of energy, as illustrated in Fig. 3.17. The goodness-of-fit metric is worst for 0γ events yet the precision on the parameters is in line with the other categories. The resolution at 1 MeV for 2γ events is consistent with the value obtained from ^{22}Na calibration and

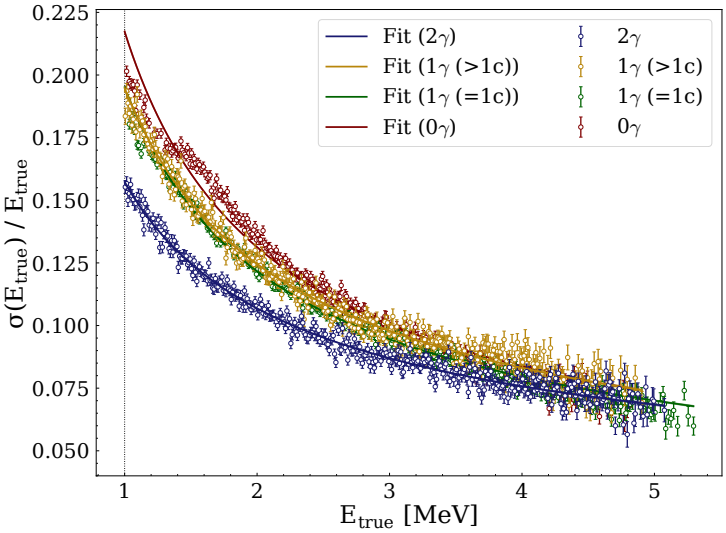


Figure 3.17: Parameterisation of the energy resolution of IBD events by fitting $\sigma(E)/E$ with equation 3.9. Each point and associated error bar represent a GaussExp fit on a rolling window. 2γ events have superior energy resolution.

Gamma Category	a^2	b^2	c^2	χ^2/ndf	σ/E @ 1 MeV (%)
0γ	0.0000(2)	0.0217(8)	0.0255(9)	6.19	21.7(1)
1γ(=1c)	0.0000(1)	0.0211(6)	0.0169(7)	2.22	19.5(1)
1γ(>1c)	0.0000(2)	0.024(1)	0.014(1)	1.72	19.5(1)
2γ	0.0007(2)	0.0187(7)	0.0053(7)	1.61	15.7(1)

Table 3.4: Fitted parameters of equation 3.9 that quantify the positron energy resolution. χ^2/ndf is the reduced chi-square goodness-of-fit metric. The constant terms are consistent with zero. The goodness-of-fit of 0γ events is very poor.

reasonably close to the target resolution of the experiment [192]. The fitted parameters are presented in Tab. 3.4. The constant terms are negligible for all gamma categories. In comparison, the stochastic and noise terms are generally at the same order of magnitude and decrease as the number of gamma-like cubes increases.

The trend of the energy resolution across the various gamma categories can be explained by the annihilation gamma deposits. The mean gamma contamination of the Cross, MAGE, and MEC estimators is calculated independently for each category and displayed in Fig. 3.18. The mean proportion of deposited gamma energy located in the MEC is 12.5% when considering all events simultaneously. Splitting the events into the gamma categories reveals that this value is dominated by the 1γ and 0γ events.The 2γ events have only a 5% contamination of gamma energy in the MEC, making them the purest available sample of IBD positron energy. As such, a 1 MeV electron is anticipated to have a similar energy resolution of 16%. The difference in mean proportion of captured gamma energy between 0γ

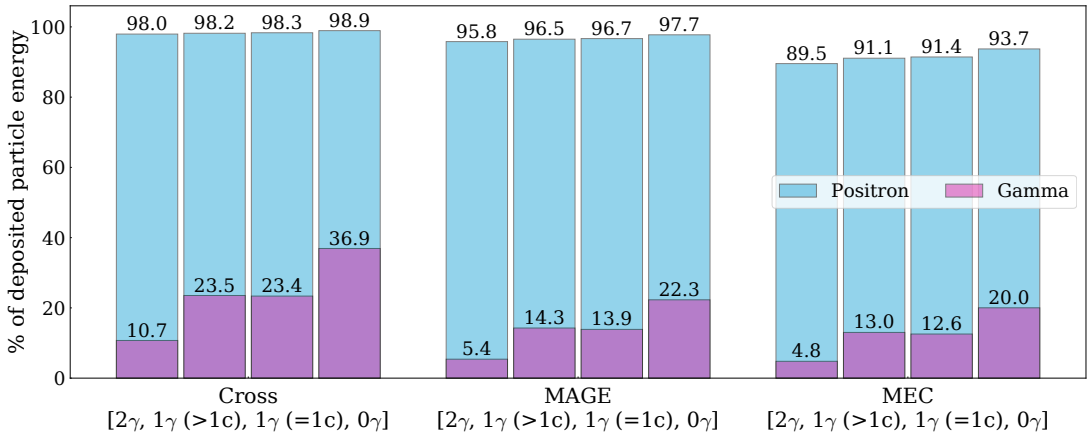


Figure 3.18: Mean proportion of deposited energy of the Cross, MAGE, and MEC estimators per gamma category. The level of annihilation gamma contamination follows the energy resolution of each category.

events and 2γ events is over 15%, which satisfactorily clarifies the differences in their energy resolution. Considering the parity between the size of the annihilation gamma component and the energy resolution of a given estimator or gamma category, it can be concluded that annihilation gamma deposits contribute significantly to the degradation of the energy resolution.

Chapter 4

Antineutrino Analysis

In this chapter, the tools developed for extracting an antineutrino signal from the data are presented. SoLid faces extremely high background rates from atmospheric particle showers, muon-induced spallation, and intrinsic radioactivity which are orders of magnitude greater than the expected antineutrino interaction rate. However, full exploitation of the spatial and temporal granularity of the detector—combined with multivariate machine learning (ML) tools—can reduce this background to a manageable level.

We begin by defining the data-quality procedures implemented to filter low-quality or useless data. The discussion proceeds to the primary background events relevant to an antineutrino signal that are observable in the data and the selection cuts used to suppress these backgrounds. Particular emphasis is given to the machine learning aspects of the background reduction. Finally, the detector simulations are validated using data-MC comparisons of a specific background that can be isolated in a high-purity region of parameter space.

4.1 Data Quality of the Phase-I Dataset

The Phase-I data-taking campaign started on 11^{th} April 2018 and ended 2^{nd} July 2020. The detector was operated near-continuously during this time, taking runs in physics-mode with the DAQ settings reported in chapter 2. Deliberate interruptions to the default data-taking were typically for in-situ source calibrations, maintenance of the server, or scheduled pressure tests of the BR2 vessel. Unintentional stoppages were caused by technical glitches in the DAQ system, power outages, or chiller failures. The collaboration has developed a set of data-quality criteria that a given run must satisfy in order to be included in the Phase-I physics dataset. Most runs that are rejected are done so based on these well-defined criteria, but other periods are excluded for practical or empirical reasons on a case-by-case basis.

The dataset analysed in this thesis is from the state-of-the-art processing (at the time of writing) that uses the horizontal-muon calibration to equalise the detector channels and

provide a uniform response over time and space. Individual DAQ files are processed according to the muon-derived system matrix (of which there are 44 in total) closest in time to when the run was taken. The processed files that pass the data-quality cuts are then merged to create single per-day files. Calibration runs taken during CROSS campaigns are not included in the general processing or merging. Overall, the filtered dataset contains 341 reactor-on dates and 258 reactor-off dates, corresponding to over 295 million NS-ES coincidences.

4.1.1 Notable Incidents

The data spanning 8$^{\text{th}}$ July 2019 to 9$^{\text{th}}$ September 2019 has been discarded in its entirety because of numerous issues including a power outage, a partially deconstructed water wall, SiPM pedestal problems in plane 38, and a BR2 pressure test. The data taken with the missing plane feasibly could be recovered and analysed using a like-for-like MC sample. Conversely, the data-taking without the full water wall occurred solely during a reactor cycle, meaning there is no equivalent reactor-off data with which to perform the subtraction. Regardless, the background rates would be intractably high.

Additionally, a chiller breakdown occurred in May 2019 that resulted in the container temperature reaching 50 °C before the electronics could be switched off and the chiller restarted. There was no evidence of damage to any equipment but it should be noted that an ^{22}Na calibration performed immediately after the event showed a distinct dip in PVT light yield, which is incompatible with the prior trend over time [207]. This can be seen in Fig. 2.23 between the 5$^{\text{th}}$ and 6$^{\text{th}}$ calibration points.

4.1.2 Reactor Schedule

The thermal power of the BR2 reactor is computed in two ways during a cycle. The first, known as MWRA4-1301, is derived from the temperature and water-flow rate in the primary cooling circuits (see Fig. 2.1), measured with resistor thermometers and Dall tubes respectively [193]. The second method, known as A3, is based on the activity of ^{16}N in the cooling water created by moderated neutrons in the reaction: ${}^{16}_{7}\text{N} + \text{n} \rightarrow {}^{16}_{7}\text{N} + {}^{1}_{1}\text{H}$. The absolute A3 powers are anchored to the MWRA4-1301 values, as the former can only track the evolution of the power over time. Measurements of the power are made every minute. Unless stated otherwise, a quoted thermal power is from the MWRA4-1301 database but missing or invalid entries are filled from the A3 table.

The thermal power over a two year period is shown in Fig. 4.1. The reactor cycles are summarised in Tab. 4.1 in terms of their duration, predicted IBD yield, and mean power. The detector was exposed to 13 reactor cycles over Phase-I, but the IBD yield has only been predicted for a subset of the cycles. The reactor is formally defined as being in a *reactor-off*

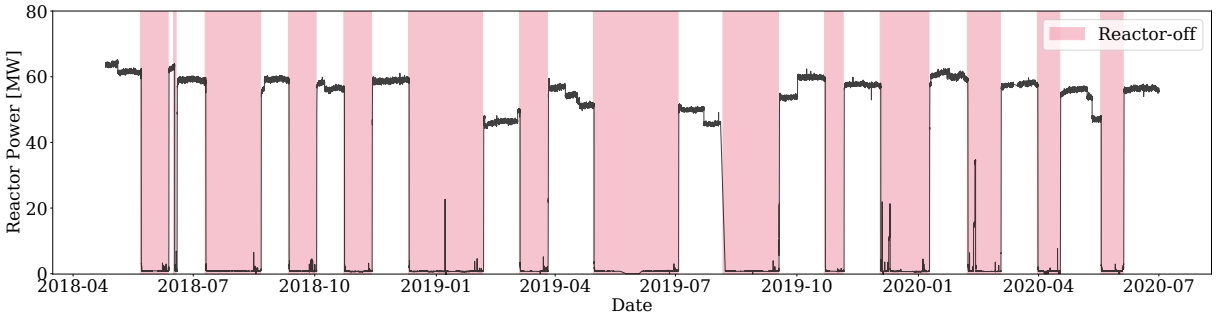


Figure 4.1: Thermal power of the BR2 reactor over the entire Phase-I data-taking campaign. The red bands denote reactor-off periods when the measured power was less than 1 MW_{th}.

No.	Dates	Days	Power [MW_{th}]	IBD Yield
1	12/06/2018 – 10/07/2018	28	60	31558.6
2	21/08/2018 – 11/09/2018	21	58	25083.3
3	02/10/2018 – 23/10/2018	21	56	24232.0
4	13/11/2018 – 11/12/2018	28	58	31639.8
5	05/02/2019 – 05/03/2019	27	45	25862.6
6	26/03/2019 – 30/04/2019	35	53	29038.9
7	03/07/2019 – 06/08/2019	35	48	–
8	17/09/2019 – 22/10/2019	35	54	–
9	05/11/2019 – 03/12/2019	28	58	–
10	09/01/2020 – 07/02/2020	29	60	–
11	03/03/2020 – 31/03/2020	28	58	–
12	17/04/2020 – 18/05/2020	31	56	–
13	04/06/2020 – 02/07/2020	27	56	–

Table 4.1: BR2 reactor schedule over the Phase-I data-taking campaign. The IBD yield is calculated using equation 2.3.

(ROff) state when the power is less than 1 MW_{th}; it is in a *reactor-on* (ROn) state if the power is greater than 40 MW_{th}. DAQ runs taken when the reactor is not in either state are excluded in order to have a clean separation between background and background-plus-signal data. Approximately 2% of the physics-mode runs fall in the transition zone between 1 MW_{th} and 40 MW_{th}, which typically occur at the boundaries of a normal cycle when the reactor is powering up or down, or during reactor scrams*.

4.1.3 Environmental Conditions

Environmental conditions within the container can affect the quality of the collected data. The performance of the electronics is directly affected by their temperature. The temperature inside the container during a run is required to fall within the range 10.5–12 °C, else the

*Tests of the emergency shutdown procedure.

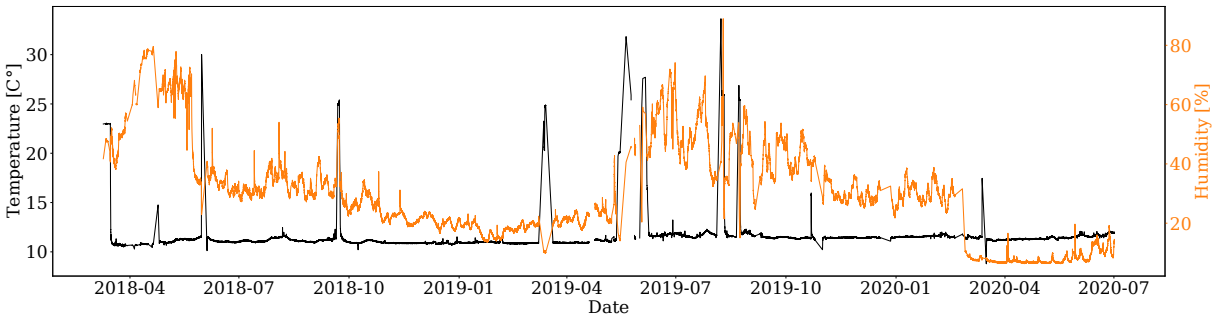


Figure 4.2: Temperature and humidity in the detector environment over the entire Phase-I data-taking campaign, as measured by sensors within the container. Large humidity variations in short time frames can be attributed to changes to the dry-air flushing rate. Temperature spikes are mostly associated with planned interventions.

run is rejected. This temperature range is large enough to capture the natural variations in chiller efficiency but exclude chiller stoppages (deliberate or otherwise). Moreover, high humidity can affect the neutron trigger rate through increased radon levels in the air, which directly impacts the number of background events from radioactive nuclei.

The humidity was measured to be persistently above 60% in April–May 2018. From 22$^{\text{nd}}$ May 2018 onwards, an external dry air supply was used to constantly flush the air inside the container, which reduced the humidity by more than 20%. A more precise flushing was introduced on 27$^{\text{th}}$ June 2019 but the flushing rate was not settled until 22$^{\text{nd}}$ August. The first reactor cycle, starting 24$^{\text{th}}$ April, coincides almost exactly with the no-flushing period so it was decided that data taken during this cycle would be dropped. The temperature and humidity inside the container as a function of time is plotted in Fig. 4.2. The initial period of high humidity is clearly visible, as well as a relatively humid interval starting in June 2019. The sudden drop in humidity in February 2020 is explained by an increase to the flushing rate from $5\,\mathrm{m}^3\,\mathrm{h}^{-1}$ to $11\,\mathrm{m}^3\,\mathrm{h}^{-1}$. Other adjustments to the flushing rate were made infrequently but did not cause such a clear effect. For the most part, the intermittent temperature spikes were caused by planned interventions.

4.1.4 Muon Bursts

Studies of reactor-off data identified sharp increases in muon rate at apparently random times [205]. It was realised that these "bursts" are correlated with occasions when the reactor pool was drained for maintenance purposes. The reactor pool blocks a fraction of the incoming muons and atmospheric neutrons, so the number of such particles incident on the detector increases when the water level is lowered. The pool level data is readily available from BR2 personnel, with a value recorded every 30 minutes. Each DAQ run is associated with the pool level closest in time and missing entries are linearly interpolated

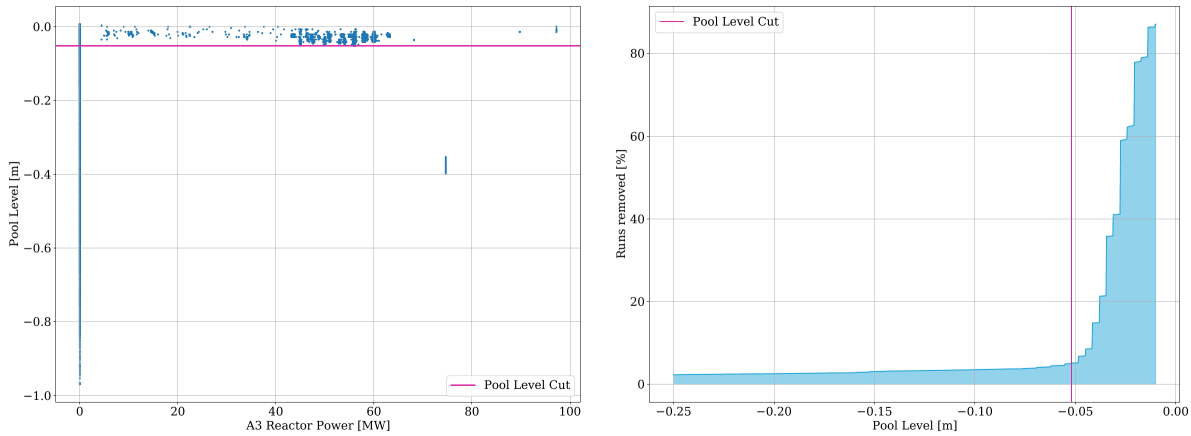


Figure 4.3: Pool height drop with respect to full against the A3 reactor power (left) and a cumulative distribution showing the percentage of DAQ runs that are rejected as a function of pool-level cut (right). The pink lines in each plot indicate the value of the applied cut.

from the neighbouring values.

The pool level as a function of A3 power is given in the left-hand side of Fig. 4.3. It decreases by 0.054 m at most when the reactor is running, but drops by up to one metre during reactor-off periods. A cut is applied to remove all runs taken when the pool was filled below −0.054 m, which retains all reactor-on data and ensures a consistent level of shielding (i.e. of background) for all data. Approximately 3% of reactor-off runs are removed by this cut, as evident in the right-hand side of Fig. 4.3.

4.1.5 DAQ & Saffron2 Parameters

The final data-quality criteria are based on a combination of five low-level DAQ parameters and three Saffron2-reconstructed parameters, as outlined in Tab. 4.2. Around 97% of the processed runs pass the cuts, which are designed to dispose of obvious outliers in the data. The evolution of the parameters over time is displayed in Fig. 4.4, where the black dotted lines indicate the accepted limits. The corresponding histograms can be seen in Fig. 4.5. A number of remarks can be made:

- The run duration plot exhibits a flat line with sharp troughs. The troughs are mostly 10 s test runs—with atypical trigger settings—which are correlated with the peaks and troughs of the random trigger.

- The mean neutron trigger and threshold trigger rates decrease slightly over time, compatible with the evolution of the PVT light yield. The remaining parameters are relatively stable over the two years.

- The threshold trigger is sensitive to changes in the reactor power as a result of increased

Parameter	Limits
Run Time [s]	(450, 500)
Dead-Time Fraction	(0, 0.06)
Random Trigger Rate [Hz]	(57, 65)
Threshold Trigger Rate [Hz]	(1800, 2500)
NS Trigger Rate [Hz]	(0, 85)
Type 1 Muon Rate [Hz]	(40, 60)
Type 2 Muon Rate [Hz]	(200, 250)
BiPonisher	(1.3, 1.4)

Table 4.2: Data-quality criteria based on DAQ-level and Saffron2 parameters that a given run must satisfy to be included in the final dataset. The parentheses indicate exclusive limits. BiPonisher is a variable constructed from the shape of the NS waveform for pulse shape discrimination.

gamma emission during reactor operation, manifesting as a double-peak structure in the threshold trigger histogram.

- There is a slight double-peak structure in the BiPonisher[†] distribution; investigations ascertain that it is correlated with humidity.
- The distribution of neutron trigger rate has a small tail at frequencies greater than 75 Hz which is also correlated with periods of high humidity.

4.2 Background Sources

4.2.1 Fast Neutron

One of the major background sources of an antineutrino signal is fast neutrons, for which there are two main production mechanisms. The first is spallation of atmospheric nuclei in collisions with primary cosmic rays. The resulting neutron flux comprises neutrons with a wide temperature spectrum that are able to reach the detector volume. Fast neutrons can create multiple ES deposits from proton recoil interactions and subsequently be captured on ^{6}Li—after reaching thermal energies—to produce the NS. The time difference between the signals ($\Delta t \equiv t_{\mathrm{NS}} - t_{\mathrm{ES}}$) follows an exponential function, thus this background is said to be *time-correlated*. The capture time for fast neutrons is very similar to that of IBD neutrons, despite the orders of magnitude difference in energy, because of the exponential energy loss experienced by the neutron in elastic collisions.

A second source of fast neutrons is cosmic-ray-induced muons that cause spallation of atomic nuclei in the detector itself or surrounding material. The former constitutes about

[†]BiPonisher is a PSD variable designed for the NS waveform, which is introduced in a later section.

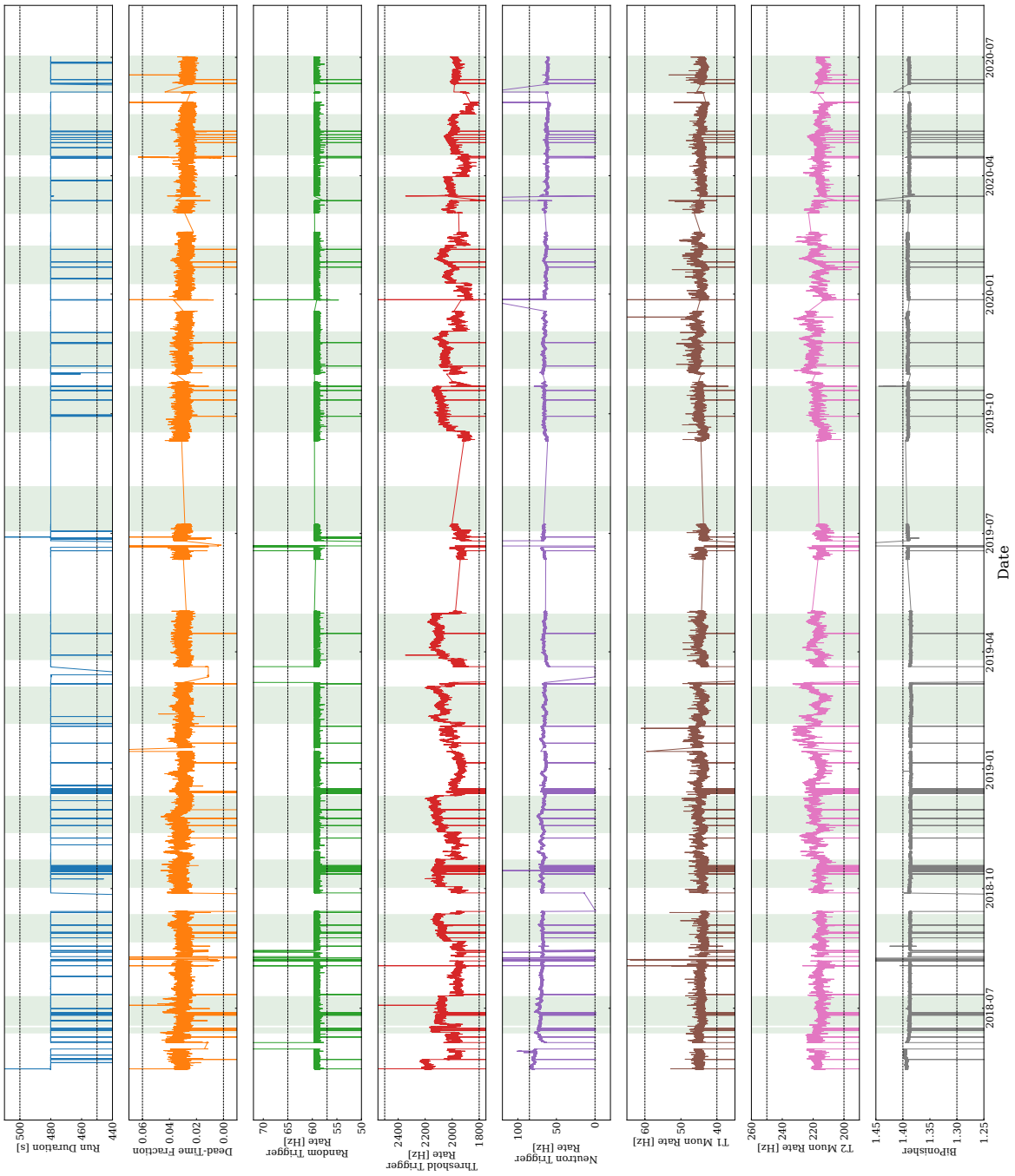


Figure 4.4: Evolution of eight DAQ-level and Saffron2 parameters over time for all Phase-I runs. The dotted black lines indicate the accepted upper and lower bounds of each parameter as in Tab. 4.2.

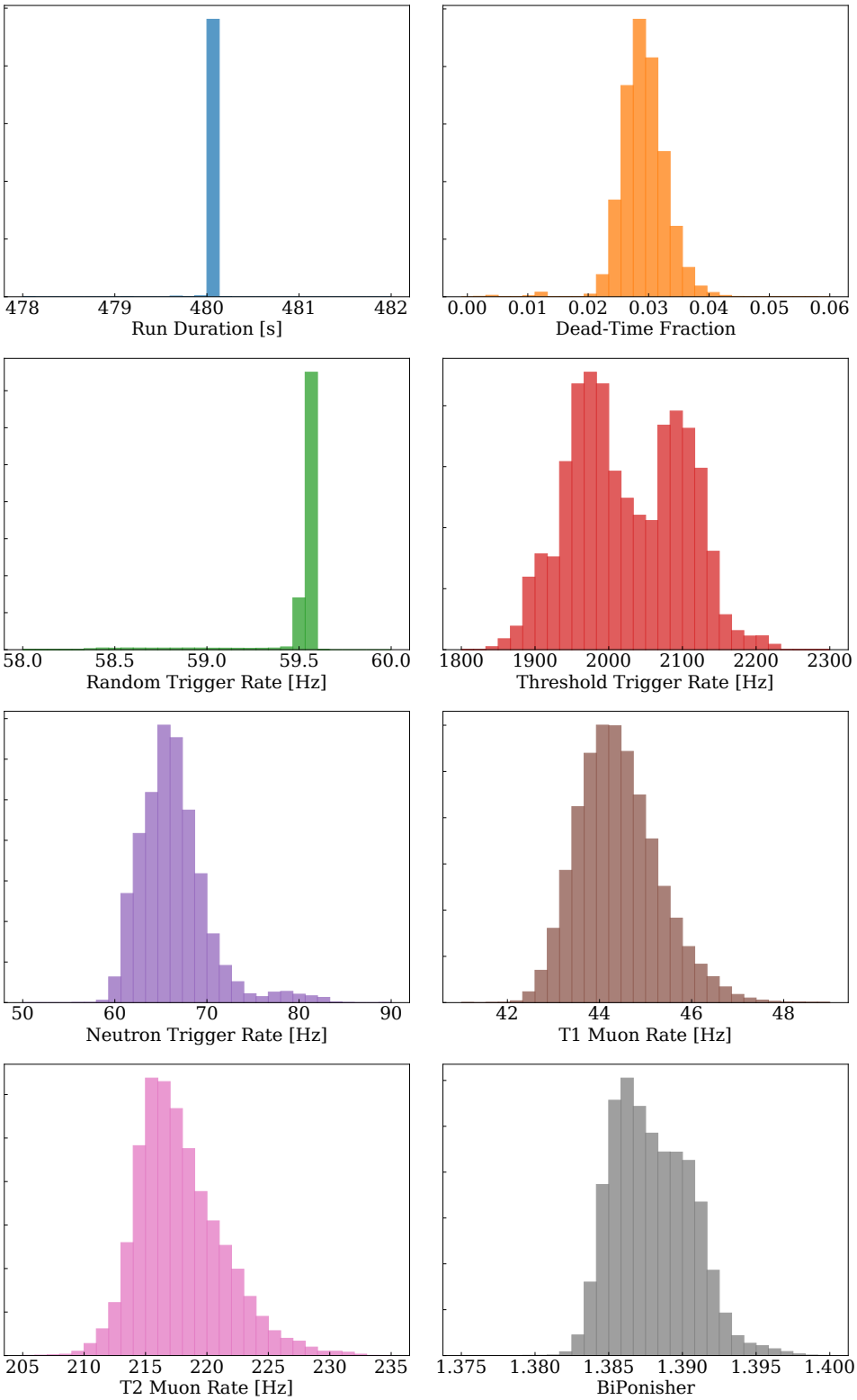


Figure 4.5: Histograms of eight DAQ-level and Saffron2 parameters for Phase-I runs that satisfy the criteria outlined in Tab. 4.2. The vertical scales are arbitrary.

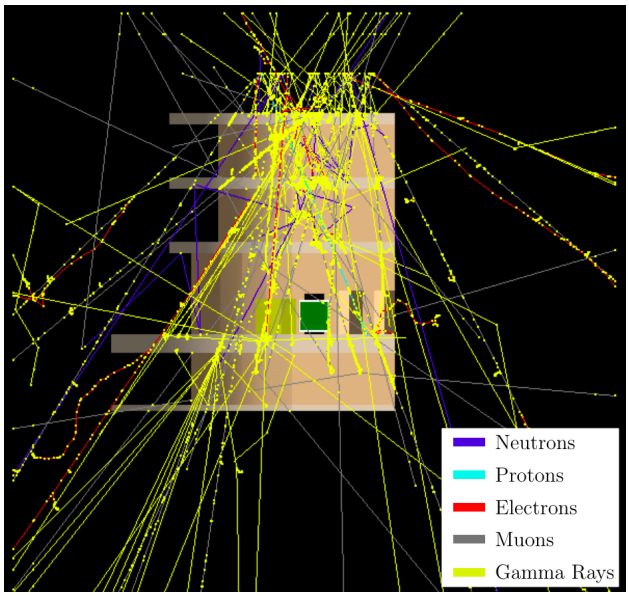


Figure 4.6: A cosmic-ray-induced shower simulated with CRY [223], courtesy of the SoLid collaboration. The secondary particles emanate from a $6 \times 6\,\mathrm{m}^2$ square situated 12 m above the first detector module. Fast neutrons are the most troublesome shower particles in terms of an IBD-like background.

a third of such events. A cosmic-ray shower simulated with the CRY software library [223] is illustrated in Fig. 4.6. The shower has been generated from a $6 \times 6\,\mathrm{m}^2$ square surface situated 12 m directly above the first detector module. The tracks of the secondary particles through the reactor building are clearly visible, including points at which they interact with matter. Despite the passive shielding, SoLid operates with little overburden and lacks neutron-gamma PSD capabilities, so the fast-neutron rate is extremely high compared to the expected IBD rate. The deposited energy in the SoLid detector from proton recoils of fast neutrons extends across the entire IBD energy range but is peaked at low energy.

4.2.2 Radioactivity

The second type of time-correlated background originates from radioactive elements in the air or detector material. The bismuth isotope $^{214}_{83}$Bi is found in the ^{238}U decay chain. It decays via beta emission of an electron (99.98% branching fraction [224]) to $^{214}_{84}$Po with a half-life of around twenty minutes and a Q-value of 3.270 MeV, which in turn decays to $^{210}_{82}$Pb after alpha particle emission. This background is known as BiPo. Approximately 20% of the bismuth nuclei decay to the ground state of $^{214}_{84}$Po, in which case only an electron is emitted resulting in a well contained scintillation signal. However, in the remaining 80% of cases a short-lived excited state of $^{214}_{84}$Po is created that releases gamma rays up to 3.18 MeV. These "radiative BiPo" events lead to complicated ES topologies. The energy spectrum of the BiPo electron is shown in Fig. 4.7, where it is compared with the spectrum of the IBD positron.

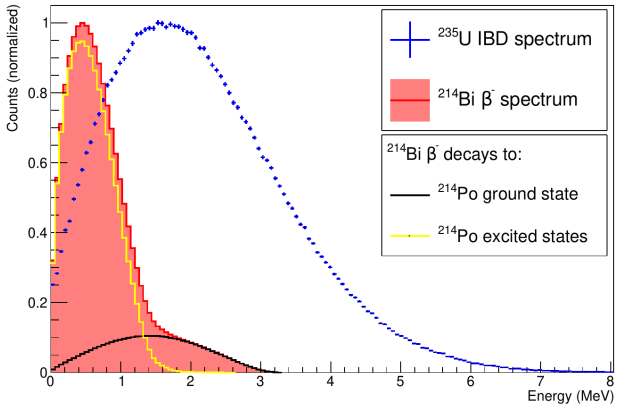


Figure 4.7: Simulated BiPo electron energy spectrum compared with the IBD positron spectrum [225]. The BiPo spectrum is decomposed into the radiative and non-radiative components. BiPo is the dominant antineutrino background below 1 MeV.

The combined BiPo spectrum is made up of the radiative and non-radiative components; the latter dominates above 1.5 MeV. BiPo background is most relevant below 1.0 MeV.

The principle source of BiPo is ^{238}U contamination in the ^{6}LiF:ZnS(Ag) neutron screens, as the beta electron can induce a signal in the PVT followed by scintillation of the alpha particle in the ZnS. Airborne nuclei (predominantly radon emanated from the concrete of the floors and walls) also contribute to the BiPo background. The isotope $^{222}_{86}$Rn is also found in the ^{238}U decay chain upstream of $^{214}_{83}$Bi, so it can create BiPo events in exactly the same way as the intrinsic component. The level of humidity-dependent radon in the container is mitigated by continuously flushing the container with an external dry air supply. The intrinsic contamination cannot be easily quelled and means the raw BiPo rate is even higher than that of fast neutrons.

The decay constant of $^{214}_{84}$Po is 235.8(6) µs [226], similar in order of magnitude to the neutron capture time of 62 µs. Hence, the two signals can be picked up by the clustering algorithm and formed into an NS-ES coincidence. Furthermore, the energy of the beta electron is in the 0–3 MeV range which heavily overlaps with the IBD energy spectrum. The NS waveforms produced by BiPo are solely from the 7.7 MeV alpha particle, as opposed to those from neutron capture events which are the consequence of scintillation by the alpha and tritium nuclei with a total energy of 4.78 MeV. The ionisation density of the latter interaction means that electrons in the ZnS are excited to higher energy levels compared to the former, so the corresponding waveform has a slower average decay time. The mean waveforms of neutron and alpha signals are depicted in Fig. 4.8.

The differences in shape between the two NS waveform types are exploited by two PSD approaches to suppressing the BiPo background. They were developed and tested by the collaboration on an AmBe calibration (neutron) dataset and a BiPo-enriched (alpha) sample [211]. The first method is a charge integration variable, known as *BiPonisher*, computed as

the ratio of a long integral over a short integral:

$$\text{BiPonisher} = \frac{\int_0^{37.5} A(t)}{\int_0^{7.5} A(t)} \,, \tag{4.1}$$

where $A(t)$ is the amplitude of the NS waveform as a function of time, in units of µs. BiPonisher allows us to reject 87% of alpha signals for 80% neutron efficiency. The second method is a one-dimensional convolutional neural network (CNN) dubbed the *BiPonator*. Compared to BiPonisher, BiPonator is able to reject 95% of alpha waveforms at the same neutron efficiency. The distributions of BiPonisher and BiPonator for the two training datasets are provided in Fig. 4.8 along with a comparison between their Receiver Operating Characteristic (ROC) curves.

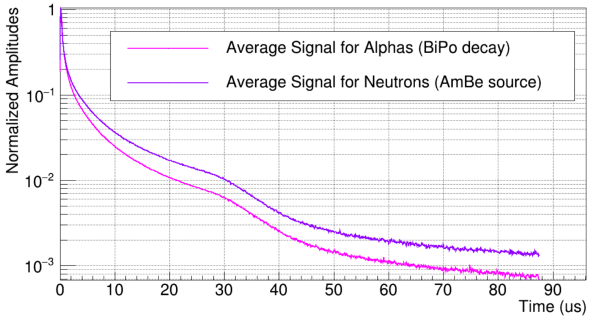
It is not currently possible to simulate NS waveforms in the ROSim framework, so MC events are assigned values for BiPonisher and BiPonator that are randomly sampled from the appropriate data distribution[‡]. The stability of BiPonator performance over time has been tested by training multiple models on data taken at different periods during Phase-I [211]. The outputs of all models are found to be statistically consistent so in this thesis the original model (trained on the September 2018 AmBe calibration dataset and equivalent BiPo-enriched reactor-off data) is used exclusively.

4.2.3 Accidental

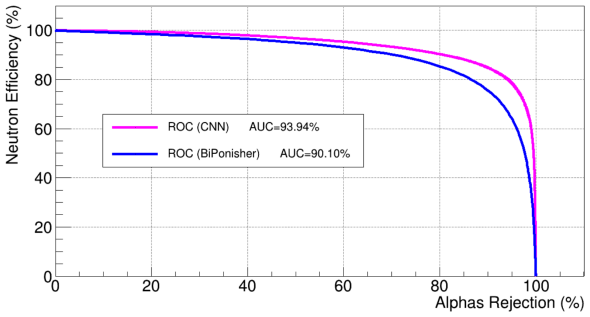
The final type of background event is known as "accidental". Accidentals originate from coincidental clustering of ES and NS waveforms that are uncorrelated in time and space. For example, the ES may be from gamma rays or Type 0 muon deposits, whilst the capture of thermalised neutrons can create the NS. The rate of accidentals is slightly higher during reactor-on periods because of neutron emission from the reactor core which causes activation in the vessel walls, in turn releasing gamma rays. The ROn-ROff difference in gamma rate has been verified by measurements taken by a dedicated sodium iodide (NaI) crystal scintillator detector.

Two methods are employed to retrieve accidental events. The simplest way is to identify NS-ES clusters in which the NS waveform was measured before the ES waveform, i.e. events in the negative $\Delta t \equiv t_{\text{NS}} - t_{\text{ES}}$ regime. Since the two waveforms are uncorrelated in time, the rate of accidentals in the negative Δt window should be equal to the rate in the positive Δt rate. However, it is observed that the Δt distribution in the accidental window has a slightly non-uniform behaviour. This is likely caused by contamination from 2.2 MeV gamma rays emitted after neutron capture on hydrogen in high-multiplicity neutron showers, which can manifest in the accidental window.

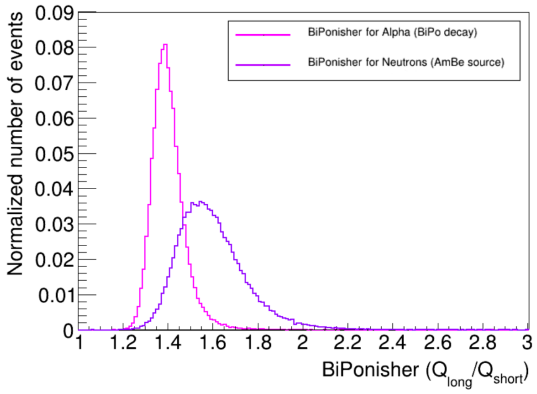
[‡]The alpha-rich dataset for BiPo MC and the AmBe calibration data for IBD MC.



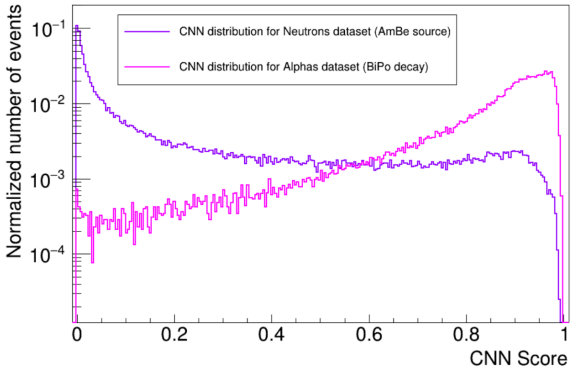
(a) Normalised average shapes of alpha- and neutron-induced NS waveforms.



(b) ROC curves of BiPonisher and BiPonator (CNN).



(c) BiPonisher distributions for alpha and neutron waveforms.



(d) BiPonator (CNN) distributions for alpha and neutron waveforms.

Figure 4.8: Comparison between BiPonisher and BiPonator (CNN). The latter offers more efficient BiPo rejection for a given neutron efficiency, as proved by their ROC curves. Figures courtesy of the SoLid collaboration.

The second method avoids this issue by directly measuring the accidentals in the positive Δt region via the False Positive Neutron Trigger (FPNT) sample. This contains events in which the neutron trigger prompts the read out of the FPGA buffer but no NS cluster is reconstructed, thus the data contains only accidental background. The majority of FPNTs are produced by the low-energy tail of muon signals that satisfy the neutron trigger conditions and the rest are from real NS clusters that are not reconstructed. The FPNT rate is roughly four times higher than the NS cluster rate, so the algorithm tries to associate each FPNT with both a muon cluster—registered in the same plane and at the same time block—and a previously reconstructed NS in the same plane not already associated with an FPNT. If this is impossible the FPNT is abandoned. A "fake" NS cluster is then created at the time of the FPNT with the properties of the reconstructed NS. Coincidences between the fake NS cluster and ES clusters that are made using the standard clustering algorithm to quantify the number of accidental events created by the real NS.

4.2.4 Composition of Reactor-off Data Before Selection

It is possible to perform a fit on the binned data to extract the absolute rates of each background, assuming no unidentified event types exist in the data in significant quantities, since the three backgrounds have distinct Δt distributions. Unless stated otherwise, in this work a "fit" of a function is always by weighted least-squares. The Δt distribution can be described by the following equation:

$$y(\Delta t) = c_0 + c_1\, e^{\Delta t/\tau_{\mathrm{BiPo}}} + c_2\, e^{\Delta t/\tau^{\mathrm{Th}}_{\mathrm{Neutron}}} + c_3\, e^{\Delta t/\tau^{\mathrm{Epi}}_{\mathrm{Neutron}}}\,, \tag{4.2}$$

where the total fast-neutron component is a double exponential comprising contributions from both standard thermal neutrons and captured epithermal neutrons. The corresponding time constants are labelled $\tau^{\mathrm{Th}}_{\mathrm{Neutron}}$ and $\tau^{\mathrm{Epi}}_{\mathrm{Neutron}}$ respectively. Their values depend on the density of ^{6}Li nuclei in the detector and have been estimated (in private SoLid studies) as 68.4(6) µs and 10.1(4) µs. The parameters c_i quantify the rate of each background type in the selected data and τ_{BiPo} is the decay constant of BiPo.

A fit is performed on reactor-off data from June 4$^{\mathrm{th}}$–8$^{\mathrm{th}}$ 2018 to evaluate the background composition before selection, as in Fig. 4.9. This sample corresponds to 3.5 d of detector exposure time and 2 899 971 events, or around 828 563 events per day. For this particular fit, the constant term c_0 parameterising the accidental rate is constrained to the value (within one-sigma uncertainty) provided by the FPNT sample. The reduced chi-square statistic (χ^2/ndf) of the fit is 40.4/26, indicating that equation 4.2 parameterises the distribution reasonably well, though improvement would be desirable.

BiPo dominates the entire Δt range, except above 40 µs where accidentals become very significant. The contribution from fast neutrons is only relevant in the region below 200 µs

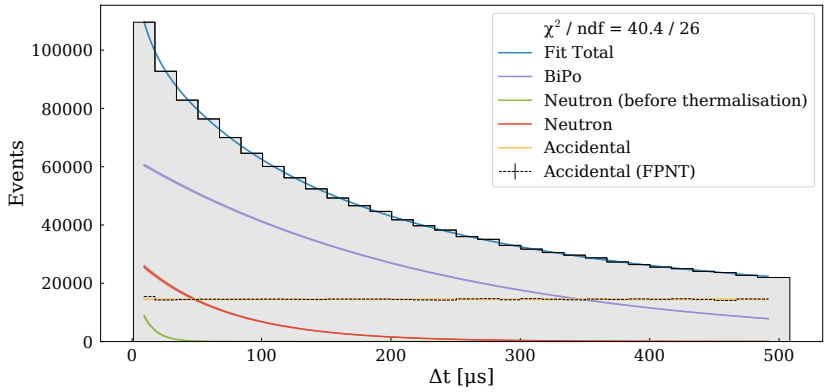


Figure 4.9: Δt distribution of June 4^{th}–8^{th} 2018 reactor-off data before selection, determined by a fit of equation 4.2. The dashed black line shows the histogram from the equivalent FPNT sample to which the flat component of the fit is constrained. The fit is decomposed into its individual components. BiPo is the dominant background across the vast majority of the Δt range up to 500 µs.

which is also populated by IBD events in reactor-on data. The relative proportions of accidentals, BiPo, and fast-neutron events in the data are roughly 32:58:10 respectively. Fewer than 1% of the fast-neutron events are from epithermal captures. The combined efficiency of neutron capture and Saffron2 reconstruction is around 40% for IBD MC events. Therefore, an initial signal-to-background (S/B) ratio of 0.00133 is estimated from the data assuming 1100 IBD interaction in the detector per day.

4.3 Preselection Cuts

At this point we have a well-defined, high-quality dataset overwhelmingly dominated by background events. A set of rectangular cuts have been designed, building on previous work [206, 211], to provide an initial reduction in background rate and enhance the proportion of IBD events in the data, prior to the deployment of ML models. The intention behind these "preselection" cuts is to retain a relatively high efficiency for IBD events, whilst rejecting the most background-dominated regions of parameter space.

A representative reactor-off sample is built by randomly selecting 15% of the data from *all* available daily reactor-off files. A sample of IBD MC events is taken from the simulation of the first reactor cycle. The number of MC events is high enough for statistical uncertainties on each bin to generally be ignored. The preselection cuts listed in Tab. 4.3 were defined by inspection of the probability density functions (PDFs) of available Saffron2 variables. The PDFs of the chosen preselection variables are displayed in Fig. 4.10.

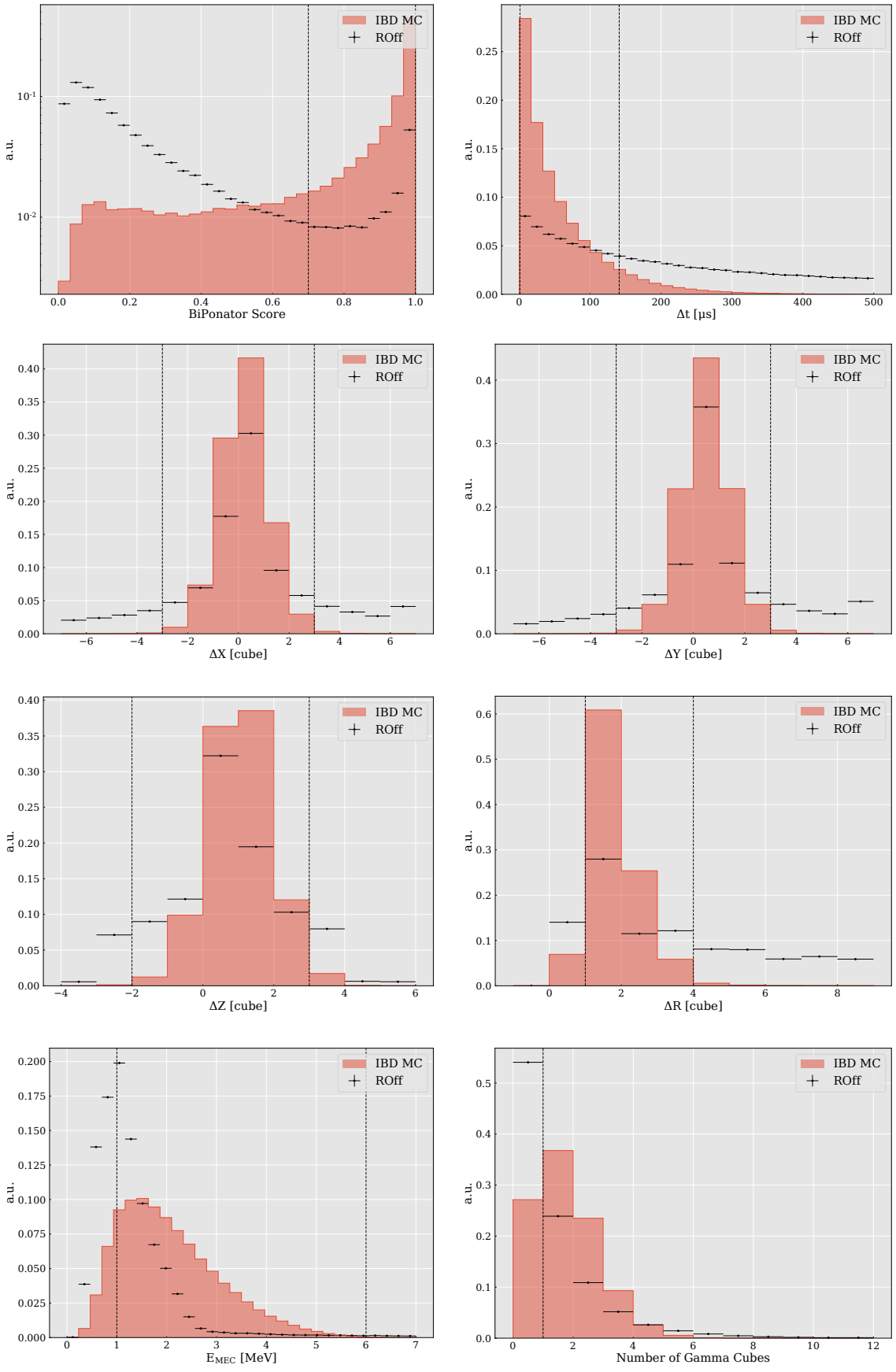


Figure 4.10: Probability density functions of the variables included in the preselection criteria, except for the muon veto. The vertical dashed lines show the cut limits of each variable.

Description	Variable and Limits	Sig. (Bck.) Efficiency
NS-ES displacements	$\Delta X \in [-3, 3]$	99.6 (68.7)
	$\Delta Y \in [-3, 3]$	99.6 (68.5)
	$\Delta Z \in [-2, 3]$	99.7 (89.2)
	$\Delta R \in [1, 4]$	92.0 (40.4)
BiPo-focused	$\Delta t \in [1, 141]$ µs	89.0 (38.6)
	BiPonator $\in [0.7, 1.0]$	75.7 (13.3)
Other	$E_{\mathrm{MEC}} \in [1.0, 6.0]$ MeV	87.2 (57.8)
	NumGammaTotal $\in [1, \infty]$	73.0 (46.1)
	tNearestMuonType1 $\in [200, \infty]$ µs	n/a (91.9)
	tNearestMuonType2 $\in [200, \infty]$ µs	n/a (91.9)

Table 4.3: Preselection criteria of the antineutrino analysis. The efficiency of a given cut in the final column is calculated when applied in the absence of the other cuts. The brackets indicate inclusive limits. The most powerful background suppression is provided by BiPonator and ΔR.

4.3.1 BiPo Suppression

BiPonator score and Δt bring no discrimination between IBD and fast-neutron events since the respective NS clusters arise from the same physical process. Thus the cuts on these variables are focused purely reducing the enormous BiPo rate. The two variables are uncorrelated so provide a combined background rejection of around 95%. The Δt <1 µs region is excluded because it contains self-correlated events in which a muon generates both the ES and the NS waveform. The lower limit on the energy of the most energetic cube (E_{MEC}) cuts 42% of the background events, of which 70% are BiPo and 28% are accidental, at the expense of 13% of the IBD events. The upper limit is simply set at the "end" of the IBD spectrum such that only a further 0.06% of events are excluded.

In analogy with Δt, the spatial difference between the NS and the ES is defined as:

$$\Delta R_{\mathrm{NS-ES}} \equiv \Delta R = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}\,, \tag{4.3}$$

where ΔX, ΔY and ΔZ are the spatial displacements of the ES cluster from the NS cluster[§]. The cuts on NS-ES spatial displacements are not specifically focused on suppressing BiPo, so they can retain high efficiency for signal events. On the other hand, although the $\Delta R = 0$ bin contains a non-trivial fraction of the IBD events, it is also estimated to be populated with 99% pure BiPo background. This is a consequence of the specific topology of BiPo events: the alpha always generates a signal in the ^{6}LiF:ZnS(Ag) screen in which it was created, so depending on the electron direction of travel the two signals will likely be displaced by zero or one cube(s). In contrast, thermalising neutrons which take a random walk over $\mathcal{O}(10)$ cm distances are unlikely to remain in their original cube. Discarding the $\Delta R = 0$ bin reduces the overall fraction of BiPo in the data and the ΔR upper bound rejects the vast majority

[§]The position/origin of the ES cluster is defined as the position of the most energetic cube.

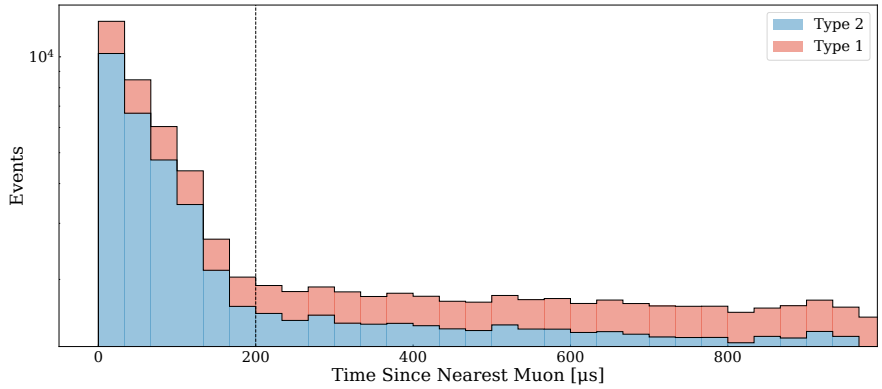


Figure 4.11: Time between reconstructed NS and nearest preceding muon (Type 1 or Type 2) for reactor-off data from June and July 2018. Only NS waveforms that are coincident with an ES are shown. The two histograms are stacked. The cut, represented by the dashed black line, was designed to remove NS-ES coincidences in which the NS was created by muon spallation.

of accidentals.

4.3.2 Topological Considerations

A little over 50% of the background events belong to the 0γ category, where NumGammaTotal is zero. These events are often located near the borders of the detector where annihilation gammas (or fast neutrons) are more likely to escape before depositing additional energy. Previous SoLid analyses [207] found that these events have extremely poor S/B ratios because they lack any topological information with which to separate signal from background. 0γ events are therefore dispensed with in the preselection, despite the relatively draconian effect on the signal efficiency.

4.3.3 Muon Veto

Finally, all clusters within a 200 µs window after a muon (Type 1 or Type 2) passes through the detector are vetoed. Type 0 muons are not included because they are not reconstructed with sufficient efficiency. Within this window there is a sharp increase in the number of NS-ES coincidences, as shown in Fig. 4.11, that is highly suggestive of muon spallation. The distributions for Type 2 and Type 1 muons above 200 µs tend slowly to zero. Raising the threshold any further would simply incur detector dead-time for no good reason.

4.3.4 Composition of Reactor-off Data After Preselection

Taking the same reactor-off sample defined in section 4.2.4, Fig. 4.12 reveals how the background composition is modified by application of the preselection cuts. Fast neutrons,

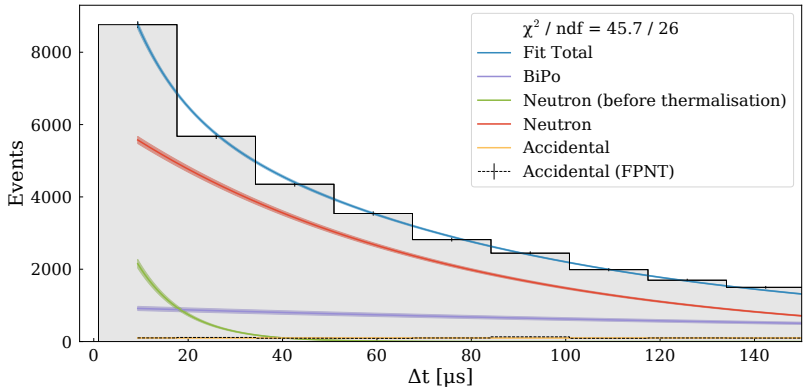


Figure 4.12: Δt distribution of June 4^{th}–8^{th} 2018 reactor-off data after preselection cuts, determined by a fit of equation 4.2. The dashed black line shows the histogram from the equivalent FPNT sample to which the flat component of the fit is constrained. The fit is decomposed into its individual components. Fast neutrons are now the dominant background across the entire Δt range up to 141 µs.

BiPo, and accidental events now exist in proportions 80:18:2 respectively. In other words, fast neutrons have become the dominant background over BiPo and the rate of accidentals is vanishingly small. The overall efficiency of the selection cuts is approximately 1.9% for background and 39% for signal, improving the overall signal-to-background ratio to around 0.0273.

4.4 Machine Learning Analysis

In principle, the preselection criteria could be optimised and expanded to other variables, particularly those related to the gamma-like and leftover cubes that would allow more of the background to be rejected. Instead, the analysis is continued by utilising machine learning tools which should allow better signal-background discrimination than a cut-based analysis alone, especially given the complex topologies captured by the detector segmentation.

4.4.1 Boosted Decision Trees

Decision trees are a type of supervised learning algorithm with a tree-like structure consisting of nodes and branches. They can be employed for both classification or regression tasks. Various algorithms have been developed to train decision trees, such as CART and C4.5 [227, 228]. In general, the *labelled* data is split according to the feature that optimises some metric, thus creating a new layer of nodes[¶] that each contain some portion of the dataset. The most widely used metrics are the Gini impurity index and the entropy. The process is

[¶]The number of layers is known as the depth of the tree.

repeated recursively on each new node until a stopping criterion is met. This approach is an example of a "greedy algorithm" because the split at each node is locally optimal, which in most cases only approximates a globally optimal model. A common stopping criterion is the maximum allowed depth of the tree, or the algorithm may simply halt because there are no additional splits that improve the metric. The specific way a model is trained depends on its numerous hyperparameters which must be tuned for a specific task.

The first node in the tree (depth zero) contains the entire training sample and is called the *root* node. The nodes that branch off from the root node are referred to as *decision* or *internal* nodes and contain some subset of the training data on which a further decision to split is made. Finally, all possible paths through the tree end at one of the *leaf* nodes which do not split the data any further. Leaf nodes are assigned a value corresponding to the label of the most frequently occurring class in the node, or alternatively some type of class probability. A trained decision tree can be used to predict the class labels of unseen data that share the same features as the training data. The predicted class of a data point depends on the leaf node at which it terminates. Compared to other ML models, decision trees are easy to interpret, do not require much data preparation, and can deal with both numerical data and categorical data. On the other hand, they can be prone to overfitting the data, are unstable with respect to small variations in the dataset, and do not handle unbalanced datasets well.

Some of these issues can be very effectively mitigated with *ensemble* methods, which can also significantly improve the accuracy of a model. Ensemble methods combine the outputs of many so-called "weak learners" to achieve a greater predictive power than any of the individual learners. A Boosted Decision Tree (BDT) is a type of ensembled machine learning model that has become very common in the field of high energy physics. One of the most powerful boosting methods is called *gradient boosting* [229, 230], named for gradient descent. In gradient boosting algorithms, one generally starts with a model F_0 that makes an initial prediction ($F_0(x_i)$) for the training labels (y_i) of each data point x_i, for $i \in \{1, 2, \ldots, n\}$. A differentiable loss function $L = L(y, F_0(x))$—dependent on the data labels and model predictions—is then defined. A second learner h is introduced which aims to minimise the loss function by training on its negative gradients:

$$-g(x_i) = \frac{\partial L(y_i, F_0(x_i))}{\partial F_0(x_i)} \,, \tag{4.4}$$

and is added to F_0 to form an updated model F_1:

$$F_1(x) = F_0(x) + \eta h_0(g(x)) \,, \tag{4.5}$$

where the learning rate η constrains the improvement provided by subsequent learners to prevent overfitting. The loss function should quantify the accuracy of the current prediction

such that the next learner can compensate for the errors of the existing model. For some loss functions the negative gradient of the loss function $L(y_i, F_m(x_i))$ is equal to the residual $y_i - F_m(x_i)$. Additional weak learners are added sequentially until some criterion is reached. At each "boosting round", i.e. the addition of each weak learner, the performance of the overall model is improved. The $(m+1)^{\text{th}}$ model can be expressed as:

$$F_{m+1}(x) = F_m(x) + \eta h_m(g(x)) \, . \tag{4.6}$$

The eXtreme Gradient Boosting (XGBoost) algorithm has been selected for this analysis [231]. It is one of the most popular and powerful models in use today because of its numerous advantages over other gradient boosting algorithms such as parallelised tree building to run more efficiently and regularisation terms to penalise overfitting. Furthermore, it is less "greedy" than other gradient-boosting algorithms because it builds each tree up to its maximum allowed depth before pruning[‖] the tree in reverse.

4.4.2 Training Variables

The categorisation of cubes depending on their relative energies and positions in the ES cluster was introduced in chapter 3. By definition, every event comprises at least one positron-like cube and the preselection criteria demand at least one gamma-like cube. Events are then sorted by the algorithmically determined number of detached gamma bunches. Six variables are provided automatically by these constructions: the total positron energy and multiplicity of positron cubes (E_{MAGE}, N_{MAGE}), the total energy and multiplicity of the 1$^{\text{st}}$ gamma bunch ($E_{\gamma 1}$, $N_{\gamma 1}$), and the total energy and multiplicity of the 2$^{\text{nd}}$ gamma bunch ($E_{\gamma 2}$, $N_{\gamma 2}$).

The relative fractions of events per gamma category for preselected events are shown in Fig. 4.13. Around half of the IBD MC events belong to the 1γ(=1c) class, comprising events with a single gamma-like cube, and the rest are split roughly equally between the 1γ(>1c) and 2γ categories. The reactor-off data has a similar proportion of 1γ(=1c) events but a smaller fraction of 2γ events compared to the signal. To advance the analysis it is necessary to formulate new variables that may provide extra signal-background discrimination and thus serve as BDT model features. The combined proportion of events containing at least two gamma cubes is around 50%, so the hope is that the topological information of such events can be maximally exploited by the BDT algorithms.

[‖]Reducing the complexity of the tree by removing redundant or detrimental splits.

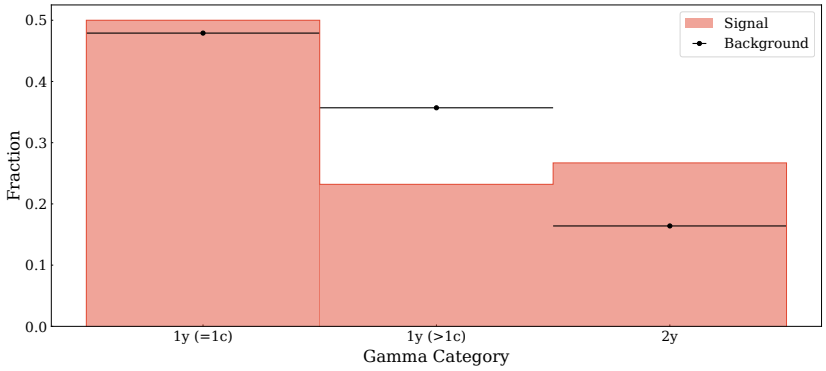


Figure 4.13: Relative fractions of each gamma category after applying preselection cuts on IBD MC (signal) events and reactor-off (background) events. Roughly half of all events after preselection belong to the least topologically complex $1\gamma(=1c)$ category, comprising events with a single gamma-like cube. 2γ events have the largest signal-to-background ratio.

Gamma Variables

Firstly, there are variables built from gamma-like cubes. For the majority of IBD events the two annihilation gammas are emitted in antiparallel directions with equal energy. Background sources produce gamma-like deposits via different processes and may differ from IBD events in energy and distance from the positron-like cubes. The position of a gamma bunch with respect to the MEC is parameterised by its energy-weighted "barycentre"'. For the j^{th} bunch which comprises $N_{\gamma j}$ cubes, the barycentre $\boldsymbol{B_{\gamma j}}$ is defined as:

$$\boldsymbol{B_{\gamma j}} = \frac{1}{E_{\gamma j}} \left(\sum_{n}^{N_{\gamma j}} E_{jn}\, \Delta X_{jn}\ ,\ \sum_{n}^{N_{\gamma j}} E_{jn}\, \Delta Y_{jn}\ ,\ \sum_{n}^{N_{\gamma j}} E_{jn}\, \Delta Z_{jn} \right) = (B^{x}_{\gamma j}, B^{y}_{\gamma j}, B^{z}_{\gamma j})\ , \tag{4.7}$$

where E_{jn} is the energy of the n^{th} cube in the bunch and ΔX_{jn}, ΔY_{jn}, ΔZ_{jn} are the displacements from the MEC of the n^{th} cube in each of the spatial directions. The index n runs over all the cubes in the bunch. Each event is assigned two barycentres regardless of the actual number of detached bunches located. The barycentre of a non-existent bunch is simply the null vector $(0, 0, 0)$. Five parameters are then defined for each event: the length of each barycentre vector, the absolute distance and dot product between the two barycentres, and the "extent" of all the gamma cubes:

$$\text{DistGamma1} = \|\boldsymbol{B_{\gamma 1}}\|\ ,\quad \text{DistGamma2} = \|\boldsymbol{B_{\gamma 2}}\|\ , \tag{4.8}$$

$$\text{DotProduct} = \frac{\boldsymbol{B_{\gamma 1}} \cdot \boldsymbol{B_{\gamma 2}}}{\|\boldsymbol{B_{\gamma 1}}\| \|\boldsymbol{B_{\gamma 2}}\|}\ , \tag{4.9}$$

$$\text{InterGammaDist} = \sqrt{\sum_{\alpha}(B^{\alpha}_{\gamma 1} - B^{\alpha}_{\gamma 2})^2}\,, \tag{4.10}$$

$$\text{Extent} = \sum_{n}^{N_{\gamma 1}} \Delta R_{1n} + \sum_{m}^{N_{\gamma 2}} \Delta R_{2n}\,. \tag{4.11}$$

where ΔR_{jn} is the absolute distance of the n^{th} cube from the MEC of the j^{th} bunch and the index α in equation 4.10 runs over the three coordinates of the vectors.

The length of the barycentre vector describes the degree to which the gamma bunch is displaced from the IBD vertex. For IBD events this should peak at the mean free path of 511 keV gamma rays in the detector (10.8 cm). The DotProduct distribution is heavily skewed towards negative values because of the emission direction of IBD annihilation gammas. This parameter is only sensible for 2γ events, otherwise the null gamma vectors would force its value to be zero. For all other categories it is set to -2 to avoid overlap with 2γ events that have perpendicular gamma vectors. The distance between the barycentres (InterGammaDist) also reflects the mean free path of annihilation gammas; it reduces to DistGamma1 if only one bunch exists. The Extent variable is a measure of the overall size of the ES cluster in terms of annihilation gamma deposits. For this reason it is highly correlated with other gamma variables such as InterGammaDist, making it somewhat redundant. Finally, there are the intra-bunch variables IntraGammaDistance, IntraGammaRatio, and IntraGammaAngle, respectively the distance, energy ratio, and angle between, the two highest energy cubes in the first gamma bunch. These focus on the correlation between cubes within the first gamma bunch, potentially improving the discrimination of $1\gamma(>1\text{c})$ events in particular.

Leftover & Other Variables

Next, variables from leftover cubes are considered. Through trial and error it is found that there are only two useful, non-trivial combinations of such cubes: the total leftover energy internal and external to the Cross. Other combinations are possible and may give equivalent results but are less well physically motivated. The equivalent variables defined by the leftover-cube multiplicity were found to provide very little additional information.

Finally, a set of "low-level" variables that are independent of the cube classification system can be defined. They comprise some of the variables used for preselection (e.g. ΔR) and other variables such as promptVolume and promptNCCube which describe the spatial "largeness" of the ES cluster, in a similar way as the Extent. The other low-level variables ESNS, ES2ES1, and ES3ES2 are legacy variables that were specifically designed to differentiate between IBD and non-radiative BiPo events, before BiPonator had been developed.

On the Training Variables

The normalised histograms of the available training variables in signal and background are plotted in Figs. 4.14 and 4.15. The full list of variables can be found in Tab. 4.4. The reactor-off E_{MEC} and E_{MAGE} spectra are dominated by low-energy fast-neutron and BiPo events. However, there is a also sizeable fast-neutron component at energies greater than 3 MeV which overlaps with the tail of the IBD spectra. Regarding the gamma energies, the background distributions extend higher in energy than for the signal events. The energy spectra of the first and second gamma bunches are not equivalent since the former is defined by the most energetic gamma-like cube, shifting the average $E_{\gamma 1}$ to higher energy than $E_{\gamma 2}$. The mode number of cubes is also higher. The two leftover energy variables appear very promising as discriminatory variables. The leftover energy external to the Cross is minimal for signal events, by construction, though there is a small amount of leftover energy within the Cross that is likely from light-leakage deposits.

The spatial distribution of events in the detector are given by X_{MEC}, Y_{MEC} and Z_{MEC}, i.e. the coordinates of the most energetic ES cube. The physical inverse beta decay vertices are distributed roughly evenly along the x- and y-axes. In contrast, the reconstructed signal distributions in X_{MEC} and Y_{MEC} exhibit a drop-off at the edges of the detector, which can be understood as the IBD neutron escaping from the fiducial volume. The Z_{MEC} signal distribution exhibits a similar decrease in efficiency at the first and last planes, superimposed on an approximately inverse-square law fall-off in rate. The background distributions of X_{MEC} and Y_{MEC} reach minima in the middle of the detector because the surrounding material effectively shields the internal voxels from incoming fast neutrons. This is not relevant for BiPo which is mostly intrinsic to the detector material itself and so cannot be shielded. The same effect is present for Z_{MEC}, but in this case the background rate increases towards the back of the detector (furthest from the reactor core) where the water-pool shielding is less substantial.

4.4.3 nuBDT Model

Given the pool of possible input features defined in the previous section, it is now possible to train a BDT model. The training dataset is assumed to have only two classes, signal and background, that are assigned the labels 1 and 0 respectively**. The background portion is the same 15% sample of reactor-off data as introduced previously. The sample contains 204662 events after application of the preselection cuts. A signal sample is created by randomly drawing the same number of events from the IBD MC simulation of the first reactor cycle. For this purpose it does not matter from which IBD MC cycle the events

**The preselection cuts augment the proportion of fast-neutron events in the reactor-off data to the extent that they make up around 80% of the background events.

Name of Variable	**Description**
EnPositron [E_{MAGE}]	Reconstructed positron energy estimator (sum of positron-like cube energies)
NumPositron [N_{MAGE}]	Number of positron-like cubes
EnHighestECube [E_{MEC}]	Energy of the most energetic ES cube (defines the local origin of the cluster)
X_{MEC}, Y_{MEC}, Z_{MEC}	Spatial coordinates of the most energetic ES cube
ΔX, ΔY, ΔZ	Displacements between the NS and the most energetic ES cube
ΔR	Absolute distance between the NS and the most energetic ES cube
EnGamma$j$ [$E_{\gamma j}$]	Total energy of the j^{th} gamma bunch
NumGamma$j$ [$N_{\gamma j}$]	Number of cubes in the j^{th} gamma bunch
DistGamma$j$ [$\|\boldsymbol{B}_{\boldsymbol{\gamma j}}\|$]	Distance of the barycentre of the j^{th} gamma bunch from the most energetic ES cube
DotProduct	Dot product between the gamma bunch unit vectors (set to -2 if either are the null vector)
Extent	Sum of gamma-like cube distances from the most energetic ES cube
EnGammaNotCrown	Sum of gamma-like cube energies external to the Crown
InterGammaDist	Absolute distance between the barycentres of the gamma bunches
IntraGammaDist	Distance between the second-most energetic gamma-like cube and most energetic cube in the 1^{st} gamma bunch
IntraGammaRatio	Ratio between the energy of the second-most energetic gamma-like cube and the energy of the most energetic cube in the 1^{st} gamma bunch
IntraGammaAngle	Angle between the second-most energetic gamma-like cube and most energetic cube in the 1^{st} gamma bunch
EnLeftoverCross	Sum of leftover cube energies internal to the Cross
EnLeftoverNotCross	Sum of leftover cube energies external to the Cross
promptVolume	Volume of the smallest box that could contain all cubes in the reconstructed ES cluster
promptNCCube	Number of cubes in the reconstructed ES cluster
ES2ES1	Ratio of the energy of the second-most energetic cube over the energy of the most energetic cube
ES3ES2	Ratio of the energy of the third-most energetic cube over the energy of the second-most energetic cube
ESNS	Energy in the cube that measured the NS cluster (set to -1 if no ES energy was deposited in the NS cube)

Table 4.4: Overview of the available BDT training variables. Alternative names or mathematical formulations of a variable are shown in brackets, where appropriate.



Figure 4.14: Probability density functions of positron and low-level training variables for signal (IBD MC) and background (reactor-off data).

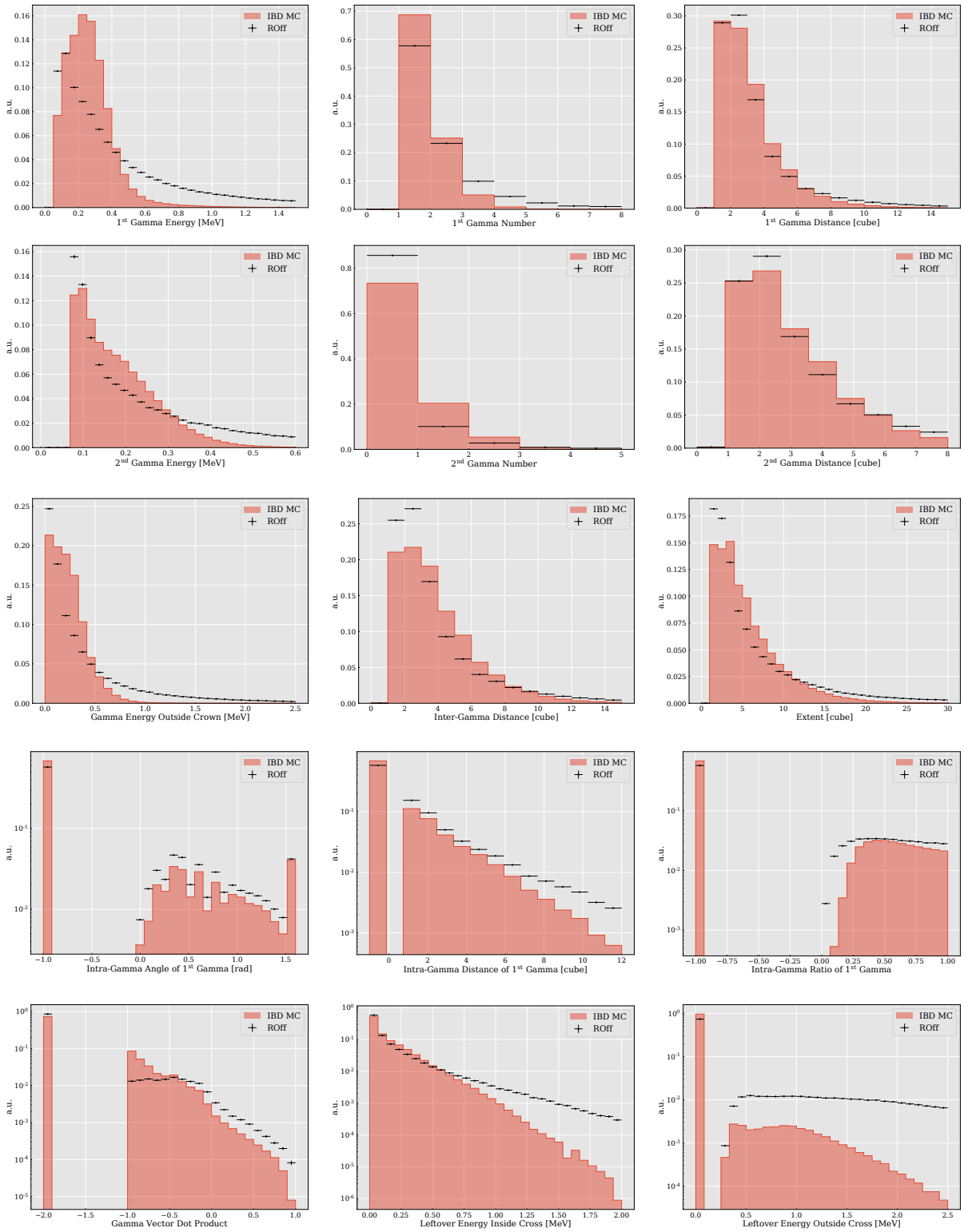


Figure 4.15: Probability density functions of gamma and leftover training variables for signal (IBD MC) and background (reactor-off data).

are taken since the spectra are found to differ significantly only by rate, not shape [103]. The model, known as the Neutrino Boosted Decision Tree (nuBDT), is trained on *all* the available events but k-fold cross-validation is used to verify that there is no overfitting of the dataset, thereby allowing the training events to be maximally utilised.

Correlations Between Variables

The final output of the analysis is the two-dimensional positron position-energy spectrum, or energy spectrum per module, which will be used to test the sterile neutrino hypothesis. For this purpose a uniform selection efficiency of signal events in the variables relevant for neutrino oscillation (E_{MAGE} and Z_{MEC}) is highly desirable to control systematic uncertainties. A dedicated boosting algorithm called uBoost that optimises some metric under the constrain of a uniform selection efficiency is freely available [232]. However, in this work the same effect is obtained by selecting BDT features that have little-to-no correlation with the positron energy or z-position. Twenty-two variables have been selected as nuBDT features, which should provide a sufficiently uniform efficiency for signal events in the desired variables. The rejected features are ES2ES1, ES3ES2, ESNS, promptNCCube, and all positron variables because of their relatively high correlation with the positron energy.

The correlation matrices of the training variables plus the two oscillation-relevant variables are visualised in Fig. 4.16. The correlations are quantified using the Pearson correlation coefficient (for a population), which describes the amount of linear correlation between two variables. Absolute correlations less than 0.2–0.3 are normally considered very weak. None of the chosen variables are correlated with Z_{MEC} in magnitude greater than 0.05 for either signal or background. This could have been anticipated from the independence of each detector module and the uniformity of the detector efficiency across the planes. The more troublesome variable with which to avoid correlations is the positron energy. As mentioned in the previous chapter, the gamma energies of IBD events have very little correlation with the positron energy, which also holds for all variables constructed from gamma-like cubes.

The largest signal correlations are found between the positron energy and the leftover energy variables. The leftover energy in the Cross (EnLeftoverCross) and the rest of the leftover energy (EnLeftoverNotCross) have Pearson coefficients with the positron energy of $+0.13$ and -0.14 respectively. The respective signs of these correlation coefficients can be attributed to the fact that leftover deposits are mostly misidentified positron deposits and to the spatial overlap in the definition of positron-like and leftover cubes. For a given initial positron energy, any energy deposited outside the Cross that is large enough to cause a cube to be identified as leftover will reduce the MAGE energy by the same amount, hence the negative correlation. Conversely, the amount of energy reconstructed in cubes adjacent to the MEC from light-leakage is proportional to the size of the initial energy deposit. These light-leakage deposits are almost always small enough to be classed as leftover. The leftover

variables have almost zero correlation with the gamma variables.

Hyperparameters

If an ML model is allowed to become too complex it learns the particular statistical fluctuations of its training data. Consequently, it will be unable to accurately generalise, thus diminishing the predictive power of the model on an unseen dataset compared to the training dataset. This is known as "overfitting". On the other hand, an overly conservative model that does not take full advantage of the available features will perform poorly. The XGBoost hyperparameters of the nuBDT model have been tuned by hand to prevent overfitting; they are outlined in Tab. 4.5. Other parameters are either not used or remain unaltered from their default values.

The chosen settings of *objective* and *booster* state that the nuBDT is a tree-based model dealing with a two-class classification problem and the function to be minimised is the Logistic Loss. The number of estimators (*n_estimators*) determines the number of boosting rounds applied; adding more trees leads to a more complex model. The learning rate shrinks the weighting of trees that are adjoined to the model at each boosting step. The model learns slower if the learning rate is small, but this can be compensated by increasing the number of estimators. The *max_depth* parameter is equal to the maximum depth of a base learner tree; a higher value gives a better-performing model but the training is more prone to overfitting. The final two parameters, *subsample* and *gamma*, are essentially regularisation terms designed to constrain the model complexity. The former is equal to the fraction of the dataset that is employed for training at each boosting stage, whilst the latter determines the minimum reduction in the loss function required to split a node.

Hyperparameter	**Value**	**Default?**
booster	gbtree	yes
objective	binary:logistic	yes
n_estimators	200	no
learning_rate	0.15	no
max_depth	2	no
subsample	0.75	no
gamma	10	no

Table 4.5: XGBoost hyperparameters of the nuBDT model. The parameters have been tuned to give the best signal-background discrimination whilst not overfitting the training sample.

Model Output

The output of the trained nuBDT model is displayed in Fig. 4.17. The distribution of scores is constructed by predicting the class probabilities of the training events. The cumulative

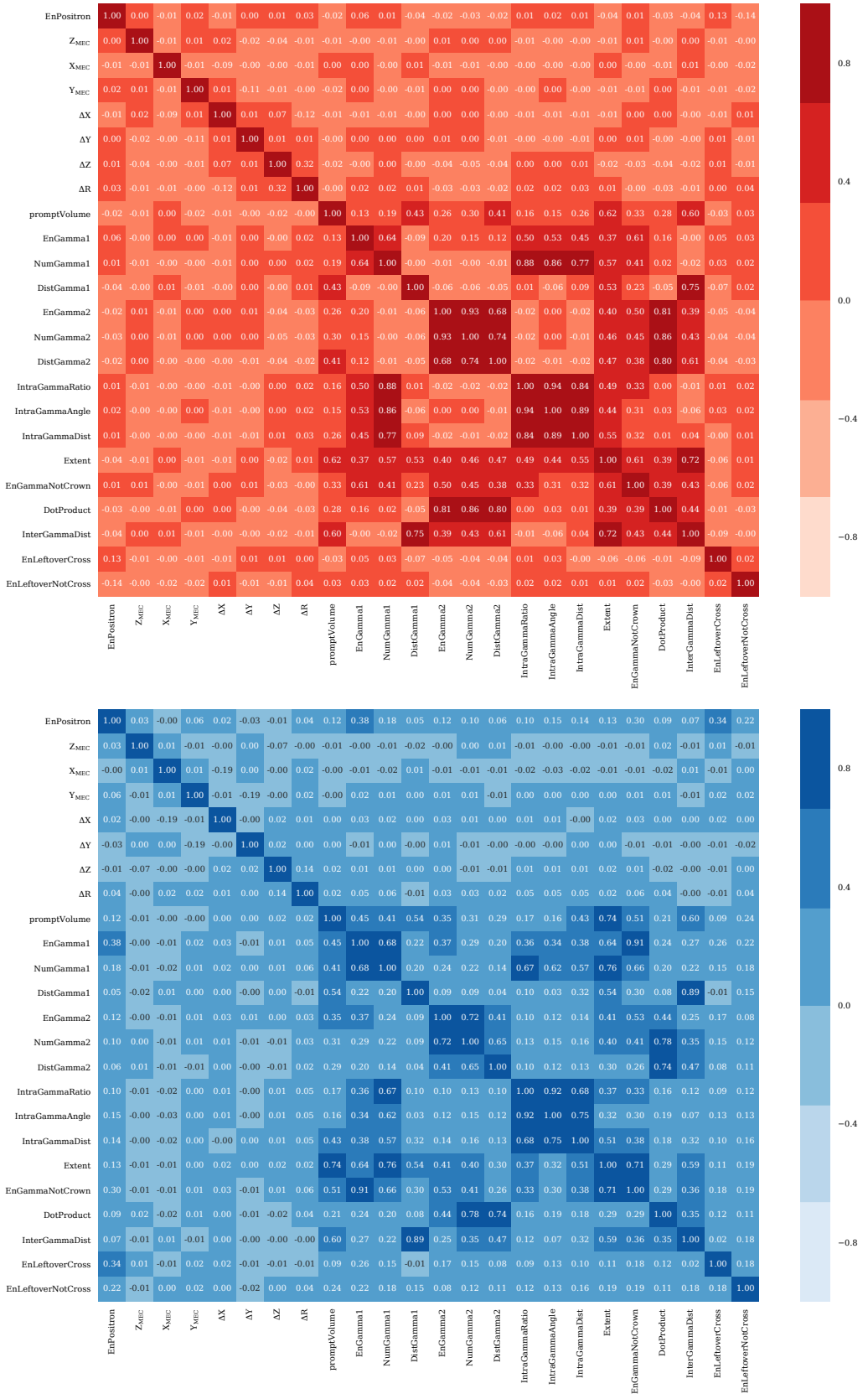


Figure 4.16: Correlation matrices of training variables for signal (red) and background (blue) events. The largest absolute correlations with positron energy are EnLeftoverNotCross for signal and EnGamma1 for background.

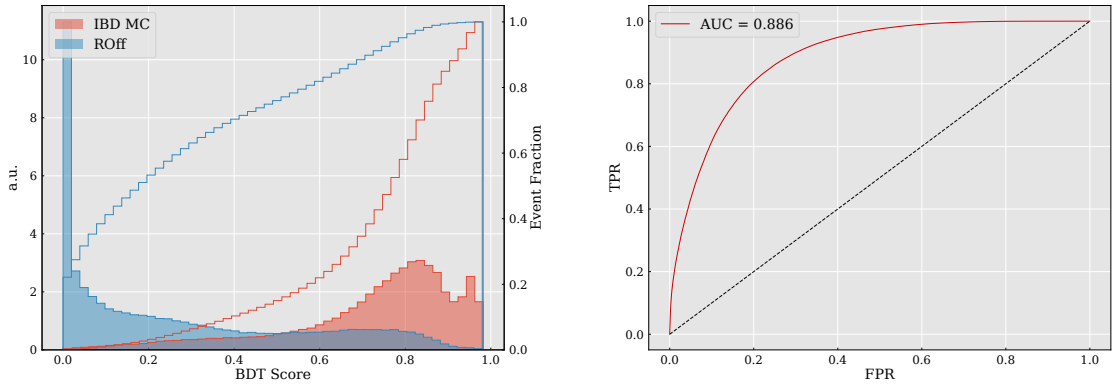


Figure 4.17: Left: Output of the nuBDT model in terms of score distributions of signal and background. The cumulative score distributions are plotted as un-filled histograms and must be interpreted in accordance with the right-hand side axis. The signal score distribution exhibits a small secondary peak composed mostly of 2γ events. Right: nuBDT ROC curve and AUC.

distributions are plotted as un-filled histograms and must be interpreted in accordance with the right-hand side axis. The nuBDT performance is evaluated at first instance by the ROC curve and corresponding Area Under the Curve (AUC) value. An AUC of 0.886 is obtained.

The signal distribution of BDT score can be analysed by dividing the range into three separate regions. The interval above 0.9 is occupied by 16% of all signal events, which exist in proportions 95:2:3 for 2γ, $1\gamma(>1c)$, and $1\gamma(=1c)$ respectively. Around 56% of the total 2γ signal population can be found in this region, manifesting as a second small peak in the distribution next to the global maximum. Similarly, background events in this region belong primarily to the 2γ category at around 91% purity, but these make up less than 1% of the total number of background events. Around 63% of the signal events are located between 0.6 and 0.9, split 15:27:58 by gamma category, as well as 18% of the background events, split 12:27:61. The remaining 21% of signal events and 81% of background events are found below 0.6, where the topologically basic $1\gamma(=1c)$ events are dominant.

Feature Importance

The importance of each model feature is evaluated via the Fischer score (F score) metric: the number of times a given feature is used to split the dataset. The Fischer scores for the nuBDT features are represented in Fig. 4.18. The variables NumGamma1 and NumGamma2 are not shown as they have Fischer scores of zero, likely because of their strong correlation with other variables. The intra-bunch variables also score very poorly, indicating that either there is little useful information at this level, or again that they are made redundant through correlations with other variables. The most important feature by the F score metric is the energy of the first gamma bunch, followed by the two leftover energy variables and ΔR.

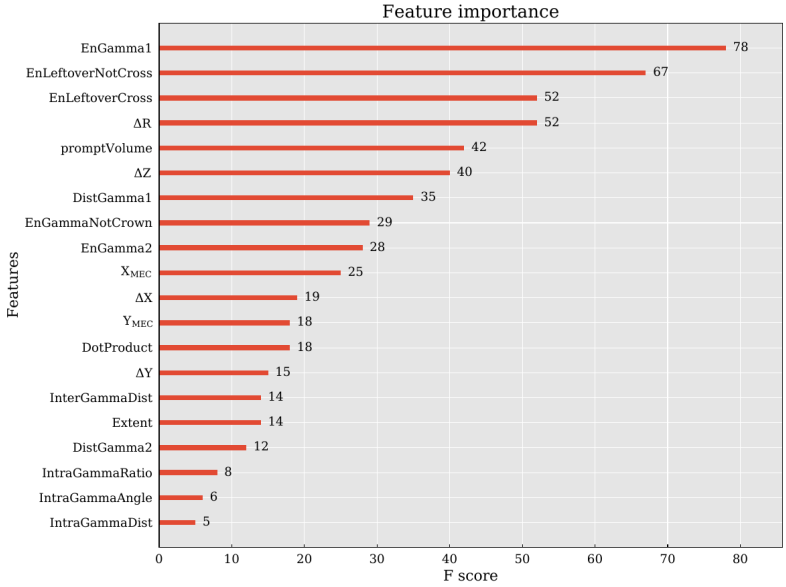


Figure 4.18: Fischer score importance of nuBDT input features, defined as the number of times a feature is used to split the data. The best variables by this metric are the energy of the first gamma, the leftover variables, and ΔR.

These four variables have relatively weak correlations amongst themselves so each bring independent signal-background separation. Despite the high-score region being dominated by 2γ events, the dataset overall contains a larger fraction of 1γ events so the energy of the second gamma bunch is relatively unimportant.

4.4.4 Overfitting Checks

The level of overfitting in the nuBDT training is assessed by k-fold cross-validation (CV). The training data is split into k equally sized test samples (s_i) in which the proportions of both classes are the same as in the overall dataset, thus neither class is under- or over-represented in any of the samples. This is called *stratified* k-fold cross-validation and it results in a more accurate estimation of the model performance than standard CV. A value of $k = 10$ is adopted in this work. An additional k training samples (S_i) are then constructed by merging all smaller samples s_j for $i \neq j$. New BDT "sub-models" (m_i) are trained on S_i using identical hyperparameters to the original model and tested on the corresponding sample s_i which contains completely unseen data. Each event is counted in exactly one test sample and $k - 1$ training samples.

Two metrics are computed for each train-test pair: the AUC and the natural logarithm of the loss (LogLoss), as in Tab. 4.6. The values of the metrics are consistent to within 0.001 over the ten training samplings and their means are equal to the respective values of the original model at the same precision. This level of agreement is not unexpected since the individual training samples S_i share 90% of the same events as the dataset on which the

Fold	Train AUC	Test AUC	Train Log. Loss	Test Log. Loss
1	0.886	0.885	0.418	0.420
2	0.885	0.887	0.419	0.417
3	0.886	0.885	0.418	0.419
4	0.886	0.883	0.418	0.421
5	0.886	0.886	0.418	0.418
6	0.886	0.884	0.418	0.421
7	0.885	0.887	0.418	0.417
8	0.886	0.883	0.417	0.423
9	0.886	0.886	0.418	0.417
10	0.886	0.884	0.418	0.421
Orig. Mod.	**0.886**	**n/a**	**0.418**	**n/a**
Sample Ave.	**0.886**	**0.885**	**0.418**	**0.419**
Sample Std.	**0.0004**	**0.001**	**0.0005**	**0.002**

Table 4.6: Overfitting check of the nuBDT model in terms of AUC and LogLoss. "Ave." refers to the arithmetic mean and "Std." to the standard deviation. The mean test score is congruent with the mean train score.

original model was trained. The standard deviations of the AUC and LogLoss test metrics are 0.001 and 0.002 respectively, and their means differ from the original model by 0.001. This is not deemed a significant decrease in performance.

Although these metrics are convenient quantifiers, information is lost when distilling the performance of an ML model to a single number. As a complementary test, we compare the score distributions of each sub-model with the original nuBDT. Fig. 4.19 shows the normalised score histograms of the ten sub-models, computed from their own training samples (left) and test samples (right). The distributions are separated into signal events and background events. The score distributions of the original model are overlaid in black. A statistical envelope is formed by taking the ratio of each sample distribution to the mean distribution of all samples and propagating the Poisson uncertainty on each bin appropriately. The black points correspond to the ratio of the original model to the same mean distributions. In both cases, the black points lie within the statistical envelope, indicating that no overfitting has occurred. It can be concluded that the chosen hyperparameters produce a model that generalises the data well.

4.4.5 Efficiency Flatness

The nuBDT is designed to achieve a uniform efficiency in positron energy and position along the z-axis. The efficiency flatness is quantified by taking the ratio of the distributions before and after cutting on BDT score, as shown in Fig. 4.20 for various cut values. The bin edges are designed such that there is an equal number of events in every bin. The oscillation

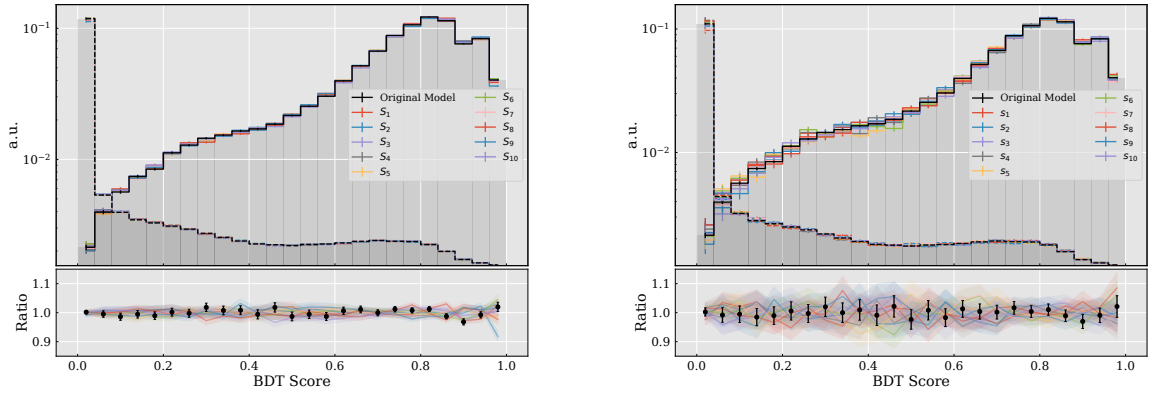


Figure 4.19: Overfitting check of the nuBDT model in terms of predicted score. Distributions are split into signal (solid lines) and background (dashed lines) for each of the ten models. The ratio plots are computed by dividing each distribution by the mean of the sample outputs for the training events (left) and the test events (right). The ratio for the original model lies within the band defined by the standard deviations of the samples. Error bars represent the Poisson statistical uncertainty on the points.

pattern of the RAA best-fit point is superimposed. The horizontal lines across each of the curves correspond to linear fits that are constrained to have zero gradient. Since the statistical uncertainty is the same for every bin, the fits are effectively un-weighted and the y-intercepts equivalent to the mean efficiency. The energy curves are typically flat to within a few percent. The largest deviation is in the first energy bin, but it is much smaller than the expected effect from a strong oscillation pattern. The Z_{MEC} curves have consistently higher efficiency towards the middle of the detector but the difference from the mean efficiency is only a few percent. The deviations are only an issue if the MC simulation is not in good agreement with data. This is the subject of the next and last section this chapter.

4.5 Validation with BiPo

BiPo events are an intrinsic source of electrons at the same energy scale as the IBD positron. They can be used to validate the detector response model and tune ROSim parameters—such as the light yield of the ZnS screens—which would otherwise be very difficult to quantify. To simulate BiPo events in the detector, vertices are generated uniformly across the neutron screens of the detector. SoLidSim then propagates the outgoing nuclei through the detector material in the same manner as for the IBD daughter particles. The resulting true-space events are processed with the reconstruction software (ROSim and Saffron2) to be comparable with data.

The most important variables to be validated are the ES energy variables such as the

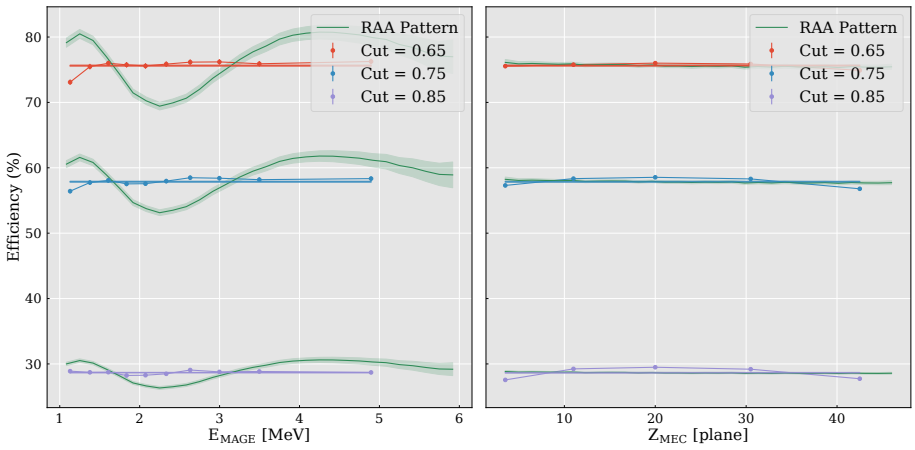


Figure 4.20: Selection efficiency on signal events in E_{MAGE} and Z_{MEC} for three cuts on nuBDT score. The same number of events populate each bin. Horizontal lines are zero-gradient fits to the points. The oscillation patterns and associated errors of the Reactor Antineutrino Anomaly best-fit point are overlaid as green lines and bands respectively; the oscillation effects are typically larger than deviations from flatness in the efficiency curves.

positron energy (essential for an oscillation analysis) and energy of the gamma bunches (important BDT features). It is also critical that the distributions of nuBDT score are in good agreement to ensure that no biases are introduced by the ML models. As the BiPo events are generated uniformly throughout the detector, unlike IBD events there is no expectation that the predicted spatial distributions of BiPo events should be representative of reality. Indeed, the main utility of BiPo is that its ES event, particularly the radiative type, acts as a proxy of an IBD event. The comparison is therefore focused on the energy variables and gamma reconstruction. Any data-MC discrepancies can be accounted for in the oscillation fit by systematic uncertainties, as described in chapter 6.

4.5.1 High-Purity BiPo Region

A set of cuts that isolates BiPo events in the data are defined in Tab. 4.7[††]. The cuts are motivated by the three distinct BiPo event types depicted in Fig. 4.21. The neutron screen in which the alpha nucleus generates the NS is always the screen containing the BiPo vertex, but the electron may travel to one of the surrounding cubes to create the ES deposit. The cuts on ΔX and ΔZ are designed to isolate the cases where the BiPo vertex is in the +x or −z screen of a voxel respectively and the electron either remains in the same voxel ($\Delta X = 0$, $\Delta Z = 0$) or in the voxel that shares a face with the neutron screen ($\Delta X = -1$, $\Delta Z = 1$). The ΔY and ΔR cuts ensure that these are the only possible scenarios. The other cuts enhance the BiPo purity by reducing the contribution from fast neutrons, i.e. at high energy or low Δt. The purity of this region calculated using a Δt fit is around 99.7%. The 0γ

[††]The NS-ES displacement cuts therein were used to create the alpha-rich dataset for training BiPonator.

Description	Variable and Limits
NS-ES displacements	$\Delta X \in [-1, 0]$
	$\Delta Y \in [0, 0]$
	$\Delta Z \in [0, 1]$
	$\Delta R \in [0, 1]$
BiPo-focused	$\Delta t \in [200, 500]$ µs
	BiPonator $\in [0.22, 0.65]$
Other	$E_{\text{Crown}} \in [1.0, 4.0]$ MeV
	NumGammaTotal $\in [1, \infty]$
	tNearestMuonType1 $\in [200, \infty]$ µs
	tNearestMuonType2 $\in [200, \infty]$ µs

Table 4.7: Summary of the cuts that define the high-purity BiPo region. The brackets indicate inclusive limits. The cuts most responsible for achieving high purity of BiPo events are on the NS-ES displacement variables and BiPonator.

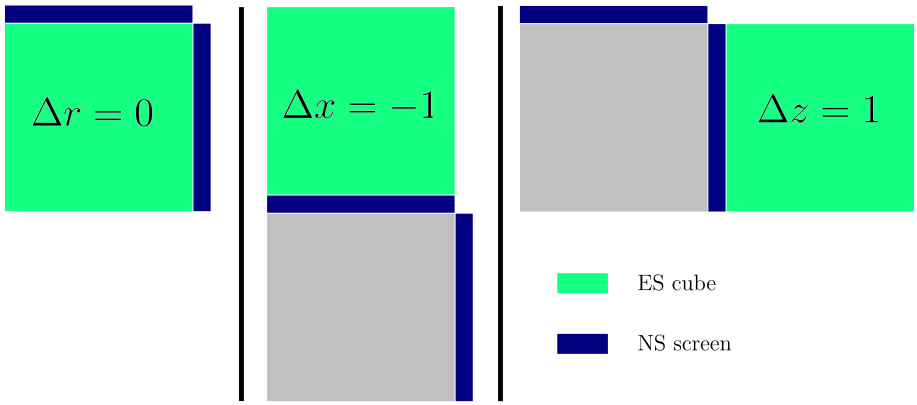


Figure 4.21: Schematic of the three types of BiPo event that are isolated by the cuts in Tab. 4.7. The green squares depict the cube in which the electron deposits its energy (ES) and the blue rectangles depict the neutron detection screens in which the alpha nucleus generates the NS. The vertical black lines simply mark the distinction between diagrams.

category is dropped because the nuBDT is not trained on such events. The BiPonator and Δt limits ensure the independence of this region from the preselected events to be used for the main analysis.

4.5.2 Data-MC Comparison

A data-MC comparison of events in the high-purity BiPo region is performed for numerous variables. The cubes on the outer edges of the detector planes were recycled from a previous prototype detector. As such, they were part of the manufacturing batch that had an order of magnitude less ^{238}U contamination. This is not accounted for in the simulation so all edge cubes in the x-y plane are dropped for the data-MC comparison. Using all available reactor-off and BiPo MC files yields 359 351 data events and 407 890 simulated events respectively,

after removing the edge cubes. The BiPonator cut defined in Tab. 4.7 is not applied to the MC events, since they are merely assigned BiPonator scores sampled from a reference distribution. The BiPo MC histograms are always normalised to the number of data events.

The comparisons are made separately for each gamma category. To evaluate the data-MC agreement, a chi-square-like metric is computed from the data-MC ratios by propagating the statistical uncertainties, assuming Poisson statistics, of both data and MC and comparing the results with unity:

$$\chi^2/\mathrm{ndf} = \frac{1}{\mathrm{ndf}} \sum_i \frac{(1 - D_i/P_i)^2}{D_i/P_i^2\,(1 + D_i/P_i)}\,, \tag{4.12}$$

where D_i and P_i are the observed and predicted count of bin i respectively and ndf is number of degrees of freedom. We also define a sub-category of $1\gamma(=1\mathrm{c})$, comprising events where the energy of the single gamma cube is below 148 keV. This value (double the cube threshold) was chosen somewhat arbitrarily. Such $1\gamma(=1\mathrm{c})$ events are said to belong to the $1\gamma(=1\mathrm{c}_{\mathrm{low}})$ sub-category and all others to $1\gamma(=1\mathrm{c}_{\mathrm{high}})$.

The nuBDT score distributions for data and MC (and their ratios) per gamma category are presented in Fig. 4.22. The total number of data and MC events allow the agreement to be probed at the 10% level for the majority of bins. The 2γ and $1\gamma(>1\mathrm{c})$ events demonstrate good data-MC agreement at this precision. Their distributions are distinctly peaked in the 0.8–1.0 interval, consisting of the most IBD-like events in categories of already relatively IBD-like events. There is also a smaller peak in both distributions at around 0.3 of unclear origin. The $1\gamma(=1\mathrm{c}_{\mathrm{high}})$ and $1\gamma(=1\mathrm{c}_{\mathrm{low}})$ events exhibit sizeable deviations from unity in the ratios. The double-peak structure in the distributions is also present for these categories but the overall shape is shifted towards lower scores.

Next, the effect of applying a BDT cut on the position and energy spectra is assessed. If any biases are introduced by the BDT they may appear as data-MC differences in the ratio of the distributions after and before cutting on the score. This test avoids any potential data-MC mismatch that could be attributed to other factors. It has been performed for one loose cut (0.5) and one tight cut (0.8), as in Fig. 4.23. The efficiency of the loose cut is greater than 90% so the effect on the spectrum shapes is relatively minor and there are no significant deviations evident in the data-MC ratios. Applying the tight cut reveals the discrepancy in the $1\gamma(=1\mathrm{c}_{\mathrm{low}})$ events that is noticeable in the nuBDT score distributions. The energy spectrum in particular exhibits a clear distortion caused by a significantly lower efficiency for data than MC. There is good agreement in the 2γ and $1\gamma(>1\mathrm{c})$ categories and adequate agreement in the $1\gamma(=1\mathrm{c}_{\mathrm{high}})$ category.

Finally, the data-MC agreement is directly assessed in Fig. 4.24, integrated over gamma category, for the majority of the nuBDT input features and the positron-related variables. Variables that are absent here are provided in appendix B. The $1\gamma(=1\mathrm{c}_{\mathrm{low}})$ events are

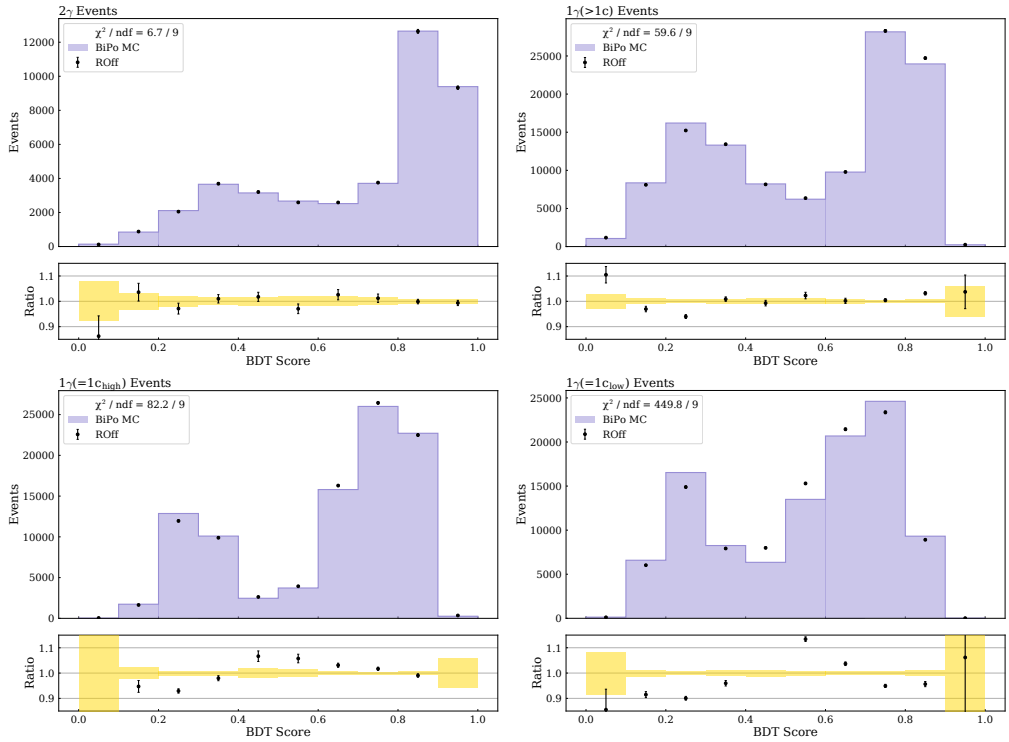


Figure 4.22: BiPo data-MC comparisons of nuBDT score for each gamma category. Error bars represent the statistical uncertainty on the points. The yellow band in each ratio plot indicates the uncertainty on the MC prediction. Overall agreement within 10% can be observed for all but the $1\gamma(=1c_{low})$ category, which comprises events whose single gamma-like deposit is less than 148 keV.

dropped at this point, given the insights obtained in the preceding paragraphs, but discussions are resumed in chapter 6 where a dedicated systematic uncertainty is derived. The positron energy spectrum exhibits a systematic shift indicative of a slight mis-tuning of the PVT light yield in the simulation, which has an error of around 1.5%. Previous studies [207] indicate that this uncertainty likely drives the discrepancy in the energy. Moreover, the small disparity in EnLeftoverCross supports the view that systematic uncertainties on the light yield have become significant at this level of precision. On the other hand, the simulation and reconstruction of BiPo gamma rays appears to be working very well, with excellent agreement in all corresponding variables.

The NS-ES displacement variables show an extraordinarily large data-MC disagreement. The error on the bin entries is very small because of the tight cuts applied on these (integer) variables. The exact cause of this issue is not known. It could arise from difficulties in simulating the passage of the electron out of the ^{6}LiF:ZnS(Ag) screens or simply from a bug in the BiPo generator code. In any case, data-MC comparisons of the neutron calibration with AmBe [206] show no such discrepancy, meaning IBD events are not affected.

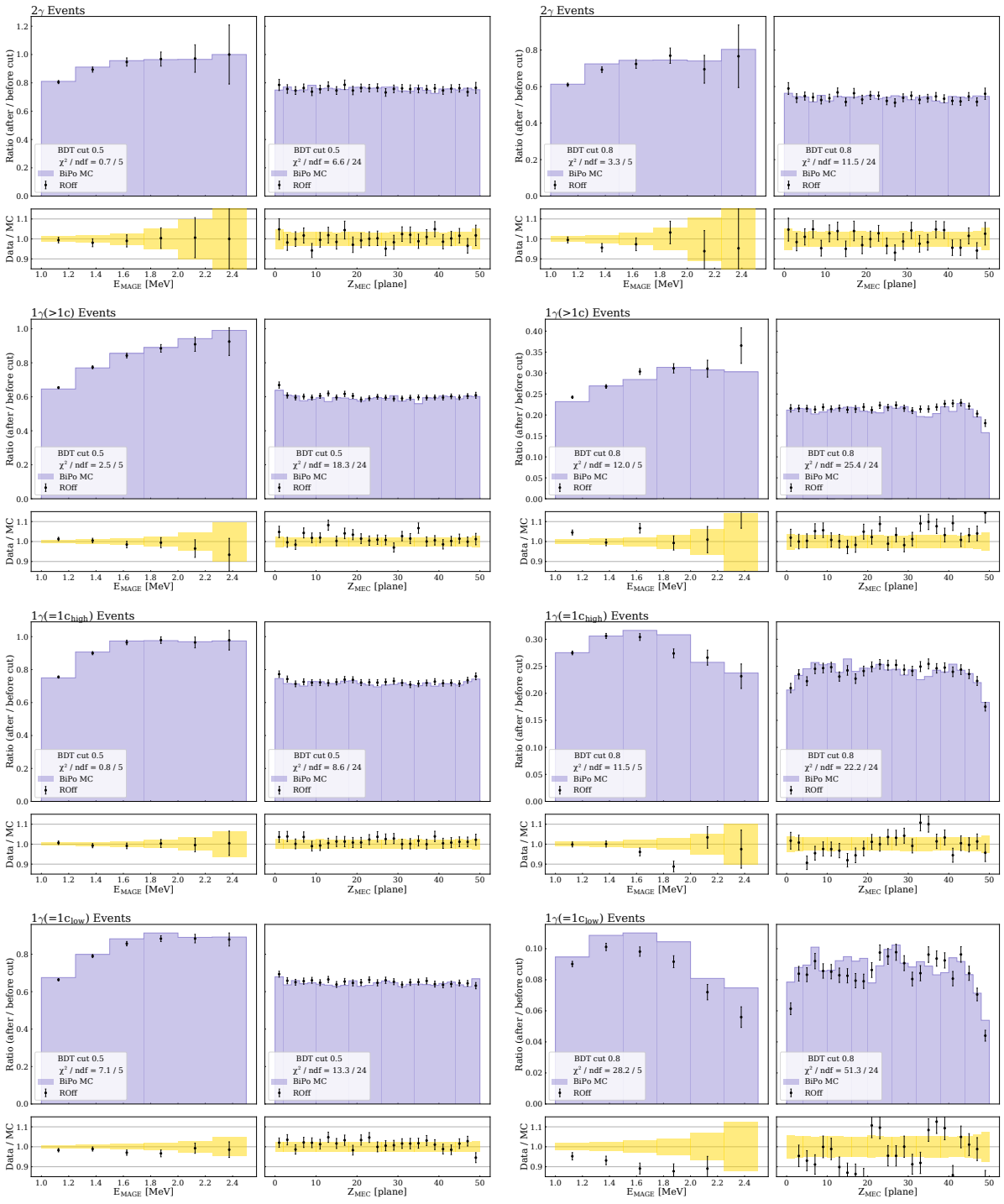


Figure 4.23: BiPo data-MC comparisons of the ratio between distributions after and before cutting on nuBDT score for a loose cut (left) and a tight cut (right). Error bars represent the statistical uncertainty on the points. The yellow band in each ratio plot indicates the uncertainty on the MC prediction. Once again, $1\gamma(=1c_{low})$ events show the worst data-MC agreement.

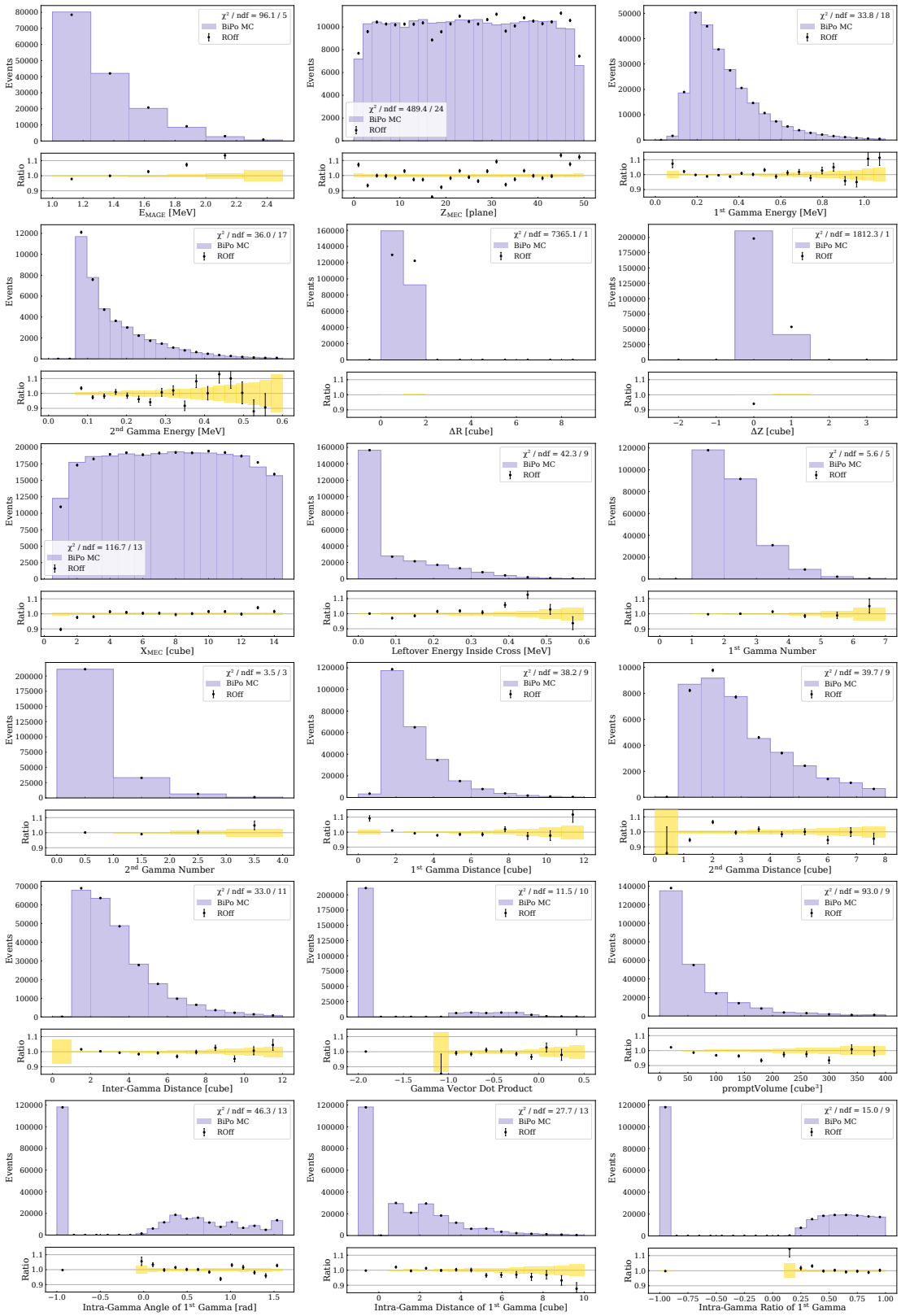


Figure 4.24: BiPo data-MC comparison plots in multiple variables integrated over gamma category, excluding $1\gamma(=1c_{\text{low}})$ events. Error bars represent the statistical uncertainty on the points. The yellow band in each ratio plot indicates the uncertainty on the MC prediction. Reasonable agreement is found for most variables, but ΔR and ΔZ are reproduced extremely poorly by the simulation.

Chapter 5

Background Subtraction

In this chapter, the strategy for extracting the antineutrino signal is described and demonstrated. Before being applied to the entire Phase-I dataset, the background subtraction is carried out on the *open dataset*, which contains the first reactor-on data that was unblinded by the collaboration for the purpose of testing and validating previous iterations of the analysis and subtraction tools. It has been employed here for the same purpose.

Prior to the subtraction, events below a certain nuBDT score threshold are discarded in order to exploit the signal-background separation provided by the model. However, as the subtraction is part of the nuBDT optimisation procedure, for now it is simply stated that a cut of 0.73 is applied. The optimisation procedure that is used to derive this value is outlined in section 5.2.5. The subtraction method is thoroughly checked and validated in the final section of this chapter.

5.1 Parameter Space Regions

The data-driven subtraction of the background remaining after analysis cuts is performed sequentially by background type and depends on a precise characterisation of the three backgrounds in dedicated hyper-volumes (event regions) of parameter space. The region of parameter space defined by the preselection cuts outlined in section 4.3 is henceforth referred to as the *signal region*. The reactor-off signal region may equally be referred to as the "neutron sideband" because of its lack of antineutrino events and relatively high proportion of atmospheric neutrons. The background hyper-volumes are independent of each other and of the signal region.

In this context, the label S refers to a sample from a particular event region, to be interpreted either as the total number of events in that sample or an N-dimensional histogram built from the sample events. The region itself is denoted by a subscript and if necessary a superscript is used to indicate the reactor state. Eight event regions are defined in this

manner:

$$\begin{array}{llll} S^{\mathrm{ROn}}_{\mathrm{Signal}}\ , & S^{\mathrm{ROn}}_{\mathrm{BiPo}}\ , & S^{\mathrm{ROn}}_{\mathrm{Signal\,FPNT}}\ , & S^{\mathrm{ROn}}_{\mathrm{BiPo\,FPNT}}\ , \\ S^{\mathrm{ROff}}_{\mathrm{Signal}}\ , & S^{\mathrm{ROff}}_{\mathrm{BiPo}}\ , & S^{\mathrm{ROff}}_{\mathrm{Signal\,FPNT}}\ , & S^{\mathrm{ROff}}_{\mathrm{BiPo\,FPNT}}\ , \end{array} \tag{5.1}$$

where S_{Signal} and S_{BiPo} denote samples from the signal region and BiPo sideband respectively, whose equivalent accidental regions, $S_{\mathrm{Signal\,FPNT}}$ and $S_{\mathrm{BiPo\,FPNT}}$, are similarly defined but comprise FPNTs rather than standard NS-ES coincidences. The events that occupy a given hyper-volume may also depend on the reactor power. For instance, ROff accidental events differ from those during reactor cycles and antineutrinos are absent in all ROff periods.

5.2 Open Dataset Subtraction

The open dataset comprises the first reactor-cycle (12^{th} June – 10^{th} July 2018) and proximate reactor-off data (7^{th} June – 12^{th} June 2018 and 17^{th} June – 21^{st} August 2018). Around 4300 ROn runs (22.7 d exposure time) and 6559 ROff runs (34.5 d exposure time) that passed the data quality controls were compiled during this time. The reactor-on component of the open dataset corresponds to approximately 8% of the total Phase-I ROn data, meaning the precision obtained on the antineutrino spectrum is not expected to allow sensitivity to any of the RAA 95% CL region.

5.2.1 Accidental

The rate of accidental NS-ES coincidences in the signal region is estimated using the corresponding FPNT sample. However, the FPNT event rate is not directly representative of the NS-ES accidental rate because of the small proportion of reconstructed NS clusters that cannot be linked with an FPNT. The number of such "orphaned" NS clusters is very low because there are around four times as many reconstructed NS clusters as there are FPNTs. Even so, a scaling factor is defined to account for this effect:

$$\epsilon = \frac{\text{No. of reconstructed NS clusters}}{\text{No. of FPNT-matched NS clusters}}\ . \tag{5.2}$$

The correction is typically very small, increasing the estimated rate by 1% at most. The accidentals in the signal region may then be subtracted as follows:

$$S^{\mathrm{ROn}}_{\mathrm{Signal-Acc}} = S^{\mathrm{ROn}}_{\mathrm{Signal}} - \epsilon^{\mathrm{ROn}} \cdot S^{\mathrm{ROn}}_{\mathrm{Signal\,FPNT}}\ , \tag{5.3}$$

$$S^{\mathrm{ROff}}_{\mathrm{Signal-Acc}} = S^{\mathrm{ROff}}_{\mathrm{Signal}} - \epsilon^{\mathrm{ROff}} \cdot S^{\mathrm{ROff}}_{\mathrm{Signal\,FPNT}}\ , \tag{5.4}$$

where the factors ϵ^{ROff} and ϵ^{ROn} are calculated independently for each period. For the open

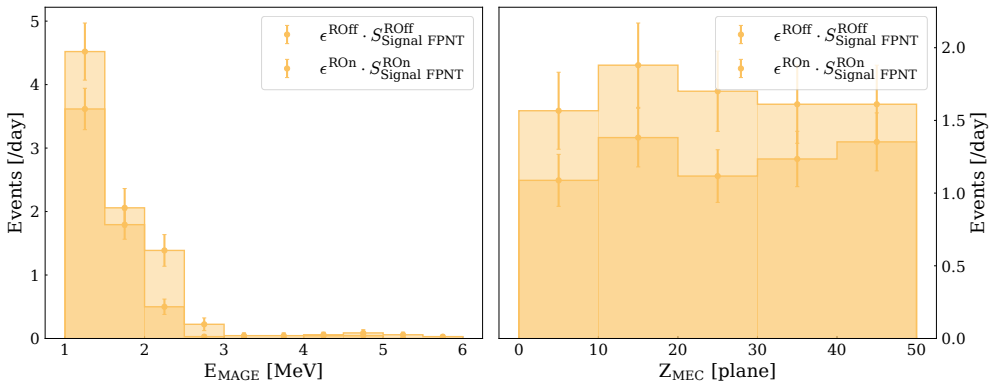


Figure 5.1: Energy and z-position spectra of events in the open-dataset FPNT sample. Error bars represent the statistical uncertainty on the bins. The energy spectrum suggests that the ES events of accidentals are predominantly made up of BiPo events. Accidentals are uniformly distributed in Z_{MEC}, within error bars.

dataset, the values are 1.008 and 1.007 respectively. The energy and z-position distributions of the open-dataset accidental background are plotted Fig. 5.1. The ROn rate of accidentals is approximately 20% larger than the ROff rate. The energy spectrum contains extremely few events above 3 MeV, suggesting that coincidences of very energetic proton recoils and random NS clusters are far less likely than between the latter and BiPo ES clusters or reactor gammas.

5.2.2 BiPo

The subtraction proceeds by dealing with BiPo background. BiPo templates are obtained from a sideband S_{BiPo} that is identical to the signal region except for the limits on BiPonator score and Δt, which are substituted by the following cuts:

- BiPonator $\in [0.0, 0.2]$
- $\Delta t \in [300, 500]$ µs

The purity of this BiPo sideband is around 80% and the remaining fraction is composed of accidental background. Accidentals in the BiPo sideband can be removed in an equivalent manner to those in the signal region:

$$S^{\mathrm{ROn}}_{\mathrm{BiPo-Acc}} = S^{\mathrm{ROn}}_{\mathrm{BiPo}} - \epsilon^{\mathrm{ROn}} \cdot S^{\mathrm{ROn}}_{\mathrm{BiPo\,FPNT}} \,, \tag{5.5}$$

$$S^{\mathrm{ROff}}_{\mathrm{BiPo-Acc}} = S^{\mathrm{ROff}}_{\mathrm{BiPo}} - \epsilon^{\mathrm{ROff}} \cdot S^{\mathrm{ROff}}_{\mathrm{BiPo\,FPNT}} \,, \tag{5.6}$$

which are directly analogous to equations 5.3 and 5.4. The implicit assumption that BiPo events in this sideband are equivalent to those in the signal region is valid for ES variables,

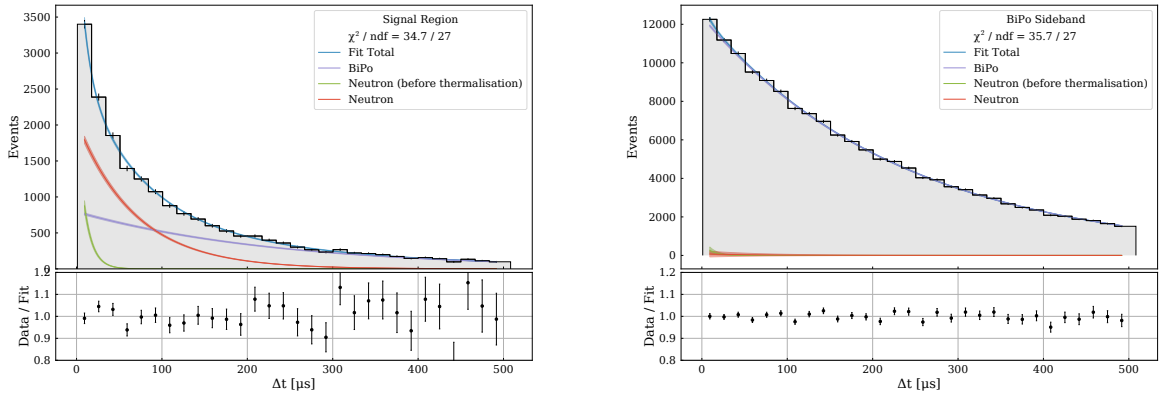


Figure 5.2: Δt distribution of open-dataset reactor-off data after analysis cuts (preselection and nuBDT) and subtraction of accidentals, in the signal region (left) and BiPo sideband (right), fit with equation 5.7. The fitted parameters for the BiPo sideband are consistent with almost 100% pure BiPo events. The relative proportions of backgrounds in the signal region are comparable to Fig. 4.12.

which are totally uncorrelated with Δt and BiPonator.

The rates of each background after subtraction of accidentals are estimated via a Δt fit. To improve the precision of the fit, the distributions are binned across an extended Δt range (1–500 µs) and the Δt limits are only enforced after the fit has been executed. The ϵ correction factors are also calculated before applying the final Δt cut. By first removing accidentals, the resulting Δt distribution can be described using one fewer parameter—reducing the chance of convergence issues in the fit—as follows:

$$y(\Delta t) = c_1\, e^{\Delta t/\tau_{\mathrm{BiPo}}} + c_2\, e^{\Delta t/\tau^{\mathrm{Th}}_{\mathrm{Neutron}}} + c_3\, e^{\Delta t/\tau^{\mathrm{Epi}}_{\mathrm{Neutron}}}\,, \tag{5.7}$$

where the parameters are unchanged from equation 4.2. The second and third fit parameters (c_2 and c_3) describe the combined rate of fast neutrons and IBD events in reactor-on data.

The ratio between the rates of BiPo in the signal region and BiPo sideband defines a quantity known as the *BiPo-to-Signal factor* (β). It is used to simultaneously normalise the templates acquired from the BiPo sideband and scale them according to the true rates in the signal region. The BiPo-to-Signal factor is calculated as:

$$\beta = \frac{\int_1^{141} c_1^{\mathrm{Signal}}\, e^{\Delta t/\tau_{\mathrm{BiPo}}}\, \mathrm{d}\Delta t}{\int_{300}^{500} c_1^{\mathrm{BiPo}}\, e^{\Delta t/\tau_{\mathrm{BiPo}}}\, \mathrm{d}\Delta t}\,, \tag{5.8}$$

where c_1^{Signal} and c_1^{BiPo} are the BiPo parameters of the fit in the signal region and BiPo sideband respectively and the integral limits represent the Δt cuts that define each region.

The Δt fits for open dataset reactor-off data are shown in Fig. 5.2. The fast-neutron component in the BiPo sideband is trivially small, even in the extended Δt range. As

a result, it is expected that the quantity of IBD events in the reactor-on version of this sideband is also irrelevant. The β factors are computed separately for the ROn and ROff components in the subtraction. The values for the open dataset are $\beta^{\mathrm{ROn}} = 0.1596(1)$ and $\beta^{\mathrm{ROff}} = 0.1687(1)$. The difference between the two quantities is statistically significant and mostly driven by the rate variations in the signal region rather than the BiPo sideband.

Finally, the combined subtraction of accidentals and BiPo events in the signal region can be expressed as:

$$\begin{aligned} S^{\mathrm{ROn}}_{\mathrm{Signal-Acc-BiPo}} &= S^{\mathrm{ROn}}_{\mathrm{Signal-Acc}} - \beta^{\mathrm{ROn}} \cdot S^{\mathrm{ROn}}_{\mathrm{BiPo-Acc}} \\ &= S^{\mathrm{ROn}}_{\mathrm{Signal-Acc}} - \beta^{\mathrm{ROn}} \cdot (S^{\mathrm{ROn}}_{\mathrm{BiPo}} - \epsilon^{\mathrm{ROn}} \cdot S^{\mathrm{ROn}}_{\mathrm{BiPo\,FPNT}}) \,, \end{aligned} \tag{5.9}$$

$$\begin{aligned} S^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}} &= S^{\mathrm{ROff}}_{\mathrm{Signal-Acc}} - \beta^{\mathrm{ROff}} \cdot S^{\mathrm{ROff}}_{\mathrm{BiPo-Acc}} \\ &= S^{\mathrm{ROff}}_{\mathrm{Signal-Acc}} - \beta^{\mathrm{ROff}} \cdot (S^{\mathrm{ROff}}_{\mathrm{BiPo}} - \epsilon^{\mathrm{ROff}} \cdot S^{\mathrm{ROff}}_{\mathrm{BiPo\,FPNT}}) \,. \end{aligned} \tag{5.10}$$

5.2.3 Fast Neutron

At this stage, fast neutrons are the only background remaining in the reactor-off signal region and the reactor-on signal region comprises only fast neutrons and IBD events. In principle, the pure IBD yield can be obtained from subtracting the former from the latter (i.e. $S^{\mathrm{ROn}}_{\mathrm{Signal-Acc-BiPo}} - S^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}}$). However, the fast-neutron rate depends on the atmospheric pressure, meaning it will differ between two periods by an amount proportional to the difference in mean pressure. The variation of the fast-neutron rate with pressure (P) is modelled by fitting the relative change in background rate from the mean as a function of the relative pressure difference:

$$S^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}} - \overline{S}^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}} = f(P, \overline{P}) \,, \tag{5.11}$$

where $\overline{S}^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}}$ is the mean rate in the ROff sample, $\overline{P}$ is the mean pressure, and f is a function that parameterises the dependence of neutron rate on pressure.

In practise, the reactor-off data is separated into 24 h time bins and the mean rate of the entire period ($\overline{S}^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}}$) is subtracted, resulting in the relative change in the number of background events per day. Similarly, the relative pressure is defined as the difference between the mean pressure of a given day and the mean value for the whole period. A first order polynomial form of f is found to sufficiently parameterise the dependence of neutron rate on pressure:

$$f(P, \overline{P}) = p_0 + p_1(P - \overline{P}) \,, \tag{5.12}$$

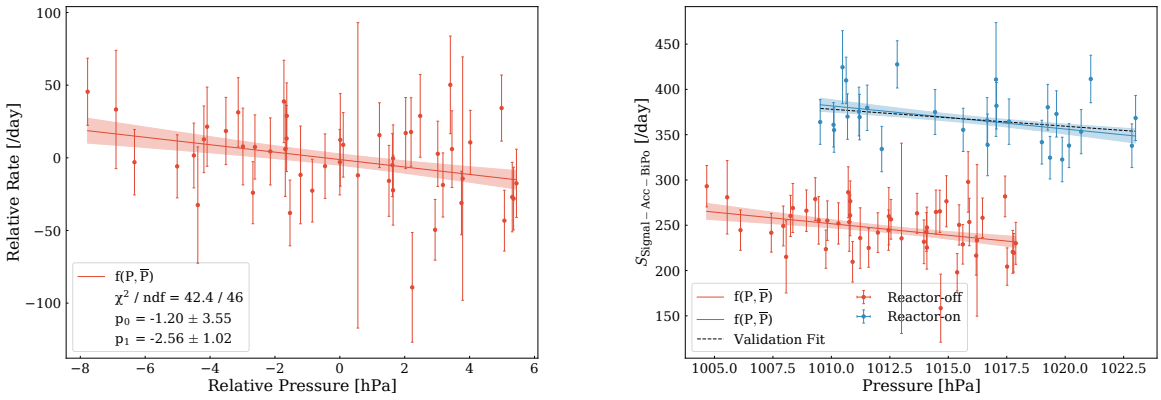


Figure 5.3: Left: Pressure model derived from open-dataset reactor-off data using equation 5.12. Right: Comparison between the open-dataset pressure model applied to reactor-off data (red) and reactor-on data (blue), scaled up by the respective mean rates. The statistical errors are displayed as translucent bands. The black dotted line depicts the fit of the pressure model on the reactor-on events per day. The antineutrino rate should be independent of pressure, which would be indicated by a constant shift between the red and blue lines.

where the coefficients p_0 and p_1 are the fitted parameters of the model. The pressure measurements are made by sensors on the outside of the BR2 building, which take readings every 30 s. The function was expanded in a previous work [207] to also include the atmospheric temperature, but this had a negligible effect.

The pressure model for the open dataset is computed using all available reactor-off data, as plotted in the left-hand side of Fig. 5.3. The dependence of neutron rate on atmospheric pressure is clearly visible. The fast-neutron rate decreases by -2.6 ± 1.0 per unit hectopascal and at zero relative pressure is consistent with zero at -1.2 ± 3.5 events per day. As displayed on the right-hand side of the figure, the relative pressures (ROn and ROff) have been fed to the model and the outputs added to the respective mean event rates and pressures. If the difference in mean rate between ROn and ROff is attributed to the presence of antineutrino events, the shift should be independent of pressure. Therefore, the ROn rates are also fit with the function f to verify that the resulting fit (dashed black line) is coherent with the ROff model. The slope lies within statistical error band of the original model, at least within the pressure range of the reactor-on period, meaning that the difference in rate is compatible with the presence of IBD events.

The pressure model is used to adjust the rates of each period in order to eliminate the correlation between fast-neutron rate and pressure. As usual, this is done independently for each sample:

$$S'^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}} = S^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}} - \sum_{i}^{I} f(P_i, \overline{P}_{\mathrm{ROff}}) \,, \tag{5.13}$$

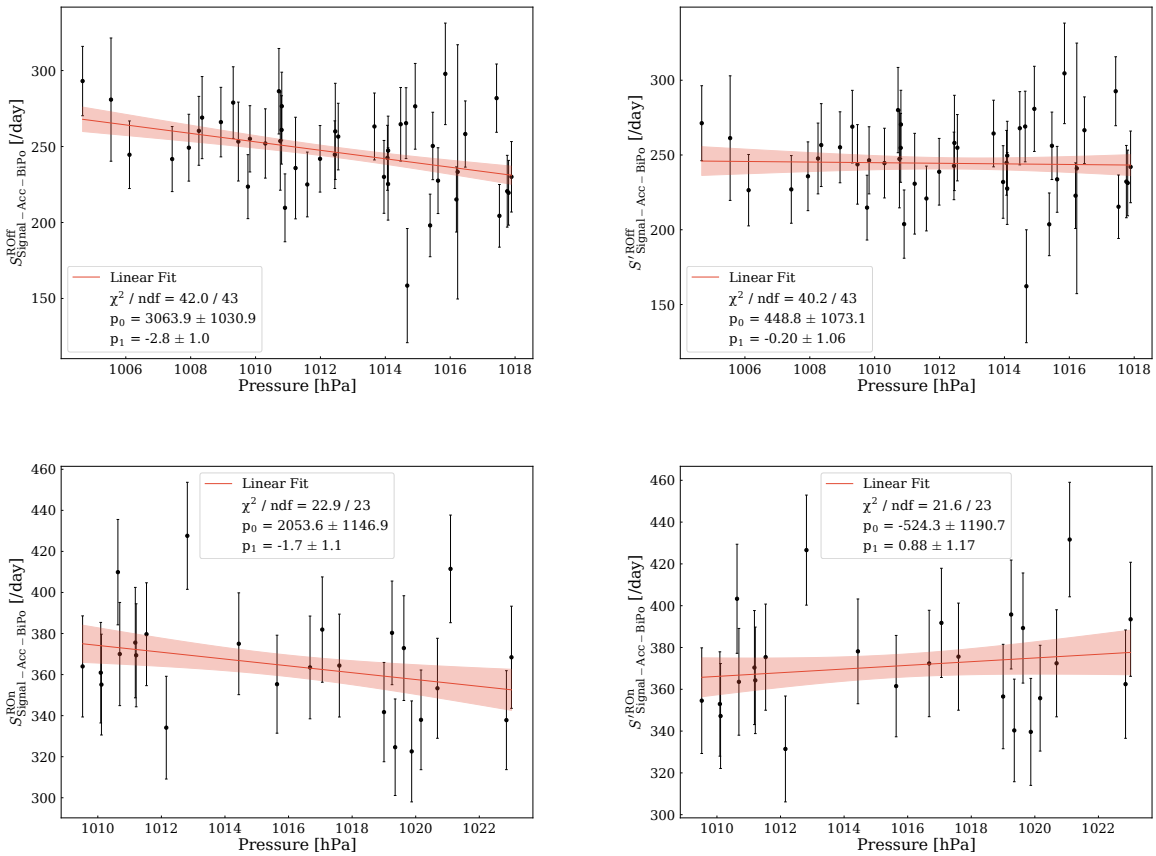


Figure 5.4: Top: Number of accidental-and-BiPo-subtracted events per day in the signal region as a function of pressure for open-dataset reactor-off data before (left) and after (right) pressure correction. Bottom: Equivalent plots for open-dataset reactor-on data. The red lines are weighted linear fits through the points, with statistical errors displayed as translucent bands. Error bars represent the statistical uncertainty on the points (black).

$$S'^{\mathrm{ROn}}_{\mathrm{Signal-Acc-BiPo}} = S^{\mathrm{ROn}}_{\mathrm{Signal-Acc-BiPo}} - \sum_{j}^{J} f(P_j, \overline{P}_{\mathrm{ROn}}) \,, \tag{5.14}$$

where I is equal to the number of reactor-off time bins, J is equal to the number of reactor-on time bins, and $\overline{P}_{\mathrm{ROff}}$ and $\overline{P}_{\mathrm{ROn}}$ are the mean pressure for each period. S' denotes a pressure-corrected sample. The final terms in each equation are the integrated pressure corrections, which for the open dataset have values of -1.38 ± 1.10 events for reactor-on and -1.16 ± 0.76 events for reactor-off. The effect of the pressure corrections is demonstrated in Fig. 5.4. For the reactor-off data—from which the model was derived—the Pearson correlation coefficient reduces from -0.39 (moderate correlation) to -0.05 (extremely weak correlation). For the reactor-on data, the model somewhat over-corrects the pressure dependence, changing from -0.29 to $+0.18$, which still equates to a reduction in the correlation but one that is not as successful as for the ROff data.

5.2.4 Reactor Subtraction

The subtraction is completed by scaling the reactor-on and reactor-off data (counts or histograms) by their respective exposure times and subtracting the latter from the former to obtain the IBD yield:

$$S_{\mathrm{IBD}} = S'^{\mathrm{ROn}}_{\mathrm{Signal-Acc-BiPo}} - \frac{t_{\mathrm{ROn}}}{t_{\mathrm{ROff}}} \cdot S'^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}} \,, \tag{5.15}$$

where t_{ROn} and t_{ROff} are the ROn and ROff exposure times. The number of events per day of each background type in the open dataset are summarised in Tab. 5.1. The higher accidental rate during the reactor-on period is due to gamma ray emission from the reactor core, but overall the accidentals are an almost negligible part of the background. The BiPo rate is expected to be independent of the reactor power so the small difference observed is likely caused by higher humidity during the ROff portion of the open dataset. The predominant difference between ROn and ROff arises from the significant excess of neutron-like events during the reactor cycle. The excess is tentatively attributed to IBD events, a hypothesis that is scrutinised in the following sections.

5.2.5 Optimisation of the nuBDT cut

As stated previously, a cut on nuBDT score at 0.73 is taken as the working point for the analysis. The cut has been optimised by scanning over a reasonable range of values and executing the subtraction steps at each point. The optimal point is the one that maximises

Event Type	**ROn [$\mathbf{d^{-1}}$]**	**ROff [$\mathbf{d^{-1}}$]**
Accidentals ($S_{\text{Signal}-\text{Acc}} \cdot \epsilon$)	8.4(6)	6.1(4)
BiPo ($S_{\text{BiPo}-\text{Acc}} \cdot \beta$)	130(1)	137(1)
Neutron-like ($S'_{\text{Signal}-\text{Acc}-\text{BiPo}}$)	363(5)	243(3)
Total	501(5)	386(4)

Table 5.1: Rates of different event types in the open dataset. The numbers in parentheses indicate the statistical uncertainty on the least significant digit of the observed value. There is a significant excess of events in the ROn data compared to the ROff data

the following figure of merit (FoM):

$$\text{FoM} = \frac{S}{\sqrt{S+B}}\,, \tag{5.16}$$

where S is equal to the number of ROn (signal) events remaining after subtraction and B is equal to the number of ROff (background) events that are subtracted. The signal-to-background ratio is also computed at every point in the scan. The results of the scan are displayed in Fig. 5.5. The FoM reaches a maximum of 5.40 ± 0.29 for a cut at 0.73, giving a mean excess per day in the open dataset of 122.5 ± 6.4 events with a signal-to-background ratio of 0.313 ± 0.016. A good agreement between data and MC is found, which is the first validation of the neutron-like excess against the IBD MC prediction.

The MC prediction is obtained by applying the same series of cuts to the IBD MC events and combining the results with the corresponding background rates from the data. The prediction curve is mapped onto the data in such a way as to avoid estimating the absolute IBD rate. This is done in order to negate any data-MC disagreement that may originate from another source such as a miscalculation of the neutron efficiency, thus only the relative effect of the BDT cuts are compared. To achieve this, the data point corresponding to the highest excess is taken as a reference and the prediction points are all shifted equally such that the highest efficiency MC point matches the reference. The prediction as a whole is then allowed to vary within the uncertainty of the reference point until the difference between data and MC is minimised. The uncertainties on the prediction are not calculated as they would be highly correlated with the data errors.

5.2.6 Open Dataset Results

The stacked energy and position background distributions for the open dataset are displayed in Fig. 5.6. Note that the subtraction is performed on the two-dimensional position-energy histograms, which are subsequently projected into the one-dimensional components along each axis. The distributions consist of the reactor-on accidentals ($\epsilon^{\text{ROn}} \cdot S^{\text{ROn}}_{\text{Signal FPNT}}$), reactor-

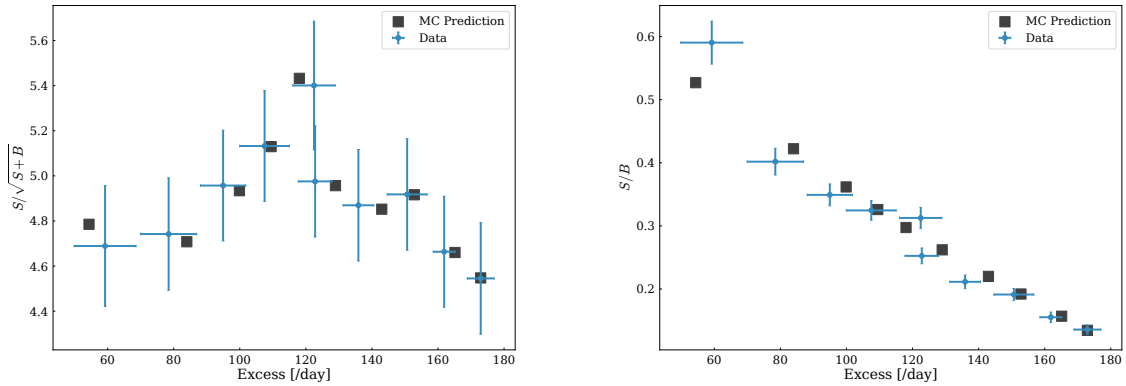


Figure 5.5: Figure of merit and signal-to-background ratio as a function of excess per day (nuBDT cut). Error bars represent the statistical uncertainty on the points; errors on the MC prediction (black squares) are not shown as they are correlated with the data. The FoM is optimised at 120 events per day which gives a signal-to-background ratio of around 0.30.

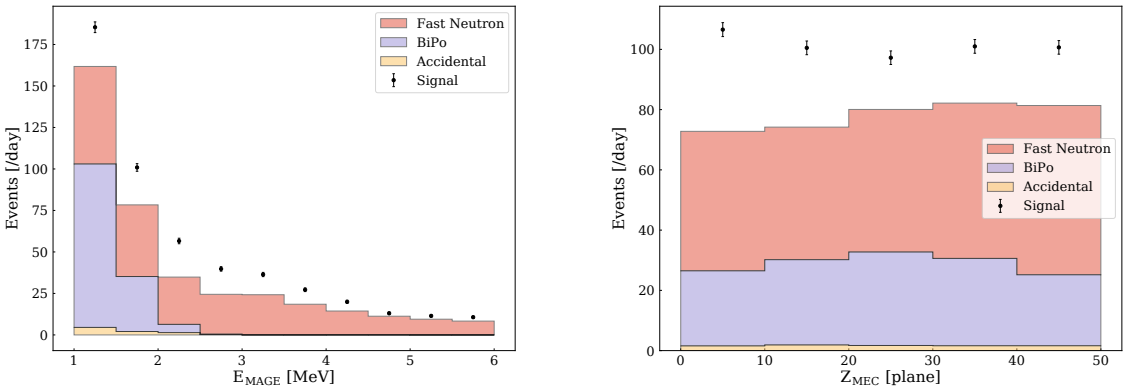


Figure 5.6: Stacked backgrounds and reactor-on signal region distributions (pressure-corrected) of the open dataset in energy and z-position. A clear excess above the background can be observed.

on BiPo ($\beta^{\mathrm{ROn}} \cdot S^{\mathrm{ROn}}_{\mathrm{BiPo-Acc}}$), and reactor-off fast neutrons ($S'^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}}$) scaled to the ROn exposure time. The total pressure-corrected reactor-on signal region distribution $S'^{\mathrm{ROn}}_{\mathrm{Signal}}$—which contains the aggregated backgrounds and the IBD candidates—is superimposed. A clear excess above the background can be observed. Neutron-like excess events are hereafter referred to IBD candidates, IBD events, or equivalent.

BiPo is the dominant background in the first energy bin. The fast-neutron energy spectrum spans the entire energy range (1–6 MeV) and it is clear that the spectrum would extend beyond 6 MeV (where IBD events are found in extremely small numbers) if not for the E_{MEC} limits. The reactor-off Z_{MEC} distribution of fast neutrons follows the same behaviour discussed in section 4.4.2 regarding the shielding provided by the reactor pool. Furthermore, the IBD yield is evaluated against the HM prediction in Fig. 5.7. The prediction is scaled to the number of data events. The total number of reconstructed IBD candidates in the open

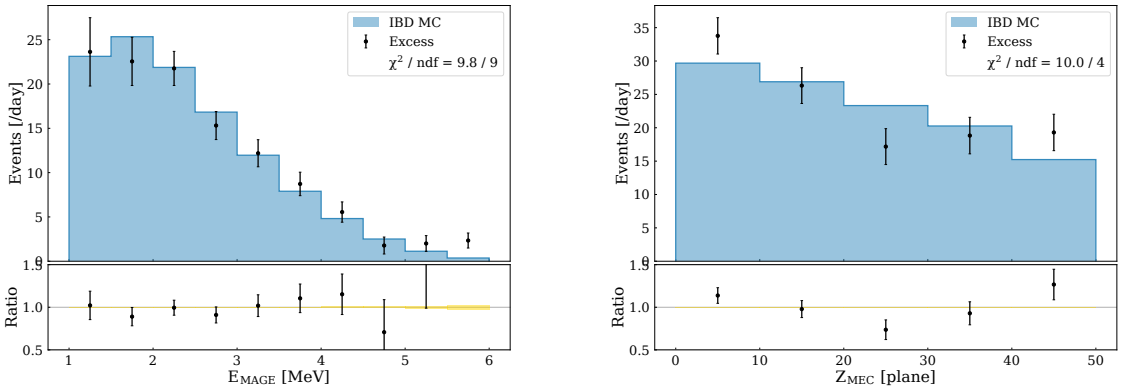


Figure 5.7: Data-MC comparison of open-dataset IBD candidates in energy and position. The one-dimensional histograms are obtained by projecting the background-subtracted two-dimensional histogram along each axis. Error bars represent the statistical uncertainty on the points. The yellow band in the ratio plots represents the statistical uncertainty on the MC prediction. The MC prediction describes the energy spectrum quite well, but there is a significant rate deficit in module three.

dataset is 2709 ± 135. The 2D data-MC comparison gives a reduced chi-square (χ^2/ndf) of 49.6/50, indicating excellent statistical agreement. However, the 1D distributions of E_{MAGE} and Z_{MEC} yield values of 9.8/9 and 10.0/4 respectively. Thus while the energy spectrum retains the statistical consistency against the prediction, the position distribution reveals a two-sigma deficit in the third module. This could be caused by an incorrect evaluation of the per-module efficiency, or simply by a relatively large statistical fluctuation. Analysis of the entire dataset should elucidate the source.

5.3 Phase-I Subtraction

The subtraction is relatively straightforward to perform with the open dataset because the runs are concentrated in a three-month interval, which is short enough that the time-evolution of detector parameters (e.g. light yield) is negligible at the current precision. The total Phase-I dataset spans approximately two years, meaning differences in detector parameters must be taken into account. Hence, a more sophisticated approach to the background subtraction has been developed to maintain the stability of the output. It is evident from both the ^{22}Na and horizontal-muon calibrations that the heating incident of May 2019 had a marked effect on the PVT light yield, as in Fig. 5.8. As well as reducing the absolute PVT light yield, it also caused the rate of degradation of the light yield to accelerate from -1.2% to -3.3% per year. For this reason, the Phase-I dataset is bisected for background subtraction purposes. The two halves are treated as independent datasets, $\mathrm{P_A}$ and $\mathrm{P_B}$, and later combined to form the final result.

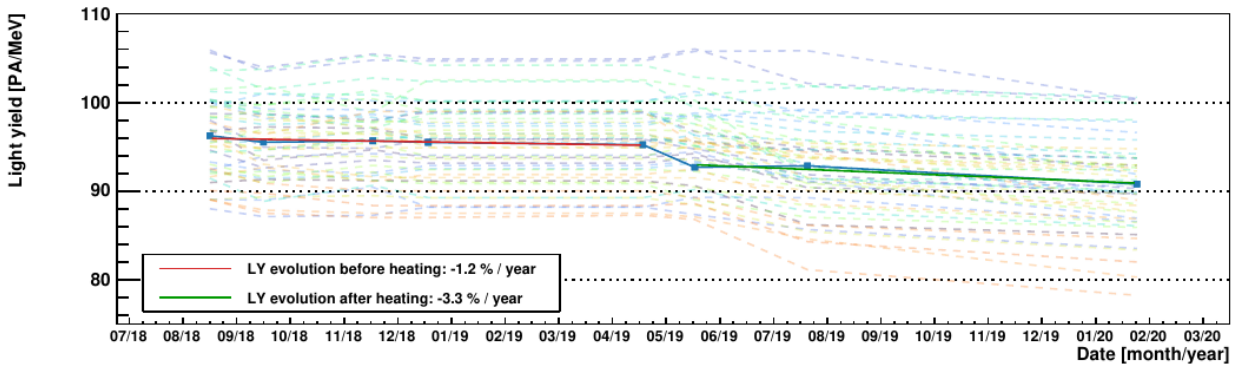


Figure 5.8: Evolution of the PVT light yield estimated by ^{22}Na calibrations [207]. Each blue square corresponds to a CROSS calibration campaign. The faded dashed lines represent the light yield of individual detector planes. The blue line joining the fifth and sixth points covers the time frame in which the heating incident occurred, showing a large, sudden drop in light yield.

5.3.1 Composite Pressure Model

Naively, one could simply compile all the ROn and ROff data, say from part P_A of the dataset, and subtract one from the other to obtain the IBD yield following the process outlined previously. This is perfectly valid for the accidental and BiPo components of the background, but the pressure model is found to be relatively sensitive to the input ROff data from which it is built. As such, one must build separate pressure models for different ROff periods rather than a single global fit. However, since the pressure models are derived from BiPo- and accidental-subtracted rates (per day), local BiPo-to-Signal (β) and ϵ factors must also be computed individually for each ROff period. The number of fast-neutron events in a given day is thus obtained by subtracting the accidentals and BiPo events, utilising the ϵ and β factors associated with the period to which that day belongs.

The reactor-on data is naturally partitioned into manageable chunks of 25–30 d each. A numbering scheme for the reactor-off data has been devised, as in Tab. 5.2, that closely follow that of the reactor cycles where each reactor-off regime is named after one of its neighbouring cycles. However, owing to frequent interruptions in the data-taking, the ROff periods are less well-defined and less regular than the ROn cycles. Notable examples are ROff0 (a small amount of data taken prior to the first reactor cycle), ROff4 and ROff5 (originally a single continuous period that was divided in two), and ROff6 (which includes all the data taken before the sixth reactor cycle and a few days after).

The size of the average reactor-on and reactor-off period is large enough to allow an accurate determination of the BiPo-to-Signal factor and FPNT correction. Shorter intervals are possible but would yield less precise results, especially for the former which are extracted from fits on binned distributions. In contrast, the number of corrections retrievable from the pressure model (once computed) is only limited by the granularity of the time-binning and the frequency of the pressure readings. Experience shows that the average ROff period

Label	Exp. Time [d]	Epoch
ROn1	22.7	1
ROn2	18.6	1
ROn3	18.0	2
ROn4	24.6	2
ROn5	23.6	3
ROn6	24.8	3
ROn7	4.0	4
ROn8	30.4	4
ROn9	23.9	4
ROn10	25.5	4
ROn11	20.7	5
ROn12	27.1	5
ROn13	16.5	5
P_A	132.2	n/a
P_B	148.1	n/a
Total	280.3	n/a

Label	Exp. Time [d]	Epoch
ROff0	4.2	1
ROff1	30.3	1
ROff2	4.1	2
ROff3	14.7	2
ROff4	20.6	2
ROff5	23.3	3
ROff6	8.0	3
ROff7	9.2	4
ROff8	6.3	4
ROff9	6.7	4
ROff10	21.5	4
ROff11	6.2	5
ROff12	12.5	5
ROff13	4.2	5
P_A	105.2	n/a
P_B	66.6	n/a
Total	171.8	n/a

Table 5.2: Delineation of the Phase-I dataset into ROn periods, ROff periods, and pressure epochs which overlap with both. The exposure times include dead-time induced by the muon veto. The double horizontal lines through the middle of each table separate P_A and P_B.

does not include enough data for a precise determination of the pressure fit parameters. Therefore, the Phase-I dataset is also split into five so-called "epochs", each containing a roughly equal amount of reactor-off data.

The breakdown of the dataset into ROff periods, ROn periods, and epochs is schematised in Fig. 5.9. Overall, there are approximately 280 d and 172 d of ROn and ROff exposure time respectively. Note that these values take into account the dead-time induced by the muon veto. The reactor-on exposure time is roughly equal in each half of the dataset. There are three epochs in part P_A, but the relative lack of ROff data only allows for two epochs in P_B. This also implies that the subtraction-based uncertainties will be larger. The first

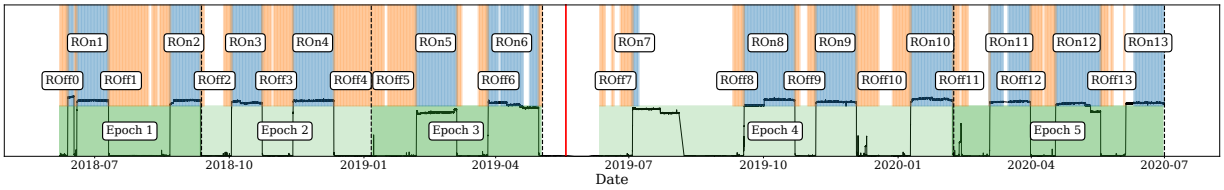


Figure 5.9: Overview of the Phase-I data-taking campaign. The boundaries of the pressure epochs are marked by vertical black dashed lines and alternating light and dark green bands. ROn days and ROff days are illustrated in blue and orange respectively. The solid black line shows the reactor power on an arbitrary scale. The vertical red line marks the heating incident.

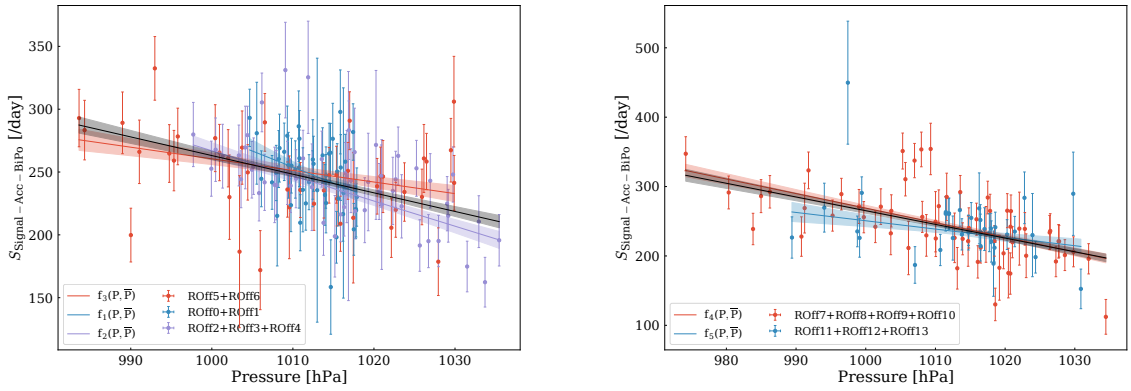


Figure 5.10: Pressure models derived from reactor-off data of different epochs for P_A (left) and P_B (right). The black lines denote the model computed from a fit on all points. Error bars represent the statistical uncertainty on the points. The gradients of the fits can be seen to change over time.

epoch corresponds to the ROff data of the open dataset. Each epoch is used to fit a separate pressure model thus allowing the pressure correction for a given day to be calculated from its local model, as for the β and ϵ factors, which limits the effect of variations in detector parameters over time. This can be condensed into a single function for the entire dataset:

$$F(t, P, \overline{P}) = \sum_{e=1}^{5} \delta(t, e)\, f_e(P, \overline{P})\,, \tag{5.17}$$

where the index e represents different epochs, f_e is the pressure model of the e^{th} epoch and δ is a function that returns 1 if the time t falls within epoch e and returns 0 otherwise.

A number of conclusions may be drawn from the pressure fits per epoch illustrated in Fig. 5.10, where the black lines represent global fits of all data points in each plot. Firstly, the magnitude of the uncertainties on the fitted parameters are coherent between the epochs, except f_4 which is slightly more precise. Also, the gradients of the fits tend to decrease over time because of a drift in detector efficiency. It is also clear that the pressure variations during the open dataset (first epoch) happened to be constrained to a relatively narrow range. The differences between the epoch fits and the global fits demonstrate the limitations of the latter if they were used for the full subtraction. The fast-neutron rates near to the beginning of P_A (or P_B) would be under-corrected on average and rates towards the end would be over-corrected. If the mean pressure over a particular subset of the data is similar to the mean pressure of the whole period, then the corrections will largely cancel out but this is not generally the case.

Overview

After deriving the five pressure models, the subtraction process is repeated twice (once for P_A and once for P_B) from equations 5.3–5.15, substituting $F(t, P, \overline{P})$ into equation 5.13, and using the total summed ROn and ROff samples. The local β and ϵ factors are replaced by versions obtained from the separately summed ROn and ROff data. In all, five partial subtractions (of accidentals and BiPo only) and two full subtractions are executed. The two full subtractions can be expressed qualitatively in terms of the four sub-datasets as P_A^{ROn} minus P_A^{ROff} and P_B^{ROn} minus P_B^{ROff}.

To summarise, the Phase-I subtraction procedure is as follows:

1. The accidental and BiPo components are subtracted separately from each ROn and ROff period outlined in Tab. 5.2 using ϵ and β correction factors derived independently for each period.

2. Five pressure models (one per epoch) are computed by fitting equation 5.12 on the per-day fast-neutron counts that result from step 1., as a function of pressure.

3. The pressure correction for each day is calculated with the combined model $F(t, P, \overline{P})$ defined in equation 5.17.

4. The ϵ and β factors are discarded and the subtraction process is repeated (once for P_A and once for P_B) using the summed pressure corrections (calculated in the previous step) and total summed ROn and ROff samples (from which values of ϵ^{ROn}, ϵ^{ROff}, β^{ROn}, and β^{ROff} are derived).

5.3.2 Time-Evolution of Signal & Background

The evolution of the IBD events and background rates over time is useful for verifying the stability and success of the signal extraction. The BiPo-to-Signal factors per period are superimposed on the container humidity as a function of time in Fig. 5.11. The actual values and summed pressure corrections per period may be found in Tabs. C.1 and C.2. The ROn β factors follow a relatively smooth curve that gradually increases over time. Conversely, the ROff values are relatively constant in P_A but exhibit a sudden jump at ROff7. The change appears to be correlated with the increased humidity at this time since the values decrease again as the humidity reduces. However, there appears to be no consistent relationship between the BiPo-to-Signal factor and the humidity. In particular, it is not clear why the ROn values do not follow the same behaviour and why the BiPo rates in the signal region and BiPo sideband should be affected differently. Unknown elements could be involved so any further conclusions drawn here would be speculative.

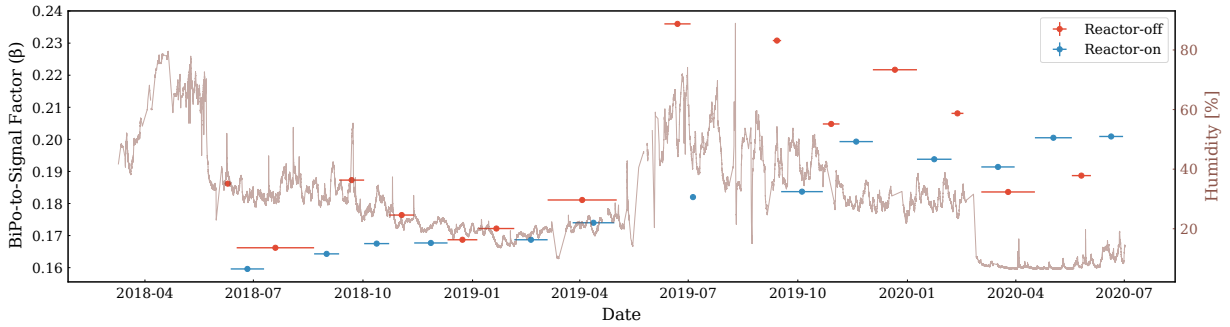
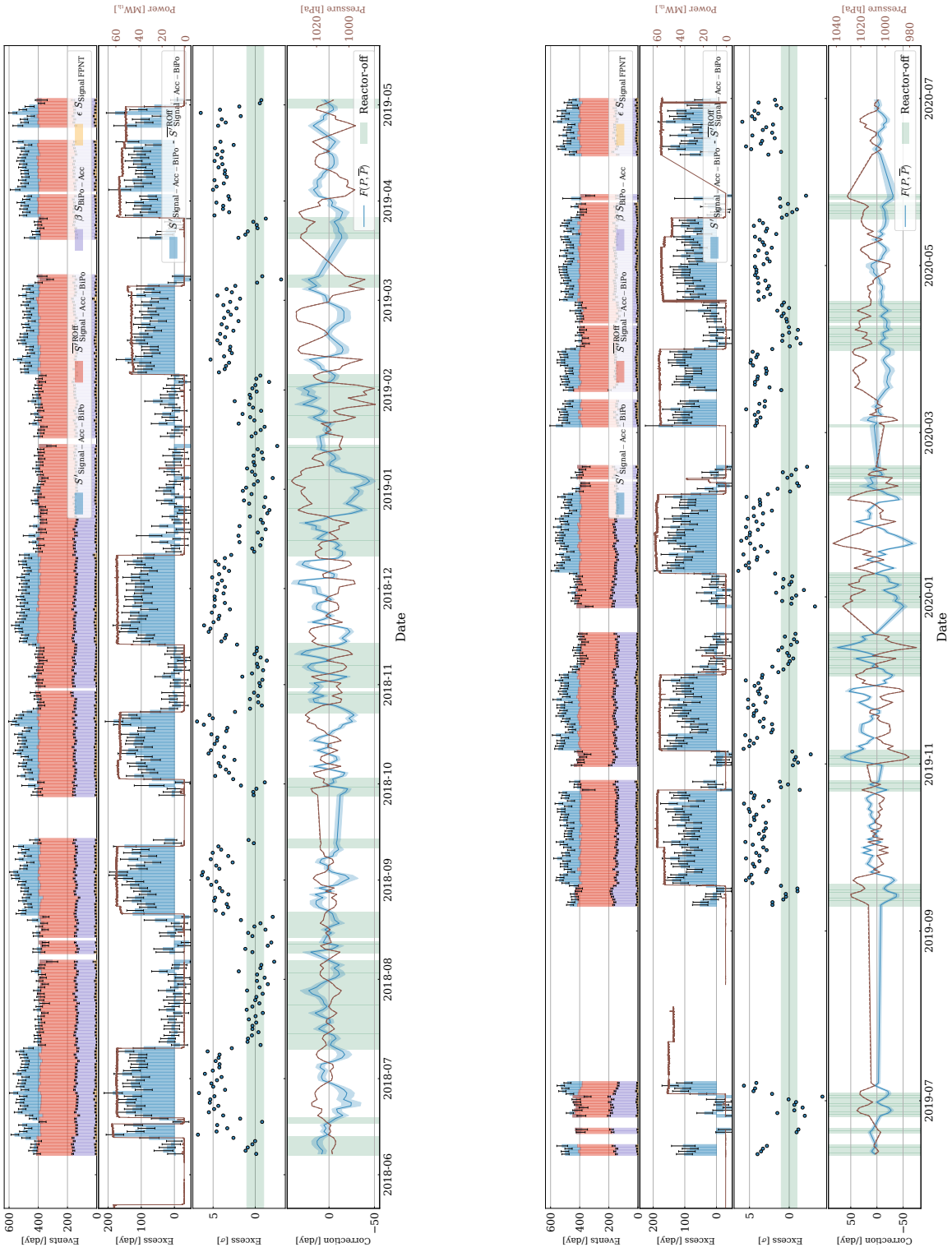


Figure 5.11: Evolution of the BiPo-to-Signal factors over Phase-I. The brown line shows the humidity as a function of time in the container. There is no clear or consistent correlation between the BiPo-to-Signal factors and the humidity, suggesting that other dependencies are involved. Horizontal error bars span the time between a given point and the two neighbouring points.

The trending plot of the IBD candidates and backgrounds is displayed in Fig. 5.12. The uppermost panels display the stacked background events per day, similar to Fig. 5.6. The number of accidental events per day is almost negligible. The BiPo rates over the P_A and P_B time frames are remarkably stable, being mostly maintained within errors at around 175 events per day, though there are occasional instances of greater variations such as during the beginning of ROn7 and throughout ROff8. The neutron-like events are presented in terms of the mean reactor-off rate (red) and the excess (IBDs) above the mean (blue). Thus the apparent variations in the height of the red bars is attributed to the BiPo variations underneath. The number of neutron-like events during reactor-on periods is significantly higher than the mean ROff rate and the number of IBD events is compatible with zero during reactor-off periods, as shown in the second and third panels.

The final panels of Fig. 5.12 track the pressure correction and measured pressure as a function of time. The vertical green bands denote the reactor-off periods from which the pressure fits were derived. The switch from f_4 to f_5 in the composite model is noticeable as a change in width of the pressure-correction error band in early February 2020. The largest corrections for a single day are ± 50 events, which depending on the reactor status can be a significant fraction of the total events. The relatively small pressure variations over the open dataset result in corrections only up to 30 events in magnitude.

The ROn-ROff separation is demonstrated more clearly in Fig. 5.13 which shows the distributions of statistical significance on the number of IBD events per day. The histograms are binned separately for ROn and ROff periods, leading to histograms that are centred on four and zero respectively. However, the extended tails of the P_B reactor-off histogram suggest that the subtraction is less stable for part P_B than P_A, as it contains a number of ROff days that are inconsistent with zero events at the ± 3–4 sigma level. Four of these outliers are located at the beginning of ROff7. A few large statistical deviations can be expected for a dataset of this size, but it is highly improbable that four ROff days with significance



(a) P_A **(b)** P_B

Figure 5.12: Phase-I trending plots of excess and background over time. Left to right in each plot: Stacked rates of each background type where the pressure-corrected neutron-like rate is split into $S'_{\text{Signal}-\text{Acc}-\text{BiPo}} - \overline{S'}^{\text{ROff}}_{\text{Signal}-\text{Acc}-\text{BiPo}}$ and $\overline{S'}^{\text{ROff}}_{\text{Signal}-\text{Acc}-\text{BiPo}}$; the number of antineutrino events per day compared to the reactor power; the number of antineutrino events per day expressed in terms of the statistical significance; the corrections per day computed from the composite pressure model against pressure.

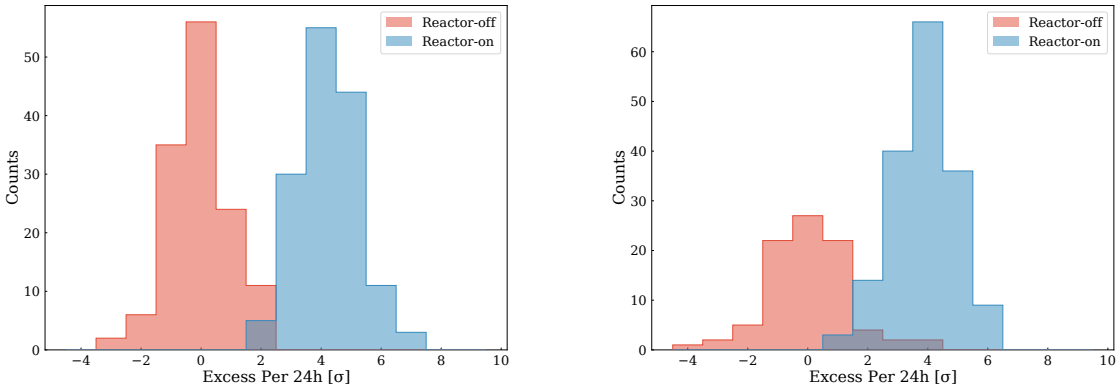


Figure 5.13: Statistical significance of the number of IBD events per day for P_A (left) and P_B (right). The distributions are generally shaped as expected, though the reactor-off P_B histogram has large tails that are suggestive of non-statistical behaviour.

greater than 2.5σ occur consecutively. It may be telling that these dates correspond to the first approved data after the heating incident, but if so the reason is unknown. In any case, these outliers constitute a very small part of the dataset and it is considered inappropriate to remove them now.

5.3.3 Phase-I Results

The Δt fits used to obtain the four BiPo-to-Signal factors referenced in step 4 of the section 5.3.1 summary are provided in appendix C, which also shows the effect of the pressure correction on the four sub-datasets. The distributions of background events in energy and position obtained from the subtractions of P_A and P_B can be found in appendix D. The results of the subtraction on the whole dataset are summarised in Tab. 5.3. The total P_B reactor-on exposure time exceeds that of P_A and consequently it contains a higher number of antineutrinos. However, the number of IBD events per day is lower in P_B because of the degraded PVT light yield. The signal-to-background ratio should be independent of light yield and indeed the two values show good agreement within their stated uncertainties. The histograms extracted from each half of the dataset are summed in quadrature to achieve the final P_{A+B} spectrum.

The data-MC comparison of the position-energy spectrum with the full dataset is presented in Fig. 5.14, separated by detector module. The prediction is normalised to the total number of data events and systematic uncertainties are not included. The deficit in the third module observed in the open dataset persists in the final result, though its significance is slightly reduced. Inspection of the separate P_A and P_B plots—located in appendix E—reveals that the deficit exists solely in the P_A spectrum. Detailed analyses of the experimental spectra are beyond the scope of this chapter. The necessary statistical tools are

Dataset	ROn Exp. Time [d]	IBDs	IBDs [d^{-1}]	S/B
P_A	132.2	$14\,165 \pm 374$	107.1 ± 2.8	0.27 ± 0.01
P_B	148.1	$15\,314 \pm 473$	103.4 ± 3.2	0.25 ± 0.01
P_{A+B}	280.3	$29\,479 \pm 603$	105.2 ± 2.1	0.26 ± 0.01

Table 5.3: Summary of the antineutrinos analysis results in terms of number of reconstructed signal and background events. The signal-to-background ratio is coherent between P_A and P_B and the two halves of the dataset have similar amount of reactor-on data.

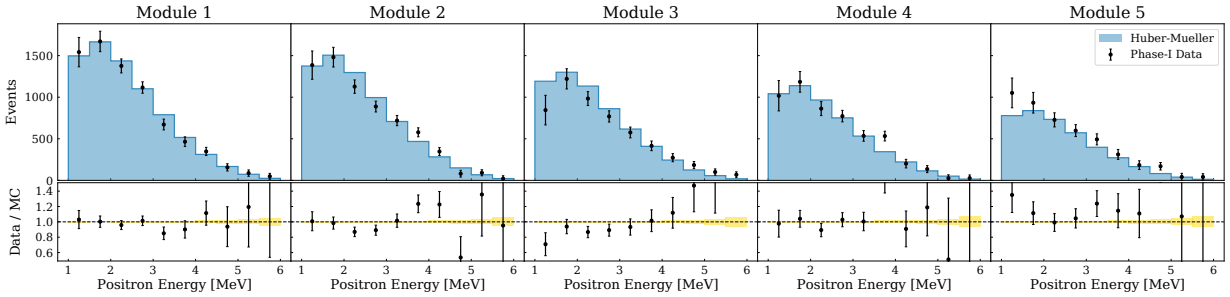


Figure 5.14: Positron energy spectrum per module of the Phase-I IBD yield, formed from the quadratic sums of the P_A and P_B histograms with statistical uncertainties propagated accordingly, compared to the HM model prediction. "Positron Energy" refers to the MAGE energy estimator. Error bars represent the statistical uncertainty on the points. The yellow band in each ratio plot indicates the uncertainty on the MC prediction.

introduced and applied in chapter 6.

5.4 Cross-Checks & Validations

The final part of this chapter is dedicated to validations of the subtraction method and cross-checks on the IBD candidates to further verify that they behave as expected, irrespective of oscillation effects.

5.4.1 Scaling with Reactor Power

Equation 2.3, which expresses the expected number of antineutrinos interacting in the detector, predicts a linear relation between the number of IBD events and the reactor power. Given the combined detection-reconstruction-analysis efficiency, the number of reconstructed IBD candidates should be roughly 10% of the number of antineutrino interactions in the detector per day, which is expected to be around 1100 at 60 MW_{th} according to the Huber-Mueller model. The number of IBD candidates as a function of power is plotted in Fig. 5.15. The individual points are computed by bunching multiple days whose mean power fall within

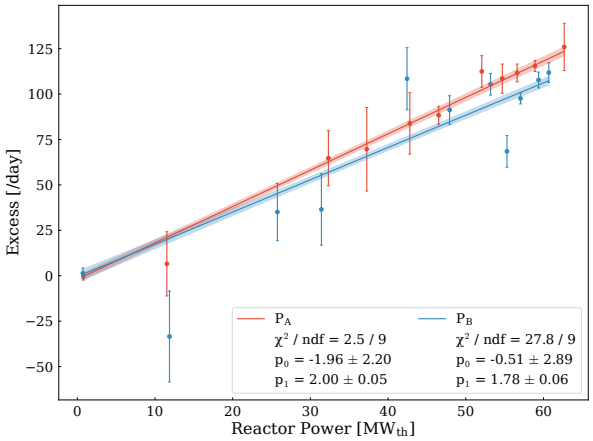


Figure 5.15: IBD rate as a function of reactor power for each half of the dataset. Error bars represent the statistical uncertainty on the points. The linear fits through the points differ in gradient which is consistent with the difference in light yield between the two periods.

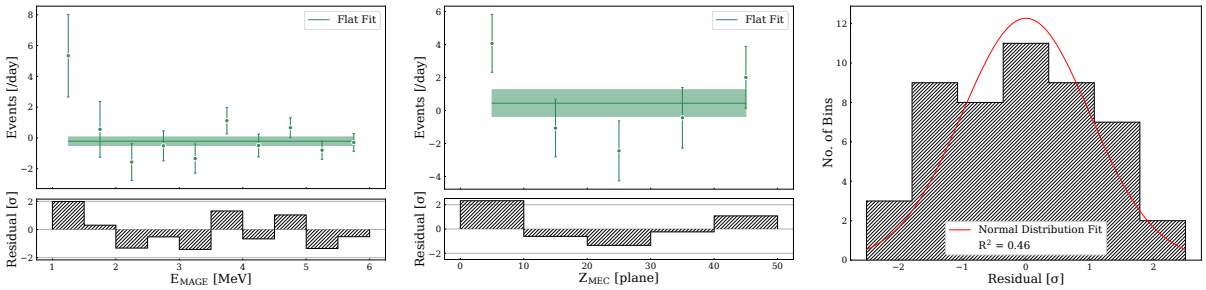
the same bin. The bins get increasingly narrow with increasing power, as follows:

$$P_{\text{lims}} = \{0, 10, 20, 30, 35, 40, 45, 50, 54, 56, 58, 60, 70\} \,, \tag{5.18}$$

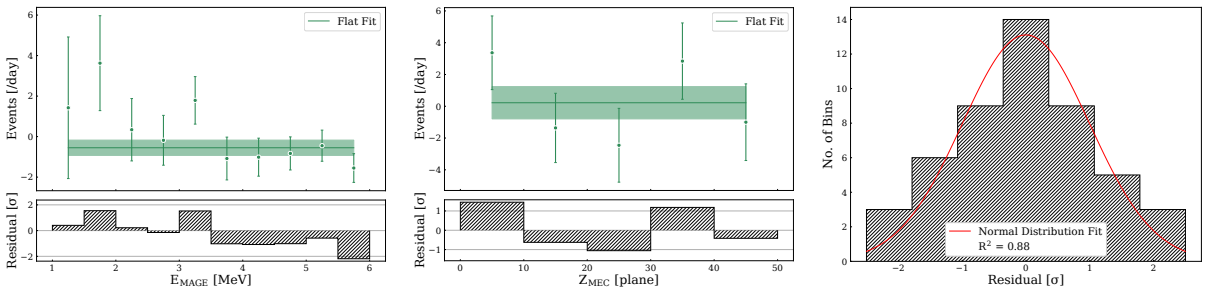
where P_{lims} expresses the boundaries of the power bins in units of MW_{th}. A linear fit describes the rate as a function of power very well for P_A with χ^2/ndf equal to 2.5/9. Also, the number of events in the 58–60 MW_{th} bin is consistent with 10% of 1100, lending additional weight to the assumption that the events are truly from antineutrino interactions. The aforementioned issues with P_B mean that the corresponding fit is not in as good agreement with the data as P_A, as indicated by the χ^2/ndf of 27.8/9. Even so, both fits intercept the y-axis at zero within error bars. The difference in gradient between the two fits can be understood in terms of the degradation of light yield of the scintillating materials, which reduces the number of reconstructed events. For example, at first order the neutron trigger efficiency is proportional to the ZnS light yield, so as time progresses the mean number of peaks over threshold will decrease, leading to fewer reconstructed antineutrinos.

5.4.2 ROff-Minus-ROff Subtraction

The next cross-check of the subtraction method is a so-called ROff-minus-ROff subtraction, where one reactor-off period is subtracted from another. If the procedure is working correctly, the number of events remaining after subtraction should be consistent with zero. The ROff data of each half of the dataset is further bisected to create two pairs of equally sized samples. This is done after randomly shuffling the data to ensure that no bias is introduced in the choice of the subtracted periods. Thus the dataset is separated into four sections: P_{A1}



(a) P_A



(b) P_B

Figure 5.16: ROff-minus-ROff subtraction for P_A and P_B. Error bars represent the statistical uncertainty on the points. Each histogram is fit with a 0^{th} degree polynomial drawn in green with the associated uncertainty band. The residuals are obtained from the two-dimensional histograms and fit with a Gaussian whose R^2 metric implies a poor fit for P_A.

(54.8 d), P_{A2} (50.4 d), P_{B1} (33.5 d), and P_{B2} (33.0 d). Two subtractions are then executed: $P_{A1}-P_{A2}$ and $P_{B1}-P_{B2}$, thus retaining the independence between P_A and P_B. For each subtraction a pressure model is computed from the subtrahend only, since the purpose of this test is to validate the ROn-minus-ROff subtraction procedure in which the reactor-on fast-neutron rate is corrected with a model derived from the reactor-off sample.

The results for E_{MAGE} and Z_{MAGE} are displayed in Fig. 5.16. The histograms are fit with a 0^{th} degree polynomial whose single parameter is equivalent to the weighted average over all the bins. Both fits are consistent with zero events at the one-sigma level. The residuals in the lower pads are computed by dividing the bin entries by their uncertainties and are unrelated to the fits. The histograms at the end of each row are the distribution of residuals from the two-dimensional subtracted histogram. Non-statistical behaviour manifests in these types of plots as deviations from a true normal distribution. The histograms are modelled by a normal distribution using an ordinary least-squares (OLS) fit with a single free parameter—the normalisation—resulting in convergence for both cases. The goodness of fit is evaluated using the coefficient of determination (R-squared), calculated as 0.46 and 0.88 for $P_{A1}-P_{A2}$ and $P_{B1}-P_{B2}$ respectively, which points to a relatively good fit for the latter and a relatively poor fit for the former. Deviations from statistical behaviour will be parameterised as a systematic uncertainty associated to each position-energy bin.

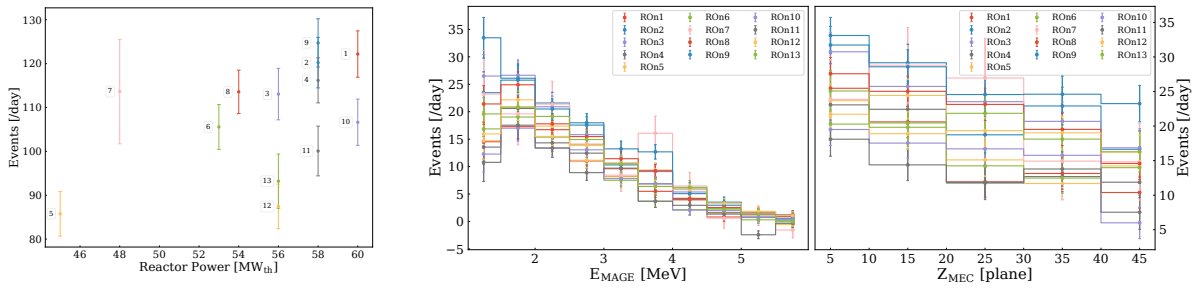


Figure 5.17: Mean IBD rates and spectra for each individually background-subtracted reactor cycle of Phase-I. Error bars represent the statistical uncertainty on the points. The rate points show two distinct trends, roughly correlated with the two halves of the dataset.

5.4.3 Individual Reactor Cycles

The penultimate check is to perform the subtraction on each reactor cycle individually and test the statistical consistency of the IBD spectra, regardless of the data-MC agreement. Some practical differences between two given ROn cycles are the operating power and the configuration and composition of fuel elements in the core. The reconstructed antineutrino rate depends on both the reactor power and PVT light yield which are difficult to disentangle. The compiled P_A reactor-off data is used to remove the background from each reactor cycle of P_A and similarly for P_B. Fig. 5.17 shows the rates and spectra of each cycle. Two distinct trends can be observed in the IBD rates as a function of power. Cycles 1–9 and cycles 10–13 fall on lines of different gradient, which may be related to the accelerated light yield degradation following the heating incident, but this is not clear.

Colleagues in the SoLid group at LPC Caen have designed a log-likelihood fit to compare the spectra of each pair of cycles. One spectrum is treated as the "reference" and one as the "model". The log-likelihood is defined as:

$$\log L(M, R, n) = -0.5 \times \sum_i \frac{(nM_i - R_i)^2}{n\,\sigma_{M_i} \cdot \sigma_{R_i}}\,, \tag{5.19}$$

where L is the likelihood, n is the normalisation, M_i and σ_{M_i} are the value and uncertainty of the i^{th} bin in the "model" spectrum, and R_i and σ_{R_i} are the same quantities for the reference spectrum. The likelihood assumes Gaussian behaviour. The normalisation n is allowed to vary randomly between zero and some arbitrary large value (N_{max}), building a distribution of $\log L(M, R, n_2)/\text{ndf}$ values that satisfy $e^{\log L(M,R,n_1) - \log L(M,R,n_2)} > r$, where r is a uniform random number between 0 and 1 and n_1 and n_2 are uniform random numbers between 0 and N_{max}.

The mean of this distribution behaves like a reduced chi-square metric. A heat map of the metric between each pair of cycles is provided in Fig. 5.18. The matrix is effectively symmetric because all uncertainties are incorporated in the fit. The most sizeable discrepan-

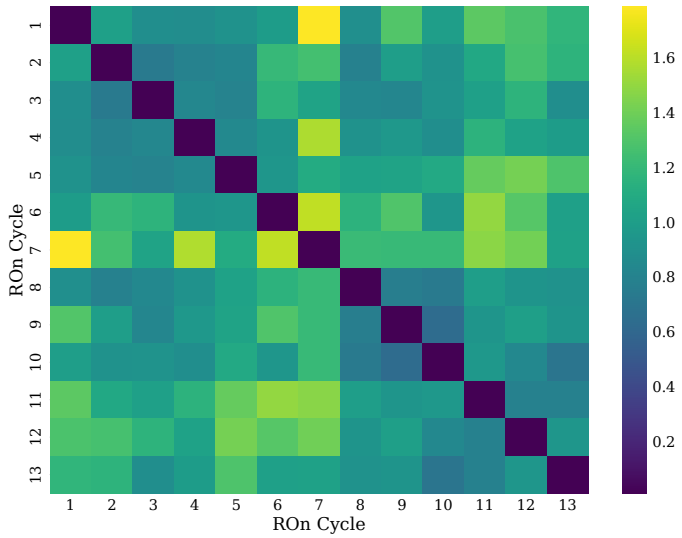


Figure 5.18: Comparison between all individually background-subtracted ROn cycles. Each square corresponds to a value of the metric computed using the LPC Caen log-likelihood fit. The ROn7 spectrum is incompatible (larger than one-sigma deviation) with that of cycles one, four, and six.

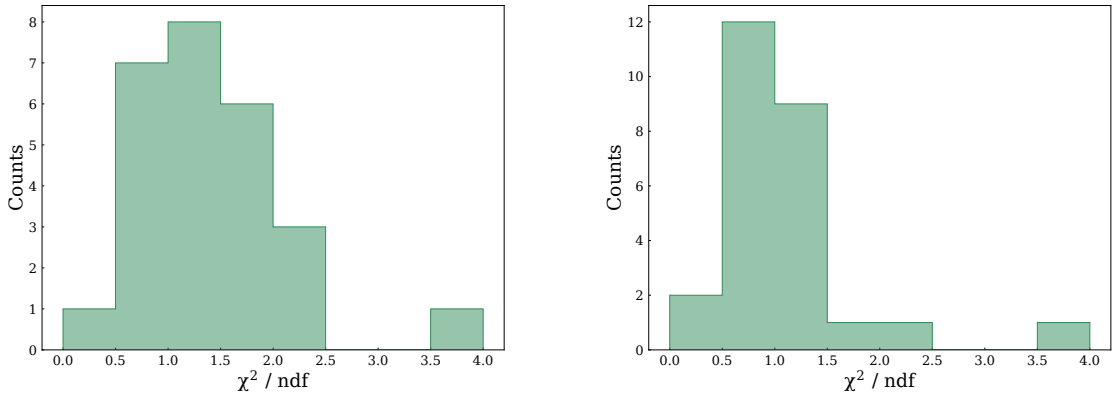
cies are between ROn7 and ROn1, ROn4, and ROn6. Since ROn7 only contributes 1.4% of the total ROn live-time and no other cycles show evidence of a systematic shape difference, these discrepancies are not considered a priority to resolve.

5.4.4 Data-MC Comparisons

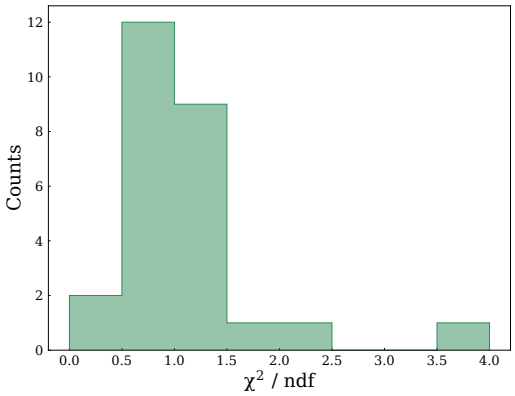
Eight control variables that are insensitive to neutrino oscillation effects have been selected to validate the IBD MC prediction with the precision of the full dataset. The variables are:

1. ΔR
2. EnGamma1
3. EnGamma2
4. EnLeftoverInCross
5. X_{MEC}
6. Y_{MEC}
7. DotProduct
8. nuBDT score

The motivations for these choices are as follows. ΔR encapsulates the displacements in all three spatial directions and is a very powerful signal-background discriminant, meaning a large enough inaccuracy in the determination of a background rate should emerge in the background-subtracted spectrum. The three energy variables are chosen because together they comprise the vast majority of non-positron ES energy. DotProduct is a relatively important nuBDT training feature that succinctly encapsulates the gamma reconstruction. The positron position variables X_{MEC} and Y_{MEC} act as proxies of the longitudinal position



(a) P_A (left) and P_B (right).



(b) The combined P_{A+B} results.

Figure 5.19: Distributions of χ^2/ndf values from data-MC comparisons of all ES variables presented in Tab. 4.4. At least one outlier is present in each distribution.

Z_{MEC}. Finally, the nuBDT score is included because it represents one of the most efficient analysis cuts in terms of background rejection.

Before considering the data-MC results of the control variables, a global view of the data-MC agreement is provided in Fig. 5.19. It shows the histograms of reduced chi-squares of *all* variables defined in Tab. 4.4. The first two histograms are for P_A and P_B separately; the third is the combined P_{A+B} results. The P_A and P_B distributions appear Gaussian-like and centred on unity, as expected from statistical fluctuations. However, there are clear outliers corresponding to Z_{MEC} (χ^2/ndf = 3.6) and ΔY (χ^2/ndf = 3.7) respectively. The reversed comparison—the Z_{MEC} value of P_B and ΔY value of P_A—gives reduced chi-squares of 1.4 and 1.3 respectively. The two outliers manifest in the P_{A+B} histogram as the single entry in the right-most bin (ΔY) and the 2.5–3.0 bin (Z_{MEC}). The latter also contains nuBDT score and NumGamma2.

Fig. 5.20 depicts the combined P_{A+B} data-MC plots for the control variables. All 29 479 IBD candidate events are included, allowing the data-MC agreement to be probed at the

$\mathcal{O}(10)\%$ level. The MC distributions are scaled to the number of data events. Some of the variables demonstrate reasonable agreement between data and prediction, (ΔR and EnGamma1, for example) yet others show a stark disagreement, especially the nuBDT score. It is highly probable that systematic uncertainties become relevant at this precision. The systematic uncertainties for the oscillation analysis are presented in the next chapter.

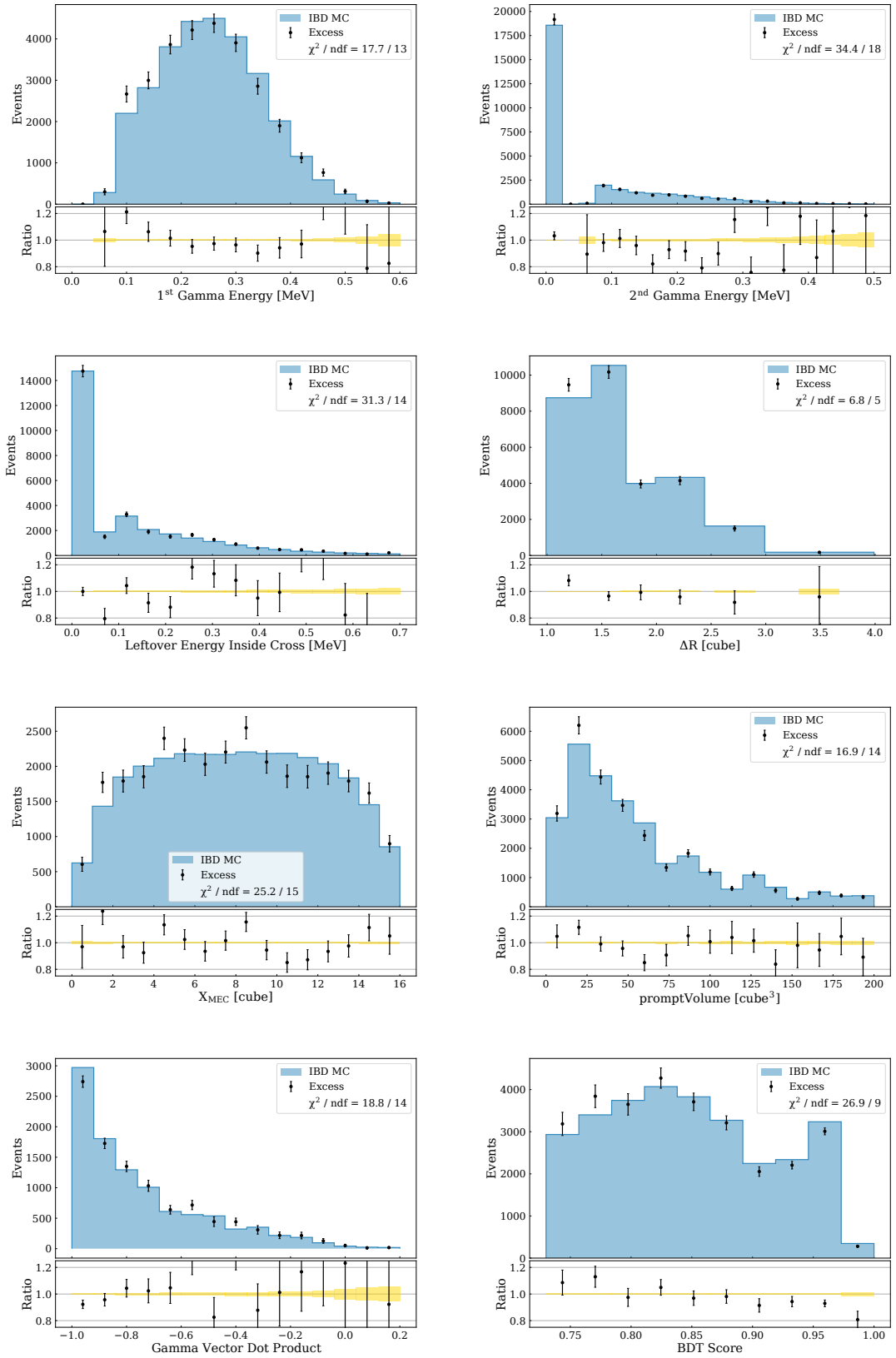


Figure 5.20: Comparisons between the IBD MC and subtracted data of eight control variables. Error bars represent the statistical uncertainty on the points. The yellow band in each ratio plot indicates the uncertainty on the MC prediction. The level of discrepancy between data and MC in some variables suggests that systematic uncertainties become important at this precision.

Chapter 6

Oscillation Analysis

If the RAA can indeed be explained by active-sterile neutrino mixing, a distortion that is consistent with an oscillation pattern should be observable in the spectrum. Hence, the sterile neutrino hypothesis of the RAA is tested by performing a *shape-only* disappearance analysis of the experimental antineutrino position-energy spectrum with respect to the HM prediction. The statistical tools employed for this purpose are primarily frequentist in nature.

In this chapter, the method of parameterising detector effects and applying them to any arbitrary physics model is introduced. Statistical and systematic uncertainties of the analysis are then evaluated, followed by a description of hypothesis testing and the construction of exclusion regions in the Δm^2_{41}-$\sin^2(2\theta_{14})$ plane. The oscillation fit is then applied to the Phase-I IBD spectrum for the first high-statistics, frequentist sterile neutrino search of the SoLid experiment.

6.1 Obtaining the Predictions

The SoLO code was introduced previously in the context of the reactor simulation. SoLO combines the distribution of fissile material in the core, the predicted antineutrino spectrum per isotope, and the detector geometry. The output is a list of antineutrino events in true-space with associated physics parameters, crucially the energy and distance travelled by the antineutrino. The events are then passed to SoLidSim and propagated through the entire analysis chain—including preselection and nuBDT cuts—to produce the prediction in reco-space. The prediction histograms in true- and reco-space of the first IBD MC cycle are illustrated in Fig. 6.1. The binning scheme of the reconstructed variables follows the IBD analysis with five position bins and ten energy bins. The key functionalities of SoLO for an oscillation analysis have been re-written for this thesis in a framework called Python3 for SoLid Oscillations (PySoLO). PySoLO drops all dependencies related to the reactor antineutrino generator, which remains the principle function of SoLO.

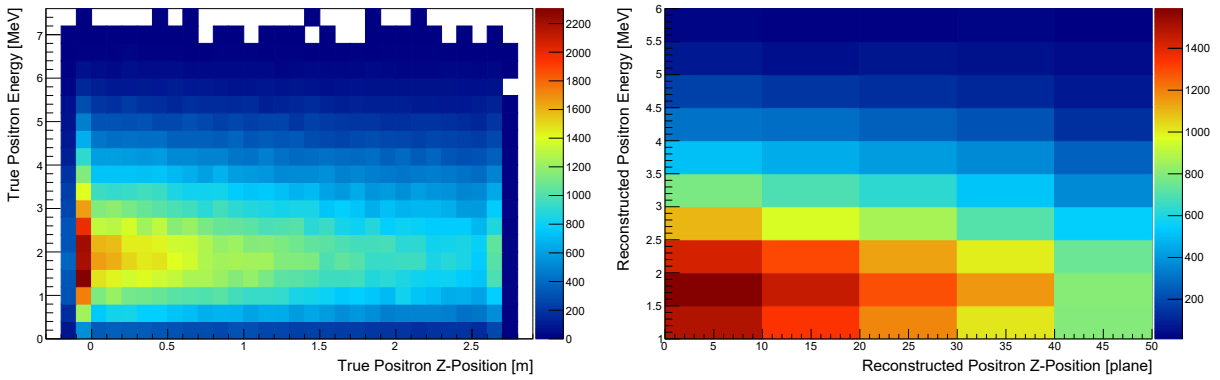


Figure 6.1: Two-dimensional IBD MC prediction in true (left) and reco space (right). The ratio between the respective total number of events reflects the overall efficiency, but the absolute scale is arbitrary. The first and last position bins in true-space are populated by IBD events in non-active material of the detector.

It is essential for an oscillation analysis that predictions from models in which the active-sterile mixing parameters are non-zero can be obtained. For the 3+1 neutrino model, this may be achieved by assigning a weight to every true-space event equal to the antineutrino survival probability under the two-flavour approximation:

$$W = 1 - \sin^2(2\theta_{14}) \sin^2\left(1.27\,\Delta m^2_{41}\,\frac{L}{E}\right) , \tag{6.1}$$

where E and L are the antineutrino energy and path length respectively. The weights determine the contribution of each event to its corresponding bin in the position-energy histogram. In general, a particular physics model must be simulated *before* detector effects are applied. Therefore, in order to test an arbitrary number of possible models, one must have a method of efficiently translating between true- and reco-space that accurately parameterises smearing and efficiency effects, without having to repeatedly run the full simulation.

6.1.1 Response Matrix

The standard method of applying detector effects quickly is to build a response matrix. An alternative technique based on single-event smearing with the GaussExp parameterisation of the reconstructed energy has also been developed and is outlined in appendix F, but the response matrix method is preferred in this analysis for various reasons. A response matrix uses the one-to-one link between the original true-space event and the final surviving reco-space event to build a probability matrix that defines a linear mapping between binned distributions. The matrix element (j, i) is the probability that an event in true bin i is measured in reco bin j. An arbitrary truth distribution that shares the same true-space

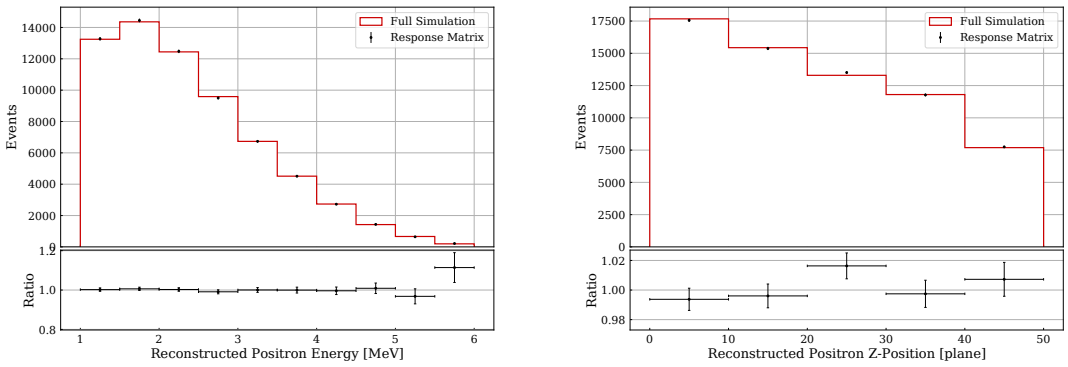


Figure 6.2: Performance of the response matrix on IBD MC events against the full simulation. Error bars represent the statistical uncertainty on the points. Agreement within 2% is observed for the vast majority of bins.

binning can then be convoluted with detector effects via matrix multiplication:

$$H_{\text{reco}} = M \times H_{\text{true}} \,, \tag{6.2}$$

where H_{reco} and H_{true} are the histograms in reco- and true-space respectively. Translation in this direction is known as "folding", as opposed to "unfolding" methods which aim to extricate the original truth distribution by removing detector effects. The histograms H_{reco} and H_{true} are first flattened along a single dimension such that a given bin represents both a position and an energy. Thus the response matrix has dimensions:

$$D_M = \prod_t N_{\text{bins}}^{V_{\text{true},\,t}} \times \prod_r N_{\text{bins}}^{V_{\text{reco},\,r}} \tag{6.3}$$

where $N_{\text{bins}}^{V_{\text{true},\,t}}$ and $N_{\text{bins}}^{V_{\text{reco},\,r}}$ are the number of bins for the t^{th} true-space variable $V_{\text{true},\,t}$ and the r^{th} reco-space variable $V_{\text{reco},\,r}$ respectively.

A response matrix can encode the detector efficiency if unreconstructed events are incorporated in the computation of the probability map. Training a response matrix requires a very large number of MC events in order to accurately reproduce the results of the full detector simulation. The number of events required depends on the binning scheme, which is a compromise between dimensionality and the statistical precision of each bin. The choice of true-space variables is flexible in that any variable upon which detector effects may depend can be included. However, every additional variable increases the size of the response matrix and consequently the number of MC events needed to generate it. PySoLO utilises the Response Matrix Utilities (ReMU) package [233, 234] to build the response matrix. To limit the dimensionality of the response matrix, true-space events are binned in initial positron energy and the z-coordinate of the IBD vertex only. The reconstructed events are binned in positron energy (E_{MAGE}) and z-position (Z_{MEC}), i.e. detector plane number.

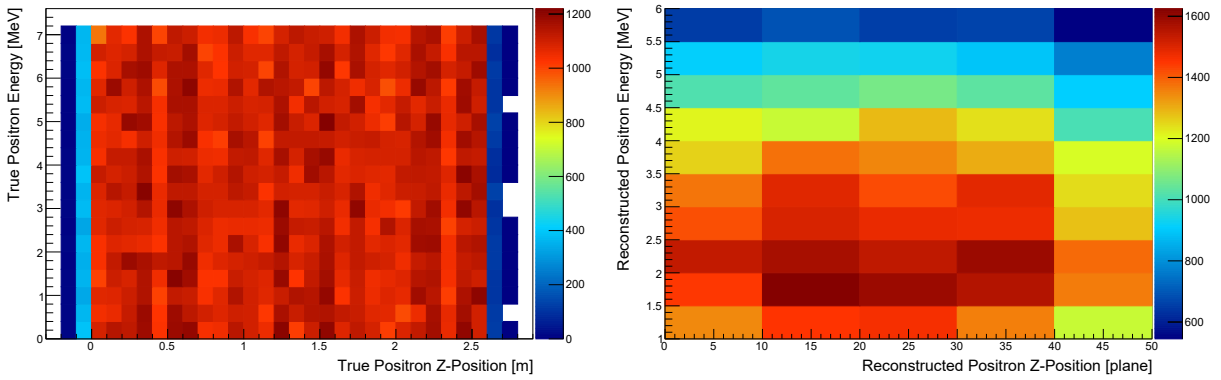


Figure 6.3: Two-dimensional Flat MC prediction in true (left) and reco space (right). The ratio between the respective total number of events reflects the overall efficiency but the absolute scale is arbitrary. The relative difference in efficiency between modules is evident in the right-hand side plot, particularly in the final module where there are significantly fewer events.

The response matrix for this analysis is built from 25 486 563 and 2 662 677 true-space and reco-space events respectively, corresponding to an overall efficiency of around 10%. The performance is evaluated at first instance by applying the matrix to the binned distribution of an independent sample of true events (numbering 280 423) and comparing the result to the corresponding reconstructed distribution (numbering 29 289) from the full simulation. The resulting one-dimensional energy and position projections are provided in Fig. 6.2. The forward-folded histograms of both variables are statistically consistent with the full simulation to within 1–2%. Moreover, the response matrix correctly accounts for the global detector efficiency and the local efficiency of each bin. Neglecting the uncertainties on the full-simulation histogram, the two-dimensional χ^2/ndf is 32.1/50, indicating excellent statistical coherence.

Furthermore, it is important that the matrix is model independent, i.e. that it parameterises detector effects with minimal bias from the input models. A test of the model independence of the response matrix is devised by applying it to a histogram containing events of a different physics model from which it was constructed. The SoLO IBD generator is used to generate a uniform distribution of IBD vertices in energy and over the detector planes. The vertex map is then processed with SoLidSim and propagated through the entire analysis chain. This is dubbed the *Flat MC*. As visualised in Fig. 6.3, the true-space event rate is highly homogeneous in all but the first two and last two position bins, which represent antineutrino interactions in non-active detector material. The decoherence of the homogeneity in both energy and position for reconstructed events is caused by bin-dependent efficiency effects. For example, the IBD efficiency of the final two planes is significantly lower than the rest of the detector because of dead channels and the relatively high probability of the neutron escaping the fiducial volume.

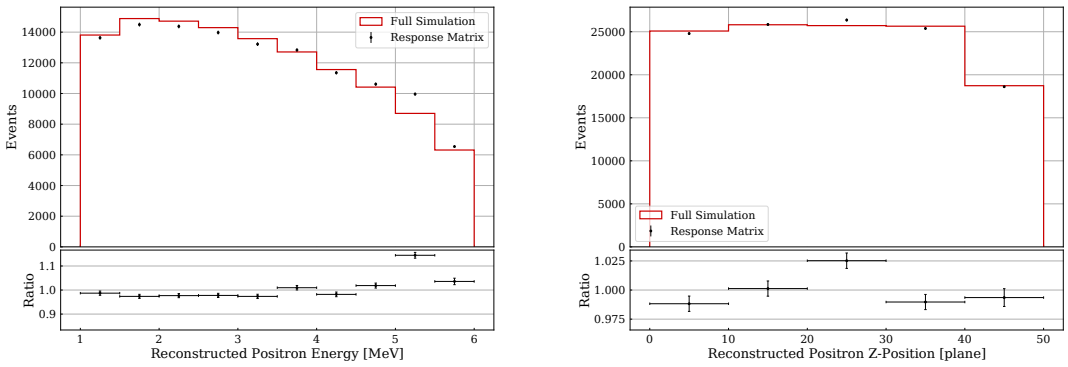


Figure 6.4: Performance of the response matrix on Flat MC events against the full simulation in energy (left) and position (right). Error bars represent the statistical uncertainty on the points. The response matrix underestimates the global efficiency so the corresponding histogram is scaled to the total number of events in the full simulation.

Applying the original response matrix (trained on IBD MC events) to the Flat MC truth distribution (522968 events) results in the histograms drawn in Fig. 6.4. The agreement between the full simulation and the forward-folded prediction is not as optimal as before. The response matrix underestimates the global efficiency by approximately 10%. As such, the forward-folded prediction in Fig. 6.4 has been scaled to the number of events of the full simulation in order to compare the spectrum shapes.

There is a systematic 2–3% deviation from the full simulation in the energy spectrum at 1.0–3.5 MeV, a relatively small effect but somewhat larger than seen previously. The 5.0–5.5 MeV bin deviates by over 10%. Regarding the position distributions, all five bins show agreement to within 2% except for the third module which is at the 2.5% level. This is similar to and only marginally worse than the performance on the IBD MC test sample. It is recommended that the model independence of the response matrix be improved in future analyses by incorporating additional truth parameters upon which the energy smearing and efficiency depend. The number of IBD MC events currently available is likely high enough to accommodate a modest expansion to the matrix dimensionality.

6.2 Error Propagation

Correct evaluation of the statistical and systematic uncertainties is crucial in a search for new physics. The statistical uncertainty on the IBD spectrum is driven by the number of signal events and the background subtraction. An estimation of the statistical uncertainties can be derived by careful consideration of these factors, which is covered in the context of the Asimov dataset. The majority of this section is dedicated to evaluation of the systematic uncertainties relevant for the oscillation analysis.

6.2.1 Statistical Uncertainties

The antineutrino analysis and background subtraction of the Phase-I dataset yields 29 479 ± 603 (stat.) IBD events with an overall signal-to-background ratio of 0.26 ± 0.01. The total reactor-on and reactor-off exposure times are 280.3 d and 171.8 d respectively. The statistical uncertainty is dominated by the background subtraction because of the high number of background events. Before subtraction, a given position-energy bin contains the background count B_{ij} and the signal count S_{ij}. Assuming Poisson statistics, a simplified estimate of the statistical uncertainty on bin ij after subtraction is given by:

$$\sigma_{ij} = \sqrt{S_{ij} + B_{ij}\left(1 + \left[\frac{t_{\mathrm{ROn}}}{t_{\mathrm{ROff}}}\right]^2\right)}\,, \tag{6.4}$$

where t_{ROn} and t_{ROff} are the total reactor-on and reactor-off exposure times respectively. An implicit assumption in this equation is that the number of background events is equal between the ROn and ROff periods. The uncertainty approaches the standard Poisson error as B_{ij} tends to zero. The Phase-I dataset is somewhat lacking in reactor-off data—particularly in 2019—so the overall statistical uncertainty is enhanced compared to the scenario in which there are equal amounts of reactor-off and reactor-on exposure time.

The *Asimov dataset* is defined as the data of an idealised experiment in which the number of signal events in each bin is exactly equal to the prediction. It is effectively the mean result one would expect from running $N \to \infty$ experiments. A background template can be utilised to estimate the statistical errors of the Asimov dataset. The template is derived from a combination of ROn and ROff data, by isolating the three background components in their respective sidebands ($S^{\mathrm{ROn}}_{\mathrm{Signal\,FPNT}}$, $S^{\mathrm{ROn}}_{\mathrm{BiPo-Acc}}$ and $S'^{\mathrm{ROff}}_{\mathrm{Signal-Acc-BiPo}}$) which are then scaled to their post-analysis-cuts rates*. The distributions are summed and normalised to yield the final shape template. By constructing the templates in this way, the effect of changing the relative rates of each background on the statistical errors per bin can be studied (e.g. if a future improvement to BiPonator were to grant more effective BiPo rejection for the same neutron efficiency). Fig. 6.5 shows the one-dimensional projections of the background templates created separately for $\mathrm{P_A}$ and $\mathrm{P_B}$. The remarkable agreement between the two energy templates is a testament to the accuracy of the ES calibration with horizontal muons. In contrast, the Z_{MEC} templates differ significantly in the first two bins.

The number of background events in a given position-energy bin for an overall signal-to-background ratio $R_{\mathrm{S/B}}$ can be predicted as:

$$B_{ij} = \frac{N_{\mathrm{IBD}}}{R_{\mathrm{S/B}}} \times T_{ij} = \frac{N_{\mathrm{IBD}}}{R_{\mathrm{S/B}}} \times \left(T^{\mathrm{Acc}}_{ij} + T^{\mathrm{BiPo}}_{ij} + T^{\mathrm{Atm}}_{ij}\right), \tag{6.5}$$

where T^{Acc}_{ij}, T^{BiPo}_{ij} and T^{Atm}_{ij} are the background template components from the aforemen-

*Equivalent to the stacked histograms of Fig D.1.

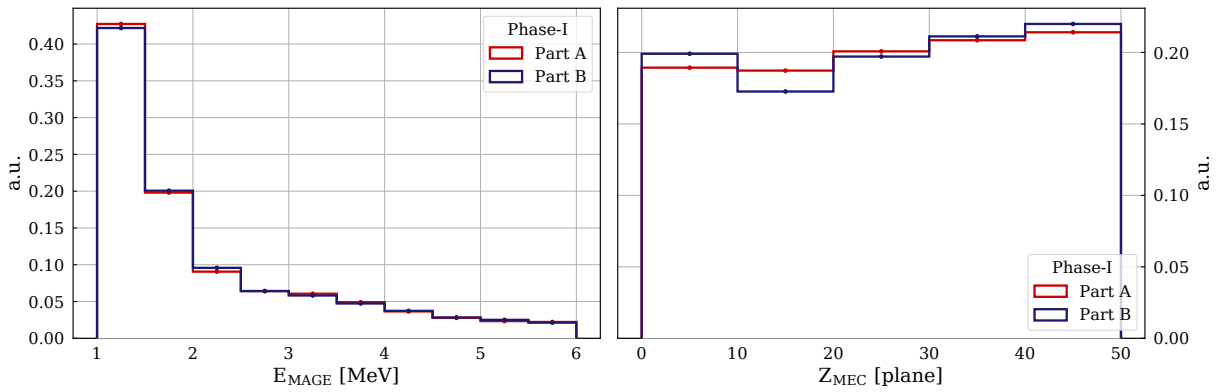


Figure 6.5: 1D projections of the background templates for each section of the Phase-I dataset. Statistical uncertainties are negligible. The two energy templates show excellent agreement; the position templates differ by roughly 10% in the third module.

tioned sidebands and N_{IBD} is the total number of IBD candidates. In this work only a single background template, equal to the normalised sum of the $\mathrm{P_A}$ and $\mathrm{P_B}$ templates, is implemented. Substituting the value of B_{ij} into equation 6.4 gives a crude estimate of the bin ij uncertainty. However, doing so assumes that all background events are subtracted simultaneously and that the error on the background is simply the square root of the number of background events. This is not the case for the real subtraction because of the high-statistics BiPo sideband which reduces the fractional uncertainty on the number of BiPo events in the signal region by a factor $\sqrt{\beta}$. The final statistical uncertainty on bin ij can then be estimated as:

$$\sigma_{ij} = \sqrt{S_{ij} + \left(1 + \left[\frac{t_{\mathrm{ROn}}}{t_{\mathrm{ROff}}}\right]^2\right)\left(\frac{N_{\mathrm{IBD}}}{R_{\mathrm{S/B}}} \times [T_{ij}^{\mathrm{Acc}} + \overline{\beta} T_{ij}^{\mathrm{BiPo}} + T_{ij}^{\mathrm{Atm}}]\right)}\,, \tag{6.6}$$

where $\overline{\beta}$ is the mean BiPo-to-Signal factor of all ROn and ROff periods (0.1876). The error on the BiPo-to-Signal factor itself is only $\mathcal{O}(0.1)\%$ so has been neglected. Contributions from other sources, such as the pressure correction, are also negligible. The accuracy of the estimation is further improved by splitting the computation into two parts corresponding to the independent subtractions of $\mathrm{P_A}$ and $\mathrm{P_B}$, in which case the error becomes:

$$\sigma_{ij} = \sqrt{\sigma_{ij,\mathrm{A}}^2 + \sigma_{ij,\mathrm{B}}^2}\,, \tag{6.7}$$

where $\sigma_{ij,\mathrm{A}}$ and $\sigma_{ij,\mathrm{B}}$ are calculated using equation 6.6. Summing the two errors on each bin in quadrature yields a total estimated uncertainty of 614, which is very close to the total statistical error of the data (603). The statistical uncertainty per bin of the Asimov dataset and experimental data spectrum are depicted in Fig. 6.6. A fractional error below 10% is reached by the most precise bins. Future analyses could also implement the uncertainties on the correction factors β, ϵ, and $F(t, P, \overline{P})$ and calculate $\sigma_{ij,\mathrm{A}}$ and $\sigma_{ij,\mathrm{B}}$ using the respective

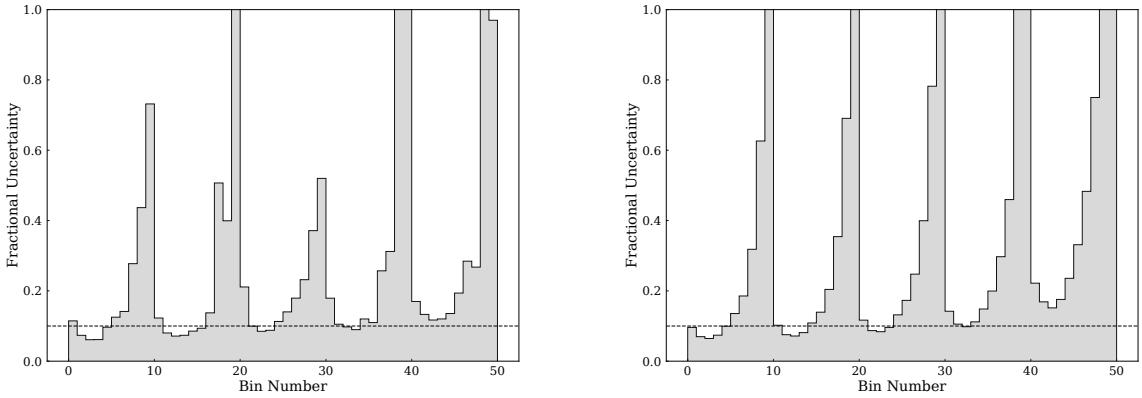


Figure 6.6: Statistical uncertainties per bin of the Phase-I IBD spectrum (left) and the Asimov dataset (right). The errors on the latter are computed using equations 6.6 and 6.7.

background templates, but the effects would be small. An alternative method of estimating the statistical errors is to create a proxy ROn sample by injecting a number of simulated IBD events into a reactor-off data sample and then to perform the subtraction procedure as normal, using a different ROff sample.

6.2.2 Systematic Uncertainties

The main systematic uncertainties relevant for the oscillation analysis have been identified. Note that the percent-level uncertainties on the HM model itself are neglected in this analysis, but since they are at the 2–3% level the effect on the result would be small. Each source of systematic uncertainty is described in turn, in the following order:

1. Absolute light yield (energy scale) of the PVT scintillator.
2. System matrix (relative calibration) elements.
3. Relative neutron reconstruction efficiency between detector modules.
4. Reactor core geometry and distribution of fissile material in relation to the detector acceptance.
5. Data-MC discrepancies in the BiPo energy spectrum.
6. Background subtraction method.

Systematics From Toys

Systematic uncertainties from sources 1–3 are estimated by repeatedly processing a single set of IBD MC events $\mathcal{O}(10^3)$ times, at each iteration varying a specific parameter in the

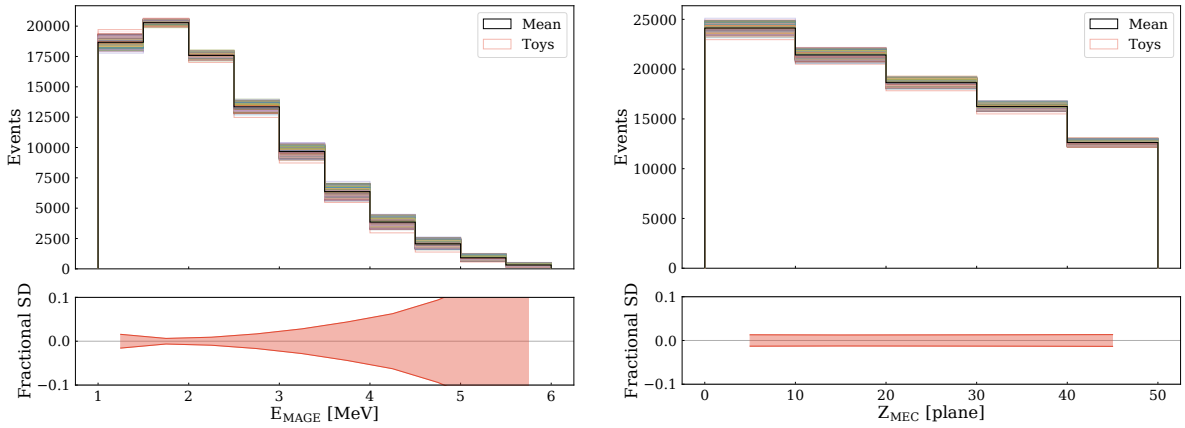


Figure 6.7: 1D projections of the toy MC distributions obtained from varying the number of photons generated per deposited MeV within their uncertainties in the readout simulation. The fractional standard deviations between the toys are depicted as red bands in the ratio plots. The overall uncertainty from the PVT light yield determination is smaller than the statistical error.

readout simulation. The parameter of interest for the PVT light yield, in terms of how the readout simulation is coded, is the number of photons generated per 1 MeV of deposited energy. The best estimate of this value originates from the AmBe pair-production calibration with an uncertainty of approximately 1.5%. Thus the ROSim parameter is sampled from a Gaussian centred on the nominal value with a 1.5% standard deviation. The IBD MC sample processed as such is referred to as a *toy*. The nominal value corresponds to the muon calibration of open-dataset ROff data. The nuBDT analysis cuts are then applied and the events are binned in two-dimensional position-energy space according to the response matrix binning scheme.

The one-dimensional projections of all 10^3 (N) samples are shown in Fig. 6.7. The 1.5% uncertainty grows significantly as it is propagated through the analysis pipeline, especially in bins at high energy where the uncertainty exceeds 10%. The position-bin effect is uniform and far less pronounced. Note that this systematic does not account for the drift in light yield over time, which is assumed to be completely accounted for by the extensive muon calibrations over Phase-I. A 50×50 covariance matrix is formed from the variances and covariances between bins. The fractional (co)variance of matrix bin ij is given by:

$$V_{ij} = \frac{1}{(N-1)} \frac{1}{\bar{x}\bar{y}} \sum_{n=1}^{N} (x - \bar{x})(y - \bar{y}) \,, \tag{6.8}$$

where N is the number of processed samples, x is the number of events in bin i, y is the number of events in bin j, and $\bar{x}$ and $\bar{y}$ are the mean values of their respective bins over all toys.

The same procedure is applied to estimate the uncertainty from the SMat elements, which parameterise other detector effects such as light attenuation, light leakage, and optical

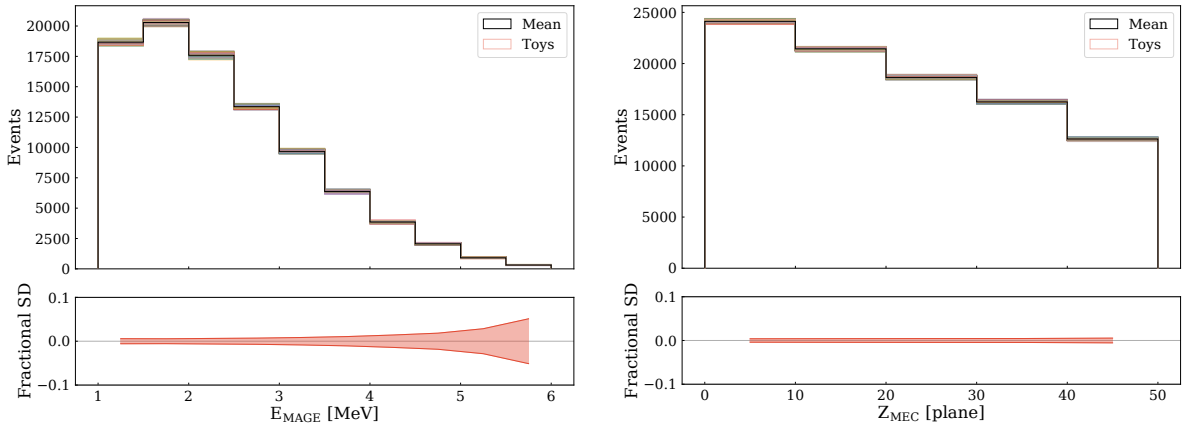


Figure 6.8: 1D projections of the toy MC distributions obtained from varying the system matrix elements within their uncertainties in the readout simulation. The fractional standard deviations between the toys are depicted as blue band in ratio plots. The overall uncertainty from the SMat elements is smaller than the total statistical error.

coupling. Each SM element is varied independently within its own uncertainty, resulting in the distributions of Fig. 6.8. The fractional uncertainties lie within a 2–3% envelope for all but the highest energy bins. The position dependence is minimal. Finally, another 10^3 toys are generated after varying the neutron reconstruction efficiency. It is varied within 1% for the first four modules and within 2% for the final module, as per the September 2018 AmBe neutron calibration results. The associated systematic is at the $\mathcal{O}(0.3)\%$ level with negligible correlation between modules.

Core Geometry & Detector Acceptance

The distribution of fission vertices in the reactor core simulation—which depends on the core geometry and fuel-loading map—affects the spatial distribution of IBD events in the detector. Between four distinct reactor cycles, it is found that the relative difference in the number of IBD vertices as a function of distance is less than 0.3% for all bins. The BR2 fuel-loading maps have also been compared to an unrealistic model in which the fission vertices are distributed uniformly within the inclined channels of the core. The relative difference in this case is found to increase from just over 0.8% at the start of the first module to around 0.7% in the final plane. A conservative fractional uncertainty of 1% is assigned to all position bins.

BiPo Data-MC

It was established in section 4.5 that BiPo events in the $1\gamma(=1c_{\text{low}})$ gamma category[†] exhibit significantly larger data-MC discrepancies than events in other categories. For this

[†] 1γ events in which the energy of the sole gamma cube is less than 148 keV.

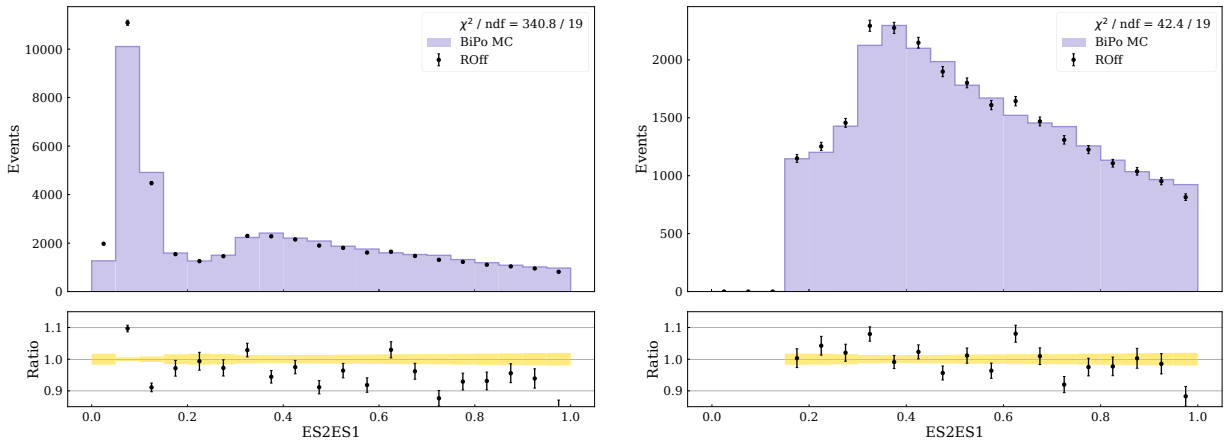


Figure 6.9: Data-MC comparison of BiPo events over the full range of ES2ES1 values (left) and only those above 0.16 (right). The goodness-of-fit is greatly improved by removing events below 0.16.

reason, these events are assigned a dedicated systematic uncertainty. A detailed inspection of the available variables indicate that most of the data-MC discrepancy is confined to the ES2ES1[‡] region below around 0.16, comprising 11% of BiPo MC events. The data and MC distributions of ES2ES1 are compared in Fig. 6.9, before and after applying a cut to remove events below 0.16. The MC distributions are scaled to the total number of data events. The data events are from the high-purity BiPo sideband of ROff data in the first and second Phase-I epochs. The size of the data-MC discrepancy reduces significantly after cutting, bringing the deviation to within 10% for almost all bins and greatly improving the goodness-of-fit. It is conjectured that BiPo events satisfying ES2ES1 < 0.16 are difficult to simulate because they are dominated by radiative BiPo events in which a significant amount of gamma radiation is deposited in the MEC, in analogy with 0γ IBD events.

A systematic uncertainty is derived for the three 500 keV energy bins spanning 1–2.5 MeV. The uncertainty of a given bin is defined as the absolute difference between the data-MC ratio and unity minus one standard deviation, as illustrated in Fig. 6.10. These are taken as fractional errors which are then multiplied by the estimated proportion of $1\gamma(=1c_{\text{low}})$ IBD events of approximately 14%. Overall, this is a relatively conservative systematic uncertainty since the data-MC discrepancy is likely correlated with the error on the absolute light yield.

Background Subtraction

The final systematic is estimated from the ROff-minus-ROff subtraction as a global error on the subtraction procedure. Two 5×10 matrices of variances M_{A} and M_{B} are obtained from the $\text{P}_{\text{A1}}-\text{P}_{\text{A2}}$ and $\text{P}_{\text{B1}}-\text{P}_{\text{B2}}$ subtractions. The variances are taken as the diagonals of two 50×50 matrices, setting the off-diagonal bins to zero. Note that unlike all previous

[‡]The ratio between the second-most and foremost highest energy cubes in the ES cluster.

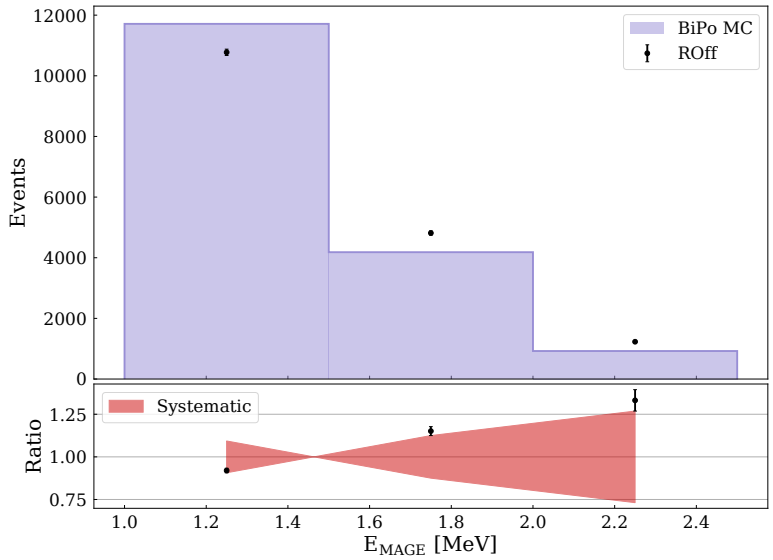


Figure 6.10: Systematic uncertainty derived from the BiPo data-MC discrepancy in the energy spectrum. The red filled region shows the size of the uncertainties on each bin. The systematic uncertainty on the 1.0–1.5 MeV bin is comparable to the statistical error on that bin.

uncertainties, this systematic is parameterised as an absolute covariance matrix. For bins that are statistically consistent with zero the associated systematic uncertainty is overly conservative. The deviations in other bins are at most two standard deviations, meaning the corresponding matrix elements return them to within at least one standard deviation of zero. The two matrices are weighted by their respective fraction of the total exposure time and summed:

$$M = w_1 M_{\mathrm{A}} + w_2 M_{\mathrm{B}} = 2\left(\frac{t_{\mathrm{A1}} + t_{\mathrm{A2}}}{t_{\mathrm{tot}}} M_{\mathrm{A}} + \frac{t_{\mathrm{B1}} + t_{\mathrm{B2}}}{t_{\mathrm{tot}}} M_{\mathrm{B}}\right), \tag{6.9}$$

where w_1 and w_2 are the weighting factors and $t_{\mathrm{tot}} = t_{\mathrm{A1}} + t_{\mathrm{A2}} + t_{\mathrm{B1}} + t_{\mathrm{B2}}$ is equal to the total reactor-off exposure time. The factor of two takes into account the fact that these subtractions are performed with roughly half the total exposure time each, meaning the overall fractional uncertainty is overestimated compared to the ROn-minus-ROff subtractions.

6.2.3 Summary

Fractional (absolute) covariance matrices can be converted to absolute (fractional) matrices by multiplying (dividing) each element ij by the corresponding product of prediction bin entries $P_i \cdot P_j$. The fractional errors of the fifty position-energy bins are plotted in Fig. 6.11 where the background-subtraction absolute variances have been transformed to fractional errors. The largest systematic uncertainty arises from the light yield calibration followed by

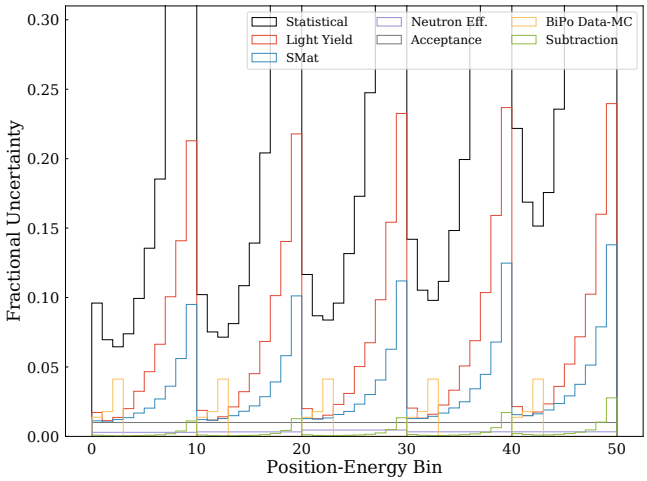


Figure 6.11: Left: Summary of the statistical and systematic uncertainties relevant to the shape-only oscillation analysis. The two-dimensional histogram has been flattened into one dimension such that each consecutive set of ten bins represents one detector module. The spectrum is statistically dominated.

the SMat elements. The BiPo data-MC discrepancy is the source of the largest systematic error on the third energy bin (2–2.5 MeV), despite only affecting 14% of the events. Overall, the total uncertainty is statistically dominated.

6.3 Frequentist Oscillation Fit

Qualitatively, the objective of frequentist hypothesis testing is to determine the likelihood of obtaining a particular measurement, given a physics model. One defines a *test statistic* that follows some distribution under the model, against which the data is evaluated to decide whether or not the model may be accepted or rejected. The null hypothesis (H_0) in this work is the Huber-Mueller reactor flux model in the absence of neutrino oscillations. In the two-neutrino approximation of the 3+1 neutrino model, there are two additional mixing parameters compared to the Standard Model. If one or both of these are non-zero then the model is considered an "alternative hypothesis" (H_1)[§]. For the oscillation analysis, the experimental data is evaluated against a large set of alternative hypotheses in order to construct an *exclusion contour* in the Δm^2_{41}-$\sin^2(2\theta_{14})$ plane.

[§]In cases where only one of the parameters is non-zero, the hypothesis is indistinguishable from the null hypothesis.

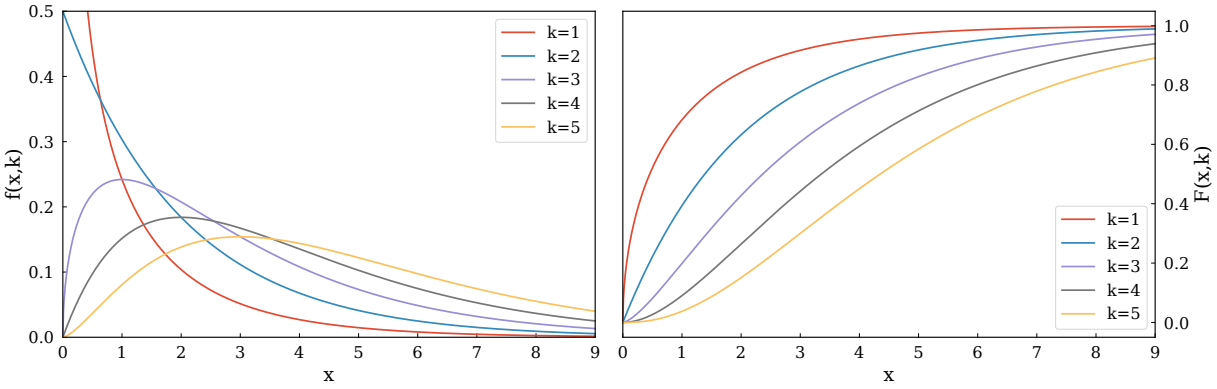


Figure 6.12: Chi-square (χ^2) probability density functions (left) and cumulative distributions (right) for number of degrees of freedom 1–5.

6.3.1 Hypothesis Testing

A chi-square distribution with k degrees of freedom is a continuous probability distribution commonly used in hypothesis testing. It results from the sum of squares of k independent variables that each follow a normal distribution and has been used extensively in this thesis for data-MC comparisons. The probability density function (PDF) of the chi-square distribution may be expressed as:

$$f(x,k) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2} \qquad x > 0\,, \tag{6.10}$$

where $\Gamma(n) = (n-1)!$ is the gamma function (for positive integer n) and x is the dependent variable representing the χ^2. The PDFs and cumulative distributions are plotted in Fig. 6.12 for various numbers of degrees of freedom. A given hypothesis is said to be excluded at some confidence level (CL) if the test statistic of the data exceeds a critical threshold. Two commons confidence levels are 90% and 95%. In a one-tailed test this may be written as the integral of f from zero up to the critical value x_{crit}:

$$\text{CL} \equiv 100 \int_0^{x_{\text{crit}}} \mathrm{d}x\, f(x,k)\,. \tag{6.11}$$

It is clear from the cumulative distributions in Fig. 6.12 that the critical threshold for a given CL increases with the number of degrees of freedom. The region below the critical threshold (or between two thresholds in two-tailed tests) is known as the "confidence interval". The interpretation of a frequentist confidence interval is that as the number of experiments performed tends to infinity, the percentage of the confidence intervals that contain the true value of the parameter being measured tends to the confidence level. The

p-value of a one-tailed test is defined as:

$$\text{p-value} \equiv \int_{x_{\text{data}}}^{\infty} \mathrm{d}x \, f(x, k) \, , \tag{6.12}$$

where x_{data} is the test statistic of the measured data. A small p-value indicates a statistically unlikely measurement if the hypothesis is true.

6.3.2 Definition of the Metric

The formulation of χ^2 is an important consideration because amongst other things it determines how systematic uncertainties are incorporated into the fit. The general form of χ^2 is:

$$\chi^2 = r^T \, V_{\text{tot}}^{-1} \, r \, , \tag{6.13}$$

where r is the residual between the vectorised data and prediction histograms and $V_{\text{tot}} = V_{\text{stat}} + V_{\text{syst}}$ is the matrix of absolute statistical and systematic (co)variances. The V_{tot} converted from the fractional variances using the null hypothesis prediction is visualised in Fig. 6.13. Statistical uncertainties dominate the on-diagonal elements of the matrix but off-diagonal correlations between bins from the systematics will degrade the sensitivity of the experiment, to a small degree.

The residual of the i^{th} position bin and j^{th} energy bin is defined in this analysis as:

$$r_{ij} = D_{ij} - \eta \, P_{ij} \, , \tag{6.14}$$

where D_{ij} and P_{ij} refer to the measured data and prediction contents of bin ij respectively. The free-floating parameter η removes the dependence of the fit on the overall normalisation. In this framework the oscillation fit is dependent on the input reactor model since the data is compared directly to the predicted spectrum. Model-independent oscillation fits typically involve taking ratios of energy spectra at different distances [177, 235, 236] to cancel out biases in the input flux model with respect to the data, at the cost of a reduced statistical precision. Thus neutrino oscillations are only tested for within the spatial range of the detector(s). Model-dependent fits are susceptible to data-MC discrepancies being interpreted either as large statistical fluctuations or evidence of non-zero oscillation parameters. The amplitude of the 5 MeV bump measured by other experiments indicates that this analysis is unlikely to be sensitive to the bump at the current statistical precision.

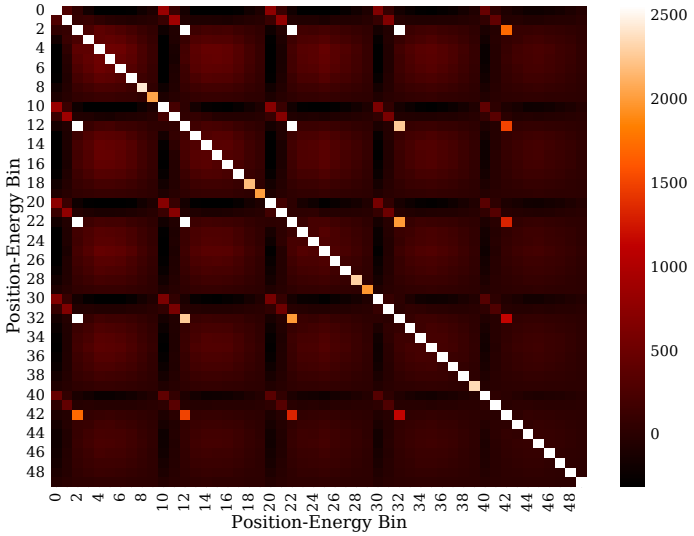


Figure 6.13: Absolute covariance matrix of statistical and systematic uncertainties. Statistical elements are calculated assuming the Asimov dataset prediction without oscillations. Off-diagonal elements (covariances) are generally very small compared to the on-diagonal elements (variances).

6.3.3 Constructing Exclusion Regions

The Feldman-Cousins method [237] is adopted for the construction of exclusion regions in parameter space. The Δm^2_{41}-$\sin^2(2\theta_{14})$ plane is discretised into a grid comprising $\mathcal{O}(10^4)$ points, where each point represents a separate hypothesis $\mathrm{H_x}$. The data is tested against the prediction at each point using the $\Delta\chi^2$ statistic, defined as the difference between $\chi^2_{\mathrm{H_x}}$ and the "best fit" value χ^2_{BF}:

$$\Delta\chi^2 = \chi^2_{\mathrm{H_x}} - \chi^2_{\mathrm{BF}}\,, \tag{6.15}$$

where χ^2_{BF} is found by minimising χ^2 over some physically allowed area A in parameter space, plus the range of the free-floating parameter η. A given point is rejected at some confidence level if the delta-chi-square of the data, $\Delta\chi^2_{\mathrm{data}}(\sin^2(2\theta_{14}), \Delta m^2_{41})$, is greater than the critical threshold $\Delta\chi^2_{\mathrm{crit}}(\sin^2(2\theta_{14}), \Delta m^2_{41})$ for that pair of oscillation parameters.

Minimisation of the test statistic over the entire Δm^2_{41}-$\sin^2(2\theta_{14})$ space is known as a *global scan* (two degrees of freedom), whereas minimisation over $\sin^2(2\theta_{14})$ with a fixed value of Δm^2_{41} is known as a *raster scan* (one degree of freedom). Raster scans identify multiple minima across the parameter-space grid, essentially making independent acceptance/rejection decisions at each Δm^2_{41} value, whilst global scans select a single minimum. Raster scans are less statistically rigorous than global scans in the context of neutrino oscillations [238], but are often preferred because they require significantly less—by a factor equal to the number of Δm^2_{41} values in the grid—computing time to run.

The determination of the critical threshold is the most important aspect of the frequentist

analysis. Wilk's theorem [239] states that under certain conditions the probability density function of the $\Delta\chi^2$ statistic follows a standard χ^2 distribution with k degrees of freedom when the null hypothesis is true. The number of degrees of freedom is the number of free parameters associated with $\mathrm{H_x}$ minus the number of free parameters associated with $\mathrm{H_0}$. As pointed out in Ref. [237], Wilk's theorem does not apply in the case of an antineutrino disappearance measurement for a number of reasons. Principally, $\sin^2(2\theta_{14})$ is bounded at zero and one but the best-fit for a given dataset may sometimes lie in the un-physical regions outside these limits, resulting in a lower $\Delta\chi^2$ which could cause the hypothesis to be falsely accepted. Additionally, in regions of high Δm^2_{41} the rapid oscillations are practically un-resolvable which reduces the number of degrees of freedom by one, making a hypothesis in that region easier to reject. A similar effect is found in low Δm^2_{41} regions. Thirdly, the χ^2 assumption of a Gaussian probability density is invalidated by the sine function in the antineutrino survival probability.

To overcome the limitations of Wilk's theorem, empirical distributions of the $\Delta\chi^2$ metric are generated via toy MC experiments (also known as pseudo-experiments) from which the critical thresholds can be derived. The method for doing so is described in the next section. For now, it is assumed that $\Delta\chi^2_{\mathrm{crit}}(\sin^2(2\theta_{14}), \Delta m^2_{41})$ is known for all hypotheses. The analysis proceeds by iterating over the Δm^2_{41}-$\sin^2(2\theta_{14})$ grid and performing the following procedure at each $\mathrm{H_x}$ point:

1. Create the true-space prediction histogram for the given oscillation parameters by calculating the survival probability of each event (equation 6.1) and weighting the events accordingly in the position-energy histogram.

2. Forward-fold the true-space prediction to reco-space by application of the response matrix.

3. Re-compute the absolute covariance matrix by combining the new reco-space signal prediction from the previous step and the total fractional covariance matrix of statistical and systematic uncertainties.

4. Calculate the χ^2 statistic between the experimental data and the resulting reco-space prediction histogram with equation 6.13

5. Find the best-fit chi-square χ^2_{BF} by scanning over the parameter-space area A and repeating steps 1–3 at each grid point therein.

6. Compute the $\Delta\chi^2$ statistic and reject the $\mathrm{H_x}$ point if it exceeds the critical threshold of that point at a given confidence level.

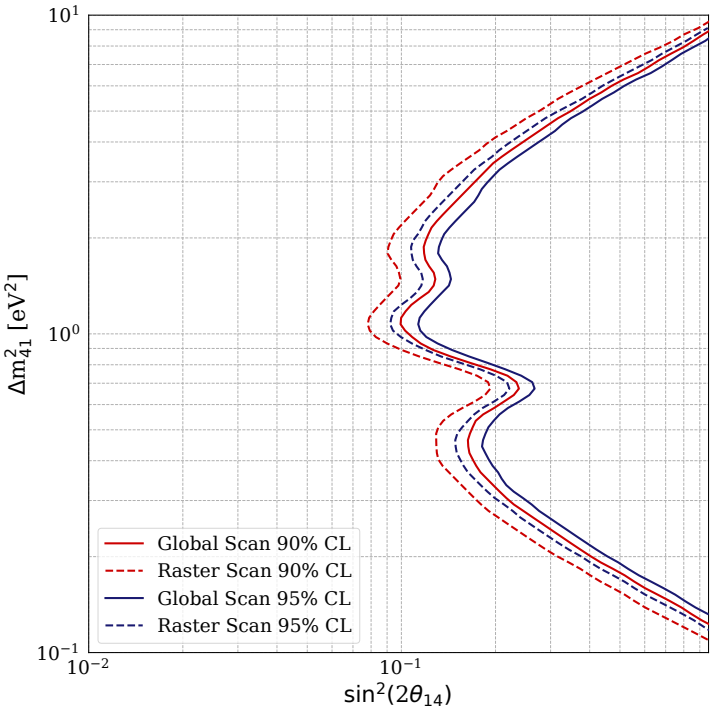


Figure 6.14: Example sensitivity contours for an Asimov dataset of arbitrary statistical precision. The regions to the right(left) of each contour represent rejected(accepted) hypotheses. The solid lines and dashed lines correspond to global scans and raster scans respectively. The latter results in greater parameter-space coverage, as does a lower confidence level.

The output of the fit is a set of accepted points defined by $\Delta\chi^2_{\text{data}}(\sin^2(2\theta_{14}), \Delta m^2_{41}) < \Delta\chi^2_{\text{crit}}(\sin^2(2\theta_{14}), \Delta m^2_{41})$, separated from the rejected points by a curve known as the exclusion contour. The sensitivity of the experiment to active-sterile neutrino mixing is estimated by running the oscillation fit with the Asimov dataset—as presented in section 6.2.1—instead of experimental data. Sensitivity contours are essentially the average exclusion contour one would obtain from running the experiment an infinite number of times. Four sensitivity contours built using Wilk's theorem are drawn in Fig. 6.14 for an Asimov dataset of arbitrary statistical precision. The contours marks the boundary between rejected hypotheses to the right and non-rejected hypotheses to the left. The area of the excluded parameter space is referred to as the *coverage*. In general, a global-scan contour has less coverage than a raster-scan contour for the same data because it is able to find a smaller minimum by scanning over the entire parameter domain, leading to smaller $\Delta\chi^2_{\text{data}}(\sin^2(2\theta_{14}), \Delta m^2_{41})$ values which are more difficult to reject. Similarly, the higher the confidence level the fewer parameter-space points are rejected, resulting in a smaller exclusion region.

6.3.4 Generating Toy Experiments

Starting from the reco-space prediction histogram, toy experiments can be generated in a computationally efficient manner using Cholesky decomposition [103]. The covariance matrix of statistical and systematic (co)variances is decomposed into the product of a lower triangular matrix ($\boldsymbol{L}$) and its conjugate transpose. Only symmetric positive-definite matrices can be decomposed in this manner. Next, a vector $\boldsymbol{u}$ of length equal to the dimensions of $\boldsymbol{L}$ is constructed by sampling random numbers from a normal distribution. Multiplying $\boldsymbol{L}$ by $\boldsymbol{u}$ yields a one-dimensional vector that is added to the flattened prediction histogram, having scaled the latter to the total number of data events. In this way, the "data" of the toy experiment, given by:

$$\boldsymbol{T} = \boldsymbol{P} + \boldsymbol{L}\boldsymbol{u}\,, \tag{6.16}$$

models the true (co)variances of the data, assuming reasonable accuracy of the original covariance matrix. This statement is validated by checking that the input covariance matrix can be reproduced by the generated toys. From the generation of 10^4 null-hypothesis toys, a new covariance matrix is extracted and compared to the original in Fig. 6.15. The on-diagonal elements are in agreement to within 2%. The off-diagonal elements vary far more depending on the bin but the general trend of the correlations appears to be modelled reasonably well. The input and output matrices have summed (covariances)variances of (434.5)629.9 and (435.4)628.7 respectively. It is assumed that the difference between all matrix elements tends to zero as the number of toys increases but this is unproven.

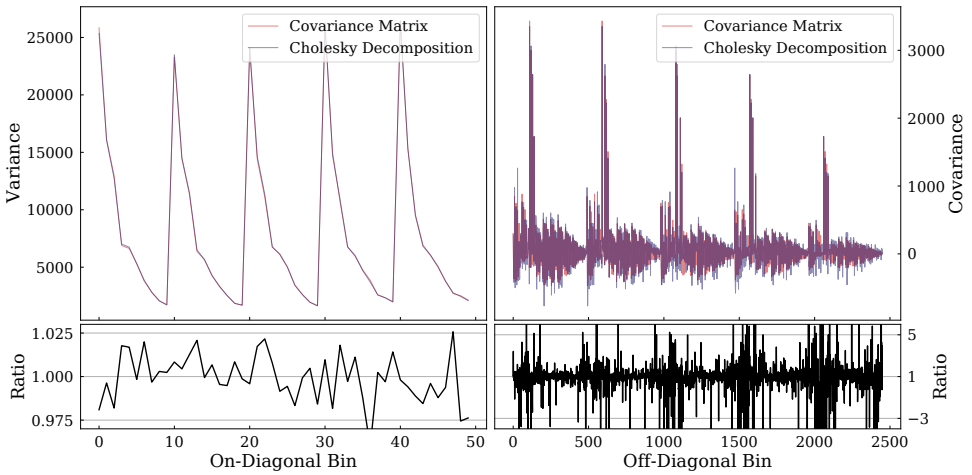


Figure 6.15: Validation of the Cholesky decomposition of the covariance matrix for on-diagonal elements (left) and off-diagonal elements (right). The red lines represent the (co)variances of the original covariance matrix and the blue lines represent the same quantities for the covariance matrix extracted from 10^4 toys. The lower pads show the ratio between the two. The variances are reproduced correctly to within 2.5%. The covariances show relatively large deviations but follow the overall trend of the original matrix.

6.3.5 Building the Test Statistic

The procedure for building the $\Delta\chi^2(\sin^2(2\theta_{14}), \Delta m^2_{41})$ distributions is almost identical to the steps outlined above for the construction of exclusion regions. Instead of using data, the process is repeated for $\mathcal{O}(10^3)$ pseudo-experiments, yielding as many $\Delta\chi^2$ values which are then binned. The number of events in the toys experiments follow a Gaussian distribution centred on the normalisation of the Asimov dataset, but these variations have no effect on the determination of the $\Delta\chi^2$ distributions because of the free-floating parameter in the residual.

The $\Delta\chi^2$ distribution of 1000 toys under an alternative hypothesis ($\Delta m^2_{41} = 2.36\,\text{eV}^2$, $\sin^2(2\theta_{14}) = 0.102$) is displayed on the left-hand side of Fig. 6.16 next to the corresponding $\Delta\chi^2_{\text{crit}}$ values as a function of confidence level. The best fit of each toy was identified via global scan of the Δm^2_{41}-$\sin^2(2\theta_{14})$ plane. Compared to a true χ^2 distribution with two degrees of freedom, the critical threshold derived from the toy-generation method is much higher, making the conditions for rejecting this hypothesis far more stringent. At 90% CL, the toy-generation critical threshold is almost double the Wilk's theorem value.

The Δm^2_{41}-$\sin^2(2\theta_{14})$ plane is discretised into a grid of 100×100 points ranging from 0.1–10 eV^2 in Δm^2_{41} and 0.01–1 in $\sin^2(2\theta_{14})$. One thousand toys are generated at each grid point and minimised via global scans to create the $\Delta\chi^2(\sin^2(2\theta_{14}), \Delta m^2_{41})$ distributions. The η parameter is initialised to the value that exactly scales the number of MC events to the number of data events (η_0). It is then allowed to vary over forty uniformly spaced

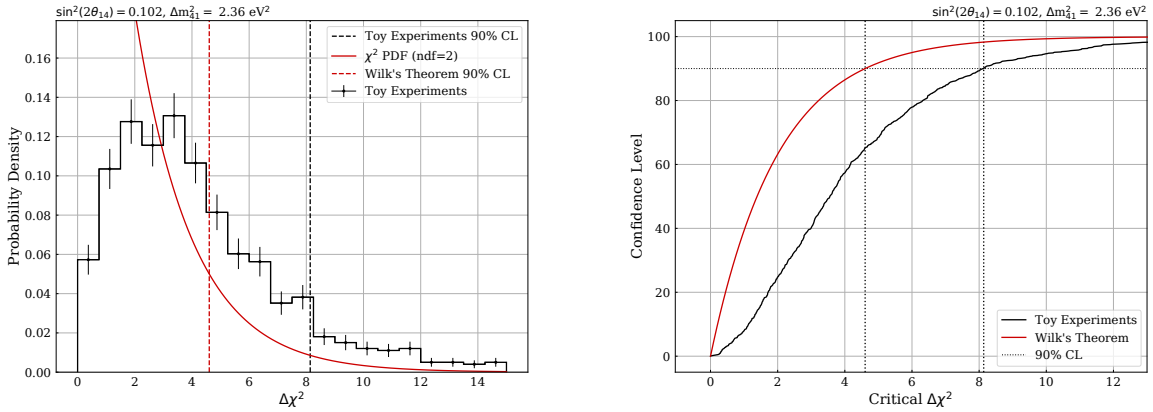


Figure 6.16: Left: $\Delta\chi^2$ distribution generated from 10^3 toys under an alternative 3+1 hypothesis with parameters $\Delta m^2_{41} = 2.36\,\mathrm{eV}^2$ and $\sin^2(2\theta_{14}) = 0.102$. The best fit of each toy was obtained via a global scan of parameter space. Right: $\Delta\chi^2_{\mathrm{crit}}$ as a function of confidence level derived from the left-hand side histogram. The dotted lines mark the 90% confidence level. A large difference between the 90% CL thresholds under the Wilk's theorem assumption and the toy-generation method is evident.

values between $0.8\eta_0$ and $1.2\eta_0$ inclusive. If the minimum chi-square is identified at either boundary—which is unlikely to be the true minimum—the limits are expanded another 20% in the direction of the boundary and the process is repeated. For example, if the chi-square is minimal at exactly $0.8\eta_0$, the limits are shifted to $0.6\eta_0$–$0.8\eta_0$ inclusive. If the previous minimum was indeed the true minimum, it will be correctly identified in this second iteration. Otherwise, the process continues for a maximum of five iterations. At least 4×10^5 separate chi-squares are calculated for each toy.

A heat map of the critical thresholds at 90% CL is provided in Fig. 6.17. Missing or invalid values caused by computing limitations are replaced by the value of their nearest-neighbour; less than 3% of the Δm^2_{41} values are affected. The green squares represent the single grid points per slice that are closest to (and within 0.5 of) the critical threshold of a true χ^2 distribution with two degrees of freedom (4.605). Points to the left of the curve traced out by the green squares are more difficult to exclude than the Wilk's theorem value would suggest[¶]. The critical value over the majority of the Δm^2_{41}-$\sin^2(2\theta_{14})$ plane is 8–9. It is evident that an exclusion region built with Wilk's theorem would overestimate the coverage, as demonstrated in appendix G.

The choice of spacing in the three-dimensional Δm^2_{41}-$\sin^2(2\theta_{14})$-η volume is driven by the available computing resources and time constraints. A finer spacing would allow the algorithm to more precisely locate the best fit of each toy, thereby increasing the exclusion threshold of all the points. Additionally, as evident in Fig. 6.16 the Poisson uncertainties

[¶]Similarly, points to the right of the green squares are easier to exclude.

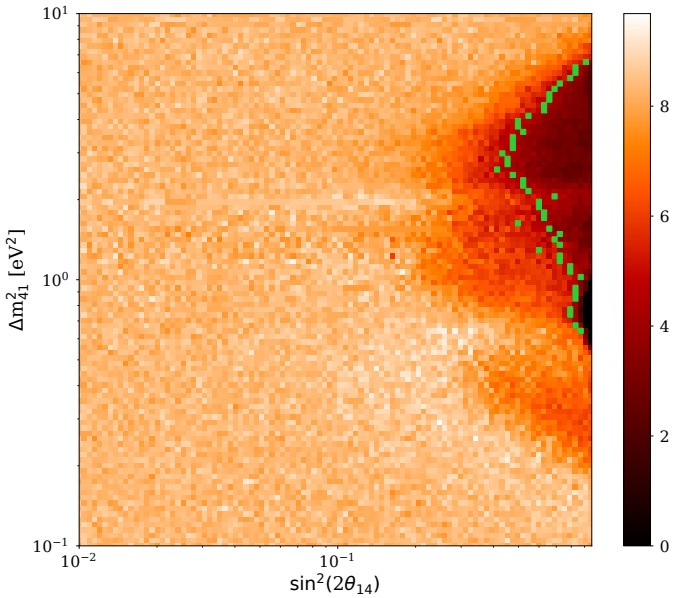


Figure 6.17: Heat map of $\Delta\chi^2_{\text{crit}}$ values at 90% CL in the Δm^2_{41}-$\sin^2(2\theta_{14})$ plane, obtained from generating $\mathcal{O}(10^3)$ toys at each grid point. The green squares mark the single grid points that are closest to (and within 0.5 of) 4.605 (the 90% CL critical threshold for two degrees freedom under Wilk's theorem) in each Δm^2_{41} slice. The critical thresholds of the points to the left of the green squares are larger than 4.605 so using Wilk's theorem in this region overestimates the coverage.

on the $\Delta\chi^2$ histogram bins are not negligible, though the critical values are not expected to change significantly with more pseudo-experiments.

The distribution of best-fit points of 7000 null-hypothesis toys is shown in Fig. 6.18. The points populate a wide band that resembles the general shape of a sensitivity contour, preferentially at high values of Δm^2_{41}, proving that even in the absence of a positive signal the data will be best-described by some alternative hypothesis. Extremely large statistical fluctuations in the toys are not probable, by definition, so the chi-square is rarely minimised by strong spectral distortions induced by pairs of parameters to the right of the best-fit band. Toys that have extremely small fluctuations, well-described by small oscillation effects on the left-hand side of the band, are similarly unlikely. From this behaviour it is deduced that the sensitivity contour should lie in very close proximity to the best-fit band.

6.3.6 Null Hypothesis

Prior to the construction of the final exclusion contour, the data is first tested against the null hypothesis. It is simultaneously a test of the IBD spatial distribution prediction (related to the simulation of the core geometry, fission vertices, etc.) and the Huber-Mueller reactor flux model. To perform the H_0 test we use the test statistic defined in equation 6.13 with the equation 6.14 residual. The Phase-I IBD spectrum yields a χ^2 of 64.9 with respect to the H_0 prediction.

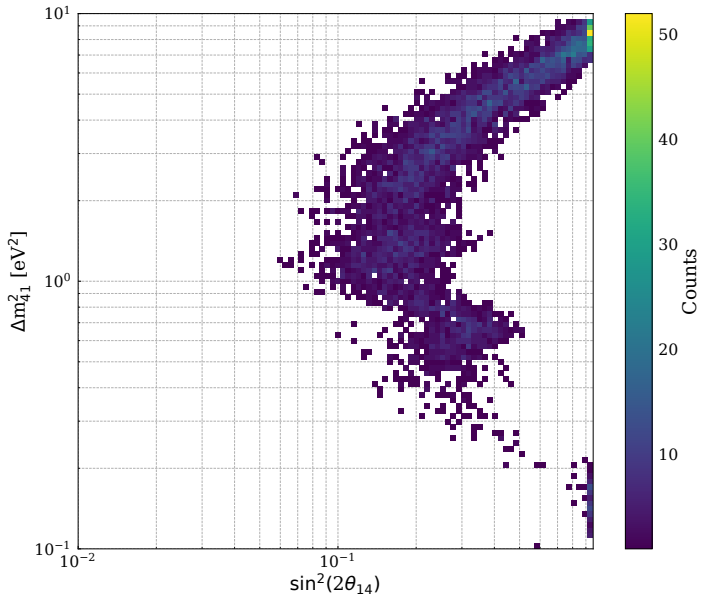


Figure 6.18: Distribution of best-fit points of 7000 pseudo-experiments generated under the null hypothesis using Cholesky decomposition. The best-fit points were located via a global scan and have a preference for high values of Δm^2_{41}. The band traces out the general shape of a sensitivity contour.

In order to avoid assuming the shape of the probability density function with which to interpret this result, the distribution can be built computationally using pseudo-experiments, as shown in Fig. 6.19. The distribution is obtained by following steps 1–4 in the exclusion contour instructions for 10^5 H_0 toys. The histogram is overlaid with a true χ^2 PDF with fifty degrees of freedom; the two PDFs agree very well. The shaded region signifies the area under the curve that is greater than the test statistic of the data. The 90% and 95% critical thresholds according to the distribution from toys are 63.6 and 68.2 respectively, depicted as vertical dashed and dotted lines in the figure, which are very similar to the theoretical values. The p-value of the data is 0.084 so it can be stated that the null hypothesis is rejected at 90% CL by this goodness-of-fit test.

6.3.7 Phase-I Sensitivity & Exclusion Contours

Exclusion contours of null-hypothesis pseudo-experiments typically extend further in parameter space than the sensitivity contour in some regions but less far in others. This is caused by statistical fluctuations that favour the rejection of certain mixing parameters and disfavour others, which occurs regardless of the presence of a positive signal. The 90% and 95% CL global-scan sensitivity contours plotted as filled regions in Fig. 6.20 indicate that the current analysis is just out of reach of the RAA best-fit point. The corresponding exclusion contours are drawn as solid lines. Mixing parameters in the $\sin^2(2\theta_{14})$ range 0.1–0.3 are excluded for some Δm^2_{41} values, including a small portion of the RAA 95% CL allowed

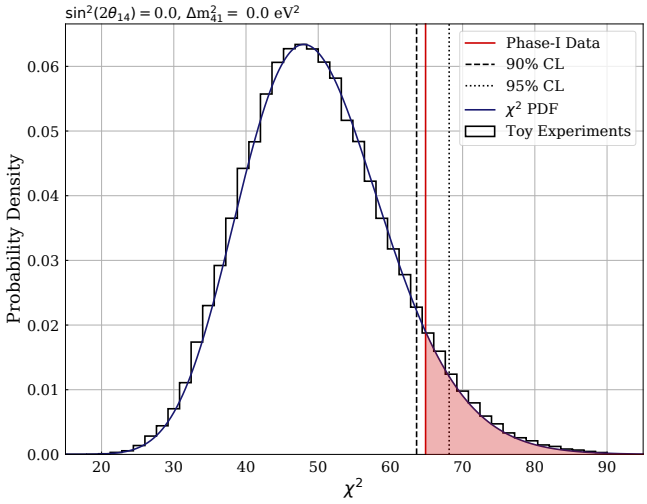


Figure 6.19: The χ^2 distribution of 10^5 null-hypothesis toy experiments generated via Cholesky decomposition. The red vertical line denotes the value for the experimental data which may be compared to the 90% and 95% confidence level lines, taken as the respective quantiles of the toy histogram. Overlaid in blue is the theoretical χ^2 distribution with fifty degrees of freedom that has been scaled to the height of the underlying histogram; it agrees very well with the toy-generated histogram. This goodness-of-fit test rejects the null hypothesis at 90% CL.

region. The fluctuations reduce the coverage of the contours in the vicinity of the RAA best-fit point which is not excluded at the given confidence levels.

We find no evidence of active-sterile neutrino oscillations according to the 3+1 model in the Phase-I IBD spectrum, but there is still plenty of room left for the existence of sterile neutrinos in the unexplored areas of the RAA 95% CL region. The data is best-described by mixing parameters $\Delta m^2_{41} = 2.154\,\text{eV}^2$ and $\sin^2(2\theta_{14}) = 0.118$ with a χ^2 of 62.6. Thus the null-hypothesis $\Delta\chi^2$ is equal to 2.3 and it cannot be rejected at 90% or 95% CL in this hypothesis test. The preference for active-sterile neutrino oscillations over the HM model is very mild. The best-fit point lies within the toy band of in Fig. 6.18. It is recommended that the result be complemented with a model-independent fit, given the current knowledge regarding the reactor flux models.

The SoLid exclusion contours are compared to the first exclusion contours published by PROSPECT [236] and STEREO [235] in Fig. 6.21. Note that their latest results exhibit much larger rejected regions. The coverage of the SoLid contour is receded at high Δm^2_{41} compared to the liquid-scintillator experiments because of their superior energy resolution. On the other hand, SoLid is able to exclude some areas of parameter space that are not reached by PROSPECT or STEREO (at least in these results). Upgrades to the analysis or inclusion of more data are required to improve the coverage of our exclusion contour. The prospects for this are discussed in the next chapter.

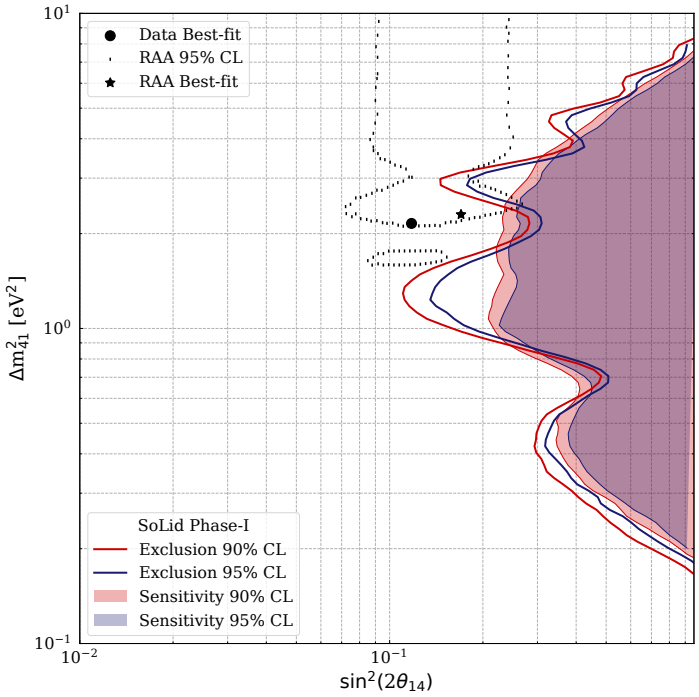


Figure 6.20: SoLid Phase-I sensitivity and exclusion contours at 90% and 95% CL. The dashed black line outlines the RAA 95% CL region, a small portion of which lies within the exclusion contours. The RAA best-fit point (black star) is not rejected at 90% CL. The Phase-I spectrum is best described by mixing parameters $\Delta m^2_{41} = 2.154\,\text{eV}^2$ and $\sin^2(2\theta_{14}) = 0.118$ (black dot).

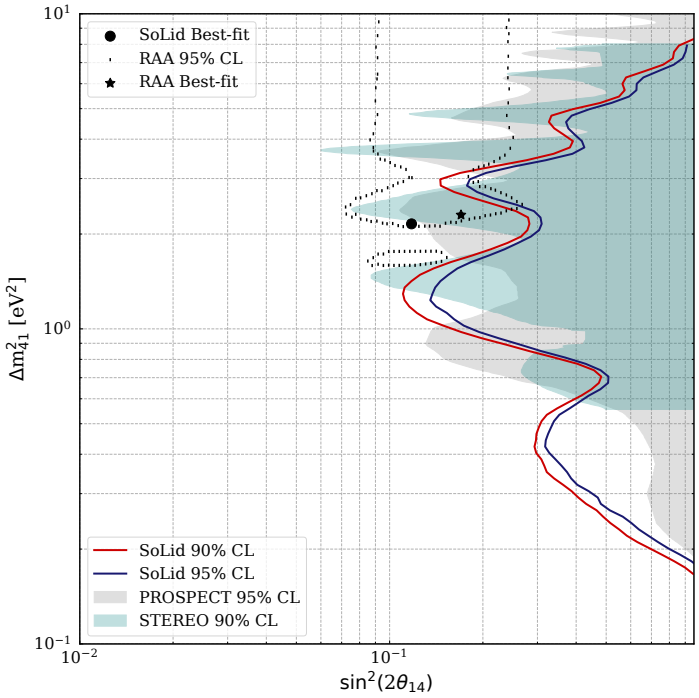


Figure 6.21: Phase-I exclusion contours compared to the first results published by PROSPECT [236] and STEREO [235] (contours courtesy of V. Pestel [206]). The SoLid exclusion regions cover some complementary regions of parameter space but does not exclude the RAA best-fit point. It also has worse coverage in the high Δm^2_{41} region because of the inferior energy resolution of SoLid.

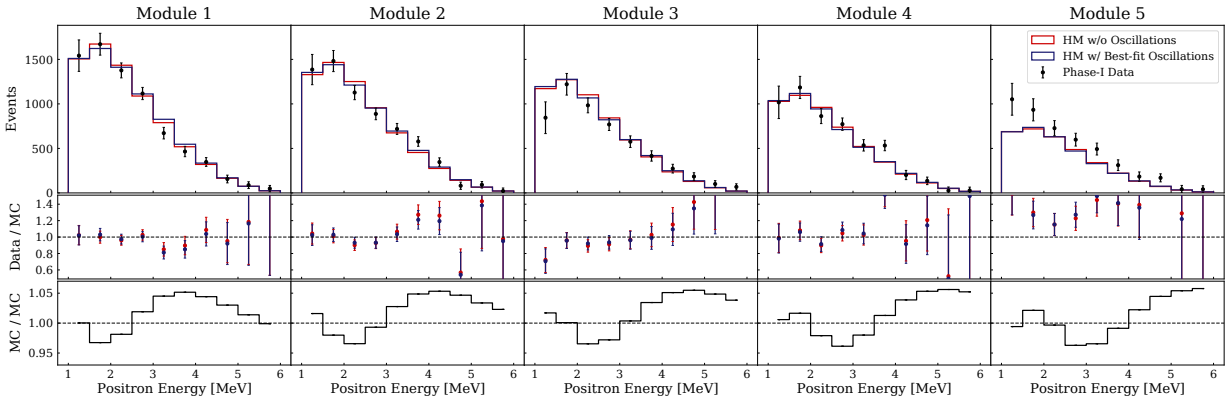


Figure 6.22: SoLid Phase-I IBD positron energy spectrum per module. "Positron Energy" refers to the MAGE energy estimator. The data is compared to the HM model in the absence of oscillations and the best-fit model with mixing parameters $\Delta m^2_{41} = 2.154\,\text{eV}^2$ and $\sin^2(2\theta_{14}) = 0.118$. The predictions are separately scaled by the factor that minimises the χ^2 with the data. Error bars represent the statistical uncertainty on the data points. The oscillation pattern induced by the best-fit mixing is displayed in the bottom row. Adding oscillations only improves the data-MC agreement to small extent.

Finally, Fig. 6.22 depicts the experimental spectrum as a function of distance against the HM model (red) and the HM model plus best-fit oscillations (blue). The two predictions are separately scaled by the η factors that yield the best agreement with the data. The lack of systematic uncertainties in the displayed error bars is the cause of the apparent excess of events in the final module. The expected normalisation-independent oscillation pattern is drawn in the lowest canvases. The best-fit mixing parameter induce a change in the spectrum shapes at the $\pm 5\%$ level, which is very small compared to the statistical fluctuations of most bins, hence the only small improvement in data-MC agreement.

Chapter 7

Prospects & Conclusion

The final chapter of this thesis is dedicated to aspects of the SoLid experiment that either have not been fully realised at the time of writing, or that are unrelated to the main oscillation analysis. These include an analysis of the pure ^{235}U energy spectrum and antineutrino directionality. The SoLid detector technology may also have potential applications in reactor monitoring and nuclear safeguards, given its suitability for stable data-taking over long periods of time. The chapter is concluded with a broad overview of the work presented in this thesis.

7.1 Analysis Improvements

The IBD yield and signal-to-background ratio granted by the nuBDT-based analysis and background subtraction (105 events per day and 0.26 respectively) is inadequate for achieving the main oscillation physics goal of the experiment, namely the testing of the RAA best-fit point. As such, some improvements to the analysis are required. One undesirable facet of the nuBDT model is that although the reactor-off data comprises fast-neutron and BiPo events—albeit a much smaller proportion of the latter—both are labelled as "background" without differentiation. An extremely pure BiPo sample is easily attainable with cuts on BiPonator score and Δt, but no equivalent cuts exist that isolate fast-neutron events with the same purity. A large sample of simulated atmospheric neutrons and cosmic muons would alleviate this issue. Then, the nuBDT could be extended from binary to multi-class classification and trained on a combination of IBD MC, BiPo, and the fast-neutron MC sample. Alternatively, separate binary classifiers could be trained for each background type (separately against IBD MC events) and later combined. Whether the supposed improvements from these options would be enough to reach the RAA best-fit point and become more competitive with the first results of PROSPECT and STEREO is an open question.

The SoLid group at LPC Clermont-Ferrand have commissioned a complementary analysis that focuses on the cleanest 2γ topologies, which is yet to be applied on the full Phase-I

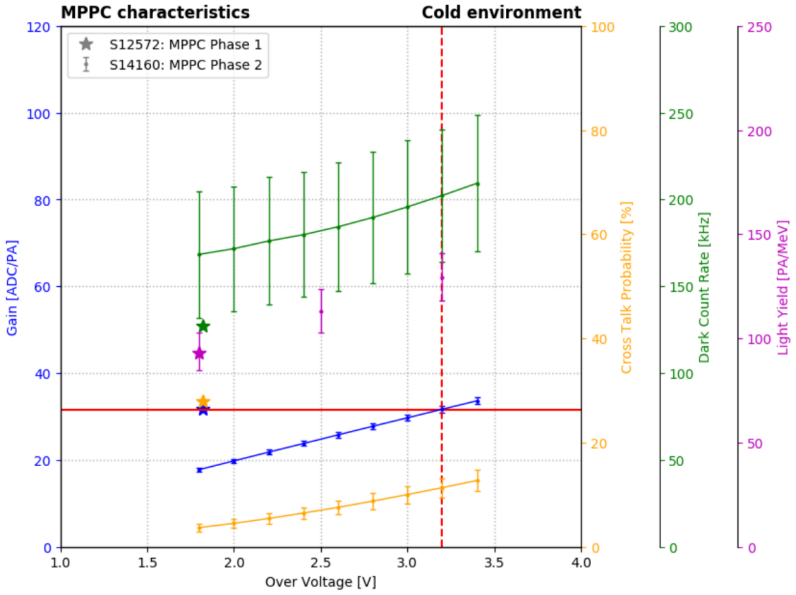


Figure 7.1: Characteristics of the S14 series MPPCs installed in the SoLid detector at BR2 [205], at nominal container temperature of 11 °C. Red lines denote the chosen working point. Error bars represent the standard deviation of the parameters across all 3200 MPPCs of the Phase-II detector. The dark-count rate, light yield, and cross-talk probability are all improved compared to Phase-I.

dataset. Rather than a background subtraction, a simultaneous ΔR-Δt fit is used to assess the rates of each background and the signal in the data. The analysis is able to achieve a signal-to-background ratio greater than unity, but this comes with the cost of an extremely low event rate of $\mathcal{O}(20)$ IBDs per day.

7.2 SoLid Phase-II

Once the Phase-I data-taking campaign ended in July 2020, work immediately began on the SoLid Phase-II upgrade. The detector upgrade and subsequent commissioning is described exhaustively in Ref. [205]. The Phase-II detector was created by replacing the existing Hamamatsu S12572-050P SiPMs with the better-performing S14160-3050HS model, which has very different properties to the S12 series. The photon detection efficiency of the S14 model peaks at roughly the same wavelength but is 10–15% higher than the Phase-I sensors. Also, the cross-talk is lower by a factor of three at the same overvoltage. The mean breakdown voltage of the S14 SiPM is 38 V compared to 65 V for the S12. On the other hand, the S14 model has slightly higher dark-count rate and gain. The higher dark-count rate can be adequately dealt with using the zero-suppression threshold of the triggers. The higher gain is not desirable because it causes the ADCs to be saturated by fewer pixel avalanches, but it can be countered by reducing the amplification factor of the analogue boards.

The dependence of numerous MPPC parameters on overvoltage—measured once the

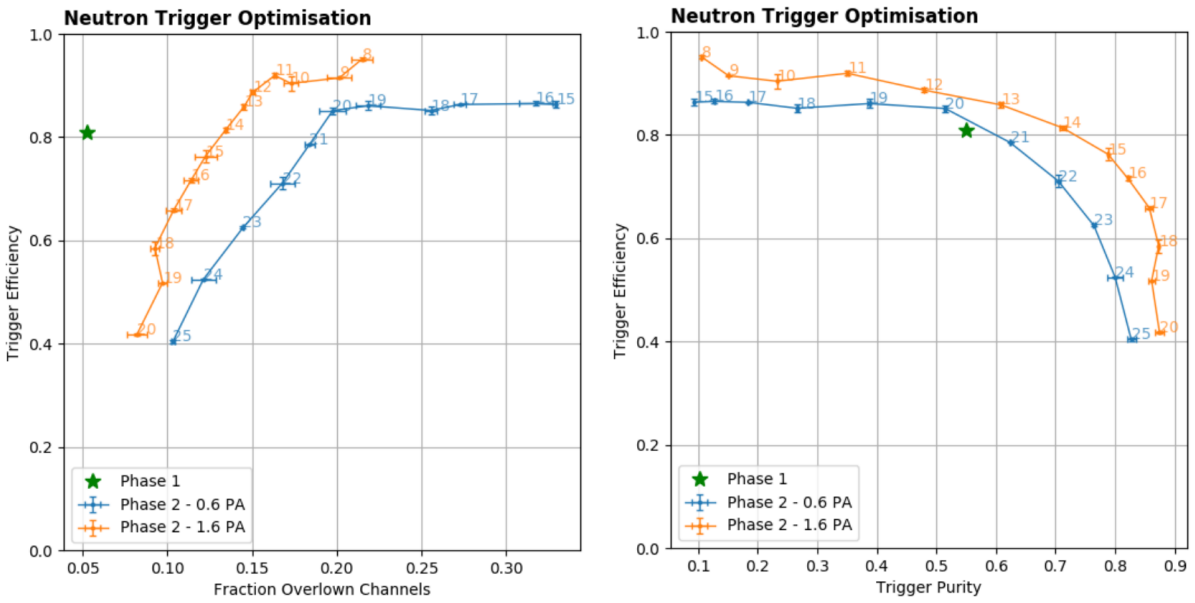


Figure 7.2: Neutron trigger optimisation of Phase-II [205]. Neutron trigger efficiency as a function of overflown channel fraction (left) and trigger purity (right). The rate of channel overflow is significantly larger in Phase-II compared to Phase-I because of the increased light yield.

sensors had been installed in the full-scale detector at BR2—is plotted in Fig. 7.1. The intersection of the two red lines signifies the chosen working point of 3.2 V overvoltage and 31.5 ADC/PA gain. The latter is made equal to the Phase-I value by lowering the analogue board amplification. At this working point the dark-count rate is slightly higher than for Phase-I, but it is more than compensated for by the strong reduction in cross-talk probability. The overall effect is an increase in light yield by around 40%, as measured by the ^{22}Na calibration procedure. A higher light yield should improve both the energy resolution and the overall IBD efficiency. The reasons for the latter are two-fold. Firstly, a higher number of detectable peaks over threshold will directly raise the neutron trigger efficiency. Secondly, the higher light yield would allow the ES cluster energy threshold to be reduced and the detector should have a greater efficiency for the low-energy deposits that are crucial for signal-background discrimination in offline analysis.

The neutron trigger settings were re-optimised for Phase-II using AmBe neutron calibration data. It was found that the extra data per trigger from the increased light yield put a significant strain on the FPGA buffers, resulting in a very large fraction of DAQ dead-time from high channel overflow rate. This was partially compensated by reducing the run duration from 480 s to 200 s, but the overall dead-time fraction is still worse than for Phase-I, as demonstrated in Fig. 7.2. In terms of the neutron trigger efficiency (number of reconstructed NS waveforms divided by the expected number from simulations of the AmBe source activity) and trigger purity (number of reconstructed NS waveforms over the number of fired triggers) the situation is slightly better than for Phase-I. The firing conditions of

the Phase-II trigger were finalised as 15 peaks over threshold (N_{peaks}), counting any local waveform with amplitude greater than 1.6 PA (T_{thr}) as a peak. Note that for Phase-I, T_{thr} was set to 0.5 PA. The number of planes read out either side of the triggered plane was also increased from three to four to capture more data per event. The ZS threshold was unchanged.

Full-time data-taking with the Phase-II trigger settings started in November 2020 and lasted until September 2022. During the first year of data-taking, an attempt was made to increase the capacity of the FPGA data buffers by re-designing the firmware. This in turn required a near-total overhaul of the DAQ software to retain compatibility with the upgraded firmware. The software update eventually had to be abandoned because of a lack of manpower, so the new firmware was never implemented and data-taking suffered from relatively high dead-time for the duration of Phase-II. Nonetheless, the Phase-II dataset could significantly bolster the statistical power of the oscillation analysis. It would also give SoLid a better precision on the ^{235}U spectrum measurement.

7.3 Uranium-235 Spectrum Measurement

The analysis of the ^{235}U spectrum is a far simpler endeavour than the oscillation analysis, at least in a shape-only form. A rate-plus-shape analysis requires extremely precise knowledge of the overall IBD efficiency, taking into account contributions from all sources. The systematic uncertainty on the spectrum shape measurement is dominated by the light yield uncertainty. All other systematics that were relevant for the oscillation analysis are still pertinent here, with the exception of the relative neutron efficiency.

Fig. 7.3 shows the statistics-only comparison between the predicted HM spectrum and the Phase-I IBD spectrum. The latter has been scaled to the number of events of the former. In the absence of systematic uncertainties, the distributions are incompatible at 95.6% CL, assuming true chi-square behaviour. Inclusion of the systematic uncertainties would reduce this significance, as would use of the free-floating parameter η in equation 6.14. Such an analysis is planned for the near future. If the slight data-MC discrepancy in the measured spectrum persists in the rigorous analysis, it is perhaps indicative of the presence of the 5 MeV spectral distortion that was observed by other reactor neutrino experiments. However, the local p-values of the spectrum should be evaluated to confirm the specific location of the disagreement, if any.

Other collaborations have provided unfolded antineutrino reference spectra to the community [184, 240]. Unfolding is a numerical technique to remove detector effects and extract the true antineutrino spectrum from the reconstructed spectrum. A common and effective unfolding method is called Wiener-SVD [241]. Unlike forward-folding which is mathematically well-defined, unfolding is difficult, in part because of the large effect that even small

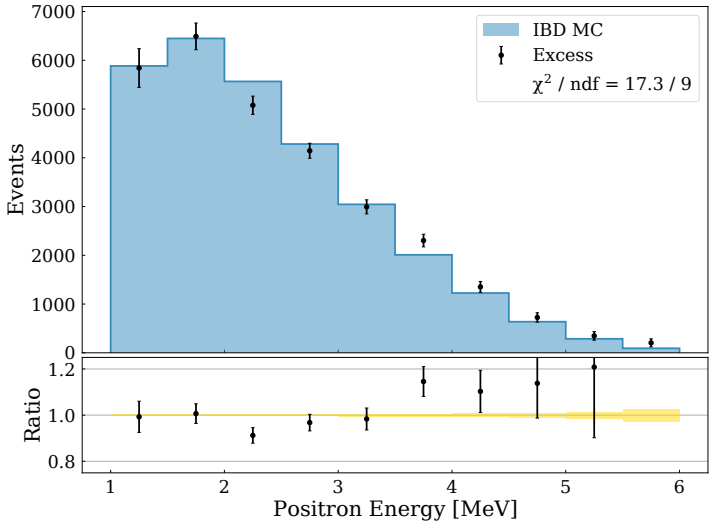


Figure 7.3: SoLid Phase-I positron energy spectrum against the HM model prediction, integrated over all modules. "Positron Energy" refers to the MAGE energy estimator. Error bars represent the statistical uncertainty on the points. The yellow band in each ratio plot indicates the uncertainty on the MC prediction. This histogram can be unfolded with a response matrix to yield the pure ^{235}U antineutrino spectrum.

statistical fluctuations can have on the unfolded spectrum, requiring some sort of regularisation to suppress. If there were no statistical fluctuations, it would be a simple inversion problem:

$$H_{\text{true}} = M^{-1} \times H_{\text{reco}} \,, \tag{7.1}$$

where H_{reco} and H_{true} are the reco-space and true-space distributions respectively and M^{-1} is an inverted response matrix. An unfolded SoLid spectrum, though lacking in statistics compared to those of other experiments, could still be valuable to the community given the alternative detector technology used for the measurement.

7.4 Bayesian Oscillation Fit

The SoLid group at LPC Caen is currently developing a parallel oscillation fit based on a Markov Chain Monte Carlo method—the Metropolis-Hastings algorithm—to construct credible regions in parameter space. A major benefit of this method is that it is far less time-consuming to run than the frequentist fit. Bayesian statistics asks a different question of the data, the probability that a particular model is true given the measured data. A Bayesian credible region from Phase-I data would be a stand-alone result that could also act as a cross-check on the frequentist confidence regions.

7.5 Reactor Monitoring Applications

The technology pioneered by SoLid has potential applications in reactor monitoring and nuclear safeguards. The successful Phase-I and Phase-II data-taking campaigns demonstrate its operational stability over long periods of time and capacity for remote control with minimal physical intervention. Note that the cause of the SiPM pedestal instabilities in the discarded data of July–September 2019 was identified (during similar issues in Phase-II) as faulty connections between the electronic boards and the power supplies. The pedestal issues were resolved after cleaning the electric plugs and did not occur again. Furthermore, the development of the ES calibration with horizontal muons reduces the required number of data-taking interruptions for in-situ source calibration. Finally, unlike liquid scintillator detectors the solid plastic construction of SoLid has no risk of leaking, which is an advantage in a high-security environment. Overall, the detector and related software could be effectively deployed at a nuclear reactor site to monitor the reactor power via measurements of its antineutrino flux, as elucidated below.

7.5.1 rBDT Model

The nuBDT model presented in chapter 4 was optimised for an antineutrino oscillation analysis. It is not trained on variables that are correlated with the positron energy or position in order to ensure a uniform selection efficiency in these variables and keep the associated systematic uncertainties small. A second model has been developed known as the Rate–only Boosted Decision Tree (rBDT) which does not have the same constraints as the nuBDT and is more suitable for alternative antineutrino measurements. The goal of the rBDT is to maximise signal–background discrimination using all available event information, including direct positron information. It is trained on exactly the same preselected events as the nuBDT. The hyperparameters are also unchanged as they are found to still provide good discrimination whilst not overfitting. The corresponding training plots may be found in appendix H.

The only criterion by which a variable is chosen as an rBDT feature is whether or not it yields an improvement in performance, as measured by the AUC metric. Correlation between variables is not as important a consideration as for the nuBDT. The nuBDT and rBDT models differ by the following variables, as defined in Tab. 4.4:

- EnPositron (E_{MAGE})
- EnHighestECube (E_{MEC})
- Z_{MEC}
- ES2ES1
- ESNS

The feature importance scores are visualised in Fig. 7.4. The most important variables by

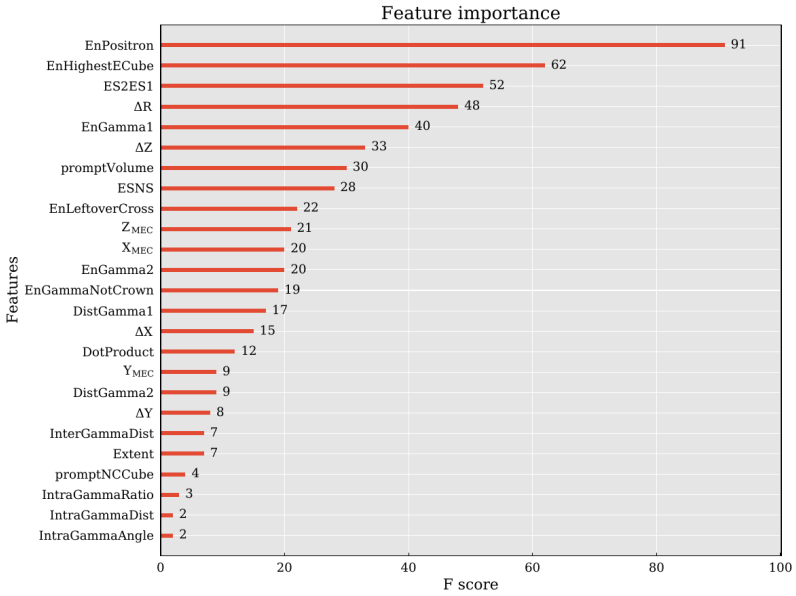


Figure 7.4: Fischer score importance of rBDT input features, defined as the number of times a feature is used to split the data. The best variables by this metric are the positron energy, the energy of the most energetic cube, ES2ES1, and ΔR.

this metric are the positron energy (E_{MAGE}), the energy of the most energetic cube (E_{MEC}), ΔR, and ES2ES1. The positron position Z_{MEC} has a comparatively modest impact. The distributions of rBDT score for signal and background events are displayed in Fig. 7.5, where the ROC curves of the two models are also drawn. Unlike the nuBDT, the rBDT score distribution for signal does not exhibit a second peak in the 2γ-rich region above 0.9. This is because the rBDT does not need to rely so heavily on topological information, as reflected in the features importance scores. By including positron information the AUC is greatly improved from 0.886 for the nuBDT to 0.920 for the rBDT.

Given the AUC disparity, the rBDT is expected to outperform the nuBDT in terms of the $S/\sqrt{S+B}$ metric. To test this, the subtraction procedure has been applied to the open dataset*. The optimisation results plotted in Fig. 7.6 indicate a modest improvement in performance over the nuBDT. The FoM is maximised at 5.51 ± 0.24 by a cut on rBDT score of 0.78, corresponding to a signal-to-background ratio of 0.320 ± 0.014 and 125 ± 5.4 events per day.

7.5.2 Reactor Power Measurement

Antineutrinos are unique amongst all fission products in that they cannot be shielded. The linear evolution of antineutrino rate as a function of power was verified with SoLid data in section 5.4.1. Once calibrated to the background rate at a particular site, the SoLid detector

*The first reactor cycle (22.7 d live data-taking) and the zeroth and first ROff periods (34.5 d live data-taking) in the Phase-I dataset.

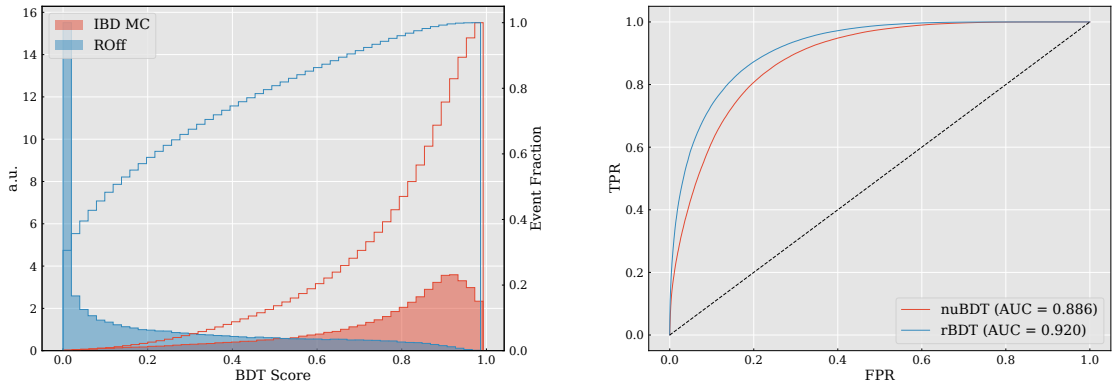


Figure 7.5: Left: Output of the rBDT model in terms of score distributions of signal and background. The cumulative score distributions are plotted as un-filled histograms and must be interpreted in accordance with the right-hand side axis. Right: rBDT and nuBDT ROC curves and AUC values. The signal-background discrimination of the rBDT surpasses that of the nuBDT, as reflected by the AUCs.

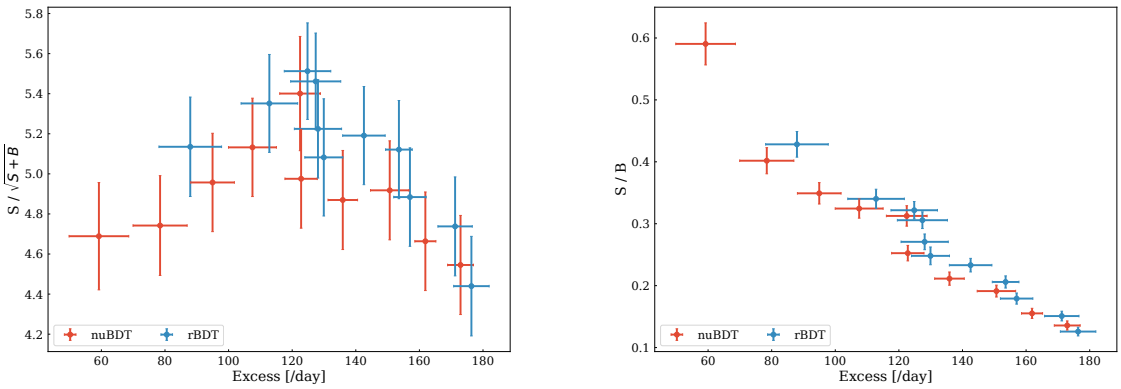


Figure 7.6: Figure of Merit and signal-to-background ratio as a function of excess per day (rBDT cut). Error bars represent the statistical uncertainty on the points. The rBDT FoM is optimised at 125 events per day which gives a signal-to-background ratio of around 0.32, slightly better than the optimal point of the nuBDT.

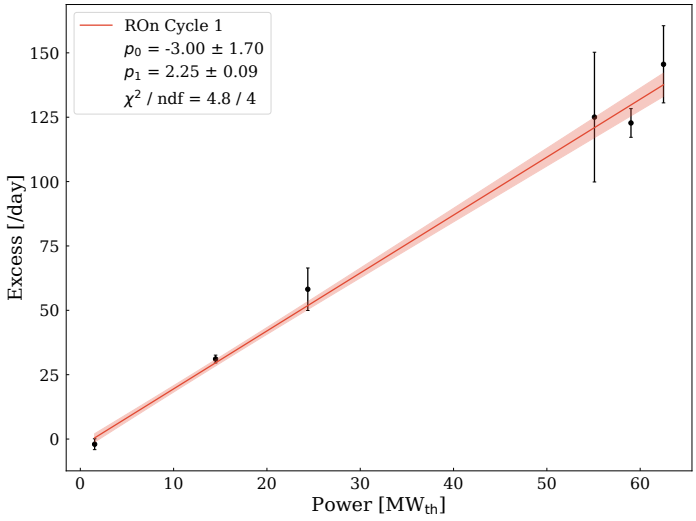


Figure 7.7: Open-dataset IBD rate as a function of reactor power obtained with the rBDT analysis. The data includes data taken during intermediate reactor-power days (1–40 $\mathrm{MW_{th}}$). The linear fit (red) through the data points (black) has a one-sigma uncertainty on the gradient of 4%. Error bars represent the statistical uncertainty on the data points.

could act as an additional check that the reactor is behaving as expected. The IBD rate as a function of power for the open dataset is plotted in Fig. 7.7. The precision on the fit parameters is improved with respect to the previous result by a) including previously discarded data taken when the reactor operated at power between 1 $\mathrm{MW_{th}}$ and 40 $\mathrm{MW_{th}}$, approximately 1.4 d exposure time, and b) using the rBDT instead of the nuBDT. It could be further improved by compiling additional reactor cycles.

The linear fit can be inverted to derive a function that outputs the mean reactor power ($\overline{P}_{\mathrm{th}}$) for a given antineutrino rate:

$$\overline{P}_{\mathrm{th}} = \frac{1}{p_1}(N_{\mathrm{IBD}} - p_0)\,, \tag{7.2}$$

where p_0 and p_1 are the y-intercept and gradient of the fit respectively and N_{IBD} is the total number of IBD candidates in the given time period. The fractional statistical uncertainty on such a measurement is given by:

$$\frac{\Delta P_{\mathrm{th}}}{P_{\mathrm{th}}} = \sqrt{\frac{\Delta N_{\mathrm{IBD}}^2 + \Delta p_0^2}{(N_{\mathrm{IBD}} - p_0)^2} + \frac{\Delta p_1^2}{p_1^2}}\,, \tag{7.3}$$

where ΔA represents the uncertainty on parameter A. The point in Fig. 7.7 in the 60–70 $\mathrm{MW_{th}}$ power bin corresponds to approximately 2.4 d of data with a fractional uncertainty of 10%, which translates to a 10.6% uncertainty on the power. The point at 58–60 $\mathrm{MW_{th}}$ corresponds to a 5% uncertainty from around seventeen days, yielding a 6.5% uncertainty.

Finally, the error on the reactor power quoted by BR2 (neglected in this analysis) is a conservative 5%. The error could be reduced by combining the three available measurements: MWRA4-1301 power (water flow rate and temperature in the cooling circuits), A3 power (nitrogen activation), and the IBD rate. However, any combined error must account for the correlation between the first two parameters, against which the IBD rate also acts as a cross-check. To summarise, a precise measurement of the reactor power can be obtained with $\mathcal{O}(10)$ days-worth of data-taking at a given power, an obvious asset for reactor monitoring purposes. One caveat is that the precision achieved is dependent on being as close to the reactor core as SoLid is at BR2, though the effect of greater distance could be unimportant if the detector were to be situated at a commercial reactor that is orders of magnitude more powerful than BR2.

7.5.3 Antineutrino Directionality

Another important measurement for reactor monitoring is the direction of the incoming antineutrino. Situations where this could be useful are conceivable at power plants that operate multiple reactors. The IBD neutron is slightly boosted in the direction of the antineutrino momentum vector. Hence, the neutron capture vertex is more likely to be displaced from the IBD vertex along the direction of this vector, creating a so-called *physical bias*. This effect can be utilised to measure the average neutron displacement in each direction and thus locate the antineutrino source. A study has been undertaken to determine the feasibility of such a measurement.

The positions of the neutron detection screens introduce an *intrinsic bias* in the measured neutron position with respect to the IBD vertex. A neutron is more likely to be captured if it travels in the −z and +x directions where the screens are placed. A sample of IBD events with homogeneous and isotropic neutron momenta has been produced with the SoLO IBD generator to characterise the intrinsic bias. This is referred to as the *Rand. MC* sample. Every neutron in the Rand. MC sample has the same magnitude of momentum. The initial positions and momenta of the neutrons are displayed in Fig. 7.8. There is a slight deficit at the centre of the momentum distributions, possibly an artefact of the momentum in a particular direction being close to zero.

The events are passed through the full simulation and reconstruction pipeline including the preselection and rBDT cuts. The Saffron2 parameters ΔX, ΔY and ΔZ describe the displacement of the neutron with respect to the highest energy reconstructed cube. Events in the Rand. MC sample do not carry the physical bias from the antineutrino direction, so their distributions in these variables reflect only the intrinsic detector bias. If the neutron travels in the positive x-direction, then the first screen it will encounter is in its own voxel, whereas the nearest screen in the opposite direction is in the adjacent voxel. The neutron

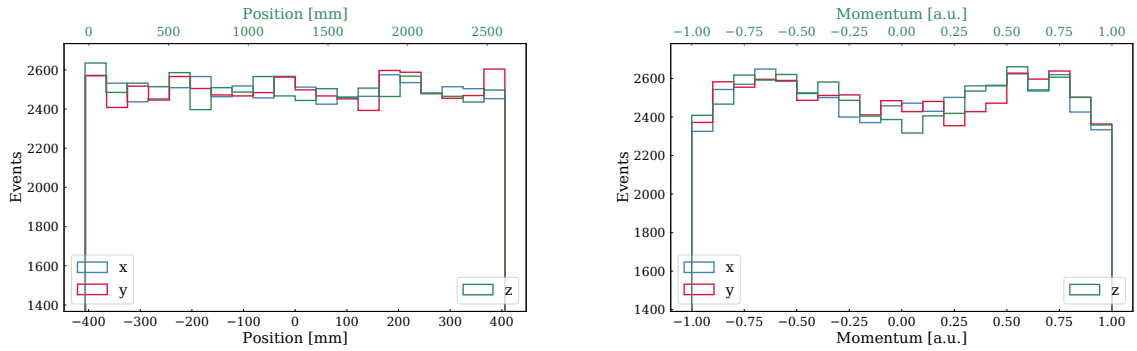


Figure 7.8: Position (left) and momentum (right) distributions of the Rand. MC sample. The IBD neutron is emitted isotropically with a constant momentum magnitude.

has an equal chance of travelling in either direction (for Rand. MC events) so a ΔX value of -1 is more likely than $+1$ and a value of zero is most probable. Similarly, the intrinsic shift in the z-direction should be positive and the lack of screens on the $\pm$y faces of the cubes means there is no shift in the y-direction

The shifts in each direction (physical and intrinsic for data and IBD MC, intrinsic only for Rand. MC) are estimated as the means (μ) of Gaussian fits on the ΔX, ΔY, and ΔZ distributions. The physical neutron shift is then calculated as:

$$\mu(\Delta X_{\text{phys}}) = \mu(\Delta X_{\text{data}}) - \mu(\Delta X_{\text{intrinsic}}) \,, \tag{7.4}$$

and similarly for $\mu(\Delta Y_{\text{phys}})$ and $\mu(\Delta Z_{\text{phys}})$. Finally, a vector is constructed that points in the direction of the average antineutrino momentum:

$$\boldsymbol{v} = (\mu(\Delta X_{\text{phys}}),\ \mu(\Delta Y_{\text{phys}}),\ \mu(\Delta Z_{\text{phys}})) \,, \tag{7.5}$$

whose norm is the magnitude of the physical shift.

The steps of this method are encapsulated in Fig. 7.9. The intrinsic biases are extracted from the Rand. MC histograms (grey) and corresponding Gaussian fits (red). The data distributions and fits are plotted in blue. The data comes from the background-subtraction of the open-dataset after applying optimised rBDT cuts, as used in section 7.5.2. As expected, the shift from the intrinsic bias in the x-direction is negative, positive in the z-direction, and zero in the y-direction. The values of the physical shifts, calculated from equation 7.5, are marked by vertical dotted lines. The only statistically significant non-zero physical shift is along the z-axis, as expected given our knowledge of the position of the reactor core in relation to the detector.

The norm and orientation in space of the vector $\boldsymbol{v}$ are shown in Fig. 7.10. The magnitude of the shift is approximately 2 cm, in excellent agreement with the predicted obtained by repeating the procedure with standard IBD MC events instead of antineutrino data. The

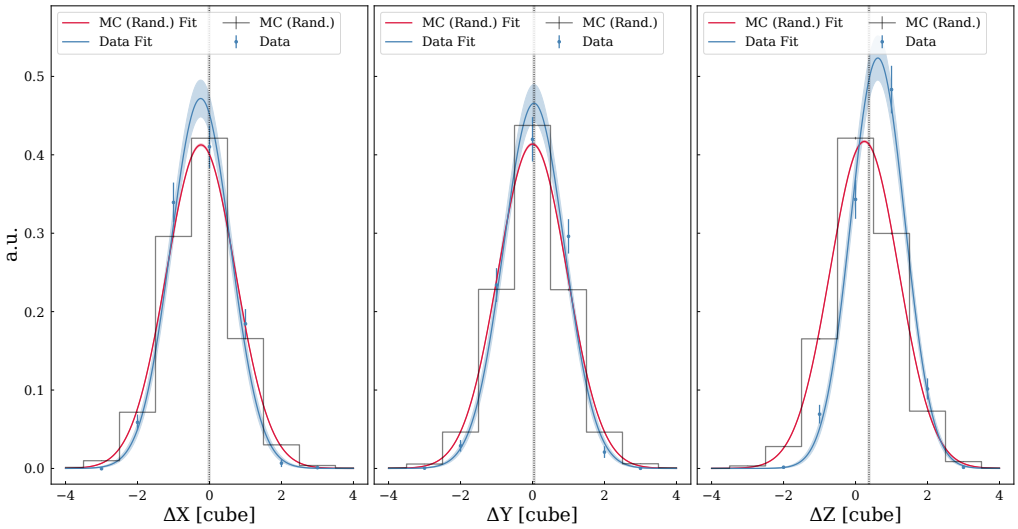


Figure 7.9: Characterisation of the intrinsic and physical neutron direction biases. ΔX, ΔY, and ΔZ histograms of the subtracted data and Rand. MC sample describe the intrinsic detector biases and data (intrinsic plus physical) biases respectively. Error bars represent the statistical uncertainty on the points. The blue bands represents the uncertainties on the Gaussian fits. The vertical dotted lines mark the calculated physical biases; the only non-zero value is obtained for the z-direction.

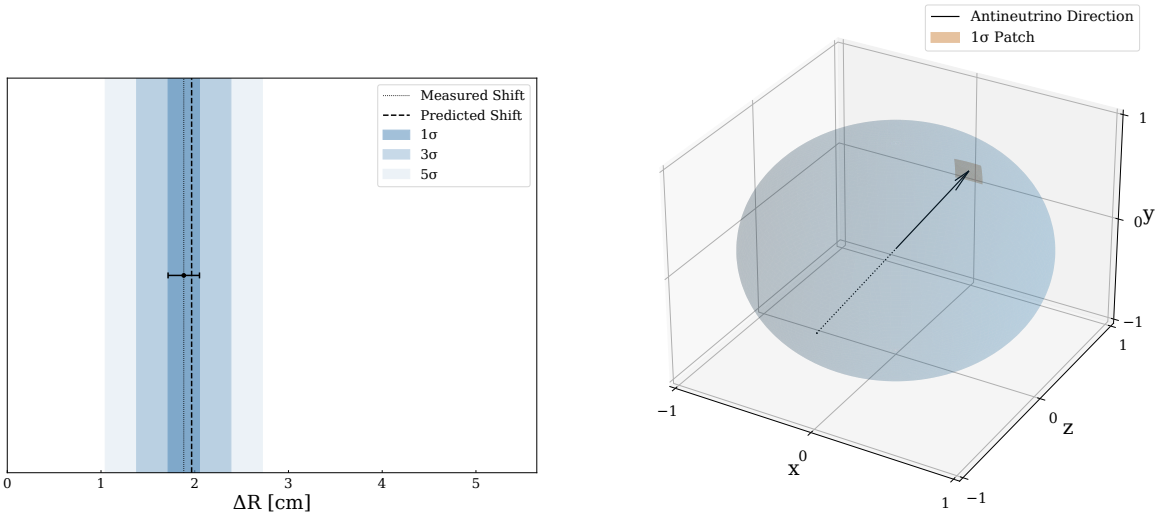


Figure 7.10: Results of the antineutrino direction measurement. Left: Magnitude of the physical shift in units of centimetres. Right: Unit vector $\boldsymbol{v}/\|\boldsymbol{v}\|$ plotted in 3D space. The orientation of the vector is consistent with the antineutrinos having emerged from the BR2 reactor core.

vector points in a direction consistent with the antineutrinos having emerged from the BR2 reactor core. The method described above should be developed either by enhancing the uniformity of the neutron momenta in the Rand. MC sample or deriving a dedicated systematic uncertainty to ensure an unbiased estimate of the intrinsic shift. Nonetheless, a five-sigma measurement is easily achieved with the 22.7 days of the open dataset. Assuming the required statistical significance for reactor monitoring would be less stringent, an acceptably precise measurement is plausible after only a few days of reactor-on data-taking.

7.6 Conclusion

The Reactor Antineutrino Anomaly that prompted a burst of activity in the field of very-short-baseline neutrino experiments is roughly twelve years old at the time of writing. Other anomalies are even older and have persisted through new independent measurements. At this time, confidence in a neutrino-based solution to the RAA is far lower than when initially suggested thanks to the wealth of data now available to inform this question. Both the experimental data and updated reactor models point cautiously at ^{235}U as the main contributor to the RAA. Despite this, the source of the RAA is still not firmly established and strong tensions exist between different measurements. The exact origin of the 5 MeV bump is also not clear. As far as the 3+1 neutrino model is concerned, the Δm^2_{41}-$\sin^2(2\theta_{14})$ parameter space has been heavily constrained by VSBL neutrino experiments, of which SoLid is one example.

The SoLid collaboration designed and built a 1.6 t plastic detector based on a duel-scintillator technology to detect antineutrinos via inverse beta decay. The high degree of segmentation allows such events to be reconstructed in all three spatial dimensions. The linear scintillator response as a function of energy is another benefit of the plastic scintillator bulk that reduces the complexity of the energy calibration. The detector was deployed at the BR2 research reactor in Belgium with the aim of accumulating enough antineutrino data to test the sterile neutrino hypothesis of the RAA and precisely measure the pure ^{235}U antineutrino spectrum. The Phase-I iteration of the experiment ran from May 2018 to July 2020 and collected approximately 280 d and 172 d of reactor-on and reactor-off data respectively.

Taking advantage of the kinematics of inverse beta decay and the detector granularity allows the energy of the IBD positron to be effectively estimated and separated from the associated annihilation gamma deposits. The subsequent classification system of cube-level energy deposits forms the basis of the machine-learning tools. The rich topological information inherent in the events and captured by the ES variables makes SoLid well-suited to ML analyses. A combination of rectangular cuts and a boosted decision tree have been used to suppress the naturally high background rates that originate from atmospheric neutron

showers and intrinsic radioactivity. The rectangular cuts include a sophisticated CNN-based PSD technique, developed prior to any work by this author, to suppress BiPo background.

The background subtraction technique provides a relatively stable IBD yield over the Phase-I data-taking period by partitioning the dataset into manageable chunks. Some instabilities are observed, especially in the second half of the dataset, but it is not clear that these are not artefacts of the data itself rather than the applied analysis techniques. The total IBD yield obtained from the Phase-I dataset is $29\,479 \pm 603$ (stat.) with an overall signal-to-background of 0.26. Numerous cross-checks have been performed to verify the behaviour of the IBD candidates and validity of the background subtraction. These include data-MC comparisons in multiple variables, the scaling of the IBD rate with reactor power, a background-background subtraction, and a shape-only test of the individual reactor-cycle spectra.

The background-subtracted IBD position-energy spectrum has been analysed in a search for evidence of sterile neutrinos in the 3+1 neutrino paradigm. A study of the relevant systematic uncertainties for a shape-only antineutrino disappearance measurement demonstrates that the data is statistically dominated. The robust Feldman-Cousins technique has been employed for the construction of confidence intervals in the space of active-sterile mixing parameters. A model-dependent, normalisation-independent fit against the HM prediction finds no evidence of antineutrino disappearance. The current sensitivity of the analysis means that the RAA best-fit point has not been properly investigated. Updates to the signal-background discrimination tools have been proposed that may prove rewarding in this regard. Aside from an oscillation analysis, the measurement and characterisation of electron antineutrinos from the BR2 reactor shows promise for reactor monitoring purposes.

Glossary

ADC analogue-to-digital converter.

APD avalanche photodiode.

AUC Area Under the Curve.

BDT Boosted Decision Tree.

BEST Baksan Experiment on Sterile Transitions.

BIDASSE BR2 Integrated Data Acquisition System for Survey and Experiments.

BNL Brookhaven National Laboratory.

BR2 Belgian Reactor 2.

BSM Beyond Standard Model.

CC charged-current.

CERN Conseil Européen pour la Recherche Nucléaire.

CL confidence level.

CNN convolutional neural network.

CP charge-parity.

CSDA Continuous-Slowing-Down Approximation.

CV cross-validation.

DAQ data acquisition.

EF Estienne-Fallot.

ES Electromagnetic Signal.

ESTAR Electron Stopping Power And Ranges.

FoM figure of merit.

FPGA field-programmable gate array.

FPNT False Positive Neutron Trigger.

GALLEX Gallium Experiment.

GWS Glashow-Weinberg-Salam.

HDPE high-density polyethylene.

HEU high-enriched uranium.

HFIR High Flux Isotope Reactor.

HKSS Hayen-Kostensalo-Severijns-Suhonen.

HM Huber-Mueller.

IBD inverse beta decay.

ILL Institut Laue Langevin.

IO inverted ordering.

KARMEN Karlsruhe Rutherford Medium Energy Neutrino.

KATRIN Karlsruhe Tritium Neutrino.

KI Kurchatov Institute.

KS Kolmogorov-Smirnov.

LArTPC liquid-argon time-projection chamber.

LEE low-energy excess.

LEU low-enriched uranium.

LL light-leakage.

LSND Liquid Scintillator Neutrino Detector.

MAGE Minimal Annihilation Gamma Estimator.

MC Monte Carlo.

MCNPX Monte Carlo N-Particle eXtended.

MEC most energetic cube.

ML machine learning.

MLEM Maximum-Likelihood Expectation-Minimisation.

MPPC Multi-Pixel Photon Counter.

MPV most probable value.

MSW Mikheyev-Smirnov-Wolfenstein.

MURE MCNP Utility for Reactor Evolution.

NC neutral-current.

NIST National Institute of Standards and Technology.

NO normal ordering.

NS Nuclear Signal.

OMP Orthogonal Matching Pursuit.

ORNL Oak Ridge National Laboratory.

OV overvoltage.

p-p proton-proton.

PA pixel avalanche.

PDE photon detection efficiency.

PDF probability density function.

PDG Particle Data Group.

PMNS Pontecorvo-Maki-Nakagawa-Sakata.

PMT photomultiplier tube.

PSD pulse shape discrimination.

PVT polyvinyl toluene.

PySoLO Python3 for SoLid Oscillations.

RAA Reactor Antineutrino Anomaly.

ReMU Response Matrix Utilities.

ROC Receiver Operating Characteristic.

ROSim Readout Simulation.

S/B signal-to-background.

Saffron2 SoLid Analysis Framework.

SAGE Soviet-American Gallium Experiment.

SBL short-baseline.

SiPM silicon photomultiplier.

SM Standard Model.

SMat System Matrix.

SMP Significant Migration Probability.

SNO Sudbury Neutrino Observatory.

SoLidSim SoLid Simulation.

SoLO SoLid Oscillations.

SSB spontaneous symmetry breaking.

SSM Standard Solar Model.

TAGS Total Absorption Gamma-ray Spectroscopy.

VSBL very-short-baseline.

WS wavelength-shifting.

XGBoost eXtreme Gradient Boosting.

ZS zero-suppression.

Bibliography

[1] James Chadwick. "Intensitatsverteilung im magnetischen Spektrum der β-Strahlen von Radium B + C". In: *Verhandlungen der deutschen physikalischen Gesellschaft* 16 (1914), pp. 383–391.

[2] James Chadwick and Charles Ellis. "A Preliminary Investigation of the Intensity Distribution in the β-Ray Spectra of Radium B and C". In: *Proceedings of the Cambridge Philosophical Society* 21 (1922), pp. 274–280.

[3] C. Ellis and W. Wooster. "The average energy of disintegration of radium E". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117 (776 Dec. 1927), pp. 109–123. ISSN: 0950-1207. DOI: 10.1098/rspa.1927.0168.

[4] L Meitner and W Orthmann. "Über eine absolute Bestimmung der Energie der primären β-Strahlen von Radium E". In: *Zeitschrift fur Physik* 60 (1930), pp. 143–155.

[5] E. Fermi. "Versuch einer Theorie der β-Strahlen. I". In: *Zeitschrift fur Physik* 88 (3-4 Mar. 1934), pp. 161–177. ISSN: 1434-6001. DOI: 10.1007/BF01351864.

[6] B. W. Sargent. "The maximum energy of the β-Rays from uranium X and other bodies". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 139 (839 Mar. 1933), pp. 659–673. ISSN: 0950-1207. DOI: 10.1098/rspa.1933.0045.

[7] F. Reines and C. L. Cowan. "A proposed experiment to detect the free neutrino". In: *Physical Review* 90 (3 1953). ISSN: 0031899X. DOI: 10.1103/PhysRev.90.492.2.

[8] F. Reines and C. L. Cowan. "Detection of the free neutrino". In: *Physical Review* 92 (3 1953). ISSN: 0031899X. DOI: 10.1103/PhysRev.92.830.

[9] C. L. Cowan et al. "Detection of the free neutrino: A confirmation". In: *Science* 124 (3212 1956). ISSN: 00368075. DOI: 10.1126/science.124.3212.103.

[10] M. Schwartz. "Feasibility of Using High-Energy Neutrinos to Study the Weak Interactions". In: *Physical Review Letters* 4 (6 Mar. 1960), pp. 306–307. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.4.306.

[11] G. Danby et al. “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos”. In: *Physical Review Letters* 9 (1 July 1962), pp. 36–44. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.9.36.

[12] J.K. Bienlein et al. “Spark chamber study of high-energy neutrino interactions”. In: *Physics Letters* 13 (1 Nov. 1964), pp. 80–86. ISSN: 00319163. DOI: 10.1016/0031-9163(64)90316-6.

[13] B. Pontecorvo. “Electron and Muon Neutrinos”. In: *Zh. Eksp. Teor. Fiz.* 37 (1959), pp. 1751–1757.

[14] M. L. Perl et al. “Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation”. In: *Physical Review Letters* 35 (22 Dec. 1975), pp. 1489–1492. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.35.1489.

[15] M.L. Perl et al. “Properties of anomalous e mu events produced in $e^+ - e^-$ annihilation”. In: *Physics Letters B* 63 (4 Aug. 1976), pp. 466–470. ISSN: 03702693. DOI: 10.1016/0370-2693(76)90399-3.

[16] M.L. Perl et al. “Properties of the proposed τ charged lepton”. In: *Physics Letters B* 70 (4 Oct. 1977), pp. 487–490. ISSN: 03702693. DOI: 10.1016/0370-2693(77)90421-X.

[17] K. Kodama et al. “Observation of tau neutrino interactions”. In: *Physics Letters B* 504 (3 Apr. 2001), pp. 218–224. ISSN: 03702693. DOI: 10.1016/S0370-2693(01)00307-0.

[18] G. Gamow and E. Teller. “Selection Rules for the β-Disintegration”. In: *Physical Review* 49 (12 June 1936), pp. 895–899. ISSN: 0031-899X. DOI: 10.1103/PhysRev.49.895.

[19] T. D. Lee and C. N. Yang. “Question of Parity Conservation in Weak Interactions”. In: *Physical Review* 104 (1 Oct. 1956), pp. 254–258. ISSN: 0031-899X. DOI: 10.1103/PhysRev.104.254.

[20] T. D. Lee and C. N. Yang. “Parity Nonconservation and a Two-Component Theory of the Neutrino”. In: *Physical Review* 105 (5 Mar. 1957), pp. 1671–1675. ISSN: 0031-899X. DOI: 10.1103/PhysRev.105.1671.

[21] L. Landau. “On the conservation laws for weak interactions”. In: *Nuclear Physics* 3 (1 Mar. 1957), pp. 127–131. ISSN: 00295582. DOI: 10.1016/0029-5582(57)90061-5.

[22] Abdus Salam. “On parity conservation and neutrino mass”. In: *Il Nuovo Cimento* 5 (1 Jan. 1957), pp. 299–301. ISSN: 0029-6341. DOI: 10.1007/BF02812841.

[23] C. S. Wu et al. “Experimental Test of Parity Conservation in Beta Decay”. In: *Physical Review* 105 (4 Feb. 1957), pp. 1413–1415. ISSN: 0031-899X. DOI: 10.1103/PhysRev.105.1413.

[24] Richard L. Garwin, Leon M. Lederman, and Marcel Weinrich. “Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon”. In: *Physical Review* 105 (4 Feb. 1957), pp. 1415–1417. ISSN: 0031-899X. DOI: 10.1103/PhysRev.105.1415.

[25] Jerome I. Friedman and V. L. Telegdi. “Nuclear Emulsion Evidence for Parity Nonconservation in the Decay Chain $\pi+$ - $\mu+$ - e+”. In: *Physical Review* 105 (5 Mar. 1957), pp. 1681–1682. ISSN: 0031-899X. DOI: 10.1103/PhysRev.105.1681.2.

[26] E. C. G. Sudarshan and R. E. Marshak. “Chirality Invariance and the Universal Fermi Interaction”. In: *Physical Review* 109 (5 Mar. 1958), pp. 1860–1862. ISSN: 0031-899X. DOI: 10.1103/PhysRev.109.1860.2.

[27] R. P. Feynman and M. Gell-Mann. “Theory of the Fermi Interaction”. In: *Physical Review* 109 (1 Jan. 1958), pp. 193–198. ISSN: 0031-899X. DOI: 10.1103/PhysRev.109.193.

[28] Sheldon L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22 (4 Feb. 1961), pp. 579–588. ISSN: 00295582. DOI: 10.1016/0029-5582(61)90469-2.

[29] A. Salam and J.C. Ward. “Electromagnetic and weak interactions”. In: *Physics Letters* 13 (2 Nov. 1964), pp. 168–171. ISSN: 00319163. DOI: 10.1016/0031-9163(64)90711-5.

[30] Steven Weinberg. “A Model of Leptons”. In: *Physical Review Letters* 19 (21 Nov. 1967), pp. 1264–1266. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.19.1264.

[31] M. Goldhaber, L. Grodzins, and A. W. Sunyar. “Helicity of Neutrinos”. In: *Phys. Rev.* 109 (3 Feb. 1958), pp. 1015–1017. DOI: 10.1103/PhysRev.109.1015.

[32] P.W. Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Physics Letters* 12 (2 Sept. 1964), pp. 132–133. ISSN: 00319163. DOI: 10.1016/0031-9163(64)91136-9.

[33] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Physical Review Letters* 13 (9 Aug. 1964), pp. 321–323. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.13.321.

[34] T. W. B. Kibble. “Symmetry Breaking in Non-Abelian Gauge Theories”. In: *Physical Review* 155 (5 Mar. 1967), pp. 1554–1561. ISSN: 0031-899X. DOI: 10.1103/PhysRev.155.1554.

[35] John N. Bahcall. "Solar Neutrinos. I. Theoretical". In: *Physical Review Letters* 12 (11 Mar. 1964), pp. 300–302. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.12.300.

[36] Raymond Davis. "Solar Neutrinos. II. Experimental". In: *Physical Review Letters* 12 (11 Mar. 1964), pp. 303–305. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.12.303.

[37] J. N. Bahcall et al. "Solar Neutrino Flux." In: *The Astrophysical Journal* 137 (Jan. 1963), p. 344. ISSN: 0004-637X. DOI: 10.1086/147513.

[38] John N. Bahcall and Giora Shaviv. "Solar Models and Neutrino Fluxes". In: *The Astrophysical Journal* 153 (July 1968), p. 113. ISSN: 0004-637X. DOI: 10.1086/149641.

[39] J.N. Bahcall et al. "Solar neutrinos and low-energy nuclear cross sections". In: *Physics Letters B* 26 (6 Feb. 1968), pp. 359–361. ISSN: 03702693. DOI: 10.1016/0370-2693(68)90610-2.

[40] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. "Search for Neutrinos from the Sun". In: *Physical Review Letters* 20 (21 May 1968), pp. 1205–1209. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.20.1205.

[41] John N. Bahcall, Neta A. Bahcall, and Giora Shaviv. "Present Status of the Theoretical Predictions for the ^{37}Cl Solar-Neutrino Experiment". In: *Physical Review Letters* 20 (21 May 1968), pp. 1209–1212. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.20.1209.

[42] John N. Bahcall. "Solar models: An historical overview". In: *Nuclear Physics B - Proceedings Supplements* 118 (2003). Proceedings of the XXth International Conference on Neutrino Physics and Astrophysics, pp. 77–86. ISSN: 0920-5632. DOI: https://doi.org/10.1016/S0920-5632(03)01306-9. URL: https://www.sciencedirect.com/science/article/pii/S0920563203013069.

[43] A. I. Abazov et al. "Search for Neutrinos from the Sun Using the Reaction ^{71}Ga(ν_e,e^-)^{71}Ge". In: *Physical Review Letters* 67 (24 Dec. 1991), pp. 3332–3335. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.67.3332.

[44] P. Anselmann et al. "Solar neutrinos observed by GALLEX at Gran Sasso". In: *Physics Letters B* 285 (4 July 1992), pp. 376–389. ISSN: 03702693. DOI: 10.1016/0370-2693(92)91521-A.

[45] K. S. Hirata et al. "Observation of ^{8}B solar neutrinos in the Kamiokande-II detector". In: *Physical Review Letters* 63 (1 July 1989), pp. 16–19. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.63.16.

[46] K. S. Hirata et al. “Real-time, directional measurement of ^{8}B solar neutrinos in the Kamiokande II detector”. In: *Physical Review D* 44 (8 Oct. 1991), pp. 2241–2260. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.44.2241.

[47] John N. Bahcall and Roger K. Ulrich. “Solar models, neutrino experiments, and helioseismology”. In: *Reviews of Modern Physics* 60 (2 Apr. 1988), pp. 297–372. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.60.297.

[48] Y. Fukuda et al. “Evidence for Oscillation of Atmospheric Neutrinos”. In: *Physical Review Letters* 81 (8 Aug. 1998), pp. 1562–1567. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.81.1562.

[49] K. S. Hirata et al. “Observation in the Kamiokande-II detector of the neutrino burst from supernova SN1987A”. In: *Physical Review D* 38 (2 July 1988), pp. 448–458. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.38.448.

[50] Bruno Pontecorvo. “Inverse beta processes and nonconservation of lepton charge”. In: *Zh. Eksp. Teor. Fiz.* 34 (Oct. 1957).

[51] Q. R. Ahmad et al. “Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory”. In: *Physical Review Letters* 87 (7 July 2001), p. 071301. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.87.071301.

[52] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. “Remarks on the Unified Model of Elementary Particles”. In: *Progress of Theoretical Physics* 28 (5 Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870.

[53] G. Bellini et al. “Neutrino Oscillations”. In: *Advances in High Energy Physics* 2014 (2014), pp. 1–28. ISSN: 1687-7357. DOI: 10.1155/2014/191960.

[54] R L Workman et al. “Review of Particle Physics”. In: *Progress of Theoretical and Experimental Physics* 2022 (8 Aug. 2022). ISSN: 2050-3911. DOI: 10.1093/ptep/ptac097.

[55] C. Giganti, S. Lavignac, and M. Zito. “Neutrino oscillations: The rise of the PMNS paradigm”. In: *Progress in Particle and Nuclear Physics* 98 (2018), pp. 1–54. ISSN: 0146-6410. DOI: https://doi.org/10.1016/j.ppnp.2017.10.001. URL: https://www.sciencedirect.com/science/article/pii/S014664101730087X.

[56] K. N. Abazajian et al. “Light Sterile Neutrinos: A White Paper”. In: *arXiv e-prints* (Apr. 2012). DOI: 10.48550/arXiv.1204.5379.

[57] Dmitry V. Naumov. “The Sterile Neutrino: A short introduction”. In: *EPJ Web of Conferences* 207 (May 2019), p. 04004. ISSN: 2100-014X. DOI: 10.1051/epjconf/201920704004.

[58] M. Sajjad Athar et al. “Status and perspectives of neutrino physics”. In: *Progress in Particle and Nuclear Physics* 124 (2022), p. 103947. ISSN: 0146-6410. DOI: https://doi.org/10.1016/j.ppnp.2022.103947. URL: https://www.sciencedirect.com/science/article/pii/S0146641022000084.

[59] R. N. Mohapatra. “Seesaw mechanism and its implications”. In: *SEESAW25: International Conference on the Seesaw Mechanism and the Neutrino Mass*. Dec. 2004, pp. 29–44. DOI: 10.1142/9789812702210_0003. arXiv: hep-ph/0412379.

[60] Manfred Lindner, Tommy Ohlsson, and Gerhart Seidl. “Seesaw mechanisms for Dirac and Majorana neutrino masses”. In: *Phys. Rev. D* 65 (5 Feb. 2002), p. 053014. DOI: 10.1103/PhysRevD.65.053014.

[61] Rabindra N. Mohapatra and Goran Senjanović. “Neutrino masses and mixings in gauge models with spontaneous parity violation”. In: *Physical Review D* 23 (1 Jan. 1981), pp. 165–180. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.23.165.

[62] Nima Arkani-Hamed et al. “Neutrino masses from large extra dimensions”. In: *Phys. Rev. D* 65 (2 Dec. 2001), p. 024032. DOI: 10.1103/PhysRevD.65.024032.

[63] Peter Minkowski. “$\mu \to e\gamma$ at a rate of one out of 109 muon decays?” In: *Physics Letters B* 67 (4 Apr. 1977), pp. 421–428. ISSN: 03702693. DOI: 10.1016/0370-2693(77)90435-X.

[64] Murray Gell-Mann, Pierre Ramond, and Richard Slansky. “Complex Spinors and Unified Theories”. In: *Conf. Proc. C* 790927 (1979), pp. 315–321. arXiv: 1306.4669.

[65] Rabindra N. Mohapatra and Goran Senjanović. “Neutrino Mass and Spontaneous Parity Nonconservation”. In: *Physical Review Letters* 44 (14 Apr. 1980), pp. 912–915. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.44.912.

[66] M. Agostini et al. “Comprehensive geoneutrino analysis with Borexino”. In: *Phys. Rev. D* 101 (1 Jan. 2020), p. 012009. DOI: 10.1103/PhysRevD.101.012009.

[67] S. Abe et al. “Abundances of Uranium and Thorium Elements in Earth Estimated by Geoneutrino Spectroscopy”. In: *Geophysical Research Letters* 49 (16 Aug. 2022). ISSN: 0094-8276. DOI: 10.1029/2022GL099566.

[68] M. G. Aartsen et al. “Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector”. In: *Science* 342 (6161 Nov. 2013). ISSN: 0036-8075. DOI: 10.1126/science.1242856.

[69] Edoardo Vitagliano, Irene Tamborra, and Georg Raffelt. “Grand unified neutrino spectrum at Earth: Sources and spectral components”. In: *Rev. Mod. Phys.* 92 (4 Dec. 2020), p. 045006. DOI: 10.1103/RevModPhys.92.045006.

[70] Ivan Esteban et al. *v5.2: Three-neutrino fit based on data available in November 2022*. 2022. URL: http://www.nu-fit.org/?q=node/256.

[71] A. Abed Abud et al. "Snowmass Neutrino Frontier Report". In: *Snowmass 2021*. Nov. 2022. DOI: 10.48550/arXiv.2211.08641. arXiv: 2211.08641.

[72] Lodewijk Nauta et al. "First neutrino oscillation measurement in KM3NeT/ORCA". In: Sissa Medialab, July 2021, p. 1123. DOI: 10.22323/1.395.1123.

[73] JUNO Collaboration. "JUNO physics and detector". In: *Progress in Particle and Nuclear Physics* 123 (Mar. 2022), p. 103927. ISSN: 01466410. DOI: 10.1016/j.ppnp.2021.103927.

[74] Hyper-Kamiokande Working Group. "A Long Baseline Neutrino Oscillation Experiment Using J-PARC Neutrino Beam and Hyper-Kamiokande". In: *arXiv e-prints* (Dec. 2014). Document submitted to 18th J-PARC PAC meeting in May 2014. 50 pages, 41 figures. DOI: 10.48550/arXiv.1412.4673.

[75] M G Aartsen et al. "IceCube-Gen2: the window to the extreme Universe". In: *Journal of Physics G: Nuclear and Particle Physics* 48.6 (Apr. 2021), p. 060501. DOI: 10.1088/1361-6471/abbd48.

[76] M Aker et al. "KATRIN: status and prospects for the neutrino mass and beyond". In: *Journal of Physics G: Nuclear and Particle Physics* 49.10 (Sept. 2022), p. 100501. DOI: 10.1088/1361-6471/ac834e.

[77] N. Aghanim et al. "Planck 2018 results". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201833910.

[78] Gia Dvali et al. "Neutrinoless double beta decay: Neutrino mass versus new physics". In: *Physical Review D* 108 (7 Oct. 2023), p. 075012. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.108.075012.

[79] R. Arnold et al. "Probing new physics models of neutrinoless double beta decay with SuperNEMO". In: *The European Physical Journal C* 70 (4 Dec. 2010), pp. 927–943. ISSN: 1434-6044. DOI: 10.1140/epjc/s10052-010-1481-5.

[80] S. Abe et al. "Search for the Majorana Nature of Neutrinos in the Inverted Mass Ordering Region with KamLAND-Zen". In: *Physical Review Letters* 130 (5 Jan. 2023), p. 051801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.130.051801.

[81] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Physical Review Letters* 13 (4 July 1964), pp. 138–140. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.13.138.

[82] K. Abe et al. "Measurements of neutrino oscillation parameters from the T2K experiment using 3.6×10^{21} protons on target". In: *The European Physical Journal C* 83 (9 Sept. 2023), p. 782. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11819-x.

[83] Patrick Huber. "Reactor antineutrino fluxes – Status and challenges". In: *Nuclear Physics B* 908 (July 2016), pp. 268–278. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2016.04.012.

[84] Soo-Bong Kim, Thierry Lasserre, and Yifang Wang. "Reactor Neutrinos". In: *Advances in High Energy Physics* 2013 (2013), pp. 1–34. ISSN: 1687-7357. DOI: 10.1155/2013/453816.

[85] Anna C. Hayes and Petr Vogel. "Reactor Neutrino Spectra". In: *Annual Review of Nuclear and Particle Science* 66 (1 Oct. 2016), pp. 219–244. ISSN: 0163-8998. DOI: 10.1146/annurev-nucl-102115-044826.

[86] Xin Qian and Jen-Chieh Peng. "Physics with reactor neutrinos". In: *Reports on Progress in Physics* 82.3 (Feb. 2019), p. 036201. DOI: 10.1088/1361-6633/aae881. URL: https://dx.doi.org/10.1088/1361-6633/aae881.

[87] F. P. An et al. "Precision Measurement of Reactor Antineutrino Oscillation at Kilometer-Scale Baselines by Daya Bay". In: *Physical Review Letters* 130 (16 Apr. 2023), p. 161802. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.130.161802.

[88] G. Bak et al. "Measurement of Reactor Antineutrino Oscillation Amplitude and Frequency at RENO". In: *Physical Review Letters* 121 (20 Nov. 2018), p. 201801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.121.201801.

[89] T. Abrahão et al. "Reactor rate modulation oscillation analysis with two detectors in Double Chooz". In: *Journal of High Energy Physics* 2021 (1 Jan. 2021), p. 190. ISSN: 1029-8479. DOI: 10.1007/JHEP01(2021)190.

[90] K. Schreckenbach et al. "Absolute measurement of the beta spectrum from 235U fission as a basis for reactor antineutrino experiments". In: *Physics Letters B* 99 (3 Feb. 1981), pp. 251–256. ISSN: 03702693. DOI: 10.1016/0370-2693(81)91120-5.

[91] F. von Feilitzsch, A.A. Hahn, and K. Schreckenbach. "Experimental beta-spectra from 239Pu and 235U thermal neutron fission products and their correlated antineutrino spectra". In: *Physics Letters B* 118 (1-3 Dec. 1982), pp. 162–166. ISSN: 03702693. DOI: 10.1016/0370-2693(82)90622-0.

[92] N. Haag et al. "Experimental Determination of the Antineutrino Spectrum of the Fission Products of ^{238}U". In: *Physical Review Letters* 112 (12 Mar. 2014), p. 122501. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.112.122501.

[93] K. Schreckenbach et al. “Determination of the antineutrino spectrum from 235U thermal neutron fission products up to 9.5 MeV”. In: *Physics Letters B* 160 (4-5 Oct. 1985), pp. 325–330. ISSN: 03702693. DOI: 10.1016/0370-2693(85)91337-1.

[94] A.A. Hahn et al. “Antineutrino spectra from 241Pu and 239Pu thermal neutron fission products”. In: *Physics Letters B* 218 (3 Feb. 1989), pp. 365–368. ISSN: 03702693. DOI: 10.1016/0370-2693(89)91598-0.

[95] O. Tengblad et al. “Integral gn-spectra derived from experimental β-spectra of individual fission products”. In: *Nuclear Physics A* 503 (1 Oct. 1989), pp. 136–160. ISSN: 03759474. DOI: 10.1016/0375-9474(89)90258-3.

[96] G. Rudstam et al. “Beta and gama spectra of short-lived fission products”. In: *Atomic Data and Nuclear Data Tables* 45 (2 July 1990), pp. 239–320. ISSN: 0092640X. DOI: 10.1016/0092-640X(90)90009-9.

[97] M. Fallot et al. “New Antineutrino Energy Spectra Predictions from the Summation of Beta Decay Branches of the Fission Products”. In: *Physical Review Letters* 109 (20 Nov. 2012), p. 202504. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.109.202504.

[98] F. T. Avignone and Z. D. Greenwood. “Calculated spectra of antineutrinos from the fission products of 235 U , 238 U , and 239 Pu , and antineutrino-induced reactions”. In: *Physical Review C* 22 (2 Aug. 1980), pp. 594–605. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.22.594.

[99] P. Vogel et al. “Reactor antineutrino spectra and their application to antineutrino-induced reactions. II”. In: *Physical Review C* 24 (4 Oct. 1981), pp. 1543–1553. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.24.1543.

[100] Th. A. Mueller et al. “Improved predictions of reactor antineutrino spectra”. In: *Physical Review C* 83 (5 May 2011), p. 054615. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.83.054615.

[101] Patrick Huber. “Determination of antineutrino spectra from nuclear reactors”. In: *Physical Review C* 84 (2 Aug. 2011), p. 024617. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.84.024617.

[102] M. Estienne et al. “Updated Summation Model: An Improved Agreement with the Daya Bay Antineutrino Fluxes”. In: *Physical Review Letters* 123 (2 July 2019), p. 022502. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.123.022502.

[103] Ianthe Michiels. “Development of the oscillation analysis framework for the SoLid experiment”. PhD thesis. Ghent University, 2020, pp. xi, 233. URL: http://hdl.handle.net/1854/LU-8686780.

[104] A. A. Sonzogni et al. “Effects of Fission Yield Data in the Calculation of Antineutrino Spectra for ^{235}U(n,fission) at Thermal and Fast Neutron Energies”. In: *Physical Review Letters* 116 (13 Apr. 2016), p. 132502. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.116.132502.

[105] A. C. Hayes et al. “Systematic Uncertainties in the Analysis of the Reactor Neutrino Anomaly”. In: *Physical Review Letters* 112 (20 May 2014), p. 202501. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.112.202501.

[106] D. A. Dwyer and T. J. Langford. “Spectral Structure of Electron Antineutrinos from Nuclear Reactors”. In: *Physical Review Letters* 114 (1 Jan. 2015), p. 012502. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.114.012502.

[107] *Evaluated Nuclear Structure Data File*. URL: https://www.nndc.bnl.gov/ensdf/.

[108] *Japanese Evaluated Nuclear Data Library*. URL: https://wwwndc.jaea.go.jp/jendl/jendl.html.

[109] C. Giunti et al. “Reactor antineutrino anomaly in light of recent flux model refinements”. In: *Physics Letters B* 829 (June 2022), p. 137054. ISSN: 03702693. DOI: 10.1016/j.physletb.2022.137054.

[110] G. Mention et al. “Reactor antineutrino anomaly”. In: *Physical Review D* 83 (7 Apr. 2011), p. 073006. ISSN: 1550-7998. DOI: 10.1103/PhysRevD.83.073006.

[111] J. H. Choi et al. “Observation of Energy and Baseline Dependent Reactor Antineutrino Disappearance in the RENO Experiment”. In: *Physical Review Letters* 116 (21 May 2016), p. 211801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.116.211801.

[112] Y. Abe et al. “Measurement of θ_{13} in Double Chooz using neutron captures on hydrogen with novel background rejection techniques”. In: *Journal of High Energy Physics* 2016 (1 Jan. 2016), p. 163. ISSN: 1029-8479. DOI: 10.1007/JHEP01(2016)163.

[113] F. P. An et al. “Measurement of the Reactor Antineutrino Flux and Spectrum at Daya Bay”. In: *Physical Review Letters* 116 (6 Feb. 2016), p. 061801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.116.061801.

[114] G. Mention et al. “Reactor antineutrino shoulder explained by energy scale nonlinearities?” In: *Physics Letters B* 773 (Oct. 2017), pp. 307–312. ISSN: 03702693. DOI: 10.1016/j.physletb.2017.08.035.

[115] H Almazán et al. “First antineutrino energy spectrum from ^{235}U fissions with the STEREO detector at ILL”. In: *Journal of Physics G: Nuclear and Particle Physics* 48 (7 July 2021), p. 075107. ISSN: 0954-3899. DOI: 10.1088/1361-6471/abd37a.

[116] J. Ashenfelter et al. "Measurement of the Antineutrino Spectrum from ^{235}U Fission at HFIR with PROSPECT". In: *Physical Review Letters* 122 (25 June 2019), p. 251801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.122.251801.

[117] H. Almazán et al. "STEREO neutrino spectrum of 235U fission rejects sterile neutrino hypothesis". In: *Nature* 613 (7943 Jan. 2023), pp. 257–261. ISSN: 0028-0836. DOI: 10.1038/s41586-022-05568-2.

[118] D. Adey et al. "Extraction of the ^{235}U and ^{239}Pu Antineutrino Spectra at Daya Bay". In: *Physical Review Letters* 123 (11 Sept. 2019), p. 111801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.123.111801.

[119] M. Andriamirado et al. "Final Measurement of the ^{235}U Antineutrino Energy Spectrum with the PROSPECT-I Detector at HFIR". In: *Physical Review Letters* 131 (2 July 2023), p. 021802. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.131.021802.

[120] A. C. Hayes et al. "Possible origins and implications of the shoulder in reactor neutrino spectra". In: *Physical Review D* 92 (3 Aug. 2015), p. 033015. ISSN: 1550-7998. DOI: 10.1103/PhysRevD.92.033015.

[121] Patrick Huber. "NEOS Data and the Origin of the 5 MeV Bump in the Reactor Antineutrino Spectrum". In: *Physical Review Letters* 118 (4 Jan. 2017), p. 042502. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.118.042502.

[122] G. Bak et al. "Fuel-Composition Dependent Reactor Antineutrino Yield at RENO". In: *Physical Review Letters* 122 (23 June 2019), p. 232501. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.122.232501.

[123] Jeffrey M. Berryman, Vedran Brdar, and Patrick Huber. "Particle physics origin of the 5 MeV bump in the reactor antineutrino spectrum?" In: *Physical Review D* 99 (5 Mar. 2019), p. 055045. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.99.055045.

[124] Mario A. Acero, Carlo Giunti, and Marco Laveder. "Limits on ν_e and $\bar{\nu}_e$ disappearance from Gallium and reactor experiments". In: *Physical Review D* 78 (7 Oct. 2008), p. 073009. ISSN: 1550-7998. DOI: 10.1103/PhysRevD.78.073009.

[125] John N. Bahcall. "Gallium solar neutrino experiments: Absorption cross sections, neutrino spectra, and predicted event rates". In: *Physical Review C* 56 (6 Dec. 1997), pp. 3391–3409. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.56.3391.

[126] W Hampel et al. "Final results of the 51Cr neutrino source experiments in GALLEX". In: *Physics Letters B* 420 (1-2 Feb. 1998), pp. 114–126. ISSN: 03702693. DOI: 10.1016/S0370-2693(97)01562-1.

[127] W. Hampel and L. P. Remsberg. "Half-life of ^{71}Ge". In: *Physical Review C* 31 (2 Feb. 1985), pp. 666–667. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.31.666.

[128] V. V. Barinov et al. “Search for electron-neutrino transitions to sterile states in the BEST experiment”. In: *Physical Review C* 105 (6 June 2022), p. 065502. ISSN: 2469-9985. DOI: 10.1103/PhysRevC.105.065502.

[129] J. N. Abdurashitov et al. “Measurement of the response of a Ga solar neutrino experiment to neutrinos from a ^{37}Ar source”. In: *Physical Review C* 73 (4 Apr. 2006), p. 045805. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.73.045805.

[130] J. Kostensalo et al. “The gallium anomaly revisited”. In: *Physics Letters B* 795 (Aug. 2019), pp. 542–547. ISSN: 03702693. DOI: 10.1016/j.physletb.2019.06.057.

[131] C. Giunti et al. “Inspection of the detection cross section dependence of the Gallium Anomaly”. In: *Physics Letters B* 842 (July 2023), p. 137983. ISSN: 03702693. DOI: 10.1016/j.physletb.2023.137983.

[132] D. Frekers et al. “The ^{71}Ga(^{3}He, t) reaction and the low-energy neutrino response”. In: *Physics Letters B* 706 (2-3 Dec. 2011), pp. 134–138. ISSN: 03702693. DOI: 10.1016/j.physletb.2011.10.061.

[133] D. Frekers et al. “Penning-trap Q-value determination of the ^{71}Ga(ν,e)^{71}Ge reaction using threshold charge breeding of on-line produced isotopes”. In: *Physics Letters B* 722 (4-5 May 2013), pp. 233–237. ISSN: 03702693. DOI: 10.1016/j.physletb.2013.04.019.

[134] S. V. Semenov. “Cross Section of Neutrino Absorption by the Gallium-71 Nucleus”. In: *Physics of Atomic Nuclei* 83 (11 Dec. 2020), pp. 1549–1552. ISSN: 1063-7788. DOI: 10.1134/S1063778820100221.

[135] J. N. Abdurashitov et al. “Solar neutrino flux measurements by the Soviet-American gallium experiment (SAGE) for half the 22-year solar cycle”. In: *Journal of Experimental and Theoretical Physics* 95 (2 Aug. 2002), pp. 181–193. ISSN: 1063-7761. DOI: 10.1134/1.1506424.

[136] F. Kaether et al. “Reanalysis of the Gallex solar neutrino flux and source experiments”. In: *Physics Letters B* 685 (1 Feb. 2010), pp. 47–54. ISSN: 03702693. DOI: 10.1016/j.physletb.2010.01.030.

[137] Jeffrey M. Berryman et al. “Statistical significance of the sterile-neutrino hypothesis in the context of reactor and gallium data”. In: *Journal of High Energy Physics* 2022 (2 Feb. 2022), p. 55. ISSN: 1029-8479. DOI: 10.1007/JHEP02(2022)055.

[138] C. Giunti et al. “Gallium Anomaly: critical view from the global picture of ν_e and $\bar{\nu}_e$ disappearance”. In: *Journal of High Energy Physics* 2022 (10 Oct. 2022), p. 164. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2022)164.

[139] Kim Goldhagen et al. "Testing sterile neutrino mixing with present and future solar neutrino data". In: *The European Physical Journal C* 82 (2 Feb. 2022), p. 116. ISSN: 1434-6044. DOI: 10.1140/epjc/s10052-022-10052-2.

[140] C. Athanassopoulos et al. "Evidence for $\bar{\nu}_\mu \to \bar{\nu}_e$ Oscillations from the LSND Experiment at the Los Alamos Meson Physics Facility". In: *Physical Review Letters* 77 (15 Oct. 1996), pp. 3082–3085. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.77.3082.

[141] A. Aguilar et al. "Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam". In: *Physical Review D* 64 (11 Nov. 2001), p. 112007. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.64.112007.

[142] B. Armbruster et al. "Upper limits for neutrino oscillations $\bar{\nu}_\mu \to \bar{\nu}_e$ from muon decay at rest". In: *Physical Review D* 65 (11 June 2002), p. 112001. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.65.112001.

[143] A.A. Aguilar-Arevalo et al. "The MiniBooNE detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 599 (1 Feb. 2009), pp. 28–46. ISSN: 01689002. DOI: 10.1016/j.nima.2008.10.028.

[144] Alexis A. Aguilar-Arevalo. "The MiniBooNE Experiment". In: *eConf* C0406271 (2004). Ed. by Ulrich Heintz, MONP01. arXiv: hep-ex/0408074.

[145] A. A. Aguilar-Arevalo et al. "Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment". In: *Physical Review Letters* 121 (22 Nov. 2018), p. 221801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.121.221801.

[146] A. A. Aguilar-Arevalo et al. "Updated MiniBooNE neutrino oscillation results with increased data and new background studies". In: *Physical Review D* 103 (5 Mar. 2021), p. 052002. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.103.052002.

[147] C. A. Argüelles et al. "MicroBooNE and the ν_e Interpretation of the MiniBooNE Low-Energy Excess". In: *Physical Review Letters* 128 (24 June 2022), p. 241802. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.128.241802.

[148] A. Ioannisian. "A Standard Model explanation for the excess of electron-like events in MiniBooNE". In: *arXiv e-prints* (Sept. 2019). Contribution to the proceedings of the 40th ICHEP2020, Prague, Czech Republic (Virtual Conference), 6 pages. DOI: 10.48550/arXiv.2012.06164.

[149] C. Giunti, A. Ioannisian, and G. Ranucci. "A new analysis of the MiniBooNE low-energy excess". In: *Journal of High Energy Physics* 2020 (11 Nov. 2020), p. 146. ISSN: 1029-8479. DOI: 10.1007/JHEP11(2020)146.

[150] Asli M. Abdullahi et al. "A panorama of new-physics explanations to the MiniBooNE excess". In: *arXiv e-prints* (Aug. 2023). DOI: 10.48550/arXiv.2308.02543.

[151] P. Adamson et al. "Search for Sterile Neutrinos in MINOS and MINOS+ Using a Two-Detector Fit". In: *Physical Review Letters* 122 (9 Mar. 2019), p. 091803. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.122.091803.

[152] P. Adamson et al. "Improved Constraints on Sterile Neutrino Mixing from Disappearance Searches in the MINOS, MINOS+, Daya Bay, and Bugey-3 Experiments". In: *Physical Review Letters* 125 (7 Aug. 2020), p. 071801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.125.071801.

[153] K. Abe et al. "Search for light sterile neutrinos with the T2K far detector Super-Kamiokande at a baseline of 295 km". In: *Physical Review D* 99 (7 Apr. 2019), p. 071103. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.99.071103.

[154] R. Acciarri et al. "Design and construction of the MicroBooNE detector". In: *Journal of Instrumentation* 12 (02 Feb. 2017), P02017–P02017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/02/P02017.

[155] P. Abratenko et al. "Search for an anomalous excess of charged-current quasielastic ν_e interactions with the MicroBooNE experiment using Deep-Learning-based reconstruction". In: *Physical Review D* 105 (11 June 2022), p. 112003. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.105.112003.

[156] P. Abratenko et al. "Search for an anomalous excess of charged-current ν_e interactions without pions in the final state with the MicroBooNE experiment". In: *Physical Review D* 105 (11 June 2022), p. 112004. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.105.112004.

[157] P. Abratenko et al. "Search for an Excess of Electron Neutrino Interactions in MicroBooNE Using Multiple Final-State Topologies". In: *Physical Review Letters* 128 (24 June 2022), p. 241801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.128.241801.

[158] P. Abratenko et al. "Search for an anomalous excess of inclusive charged-current ν_e interactions in the MicroBooNE experiment using Wire-Cell reconstruction". In: *Physical Review D* 105 (11 June 2022), p. 112005. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.105.112005.

[159] P. Abratenko et al. "Search for Neutrino-Induced Neutral-Current Δ Radiative Decay in MicroBooNE and a First Test of the MiniBooNE Low Energy Excess under a Single-Photon Hypothesis". In: *Physical Review Letters* 128 (11 Mar. 2022), p. 111801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.128.111801.

[160] The ALEPH Collaboration et al. “Precision electroweak measurements on the Z resonance”. In: *Physics Reports* 427.5 (2006), pp. 257–454. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2005.12.006. URL: https://www.sciencedirect.com/science/article/pii/S0370157305005119.

[161] S Gariazzo et al. “Light sterile neutrinos”. In: *Journal of Physics G: Nuclear and Particle Physics* 43 (3 Mar. 2015), p. 033001. ISSN: 0954-3899. DOI: 10.1088/0954-3899/43/3/033001.

[162] Carlo Giunti and Thierry Lasserre. “eV-Scale Sterile Neutrinos”. In: *Annual Review of Nuclear and Particle Science* 69 (1 Oct. 2019), pp. 163–190. ISSN: 0163-8998. DOI: 10.1146/annurev-nucl-101918-023755.

[163] A. Diaz et al. “Where are we with light sterile neutrinos?” In: *Physics Reports* 884 (2020). Where are we with light sterile neutrinos?, pp. 1–59. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2020.08.005. URL: https://www.sciencedirect.com/science/article/pii/S0370157320302921.

[164] Sebastian Böser et al. “Status of light sterile neutrino searches”. In: *Progress in Particle and Nuclear Physics* 111 (2020), p. 103736. ISSN: 0146-6410. DOI: https://doi.org/10.1016/j.ppnp.2019.103736. URL: https://www.sciencedirect.com/science/article/pii/S0146641019300717.

[165] J. Ashenfelter et al. “The PROSPECT reactor antineutrino experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 922 (Apr. 2019), pp. 287–309. ISSN: 01689002. DOI: 10.1016/j.nima.2018.12.079.

[166] M. Andriamirado et al. “Improved short-baseline neutrino oscillation search and energy spectrum measurement with the PROSPECT experiment at HFIR”. In: *Physical Review D* 103 (3 Feb. 2021), p. 032001. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.103.032001.

[167] H. Almazán et al. “Improved sterile neutrino constraints from the STEREO experiment with 179 days of reactor-on data”. In: *Physical Review D* 102 (5 Sept. 2020), p. 052002. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.102.052002.

[168] A. P. Serebrov et al. “Search for sterile neutrinos with the Neutrino-4 experiment and measurement results”. In: *Physical Review D* 104 (3 Aug. 2021), p. 032003. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.104.032003.

[169] A. P. Serebrov et al. “Search for sterile neutrinos with the Neutrino-4 experiment and measurement results”. In: *Physical Review D* 104 (3 Aug. 2021), p. 032003. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.104.032003.

[170] H. Almazán et al. "Note on arXiv:2005.05301, 'Preparation of the Neutrino-4 experiment on search for sterile neutrino and the obtained results of measurements'". In: *arXiv e-prints* (June 2020). DOI: 10.48550/arXiv.2006.13147.

[171] C. Giunti et al. "Neutrino-4 anomaly: Oscillations or fluctuations?" In: *Physics Letters B* 816 (May 2021), p. 136214. ISSN: 03702693. DOI: 10.1016/j.physletb.2021.136214.

[172] I. Alekseev et al. "DANSS: Detector of the reactor AntiNeutrino based on Solid Scintillator". In: *Journal of Instrumentation* 11 (11 Nov. 2016), P11011–P11011. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/11/P11011.

[173] Youngju Ko et al. "NEOS Experiment". In: *Journal of Physics: Conference Series* 1216 (Apr. 2019), p. 012004. ISSN: 1742-6588. DOI: 10.1088/1742-6596/1216/1/012004.

[174] I. Alekseev et al. "Search for sterile neutrinos at the DANSS experiment". In: *Physics Letters B* 787 (Dec. 2018), pp. 56–63. ISSN: 03702693. DOI: 10.1016/j.physletb.2018.10.038.

[175] Z. Atif et al. "Search for sterile neutrino oscillations using RENO and NEOS data". In: *Physical Review D* 105 (11 June 2022), p. L111101. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.105.L111101.

[176] Mikhail Danilov. "Review of sterile neutrino searches at very short-baseline reactor experiments". In: *Physica Scripta* 97 (9 Sept. 2022), p. 094001. ISSN: 0031-8949. DOI: 10.1088/1402-4896/ac81fd.

[177] Jeffrey M. Berryman and Patrick Huber. "Sterile neutrinos and the global reactor antineutrino dataset". In: *Journal of High Energy Physics* 2021 (1 Jan. 2021), p. 167. ISSN: 1029-8479. DOI: 10.1007/JHEP01(2021)167.

[178] N. A. Skrobova. "Estimation of the DANSS Experiment Sensitivity to Neutrino Oscillations". In: *Bulletin of the Lebedev Physics Institute* 47 (4 Apr. 2020), pp. 101–104. ISSN: 1068-3356. DOI: 10.3103/S1068335620040077.

[179] M Andriamirado et al. "PROSPECT-II physics opportunities". In: *Journal of Physics G: Nuclear and Particle Physics* 49 (7 July 2022), p. 070501. ISSN: 0954-3899. DOI: 10.1088/1361-6471/ac48a4.

[180] M. Aker et al. "Improved eV-scale sterile-neutrino constraints from the second KATRIN measurement campaign". In: *Physical Review D* 105 (7 Apr. 2022), p. 072004. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.105.072004.

[181] F. P. An et al. “Evolution of the Reactor Antineutrino Flux and Spectrum at Daya Bay”. In: *Physical Review Letters* 118 (25 June 2017), p. 251801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.118.251801.

[182] F. P. An et al. “Improved Measurement of the Evolution of the Reactor Antineutrino Flux and Spectrum at Daya Bay”. In: *Physical Review Letters* 130 (21 May 2023), p. 211801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.130.211801.

[183] C. Buck et al. “Investigating the spectral anomaly with different reactor antineutrino experiments”. In: *Physics Letters B* 765 (Feb. 2017), pp. 159–162. ISSN: 03702693. DOI: 10.1016/j.physletb.2016.11.062.

[184] F. P. An et al. “Antineutrino energy spectrum unfolding based on the Daya Bay measurement and its applications”. In: *Chinese Physics C* 45 (7 July 2021), p. 073001. ISSN: 1674-1137. DOI: 10.1088/1674-1137/abfc38.

[185] F. P. An et al. “Joint Determination of Reactor Antineutrino Spectra from ^{235}U and ^{239}Pu Fission by Daya Bay and PROSPECT”. In: *Physical Review Letters* 128 (8 Feb. 2022), p. 081801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.128.081801.

[186] A. Letourneau et al. “Origin of the Reactor Antineutrino Anomalies in Light of a New Summation Model with Parametrized β^- Transitions”. In: *Physical Review Letters* 130 (2 Jan. 2023), p. 021801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.130.021801.

[187] Lorenzo Perissé et al. “A comprehensive revision of the summation method for the prediction of reactor antineutrino fluxes and spectra”. In: *Phys. Rev. C* 108 (5 Nov. 2023), p. 055501. DOI: 10.1103/PhysRevC.108.055501.

[188] L. Hayen et al. “First-forbidden transitions in reactor antineutrino spectra”. In: *Physical Review C* 99 (3 Mar. 2019), p. 031301. ISSN: 2469-9985. DOI: 10.1103/PhysRevC.99.031301.

[189] Dong-Liang Fang and B. Alex Brown. “Effect of first-forbidden decays on the shape of neutrino spectra”. In: *Physical Review C* 91 (2 Feb. 2015), p. 025503. ISSN: 0556-2813. DOI: 10.1103/PhysRevC.91.025503.

[190] A. A. Sonzogni, E. A. McCutchan, and A. C. Hayes. “Dissecting Reactor Antineutrino Flux Calculations”. In: *Physical Review Letters* 119 (11 Sept. 2017), p. 112501. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.119.112501.

[191] V. Kopeikin, M. Skorokhvatov, and O. Titov. “Reevaluating reactor antineutrino spectra with new measurements of the ratio between ^{235}U and ^{239}Pu β spectra”. In: *Physical Review D* 104 (7 Oct. 2021), p. L071301. ISSN: 2470-0010. DOI: 10.1103/PhysRevD.104.L071301.

[192] Y. Abreu et al. "A novel segmented-scintillator antineutrino detector". In: *Journal of Instrumentation* 12 (04 Apr. 2017), P04024–P04024. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/04/P04024.

[193] Y. Abreu et al. "SoLid: a short baseline reactor neutrino experiment". In: *Journal of Instrumentation* 16 (02 Feb. 2021), P02025–P02025. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/02/P02025.

[194] SCK CEN. *Infrastructure*. URL: https://www.sckcen.be/en/Research/Infrastructure/BR2.

[195] P. Vogel and J. F. Beacom. "Angular distribution of neutron inverse beta decay, $\bar{\nu}_e + p \rightarrow n + e^+$". In: *Physical Review D* 60 (5 July 1999), p. 053003. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.60.053003.

[196] DuPont. *DuPont Tyvek 1082D*. URL: https://www.dupont.cn/content/dam/dupont/amer/us/en/microsites/tyvek-design/images/documents/2019-C&I_Tyvek_1082D_Datasheet.pdf.

[197] Y. Abreu et al. "Optimisation of the scintillation light collection and uniformity for the SoLid experiment". In: *Journal of Instrumentation* 13 (09 Sept. 2018), P09005–P09005. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/09/P09005.

[198] Eljen Technology. *GENERAL PURPOSE PLASTIC SCINTILLATOR EJ-200, EJ-204, EJ-208, EJ-212*. URL: https://eljentechnology.com/images/products/data_sheets/EJ-200_EJ-204_EJ-208_EJ-212.pdf;%20https://eljentechnology.com/products/plastic-scintillators/ej-200-ej-204-ej-208-ej-212.

[199] Varley F. Sears. "Neutron scattering lengths and cross sections". In: *Neutron News* 3 (3 Jan. 1992), pp. 26–37. ISSN: 1044-8632. DOI: 10.1080/10448639208218770.

[200] Scintacor. *Neutron Screens*. URL: https://scintacor.com/wp-content/uploads/2015/09/Datasheet-Neutron-Screens-High-Res.pdf.

[201] Saint-Gobain. *Plastic Scintillating Fibers*. URL: https://www.crystals.saint-gobain.com/sites/hps-mac3-cma-crystals/files/2021-11/Fiber-Product-Sheet.pdf.

[202] F Joppen, E Koonen, and S Van Dijck. "Periodic Safety Review of the BR2 Reactor". In: IAEA, 2012. ISBN: 978-92-0-184610-5. URL: https://www-pub.iaea.org/MTCD/publications/PDF/P1575_CD_web/datasets/papers/D23%20Joppen.pdf.

[203] XILINX. *Artix-7 FPGAs Data Sheet: DC and AC Switching Characteristics*. URL: https://docs.xilinx.com/v/u/en-US/ds181_Artix_7_Data_Sheet.

[204] Y. Abreu et al. “Commissioning and operation of the readout system for the SoLid neutrino detector”. In: *Journal of Instrumentation* 14 (11 Nov. 2019), P11003–P11003. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/11/P11003.

[205] Giel Vandierendonck. “The SoLid neutrino experiment : stabilising, operating, understanding and upgrading the detector”. PhD thesis. Ghent University, 2022, pp. X, 178, VI. URL: http://hdl.handle.net/1854/LU-8744666.

[206] Valentin Pestel. “Détection de neutrinos auprès du réacteur BR2: analyse des premières données de l’expérience SoLid”. PhD thesis. Normandie Université, Oct. 2019. URL: https://theses.hal.science/tel-02469046.

[207] Noe Roy. “Recherche d’oscillation vers un neutrino stérile auprès de l’expérience SoLid située au réacteur BR2: calibration en énergie du détecteur et extraction du signal antineutrino”. PhD thesis. l’Université Paris-Saclay, Dec. 2021. URL: https://theses.hal.science/tel-03548942.

[208] S.G. Mallat and Zhifeng Zhang. “Matching pursuits with time-frequency dictionaries”. In: *IEEE Transactions on Signal Processing* 41 (12 1993), pp. 3397–3415. ISSN: 1053587X. DOI: 10.1109/78.258082.

[209] A. Simón, P. Ferrario, and A. Izmaylov. “Event reconstruction in NEXT using the ML-EM algorithm”. In: *Nuclear and Particle Physics Proceedings* 273-275 (Apr. 2016), pp. 2624–2626. ISSN: 24056014. DOI: 10.1016/j.nuclphysbps.2015.10.010.

[210] S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506 (3 July 2003), pp. 250–303. ISSN: 01689002. DOI: 10.1016/S0168-9002(03)01368-8.

[211] Savitri Gallego. “Analyse des données de l’expérience SoLiD : mesure du spectre antineutrino à BR2 et recherche d’oscillations vers un neutrino stérile”. PhD thesis. Normandie Université, Nov. 2022. URL: https://theses.hal.science/tel-03946790.

[212] David Henaff. “Toward the search of a sterile neutrino with the SoLid experiment at the SCK-CEN BR2 reactor : detector calibration and design of a multivariate topological selection of antineutrino events”. PhD thesis. Ecole nationale supérieure Mines-Télécom Atlantique, June 2021. URL: https://theses.hal.science/tel-03336155.

[213] H. Bethe. “Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie”. In: *Annalen der Physik* 397 (3 1930), pp. 325–400. ISSN: 00033804. DOI: 10.1002/andp.19303970303.

[214] H. Bethe. "Bremsformel fur Elektronen relativistischer Geschwindigkeit". In: *Zeitschrift fur Physik* 76 (5-6 May 1932), pp. 293–299. ISSN: 1434-6001. DOI: 10.1007/BF01342532.

[215] R. M. Sternheimer. "The Density Effect for the Ionization Loss in Various Materials". In: *Physical Review* 88 (4 Nov. 1952), pp. 851–859. ISSN: 0031-899X. DOI: 10.1103/PhysRev.88.851.

[216] R. M. Sternheimer, S. M. Seltzer, and M. J. Berger. "Density effect for the ionization loss of charged particles in various substances". In: *Physical Review B* 26 (11 Dec. 1982), pp. 6067–6076. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.26.6067.

[217] Stephen M. Seltzer and Martin J. Berger. "Bremsstrahlung spectra from electron interactions with screened atomic nuclei and orbital electrons". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 12 (1 Aug. 1985), pp. 95–134. ISSN: 0168583X. DOI: 10.1016/0168-583X(85)90707-4.

[218] M.J. Berger et al. "Stopping-Power & Range Tables for Electrons, Protons, and Helium Ions". In: *ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3). [Online] Available: http://physics.nist.gov/Star [Accessed 08/03/2023]* (2005).

[219] M.J. Berger et al. "XCOM: Photon Cross Section Database (version 1.5)". In: *[Online] Available: http://physics.nist.gov/xcom [Accessed 12/03/2023]. National Institute of Standards and Technology, Gaithersburg, MD.* (2010).

[220] J.H Hubbell and S.M Seltzer. "X-Ray Mass Attenuation Coefficients". In: *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients (version 1.4). [Online] Available: http://physics.nist.gov/xaamdi [year, month day]. National Institute of Standards and Technology, Gaithersburg, MD.* (2004).

[221] Souvik Das. "A simple alternative to the Crystal Ball function". In: *arXiv e-prints* (Mar. 2016). DOI: 10.48550/arXiv.1603.08591.

[222] M. Oreglia et al. "A Study of the Reactions $\psi' \to \gamma\gamma\psi$". In: *Physical Review D* 25 (9 May 1982), pp. 2259–2277. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.25.2259.

[223] Chris Hagmann, David Lange, and Douglas Wright. "Cosmic-ray shower generator (CRY) for Monte Carlo transport codes". In: IEEE, 2007, pp. 1143–1146. ISBN: 978-1-4244-0922-8. DOI: 10.1109/NSSMIC.2007.4437209.

[224] R R Kinsey et al. *The NUDAT/PCNUDAT program for nuclear data.* Report Number: BNL-63351, CONF-961010-5, ON: DE97000167, TRN: 96:028428. Oct. 1996. URL: https://www.osti.gov/biblio/380330.

[225] Simon Vercaemer. "Commissioning of the SoLid experiment for the observation of electron antineutrinos at the BR2 reactor". PhD thesis. University of Antwerp, Nov. 2018. ISBN: 978-94-93079-02-1. URL: https://hdl.handle.net/10067/1571810151162165141.

[226] E. N. Alexeyev et al. "Results of a search for daily and annual variations of the ^{214}Po half-life at the two year observation period". In: *Physics of Particles and Nuclei* 47 (6 Nov. 2016), pp. 986–994. ISSN: 1063-7796. DOI: 10.1134/S1063779616060034.

[227] Leo Breiman et al. *Classification and Regression Trees*. Taylor & Francis, Oct. 1984. ISBN: 9780412048418. DOI: 10.1201/9781315139470.

[228] Steven L. Salzberg. "C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993". In: *Machine Learning* 16 (3 Sept. 1994), pp. 235–240. ISSN: 0885-6125. DOI: 10.1007/BF00993309.

[229] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. 2nd ed. Springer New York, NY, 2009, pp. 353–367. ISBN: 978-0-387-84858-7. DOI: https://doi.org/10.1007/978-0-387-84858-7.

[230] Yann Coadou. "Boosted Decision Trees". In: *Artificial Intelligence for High Energy Physics* (Mar. 2022), pp. 9–58. DOI: 10.1142/9789811234033_0002.

[231] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. URL: http://doi.acm.org/10.1145/2939672.2939785.

[232] J Stevens and M Williams. "uBoost: a boosting method for producing uniform selection efficiencies from multivariate classifiers". In: *Journal of Instrumentation* 8 (12 Dec. 2013), P12013–P12013. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/12/P12013.

[233] L. Koch. *Response Matrix Utilities*. DOI: doi:10.5281/zenodo.1217572. URL: http://github.com/ast0815/remu.

[234] L. Koch. "A response-matrix-centred approach to presenting cross-section measurements". In: *Journal of Instrumentation* 14.09 (Sept. 2019), P09013. DOI: 10.1088/1748-0221/14/09/P09013. URL: https://dx.doi.org/10.1088/1748-0221/14/09/P09013.

[235] H. Almazán et al. "Sterile Neutrino Constraints from the STEREO Experiment with 66 Days of Reactor-On Data". In: *Physical Review Letters* 121 (16 Oct. 2018), p. 161801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.121.161801.

[236] J. Ashenfelter et al. “First Search for Short-Baseline Neutrino Oscillations at HFIR with PROSPECT”. In: *Physical Review Letters* 121 (25 Dec. 2018), p. 251802. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.121.251802.

[237] Gary J. Feldman and Robert D. Cousins. “Unified approach to the classical statistical analysis of small signals”. In: *Physical Review D* 57 (7 Apr. 1998), pp. 3873–3889. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.57.3873.

[238] Louis Lyons. “Raster scan or 2-D approach?” In: *arXiv e-prints* (Apr. 2014). DOI: 10.48550/arXiv.1404.7395.

[239] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *The Annals of Mathematical Statistics* 9 (1 Mar. 1938), pp. 60–62. ISSN: 0003-4851. DOI: 10.1214/aoms/1177732360.

[240] H. Almazán et al. “Joint Measurement of the ^{235}U Antineutrino Spectrum by PROSPECT and STEREO”. In: *Physical Review Letters* 128 (8 Feb. 2022), p. 081802. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.128.081802.

[241] W. Tang et al. “Data Unfolding with Wiener-SVD Method”. In: *Journal of Instrumentation* 12 (10 Oct. 2017), P10002–P10002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/10/P10002.

[242] Rene Brun and Fons Rademakers. “ROOT — An object oriented data analysis framework”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389 (1-2 Apr. 1997), pp. 81–86. ISSN: 01689002. DOI: 10.1016/S0168-9002(97)00048-X.

Appendix A

Positron Energy Estimators

Figs. A.1 and A.2 below show the fractions of positron and annihilation gamma energy that are captured by each energy estimator—as a function of both initial particle energy and deposited particle energy—from which the mean values displayed in Fig. 3.9 of section 3.2.4 are derived. The histograms are built from 232 000 simulated IBD MC events which yield 375 202 and 2 650 408 cube-level deposits from the positron and annihilation gammas respectively. The probability that the positron does not deposit all of its kinetic energy is not negligible and caused by a combination of decay in flight (i.e. leftover positron kinetic energy converted to annihilation gamma energy), deposits in non-active detector materials, and the positron leaving the detector volume.

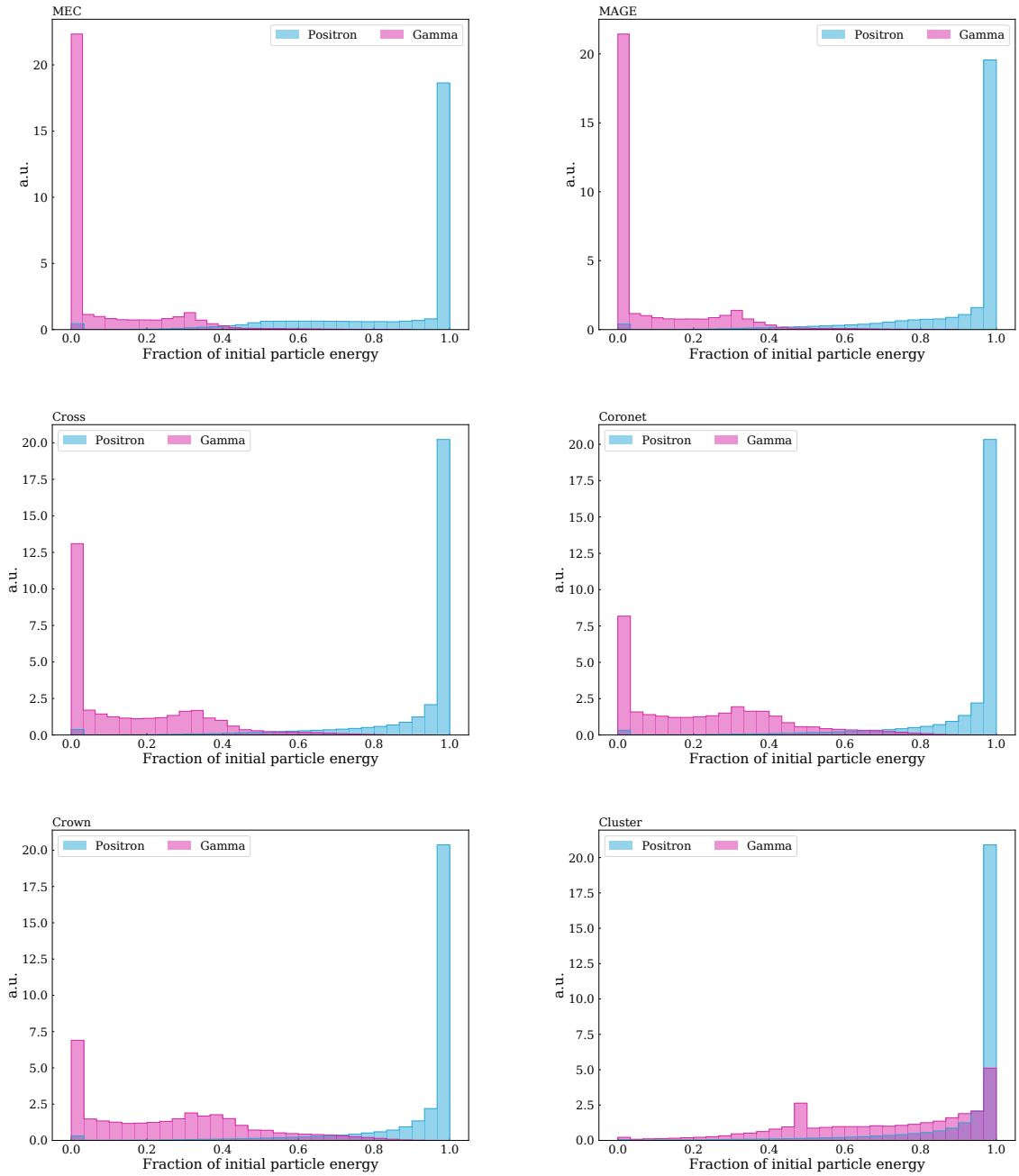


Figure A.1: Proportions of initial particle energy captured by each estimator for the IBD positron (blue) and the combined annihilation gammas (pink). It is hypothesised that the small peak in the Cluster distribution of annihilation gamma deposits (and to a lesser extent for the other estimators) is caused by events that are located at the periphery of the detector where one of the gamma rays escapes without scattering.

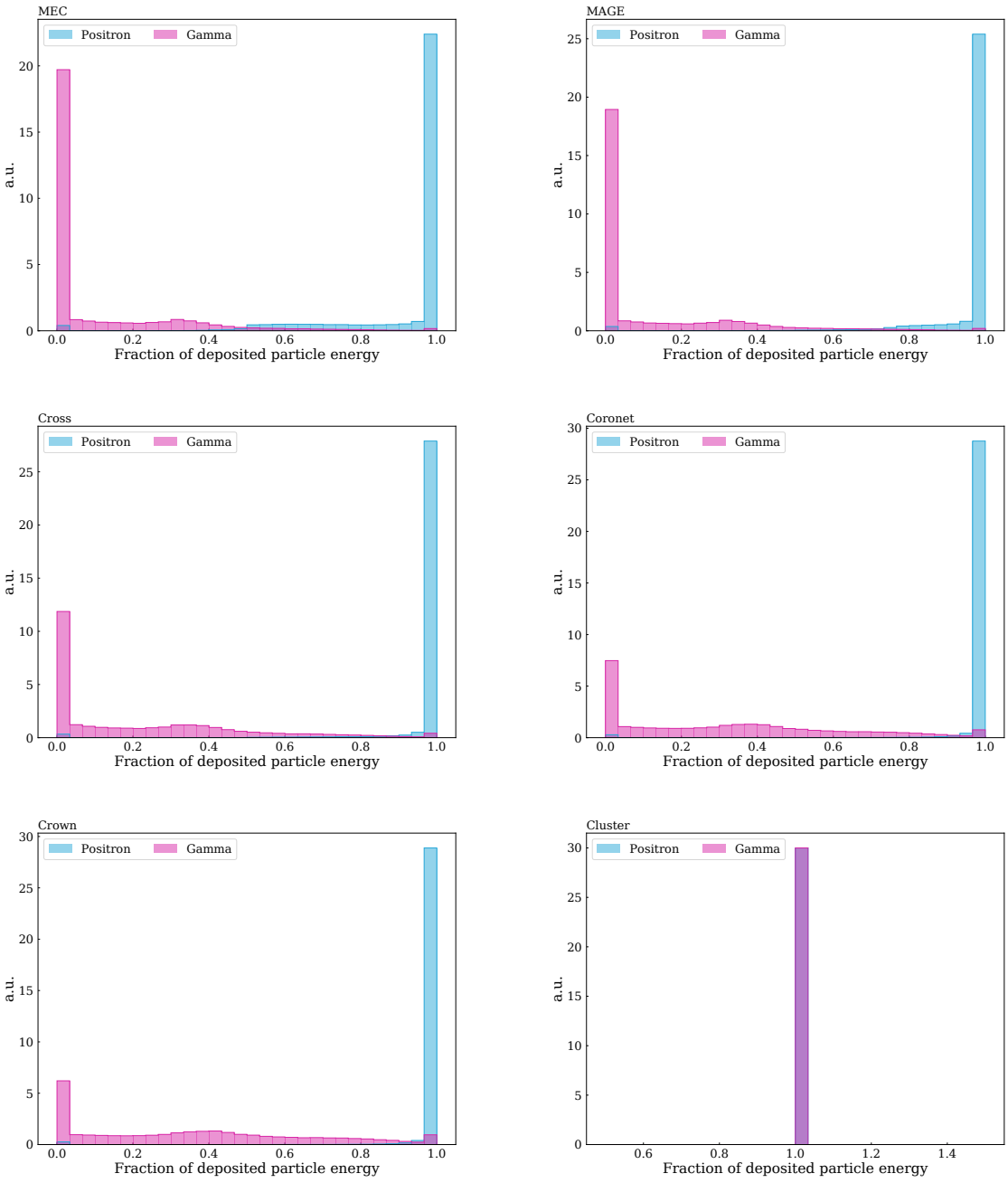


Figure A.2: Proportions of deposited particle energy captured by each estimator for the IBD positron (blue) and the combined annihilation gammas (pink). There are extremely rare instances in which the positron deposits less than 3% of its energy in the detector.

Appendix B

BiPo Data-MC

This appendix contains BiPo data-MC plots that were not previously displayed in section 4.5.2. The variables provided below are:

- Y_{MEC}
- Extent
- ΔX
- EnGammaNotCrown
- promptNCCube
- ES2ES1
- ES3ES2
- ESNS

Notable exclusions are Y_{MEC} and EnLeftoverNotCross. The former is left out because it can only take a single value in this high-purity BiPo region and the latter because it is identically zero for all BiPo events in the sample.

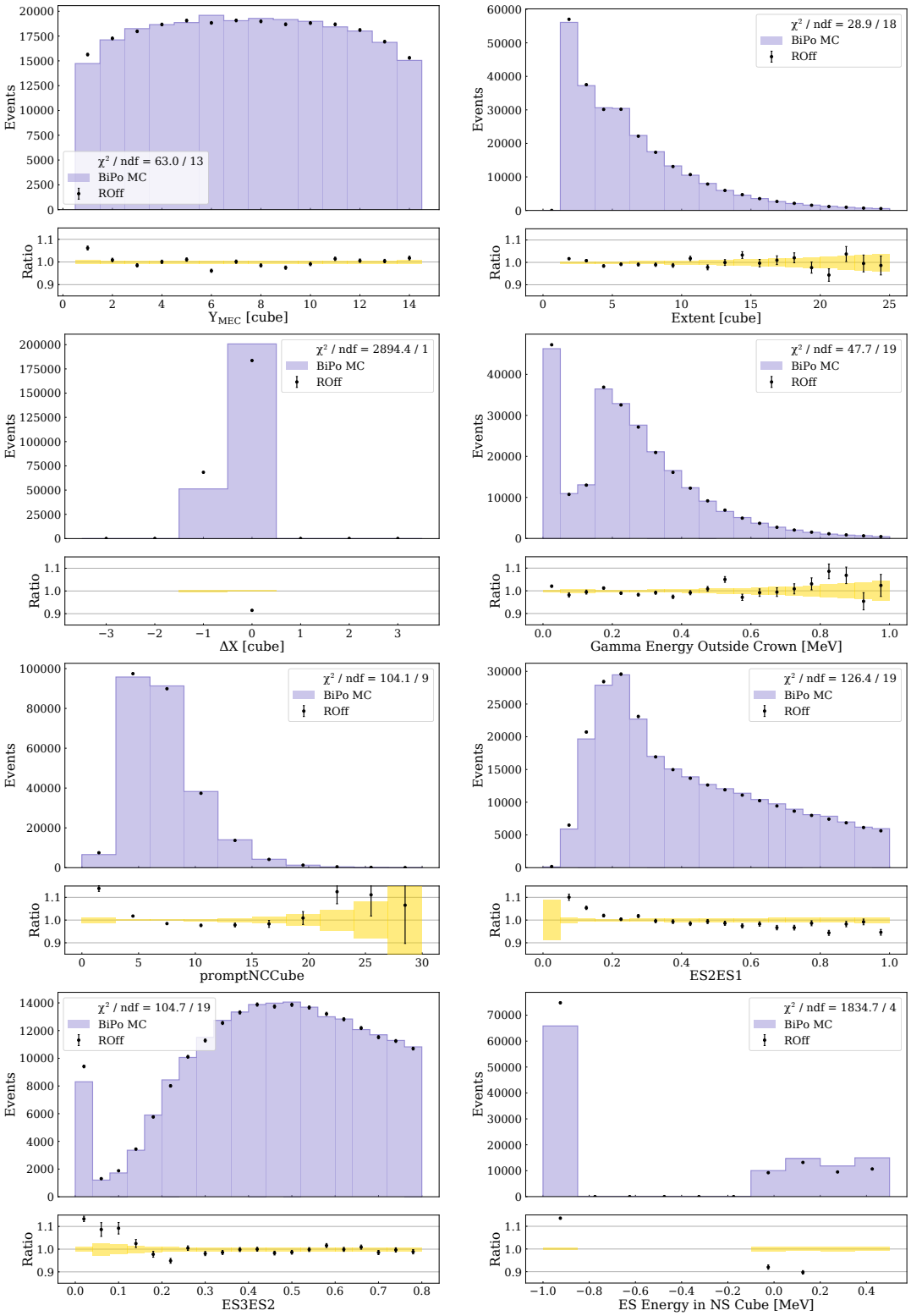


Figure B.1: BiPo data-MC comparison plots in variables not previously displayed. Error bars represent statistical uncertainties of the data. Yellow bands in ratio plots indicate uncertainty on the MC prediction.

Appendix C

Background Corrections

This appendix supports the background subtraction of the full Phase-I dataset in section 5.3. The BiPo-to-Signal factors and summed pressure corrections per period are provided in Tabs. C.1 and C.2. The β values in the final two rows are the global values from Δ fits on the histograms over the aggregated cycles in each "global" period (P_A reactor-off, P_A reactor-on, P_B reactor-off, and P_B reactor-on). The ROn and ROff fast-neutron rates before and after pressure correction for these periods are shown in Fig. C.1. Also, the equivalent BiPo-to-Signal factor fits (which correspond to the final two rows of Tabs. C.1 and C.2) are plotted in Fig. C.2.

Label	β	$\sum f(P, \overline{P})$
ROn1	0.1596(2)	-115 ± 30
ROn2	0.1643(1)	33 ± 26
ROn3	0.1675(1)	-81 ± 18
ROn4	0.1677(1)	160 ± 25
ROn5	0.1687(1)	-46 ± 29
ROn6	0.1740(1)	28 ± 26
ROn7	0.1820(2)	-57 ± 8
ROn8	0.1837(1)	144 ± 22
ROn9	0.1993(1)	404 ± 22
ROn10	0.1938(1)	-377 ± 24
ROn11	0.1914(1)	-321 ± 36
ROn12	0.2005(1)	-242 ± 32
ROn13	0.2009(1)	-155 ± 23
P_A	0.16817(4)	-20 ± 64
P_B	0.1948(4)	-604 ± 67

Table C.1: BiPo-to-Signal factors and summed pressure corrections for every reactor-on period. The double horizontal lines through the middle of the table separate P_A and P_B.

Label	β	$\sum f(P, \overline{P})$
ROff0	0.1862(2)	57 ± 12
ROff1	0.1662(1)	-13 ± 32
ROff2	0.1873(2)	-91 ± 10
ROff3	0.1764(1)	141 ± 21
ROff4	0.1687(1)	-397 ± 24
ROff5	0.1722(1)	150 ± 28
ROff6	0.1811(2)	-24 ± 20
ROff7	0.2360(2)	-51 ± 14
ROff8	0.2308(2)	-152 ± 11
ROff9	0.2048(2)	237 ± 16
ROff10	0.2217(2)	-23 ± 26
ROff11	0.2081(2)	20 ± 25
ROff12	0.1836(1)	-207 ± 24
ROff13	0.1887(2)	-133 ± 18
P_A	0.17290(3)	-178 ± 59
P_B	0.2147(1)	-309 ± 53

Table C.2: BiPo-to-Signal factors and summed pressure corrections for every reactor-off period. The double horizontal lines through the middle of the table separate P_A and P_B.

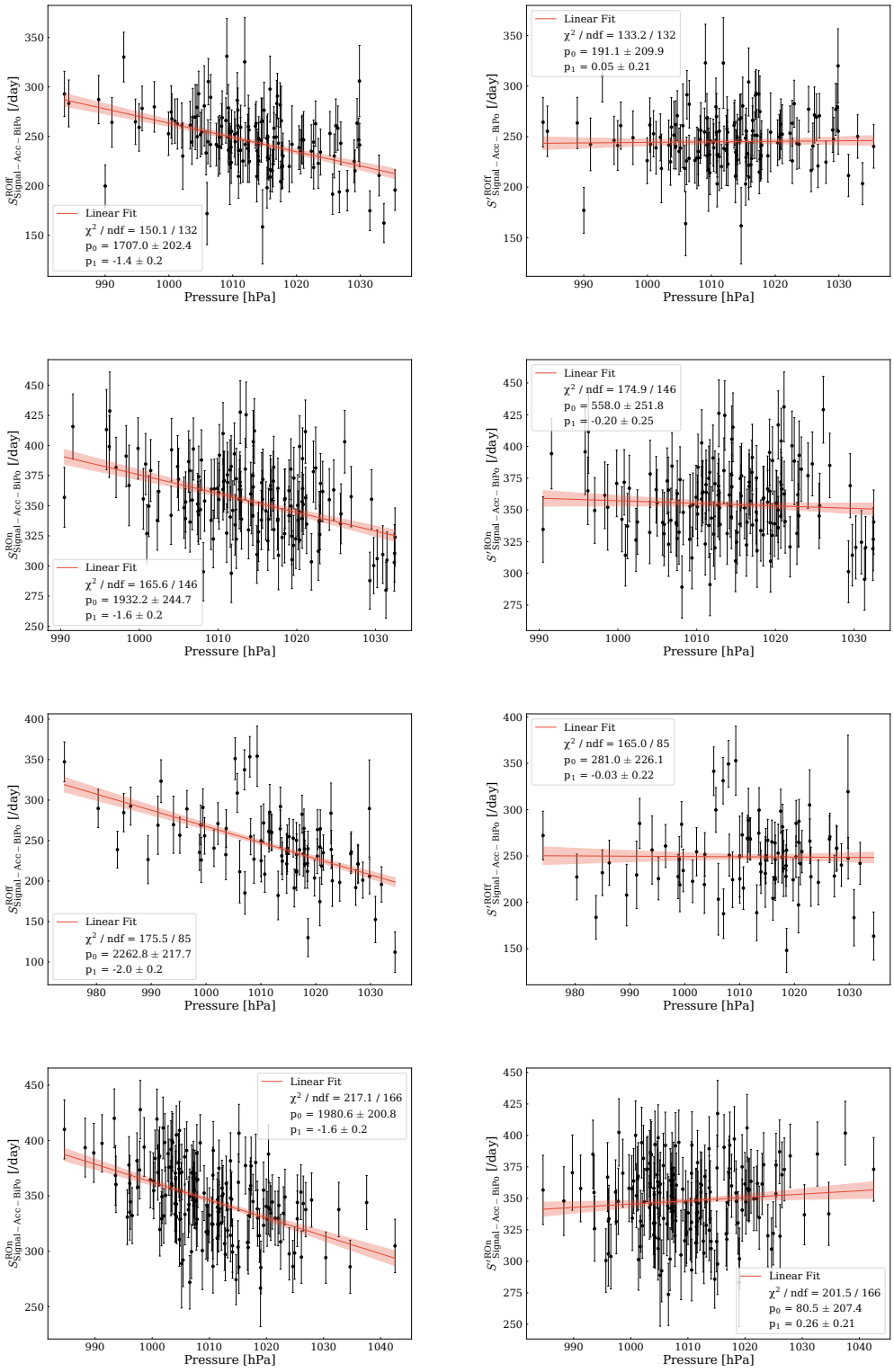


Figure C.1: Effects of the pressure correction on the fast-neutron rates of the Phase-I sub-datasets. Top to bottom: P_A reactor-off, P_A reactor-on, P_B reactor-off, and P_B reactor-on. The linear fits (red) are used to quantify the dependence of fast-neutron rate on pressure before (left) and after (right) applying the correction corrections with equation 5.17.

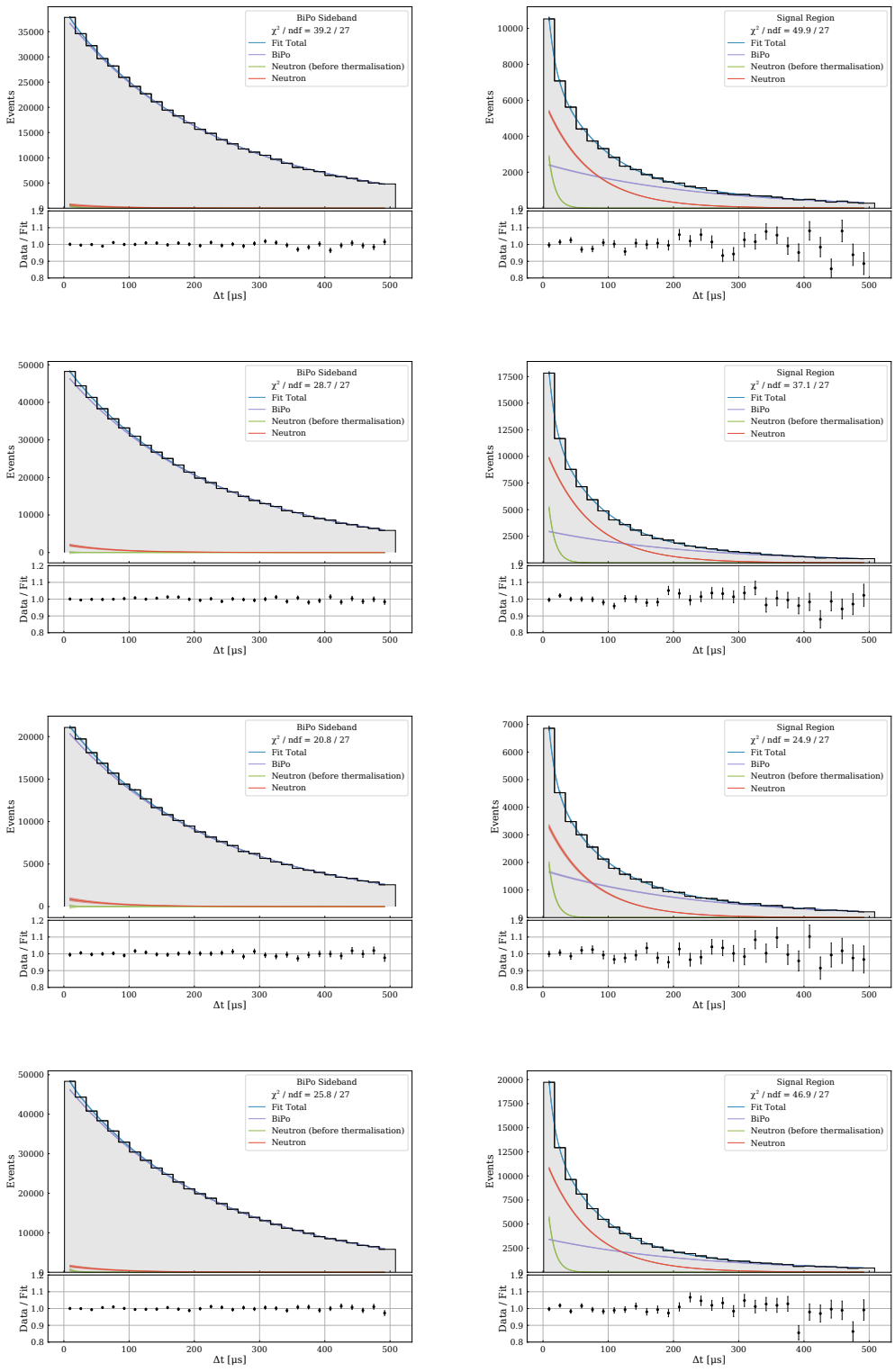


Figure C.2: Fits on the Δt distributions for the determination of the BiPo-to-Signal factors of the Phase-I sub-datasets, as expressed in the bottom two rows of Tabs. C.2 and C.1. Top to bottom: P_A reactor-off, P_A reactor-on, P_B reactor-off, and P_B reactor-on.

Appendix D

Background Characterisation

This appendix consists of the distributions of background events that can be obtained at each stage of the background subtraction procedure (see Tab. 5.1):

- $\boldsymbol{S^{\mathbf{ROn}}_{\mathbf{Signal-Acc}}} \cdot \boldsymbol{\epsilon}$: Accidentals from the FPNT tree, scaled by ϵ.
- $\boldsymbol{S^{\mathbf{ROn}}_{\mathbf{BiPo-Acc}}} \cdot \boldsymbol{\beta}$: Accidental-subtracted distributions of events in the BiPo sideband, scaled by the BiPo-to-Signal factor β.
- $\boldsymbol{S'^{\mathbf{ROff}}_{\mathbf{Signal-Acc-BiPo}}}$: Accidental-and-BiPo-subtracted histograms of fast neutrons in the signal region, corrected for pressure variations and the difference between ROn and ROff exposure time.

The distributions derived for the independent P_A and P_B subtractions are displayed below in three forms. The first is the stacked histograms of ROn (accidental and BiPo) and ROff (pressure-corrected fast neutron) events with overlaid pressure-corrected total ROn signal region distributions ($\boldsymbol{S'^{\mathbf{ROn}}_{\mathbf{Signal}}}$). These are equivalent to Fig. 5.6. The second is overlaid ROn and ROff background histograms—separated by event type—for the purpose of assessing the shape and rate differences between the two. The third form displays the same information as the second, but each background type is given its own canvas and y-axis scale in order to better compare the differences between ROn events and ROff events.

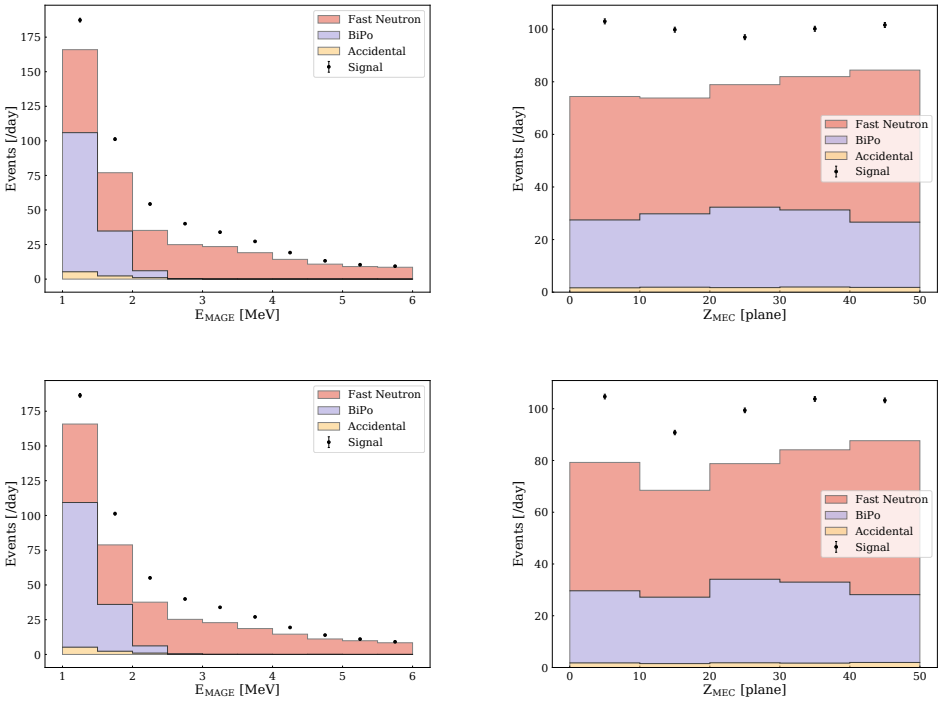


Figure D.1: Stacked histograms of ROn (accidental and BiPo) and ROff (pressure-corrected fast neutron) events with pressure-corrected total ROn signal region distributions ($\boldsymbol{S'^{\mathbf{ROn}}_{\mathbf{Signal}}}$) overlaid. Top row is for P_A and bottom row for P_B.

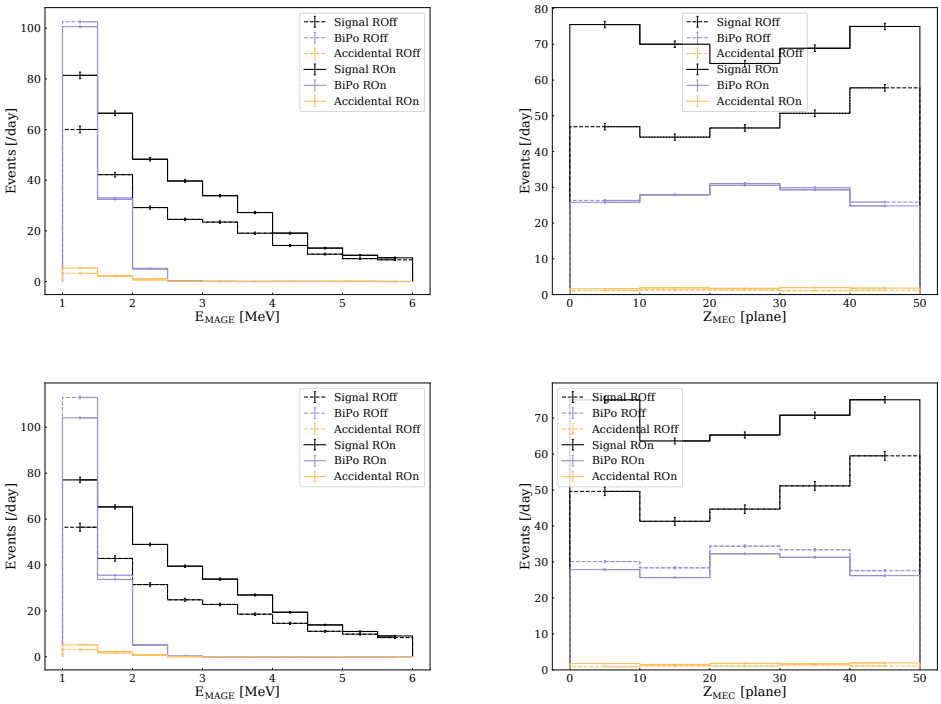


Figure D.2: Normalised histograms of accidental, BiPo, and pressure-corrected fast-neutron events (as defined by the sidebands at each stage in the background-subtraction procedure) on the same axes, separated by reactor status. Top row is for P_A and bottom row for P_B.

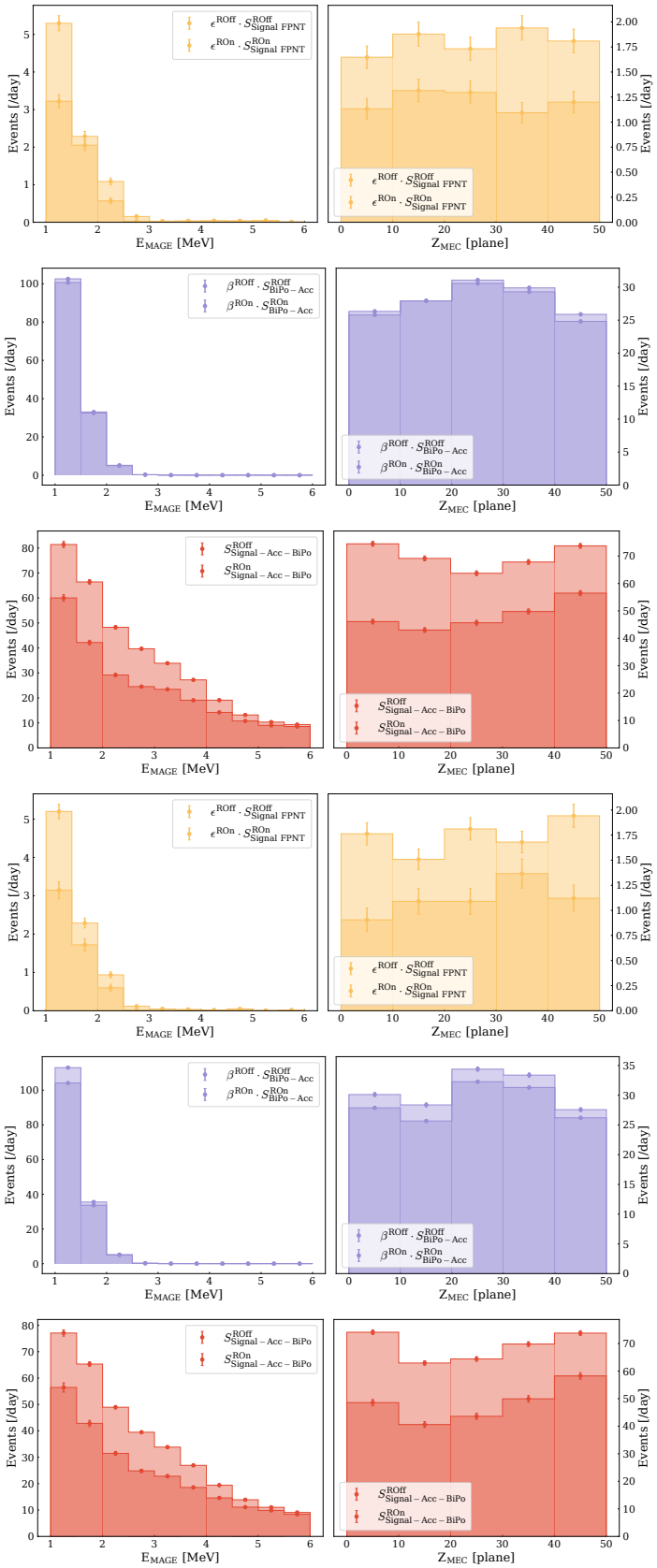
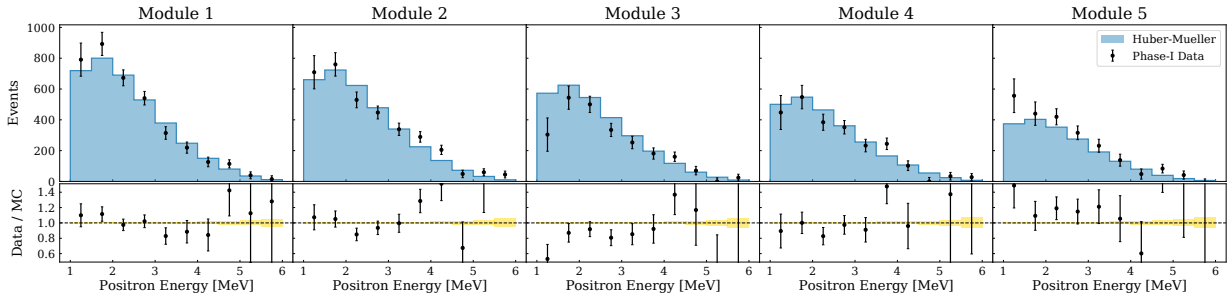


Figure D.3: Normalised histograms of accidental, BiPo, and pressure-corrected fast-neutron events (as defined by the sidebands at each stage in the background-subtraction procedure) on different axes, separated by reactor status. Top three rows are for P_A P_A and bottom three rows for P_B.

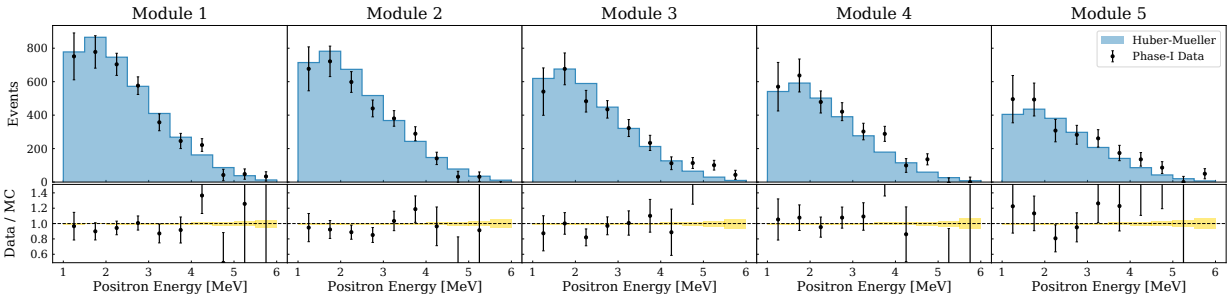
Appendix E

Positron Spectra

Presented below are the separate P_A and P_B spectra that are later combined to give the final Phase-I result. The two-dimensional histograms are projected into the energy spectrum of each module, as per Fig. 5.14. The event deficit observed in the summed spectrum can be almost entirely attributed to the first half of the dataset and the P_B spectrum has better overall statistical agreement with the prediction.



(a) P_A



(b) P_B

Figure E.1: Positron energy spectrum per module for the independent halves of the Phase-I dataset, compared to the HM model prediction. Error bars represent the statistical uncertainty on the points. The yellow band in each ratio plot indicates the uncertainty on the MC prediction."Positron Energy" refers to the MAGE energy estimator.

Appendix F

Single-Event Energy Smearing

The smearing of the energy spectrum was modelled in section 3.4.1 with reasonable success using the GaussExp function. For a given true positron energy (E_{true}), the probability density function of reconstructed energy (E_{reco}) is given by:

$$P(E_{\text{true}}, \mu, \sigma, \alpha, N) = \frac{1}{N} \begin{cases} e^{-\frac{1}{2}(\frac{E_{\text{true}}-\mu}{\sigma})^2} & \text{for } \frac{E_{\text{true}}-\mu}{\sigma} \geq -\alpha \\ e^{\frac{k^2}{2}+k(\frac{E_{\text{true}}-\mu}{\sigma})} & \text{for } \frac{E_{\text{true}}-\mu}{\sigma} < -\alpha \end{cases} \tag{F.1}$$

where μ and σ are the mean and standard deviation of the Gaussian respectively, N is a normalisation factor, and α determines the transition point between the Gaussian and the exponential curve.

The fit parameters as a function of true positron energy are computed by the rolling-window method detailed in section 3.4.2. Running it on post-analysis IBD events yields the results in Fig. F.1. The reduced chi-square (χ^2/ndf) values of the GaussExp fits (lower right panel) are all relatively high—especially in the low-energy region—owing to the presence of 1γ events that are not described very well by equation F.1. Note that the goodness of fits of the 0γ events were found to be even worse, but those events are excluded by the preselection cuts.

The fit parameters are then themselves modelled using some simple functions. The mean (μ), standard deviation (σ) and α are fit with a 1$^{\text{st}}$ order polynomial, equation 3.9, and a 3$^{\text{rd}}$ order geometric series respectively:

$$\mu(E_{\text{true}}) = p_0 + p_1 E_{\text{true}} \,, \tag{F.2}$$

$$\sigma(E_{\text{true}}) = \sqrt{\left(p_0^2 + \left[\frac{p_1}{\sqrt{E_{\text{true}}}} \right]^2 + \left[\frac{p_2}{E_{\text{true}}} \right]^2 \right)} \,, \tag{F.3}$$

$$\alpha(E_{\text{true}}) = p_0 + p_1 E_{\text{true}}^{-1} + p_2 E_{\text{true}}^{-2} + p_3 E_{\text{true}}^{-3} \,. \tag{F.4}$$

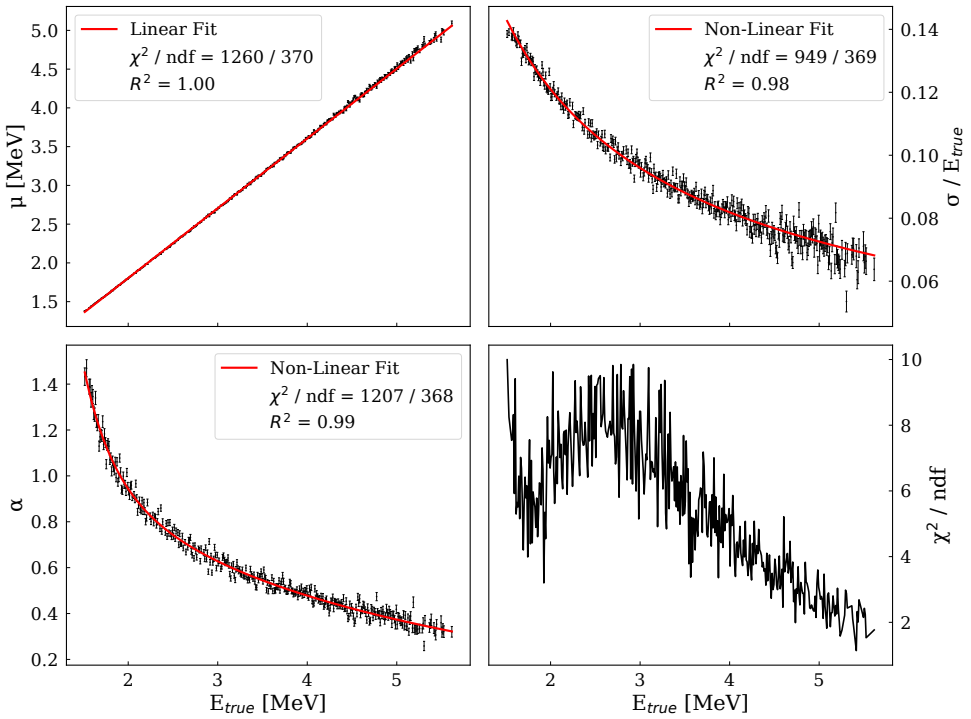


Figure F.1: GaussExp parameters as a function of true positron energy. The red lines denote the "fits of fits" that are used to describe the behaviour of each parameter.

Note that there is no specific physical motivation for the α parameterisation. The goodness of fit for each case is evaluated via both the χ^2/ndf and R^2 (R-squared) metrics. The latter is used as a cross-check that is insensitive to points with vanishingly small uncertainties, as has been observed to arise infrequently from the GaussExp fits. The χ^2/ndf values do not suggest good fits, in contrast to R^2 which are all extremely close to unity.

Regardless, these so-called "fits of fits" can be co-opted to facilitate the quick application of detector smearing effects to a given true-space event, via the following steps:

1. Compute the three GaussExp parameters for the initial positron energy of the event using the "fits of fits".

2. Define a GaussExp probability density function with the parameters obtained in step 1.

3. Sample a random number from the PDF to obtain a value of E_{reco}.

In practise, the random sampling from the GaussExp PDF is achieved using the in-built TF1 GetRandom method of ROOT [242], which employs inverse transform sampling to generate pseudo-random numbers. The reconstructed energy distribution obtained via this method is plotted in Fig. F.2. It exhibits a clear distortion with respect to the full simulation. The distortion is more prominent at high energy where it reaches the 20% level. In its current implementation the method does not parameterise the overall detector efficiency, nor the efficiency per energy bin.

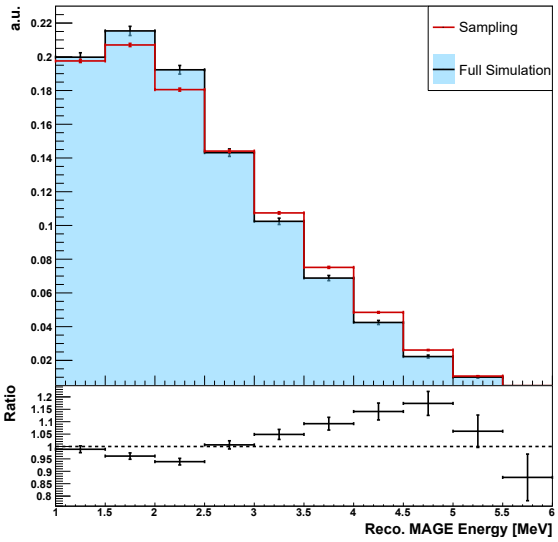


Figure F.2: Reconstructed positron spectra from the full simulation and the single-event sampling method. The histograms are scaled to the same number of events. Error bars represent the Poisson error on each bin. A significant deviation between the two can be observed in the ratio plot, particularly at high energy.

It can be concluded that this method is presently too simplistic to accurately capture the energy smearing effects of the detector. The sampling also fails to account for position smearing as events are assumed to be reconstructed in the same voxel in which they were created. Finally, the method is also sensitive to the chosen forms of equations F.2–F.4 and sampling a random number for every event quickly becomes impractical for a large number of events. The computing times scales as $\mathcal{O}(n)$ for n true events. Hence, compared to the response matrix method, single-event smearing is less accurate and orders of magnitude slower. On the other hand, it offers greater flexibility for varying individual parameters such as the energy resolution, which would be useful for studying their effects on the reconstructed spectrum.

Appendix G

Wilk's Theorem vs Pseudo-Experiments

The 90% CL sensitivity and exclusion contours drawn in Fig. G.1 illustrate the difference in coverage between a contour constructed using $\Delta\chi^2_{\text{crit}}(\sin^2(2\theta_{14}),\ \Delta m^2_{41})$ calculated from pseudo-experiments and one made by assuming the critical values indicated by Wilk's theorem. The exclusion contours are created from the experimental Phase-I IBD spectrum. The Wilk's theorem exclusion contour rejects a significantly larger area of the parameter space, especially in the vicinity of the data best-fit point. The effect on the sensitivity contours is not as large.

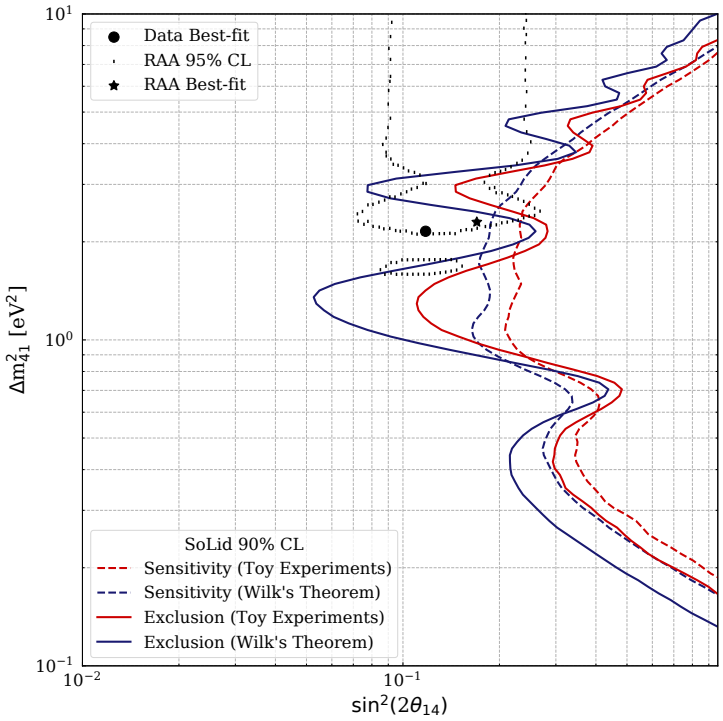


Figure G.1: SoLid Phase-I exclusion and sensitivity contours at 90% CL from Wilk's theorem (blue) and the toy-generation method (red). Wilk's theorem causes the coverage to be overestimated.

Appendix H

rBDT Training

The rBDT is identical to the nuBDT in terms of hyperparameters and training events, but differs by six variables on which the nuBDT is not trained. The rBDT is checked for overfitting via stratified k-fold cross-validation, as described in section 4.4.4. No evidence of overfitting is observed for the rBDT model.

Fold	Train AUC	Test AUC	Train Log. Loss	Test Log. Loss
1	0.920	0.920	0.357	0.358
2	0.920	0.918	0.357	0.361
3	0.919	0.921	0.358	0.356
4	0.920	0.919	0.358	0.359
5	0.920	0.919	0.356	0.357
6	0.920	0.918	0.357	0.361
7	0.920	0.919	0.357	0.360
8	0.920	0.921	0.357	0.355
9	0.920	0.921	0.358	0.357
10	0.920	0.917	0.357	0.363
Orig. Mod.	**0.920**	**n/a**	**0.358**	**n/a**
Sample Ave.	**0.920**	**0.919**	**0.357**	**0.359**
Sample Std.	**0.0005**	**0.001**	**0.0006**	**0.002**

Table H.1: Overfitting check of the rBDT model in terms of AUC and LogLoss. "Ave." refers to the arithmetic mean and "Std." to the standard deviation. The mean test score is congruent with the mean train score.

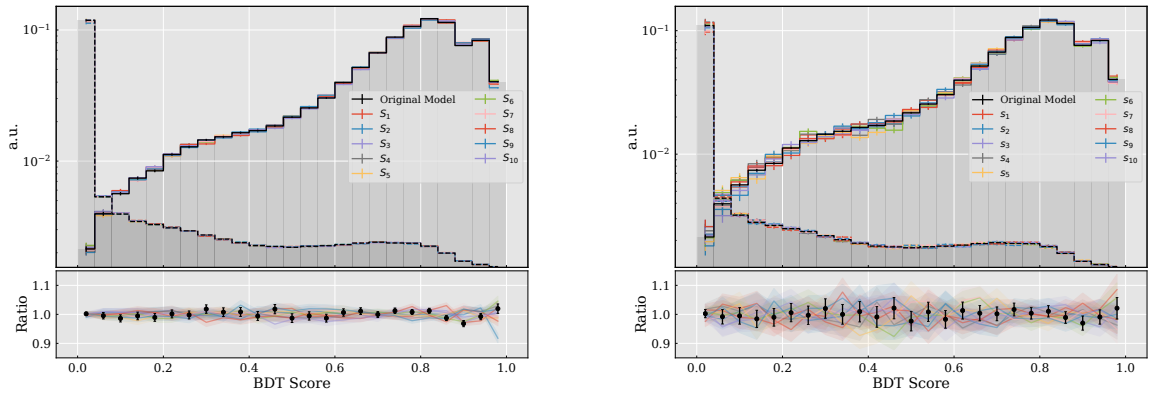


Figure H.1: Overfitting check of the rBDT model in terms of predicted score. Distributions are split into signal (solid lines) and background (dashed lines) for each of the ten models. The ratio plots are computed by dividing each distribution by the mean of the sample outputs for the training events (left) and the test events (right). The ratio for the original model lies within the band defined by the standard deviations of the samples. Error bars represent the Poisson statistical uncertainty on the points.