MEAUREMENT OF THE LOW ENERGY ^{8}B SOLAR NEUTRINO FLUX AT SNO+ USING A DATA DRIVEN BACKGROUNDS MODEL

Samuel Naugle

A DISSERTATION

in

Physics and Astronomy

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2026

Supervisor of Dissertation

Joshua R. Klein, Edmund J. and Louise W. Kahn Term Professor, Physics and Astronomy

Graduate Group Chairperson

Cullen Blake, Associate Professor of Physics and Astronomy,

Dissertation Committee

Christopher Mauger, Associate Professor of Physics and Astronomy
Dylan Rankin, Assistant Professor of Physics and Astronomy
Jonathan Heckman, Professor of Physics and Astronomy
James Aguirre, Associate Professor of Physics and Astronomy

MEAUREMENT OF THE LOW ENERGY ^{8}B SOLAR NEUTRINO FLUX AT SNO+ USING A DATA DRIVEN BACKGROUNDS MODEL

COPYRIGHT

2026

Samuel Naugle

For my family.

ACKNOWLEDGMENT

While you are presumably reading this first, this is the last section I will write. Trying to enumerate all the people that have supported me, educated me, agreed with me, disagreed with me, is of course an impossible task. But I will try.

Let me begin with the always present preemptive apology for all the people I do not mention in this section. I do truly believe that every interaction, regardless of the depth or duration, leaves some sort of lasting mark on our person. These marks build up, like the indentations in the marble staircases of the row-homes of Philadelphia; caused by individual steps whose effects are unnoticeable, but when viewed in totality create something perceptible. These marks have inevitably worked their way into my thesis, perhaps in a specific choice of word in some section or in the implementation details of my analysis cuts. Regardless, this is all to say that what I have created here in this thesis is a literal reflection of everyone I have ever interacted with, and so I want to thank those that I have been fortunate to cross paths with — regardless of how insignificant the interaction may have appeared.

Next, I want to thank my adviser, Josh Klein. Josh has a gift for knowing the correct way to advise a student given their current needs, which will inevitably change over the course of a PhD. Josh provided the space to grow in the moments where fumbling through the depths of confusion would result in a deeper understanding of our research, while also knowing when a gentle nudge in the correct direction was required. Josh also has a defiant optimism coupled with an unwavering grit, which would temper my mettle in moments of doubt or overwhelm. Josh also approached our work with a refreshing candor wherein he gave each person's thoughts the same weight coupled with honest feedback, which I have both appreciated as a student and tried to adopt as a mentor. I could not have asked for a better adviser.

I want to also thank the people who have advised and taught me, without the title of adviser. Tereza Kroupova taught me the language of statistics in high energy particle physics, which involved many discussions of likelihoods, PDFs, and confidence intervals. Tereza also taught me how to do

reconstruction in large scintillator detectors, which is now one of favorite parts of my research. I will always appreciate Tereza's gentle teaching, which made me feel comfortable to ask her any question, regardless of how trivial it may be. Tony Zummo shared with me both his deep knowledge of the SNO+ electronics, as well as his knowledge regarding the actual "doing" of a SNO+ analysis. Tony, like Tereza, also has a calm and kind demeanor, making me feel immediately comfortable in our group. Logan Lebanowski also provided useful feedback on multiple occasions on all kinds of topics. When Logan would reply with "that's a good idea" to something I mentioned, I knew that it was indeed a good idea. I have also had many great discussions with Ziping Ye, who made me realize that using the direction of the outgoing electron in solar neutrino events could allow us to more precisely determine the energy of the incoming neutrino. Gene Beier provides a wealth of knowledge in our group meetings, and pointed out to me that the kinematics of neutrinoless double beta decay are highly sensitive to the short-range physics involved in the process. I also want to thank the other PI in the neutrino group at Penn, Christopher Mauger, who was always was willing to patienly explain "big picture" ideas in neutrino physics, like detector design, funding politics, and the physics involved in experiments outside of SNO+.

Everyone in the Penn neutrino group made it a delightful place in which to learn and do physics. It was always great to make the journey to the other side of the office and say hi to Jon Sensenig and Eric Chong, and hearing about their analyses which were much different than mine. I want to apologize for Amanda Bacon for never knowing what was going on with anything related to scheduling or administration, and thank her for always being willing answer my questions about site approvals, reimbursements, etc. Amanda and I also learned about the SNO+ electronics together, and having her by my side made the learning process both easier and more fun. Ben Harris and James Shen provided me with my first opportunity at some small form of mentorship, which is something I will always remember and reflect upon. Working with Ben was always a joy, and his willingness to ask questions undoubtedly sharpened my understanding of our research. The intensity with which James approached our work is also something I have come to admire and appreciate, and I must mention that James reignited my interest in computing and coding from an aesthetic perspective. I cannot mention Ravi Pitelka without also mentioning his kindness

— I might even say that he is too kind as he agreed to being both the SNO+ electronics and reconstruction czar, two terrifying positions to hold. While this is someone with whom I did not spend much time working with directly, I want to also thank Ben Land for writing much of the code I used in my first few years of grad school. This eventually provided me with the inspiration for writing my own analysis tools, namely the ones used in the analysis presented here.

I also received indispensable advice from many of my SNO+ collaborators. First, I am convinced that the SNO+ experiment would not be able to function without the institutional knowledge provided by Valentina Lozza. Valentina knows the number of atoms of every isotope in the SNO+ detector, how they change over time, where they come from, and where they are going (well…with one notable exception). She also is always willing to not only hear, but also consider, all my crazy explanations for the "low energy excess" that we both have spent far too long trying to understand. I have also recieved helpful feedback from many other SNO+ collaborators, including: Tanner Kaptonoglu, José Maneira, Ed Callaghan, Daniel Cookman, Ana Sofia Ignacio, Serena Riccetto, Will Parker, and many others. Matt Depatie and Mark Ward also saved me many times during SNO+ shifts due to their encyclopedic knowledge of the SNO+ electronics.

I must also thank Gabriel Orebi Gann, my undergraduate adviser and where my journey in physics research began. I am sure my present situation would look quite different without Gabriel taking a chance on me. Gabriel has a knack for finding solutions regardless of the present situation, something that I admire and try to emulate. Now that I have returned to her group as a post-doc I am hopeful that some of her talent will rub off on me. Another important piece in the beginning of my work as a researcher includes Mich Sakai, who patiently answered every single question I asked as a budding researcher. I recall him spending days getting me up to speed when I first started working with him. Now that I am a post-doc myself I understand how lucky I was to receive that kind of detailed mentorship.

I also want to mention and thank those who have supported me through their friendship. I want to thank Steven Gassner for always being willing to discuss a myriad of different subjects, including physics, mathematics, and video games. I have also enjoyed many fun days and nights with Andie

and Mitch, Lucas and Allie, and Nat and Vernon. I am still in disbelief regarding the chance meeting of lifelong friends through UPenn's club water polo team in Lenny, Declan and Aidan, and Juan. And many friends have supported me by visiting the East coast while I am far from where I grew up, including Tyler, Ari, Cole, Jack, and Davis.

Finally, I have to thank the ones that have undoubtedly had the largest impact on the person I am today, my family. My Dad for being the one who first instilled in me a curiosity about the natural world. My Mom for always believing in me. Sarah for always supporting and fighting for me no matter the circumstances. Cameron for being my best friend and setting an example for me. And Chloe for knowing, understanding, and seeing me in way that only a sibling could. And, of course, I am the person I am today because of my fiance, Jackie. Jackie has been with me through every step of this process, from editing my first email to Gabriel to running to find champagne before my defense. Jackie, I cannot even begin to understand the full weight of what you have done for me over the last 9 years. Having a safe, cozy, loving home to return to everyday has allowed me to flourish during graduate school. While I have grown considerably professionally, the personal growth that you have nurtured in me is infinitely more profound and substantial than any of my professional accomplishments. We have grown into adults together, and I love the adults that we have become.

Now we begin....

ABSTRACT

MEAUREMENT OF THE LOW ENERGY ^{8}B SOLAR NEUTRINO FLUX AT SNO+ USING A DATA DRIVEN BACKGROUNDS MODEL

Samuel Naugle

Joshua R. Klein

This dissertation presents results measuring the ^{8}B solar neutrino rate in two separate energy regions, a low energy region (LER) and a high energy region (HER). The low energy region spans from 2.4 to 3.25 MeV and the high energy region ranges from 3.25 to 7.0 MeV. The LER probes populations of ^{8}B solar neutrinos not previously measured.

Likelihood fits using an empirical backgrounds model are used to measure the rates in each of the regions defined above. The fits are done to a 237.4 day dataset taken during the SNO+ basic scintillator phase. The rates are converted to ν_e survival probabilities using the ^{8}B flux calculated by the standard solar model (B16-GS98). The survival probability of the LER is $P_{ee}(\bar{E}_\nu = 7.11 \text{ MeV}) = 0.67^{+0.26}_{-0.22}$ and for the HER we measure $P_{ee}(\bar{E}_\nu = 8.16 \text{ MeV}) = 0.30^{+0.07}_{-0.06}$. The expected values in each region calculated using the MSW-LMA solution are 0.36 and 0.35 in the LER and HER, respectively. Large uncertainties are present in the LER due to low statistics, but there is a slight 1.4σ tension between the result and the MSW prediction. We observe good agreement between the MSW solution and the survival probability in the HER.

Additionally, we make predictions of the sensitivities of the SNO+ detector to the $0\nu\beta\beta$ process during the Te-phase. The predicted half life limits are 1.7×10^{26} yrs for 3 years of livetime at 0.5%, and 4.3×10^{26} yrs for 5 years of livetime at 1.5% loading. The resulting $m_{\beta\beta}$ limits are 35 - 164 meV for the 0.5% dataset, and 20 - 95 meV for the 1.5% dataset. At 1.5% loading SNO+ can set a world leading result in under 5 years of livetime.

TABLE OF CONTENTS

ACKNOWLEDGMENT iv

ABSTRACT viii

LIST OF TABLES xiii

LIST OF FIGURES xiv

CHAPTER 1 : Introduction 1

CHAPTER 2 : Neutrino Physics, What We Know and How We Know It 3

2.1 First detection 3

2.2 Solar neutrinos and the solar neutrino problem 4

2.3 Theoretical descriptions of neutrinos 9

2.3.1 Neutrinos in the Standard Model 9

2.3.2 Massive neutrinos 11

2.3.3 Neutrino oscillations in vacuum 13

2.3.4 Neutrino oscillations in matter 15

2.3.5 MSW effect for solar neutrinos 17

2.3.6 Neutrinoless double beta decay 21

2.4 Neutrino experiments 23

2.4.1 Solar neutrinos 23

2.4.2 Neutrinoless double beta decay 26

CHAPTER 3 : The SNO+ Experiment 30

3.1 Neutrino observatories, in general 30

3.2 The SNO+ detector 31

3.2.1 Chemical processing plants 34

3.3 Liquid scintillators 35

3.3.1 Quenching . . . 37
3.3.2 Te loading . . . 38
3.3.3 Cherenkov radiation in liquid scintillator detectors . . . 38
3.4 Phases of the SNO+ experiment . . . 39
3.4.1 Water phase . . . 39
3.4.2 Scintillator phase . . . 40
3.4.3 Tellurium phase . . . 42
3.5 SNO+ Electronics . . . 43
3.5.1 Deadtime and re-triggering . . . 48

CHAPTER 4 : Simulation and Reconstruction in SNO+ . . . 51
4.1 Monte Carlo simulation of the SNO+ detector . . . 51
4.1.1 Run by run productions . . . 53
4.2 Reconstruction . . . 53
4.2.1 Position and time reconstruction . . . 54
4.2.2 Energy fitting . . . 59
4.2.3 Classifiers . . . 62
4.3 Calibrations and corrections . . . 64
4.3.1 ^{214}Bi^{214}Po coincidence tagging . . . 64
4.3.2 Scintillator timing calibrations . . . 66
4.3.3 Energy and ITR corrections . . . 67
4.3.4 Acrylic vessel offset correction . . . 69

CHAPTER 5 : Processing and Fitting of the Basic Scintillator Phase Dataset . . . 71
5.1 Backgrounds in liquid scintillator detectors . . . 72
5.1.1 Externals . . . 75
5.1.2 Background classes . . . 76
5.2 Cuts and coincidence tagging . . . 78
5.2.1 Event level cuts . . . 79

5.2.2 General coincidence tag . . . 80
5.2.3 Atmospheric $\bar{\nu}$ tag . . . 80
5.2.4 Hot-spot cuts . . . 81
5.2.5 Dataset livetime and cut efficiencies . . . 82
5.2.6 Data tagging and processing software . . . 82
5.3 The likelihood function . . . 84
5.3.1 Deriving the likelihood function . . . 85
5.3.2 Priors and constraints . . . 86
5.3.3 Systematic parameters and constraints . . . 87
5.3.4 Background rate constraints . . . 90
5.4 Fitter . . . 92
5.4.1 Why Zig? . . . 93
5.4.2 snopfit dependencies . . . 95
5.4.3 Application of systematics . . . 96
5.4.4 Runtime fitter configuration . . . 100
5.4.5 Tracking cut efficiencies . . . 101

CHAPTER 6 : Constructing a Backgrounds Model . . . 102
6.1 The backgrounds model . . . 102
6.2 Low energy excess investigations . . . 106
6.2.1 General coincidence tagging studies . . . 106
6.2.2 Instrumentals and mis-reconstructions . . . 107
6.2.3 Neutron captures . . . 108
6.2.4 Contamination from external events . . . 112
6.2.5 Rare interaction or process . . . 115
6.2.6 Rare isotopes . . . 116
6.2.7 ^{102}Rh and ^{102m}Rh as potential explanations of the low energy excess . . . 118
6.3 A generic backgrounds model . . . 120
6.3.1 Bayesian model comparison . . . 123

6.3.2 Nested sampling 124
6.3.3 PDF generation and interpolation 126
6.3.4 Fits including empirical backgrounds 133

CHAPTER 7 : A Measurement of the Low Energy ^{8}B Neutrino Survival Probability 136
7.1 Defining the region of interest 137
7.2 Model comparisons and marginalized results 139

CHAPTER 8 : Tellurium Phase Fits and Sensitivity Projections 145
8.1 Expected background rates 145
8.2 Calibrations 146
8.3 Coincidence and classifier cuts 147
8.4 Backgrounds in the region of interest 148
8.5 Half life and $m_{\beta\beta}$ limits 150

CHAPTER 9 : Conclusions 153

APPENDIX A : The Likelihood Function and Bayesian Inference 155

APPENDIX B : Nested Sampling 159

APPENDIX C : Measure Theory, Optimal Transport, and PDF Interpolation 169

APPENDIX D : Example YAML snopfit configuration file 178

BIBLIOGRAPHY 181

LIST OF TABLES

TABLE 4.1 *Top:* The cuts applied to the ^{214}Bi^{214}Po pairs. The details of each of these cuts will be discussed in a later section. *Bottom:* The run ranges used for calibration and testing applied to the ^{214}Bi^{214}Po pairs 68

TABLE 5.1 The ^{238}U and ^{232}Th decay chain isotopes, decay modes, half-lives, and Q-values for decays relevant to this work. 74
TABLE 5.2 Typical cuts used to select events in the scintillator phases. 79
TABLE 5.3 Coincidence criteria for the general coincidence tag. 80
TABLE 5.4 Coincidence criteria to remove followers of atmospheric $\bar{\nu}$ events. 81
TABLE 5.5 Coincidence criteria to remove followers after flashes from the hotspots. 81
TABLE 5.6 Cut efficiencies for the various signals expected in the SNO+ dataset, calculated using the MC. It can be seen that externals and ^{214}Bi and ^{212}Bi events are effectively removed by the cuts relative to signals like the ^{8}B solar neutrinos. . 82
TABLE 5.7 Full volume pre-cut rates per year for the expected backgrounds in the basic scintillator phase. Columns where the uncertainties are left blank signifies that the uncertainty used in the fit is the Poisson error of the predicted rate after applying all cuts and a livetime scaling, or in other words the error on the rate shown here is small relative to the post-cuts Poisson error. For columns with uncertainties, the error is propagated into the uncertainty on the background normalization fit parameter. ^{234m}Pa has an infinite error since the rate might not be in equilibrium with ^{214}Bi tagging results because of long lived isotopes seperating the bottom and top of the ^{238}U chain. 92

TABLE 6.1 Comparison of systematic priors and their optimal values from the likelihood fit shown in Fig. 6.1. Large tension between the prior constraints and the fitted values can be seen, summarized in the last column. Note that the constraint values listed in the table here differ from those discussed in the previous chapter since this fit uses prior constraints obtained from ^{214}Bi events in the 3.3 m FV. With that said, the differences between the priors in the two FVs are small. . . 103

TABLE 7.1 The log of the evidence calculated for each of the models considered in this work. Note that the models lacking a β/γ background had to be ran with the systematic constraint on the energy smear increased by a factor of 5 and the constraint on the ITR shift increased by a factor of 100 for the slice sampling algorithm in those dimensions to not break down. This breaking down of the algorithm was due to the fitted value for those systematics being so far from MLE, leading to numerical instability since the probability of being that far away from the prior mean becomes vanishingly small. 140

TABLE 8.1 Expected backgrounds rates in the 0.5% (*left*) and 1.5% (*right*) phases. 146

LIST OF ILLUSTRATIONS

FIGURE 2.1 The fluxes of neutrinos from the various neutrino emitting processes in the Sun. Taken from [8] 5

FIGURE 2.2 The angle between the direction of events in the Kamiokande-II detector and the direction to the Sun. The dashed line represent the expected signal based on the SSM, and the dashed histogram represents the fitted flux from the signal. From [11]. 7

FIGURE 2.3 The measured fluxes of ν_e and $\nu_{\mu/\tau}$ neutrinos in SNO+ alongside the measured ν_e flux by Super-K and the prediction by the standard solar model (SSN). It can be seen that SNO can place limits on the $\nu_{\mu/\tau}$ flux by comparing the rates in the different interaction channels, while Super-K has no sensitivity to the $\nu_{\mu/\tau}$ flux. Taken from [14]. 8

FIGURE 2.4 *Left:* Charged-current vertex. *Right:* Neutral-current vertex. 10

FIGURE 2.5 This caption is taken directly from the reference [14], Figure 14.3. "Electron neutrino survival probability as a function of neutrino energy. The points represent, from left to right, the Borexino pp, 7Be, pep, and 8B data (red points) and the SNO+SK 8B data (black point). The three Borexino 8B data points correspond, from left to right, to the low- energy (LE) range, LE+HE range, and the high-energy (HE) range. The electron neutrino survival probabilities from experimental points are determined using a high metallicity SSM from [26]. The error bars represent the $\pm 1\sigma$ experimental + theoretical uncertainties. The curve corresponds to the $\pm 1\sigma$ prediction of the MSW-LMA solution using the parameter values given in [27]. This figure is provided by A. Ianni." 19

FIGURE 2.6 The MSW transition region for some range of NSI parameters ϵ_1 and ϵ_2. Taken from [28] 21

FIGURE 2.7 $0\nu\beta\beta$ Feynman diagram. 22

FIGURE 2.8 From [30]. Shows the ^{8}B flux across the different SK phases. The gray band shows the combined value. Note that the DATA/MC is higher than the P_{ee} since they assume a pure ν_e flux, which includes the elastic scatters from the other neutrino flavors as well. 24

FIGURE 2.9 *Left:* The detection efficiency over the two different energy regions (HER I and HER II). The detection efficiency remains 1 until the 8500 p.e. upper edge in HER II. *Right:* The neutrino energies corresponding to HER I (red), HER II (blue), and the total HER (green). Both figures from [32]. 25

FIGURE 2.10 The energy spectrum of the full KamLAND-Zen dataset, after applying some backgrounds removing cuts. *Top:* The singles data (SD), which is supposed to be a dataset with a lower leakage of cosmogenic events. *Bottom:* The long-lived data, the remaining data aside from the singles data. From [35]. . 27

FIGURE 2.11 The energy spectrum of the first LEGEND-200 dataset. From [34]. The much finer energy resolution clearly be seen relative to Fig. 2.10. 28

FIGURE 2.12 The phase space scaled half-life limits of various experiments and the corresponding $m_{\beta\beta}$ limits, calculated for various matrix elements. Figure is courtesy of S. Biller [38]. 29

FIGURE 3.1 A rendering of the SNO+ detector. 32
FIGURE 3.2 A picture taken from between the PSUP and the AV during the water phase. The hexagonal shapes that can be seen through the AV are the PMTs and concentrators. 34
FIGURE 3.3 A schematic of the PMT and concentrator assembly in the SNO and SNO+ detectors [12]. Dimensions are in cm. 35
FIGURE 3.4 *Left:* The molecular form of LAB, which is the liquid scintillator used in the SNO+ experiment. *Right:* The molecular form of PPO, which is the primary fluor used in the SNO+ experiment 36
FIGURE 3.5 A comparison of the calibrated timing models used for 0.6 g/L PPO and 2.2 g/L PPO 37
FIGURE 3.6 The phases of the SNO+ detector. Adapted from A. Inacio [54] 40
FIGURE 3.7 The change in scintillator light yield as a function of Te concentration. From [44]. 42
FIGURE 3.8 A schematic of a dynode chain like the ones found in the SNO+ PMT. From [56] 43
FIGURE 3.9 A schematic of the SNO+ electronics with components added during the upgrade from SNO to SNO+ in blue. Taken from T. Zummo [57]. Some of the components not involved during standard data taking will not be discussed in this section. 45
FIGURE 3.10 A cartoon representing different classes of pileup events based on the summed PMT pulses. Taken from A. Sofia [54] 49

FIGURE 4.1 A plot showing how the hit time residuals change in different radial shells. A similar effect is seen 1-10 MeV energy range as well. Plot is courtesy of Will Parker [60]. 58
FIGURE 4.2 A plot showing the length between reconstructed and true positions of simulated $0\nu\beta\beta$ events (E$\sim$2.5 MeV) in 2.2g/L LABPPO. Most events reconstruct within a few 10s of cm of their true position. 59
FIGURE 4.3 The H parameter (Sec. 4.2.2) is able to effectively reproduce an estimator that is linear in energy. From [61]. 60
FIGURE 4.4 Reconstructed energies of $0\nu\beta\beta$ events in the 2.2 g/L LABPPO optics. 61
FIGURE 4.5 A plot comparing the expected ITR distributions of three distinct event classes. The ^{212}Bi^{212}Po are mostly in window β/α coincidences, giving them a distinct double peaked hit time residual distribution, resulting in a very different ITR distribution. ^{214}Bi events on the other hand are β decays that result in γ cascades that which occur on pico-second time scales relative to the initial β emission. Since the γ's will Compton scatter electrons 10s of cm away from the primary β emission, these events result in hit time residual distributions that are slightly broader than those from ^{8}B ν_e induced electron scatters, which are comprised of only one pure β. 63

FIGURE 4.6 Various distributions of tagged ^{214}Bi^{214}Po pairs from the basic scintillator phase. Candidates from tagging on both data and simulation are shown. Δt and Δr are the separation between the events in position and time, respectively. There are small but clear systematic disagreements between the energy distributions for both the prompt and delayed events, which the calibration procedures discussed in this section are intended to fix. The tagging criteria used to select these events can be found in Tbl. 4.1, aside from a more aggressive r cut that only accepted events out to 4.5 m. 65
FIGURE 4.7 The hit time residual distributions of internal ^{214}Bi and ^{214}Po events before and after tuning the scintillator emission time profile. 67
FIGURE 4.8 Mean energy (left) and ITR (right) distributions of tagged ^{214}Bi events before and applying the data driven corrections. 69
FIGURE 4.9 AV z-offset over the basic scintillator phase. 70

FIGURE 5.1 The backgrounds spectrum of events from the "singles data" part of the full Kamland-Zen dataset. From [62]. 71
FIGURE 5.2 *Left:* Uranium-238 decay chain. *Right:* Thorium-232 decay chain. 73
FIGURE 5.3 Results of the externals simulation for the external backgrounds considered in this analysis. 76
FIGURE 5.4 Energy (left) and ITR (right) distributions of some relevant backgrounds to this work. 77
FIGURE 5.5 *Left:* Energy marginal of systematic constraint fits. *Right:* ITR marginal of systematic constraint fits. 88
FIGURE 5.6 -2ln$\mathcal{L}$ profile curves for the systematic effects considered in this analysis. λ represents a scale, δ represents a shift, and σ represents a smear. The numbers in the legend are the mean and standard deviation derived from a quadratic fit to the likelihood profile points. 89
FIGURE 5.7 *Top:* The g/g concentration of ^{238}U based on the rate of tagged ^{214}Bi^{214}Po events. *Bottom:* The same but for ^{232}Th based on the rate of tagged ^{212}Bi^{212}Po events. Both images courtesy of G. Milton [65]. 91
FIGURE 5.8 Posterior profiles of the systematic parameters when fitting to a manually transformed version of the PDFs used in the fit. The values used in the transformations are $\lambda_{\text{Energy}} = 1.1$, $\delta_{\text{ITR}} = 0.001$, $\sigma_{\text{Energy}} = 0.03$ MeV, and $\sigma_{\text{ITR}} = 0.01$. 98
FIGURE 5.9 *Top:* Energy marginal of systematic validation fits. *Bottom:* ITR marginal of systematic validation fits. 99

FIGURE 6.1 *Top:* Energy marginal of fit to SNO+ data between 1.5 and 7.0 MeV. *Bottom:* ITR marginal. 104
FIGURE 6.2 *Top:* Energy marginal of the same fit as shown in Fig. 6.1 with the systematic parameters held constant at their optimal values from the systematic constraint fits. *Bottom:* ITR marginal. 105

FIGURE 6.3 A search 1 second into the past and into the future for any coincidence signals with Δr less than a meter. There is no clear preference for low Δr or Δt values, which is what we would expect for some kind of coincidence signature — rather the results seem entirely consistent with accidental coincidences from high rate backgrounds like ^{210}Po and ^{14}C. 107
FIGURE 6.4 A comparison between a "normal" physics event and a "pileup" physics event. The event display shown here is custom. The black, green, and red points represent the x, y, and z projections of hit PMTs in an event, and the blue points are the true 3D locations of the PMTs. The pileup event (right) clearly has a distinct hit topology, which is captured by the β_{14} classifier. 108
FIGURE 6.5 The summed hit time residuals of the excess events (red histogram labeled "2.2 MeV") relative to the summed hit time residuals of tagged ^{214}Bi events. The two agree with one another, suggesting that the low energy excess events appear to be physics events. 109
FIGURE 6.6 A fit with an additional free neutron capture component. 110
FIGURE 6.7 Schematic representing reactor antineutrino interactions in SNO+. From [57]. 110
FIGURE 6.8 *Top:* Energy marginal of a 3D fit out to 4.5 meters in energy, ITR, and $(r/r_{AV})^3$. *Bottom:* $(r/r_{AV})^3$ marginal. 113
FIGURE 6.9 A fit with the ^{214}Bi AV events freely floated without using any radial information to constrain the rate. Since external events that make it all the way into the 3.3 m FV are rare, it is difficult to generate a large enough simulation dataset to have a high statistics sample of events close to the center of the AV. Instead, the external PDF used in this plot applies a radial cut of 4.5 meters to select the events to load into the PDF, while using the 3.3 meter cut for all the other backgrounds. 114
FIGURE 6.10 The rate of events in the low energy excess in two different fiducial volumes. The red histogram goes out to larger radius, and the fluctuation in rate of the external events can be clearly seen in the histogram. The black histogram does not show any correlation with the red histogram, which one would expect to see if the low energy excess events were caused by mis-reconstructed externals. 115
FIGURE 6.11 The energy spectrum seen by Borexino in their CNO results paper with red lines superimposed to represent the rates Borexino would see if the low energy excess was present in their dataset. From [45], with red lines representing rates seen in SNO+ artfully drawn by the author. 116
FIGURE 6.12 Energy marginal of 2D fit in the 3.3 m FV with ^{102}Rh and ^{102m}Rh included in the backgrounds model. 118
FIGURE 6.13 *Top left:* The time separation between the two vertices in the multi-vertex fitter from simulations of ^{102m}Rh decays. *Top right:* The time separation between the two vertices in the multi-vertex fitter when applied to tagged reactor $\bar{\nu}$ events. An exponential fit to a histogram of the time differences agrees nicely with the expectation based on the ortho-positronium half life. *Bottom:* The results of applying the multi-vertex fitter on the low energy excess events. There is no clear signal of the formation of ortho-positronium. 121

FIGURE 6.14 The results of a nested sampling run with three parameters. The priors for every parameter are uniform. The likelihood is the product of Gaussian functions, one for each parameter with the mean and standard deviation defined in the legend. 126
FIGURE 6.15 Energy and multisite distributions of β/γ backgrounds simulated with different endpoints and beta fractions. Note that multisite is essentially a more advanced version of ITR that essentially carries the same exact information. . 128
FIGURE 6.16 Comparison between linear and McCann interpolation of two Gaussians with different means and standard deviations at 4 different interpolation points. *Left*: Linear interpolation showing the "ghosting" effect that can occur when PDFs do not lie on top of one another. *Right*: McCann interpolation. 128
FIGURE 6.17 A representation of the interpolation space V and the coordinates on the PDFs x. Each black dot is a coordinate in V, and we associate each coordinate in v with a distinct PDF. In this figure both the interpolation space and the PDFs are 2-dimensional, but in general the dimensionalities of each can be different. 130
FIGURE 6.18 Weights assigned to vertices of the bounding rectangle in bilinear interpolation. From [81]. 132
FIGURE 6.19 *Left*: Illustration of the W_b part of the interpolation algorithm. For each dimension in the PDF we calculate the marginals and then use the weighted QFs of the marginals to calculate W_b. *Right*: Illustration of the application of the map found in the previous step to each 1D slice in PDFs. 133
FIGURE 6.20 *Top:* An example of the simulated PDFs for a given empirical backgrounds (left) and the interpolated empirical background (right) with interpolation points at 2.25 and 2.5 MeV in energy and 0.40 and 0.425 in beta fraction. *Right:* The residuals between the two distributions on top. 134
FIGURE 6.21 *Left:* A fit to the basic scintillator phase dataset in the 3.3 m FV where the shapes of the two empirical backgrounds are floated. The fit prefers a γ endpoint of 2.18 MeV and a β/γ endpoint of 2.39 MeV with a beta fraction of 0.34. *Right:* A fit to the enhanced scintillator phase dataset using empirical backgrounds simulated only at the optimal shape parameters found in the fit to the basic scinitllator phase. Good agreement between the data and MC can be seen in both fits, suggesting that the model predicts the empirical backgrounds present in the enhanced scintillator phase. 135

FIGURE 7.1 The energy marginal of the Asimov dataset and the resulting fit used to maximize the ^{8}B ROI. 138
FIGURE 7.2 The true neutrino energies corresponding to the low energy region and high energy region in this work. The black point is the mean $\pm$ the numerical standard deviation of all the energies in the black histogram, and the same is true for the yellow point and the yellow histogram. 139
FIGURE 7.3 Marginals in each dimension of the 3D fit with an addition γ and β/γ empirical background. As can be seen in the legend, the resulting β/γ parameters are an endpoint of 2.42 MeV and a beta fraction of 0.35. The energy of the γ is fit to 2.18 MeV. 141
FIGURE 7.4 Posterior distributions of the three solar neutrino energy regions. 142

FIGURE 7.5 The survival probabilities measured by Borexino [32, 31], as well as the ones measured in this work. 144

FIGURE 8.1 *Left:* The results of the energy correction applied to the tagged Te production ^{214}Bi events. *Right:* The results of the ITR correction applied to the same events. 147
FIGURE 8.2 *Left:* The ROC curve of the ITR cut for the 0.5% Te production where the signal is just the 0ν events, and the background is only the in-window ^{212}Bi^{212}Po events. Also shown is the optimal ITR cut based on the derivative of the ROC curve. *Right:* The same but for the 1.5% production. 148
FIGURE 8.3 Plots showing the expected background contributions in a region near the 0ν ROI for 0.5% (*left*) and 1.5% (*right*) productions. Note the difference in y-axis scales. 149
FIGURE 8.4 Pie charts showing the expected background contributions in the 0ν ROI for 0.5% (*left*) and 1.5% (*right*) productions. The ROI is somwhat arbitrarily defined as -0.5 σ to 1.5 σ around the $0\nu\beta\beta$ peak where σ is the energy resolution at the 0ν endpoint. 150
FIGURE 8.5 $-2\ln\mathcal{L}$ constructed by profiling the log-likelihood with the profile points shown in black. The blue curve is a parabolic fit, which is then converted to a normal distribution, the parameters of which are shown in the legend. These parameters are then used to analytically calculate the 90% C.L. 151

FIGURE B.1 Plot of the 1-dimensional function $f(X)$ and the integral Z. Taken from [79]. . 160
FIGURE B.2 The mapping of θ to X. From [79] 161
FIGURE B.3 A summary of slice sampling. Process begins with a step width w and a point x_0. There is some allowed region over which we would like to sample, represented in blue. c shows the result of the stepping out procedure where upper and lower bounds have been found where the allowed region ends. In d a random sample as has been taken over the range in black, resulting in a point that falls in the region between the true lower bound of the allowed region and the lower bound of the region found in the stepping out procedure. At this point we update the lower bound of the sampling region to be x_1, shown in step e. Step f shows a sample in the allowed region being selected, which results in the termination of the slice sampling loop. From [87]. 166

CHAPTER 1

Introduction

This dissertation contains work mostly related to the development of a backgrounds model for the SNO+ detector. Typically the development of a backgrounds model is straightforward and quick, but due to some peculiarities of this analysis and experiment, we will see that developing such a backgrounds model is indeed a challenging problem. Once equipped with the backgrounds model though, we will make a measurement of solar neutrinos using the SNO+ detector, specifically attempting to probe a region of the ^{8}B solar neutrino spectrum that no detector has directly probed to date.

Chapter 2 will begin with a discussion of neutrino physics, focused on the solar neutrino problem and its resolution. The chapter will also attempt to provide some context on the current status of neutrino experiments, again focused on solar neutrinos and neutrinoless double beta decay ($0\nu\beta\beta$).

Chapter 3 provides an overview of the design of neutrino detectors in general, and also the design of the SNO+ experiment. We will discuss both the mechanical design of the detector itself, and the design of the electronics used to read out the data collected by the detector. Following that, the phases of the SNO+ experiment will be discussed This includes the basic scintillator phase which is the phase during which the dataset used in this analysis was taken.

Chapter 4 will then cover how we simulate and reconstruct events within the SNO+ detector. Simulation and reconstruction will play crucial roles in the work presented here since they are heavily relied upon to construct and interpret the backgrounds models of the SNO+ experiment.

Chapters 5 and 6 will discuss the expected backgrounds in the basic scintillator phase dataset, as well as how we can remove these backgrounds through event level cuts and coincidence tags. We will find that our assumptions regarding the content of the backgrounds model are not correct, which can be clearly seen in preliminary fits to the basic scintillator phase dataset. Some additional features must be added to the backgrounds model to fully explain what is seen in the data. Chapter

6 will build up the required machinery to generate a "empirical generic backgrounds model", which will model the backgrounds in the SNO+ experiment much more effectively.

Chapter 7 will then use the analysis techniques developed in Chapter 6 to make a measurement of the ^{8}B solar neutrino spectrum all the way down to energies of 2.4 MeV, probing regions of the ^{8}B neutrino spectrum that have not been measured before.

Chapter 8 will then cover the expected sensitivities of SNO+ to neutrinoless double beta decay, the main physics goal of the SNO+ experiment. Since SNO+ is nearing the Te phase, it is helpful to re-evaluate the detector sensitivities using optics and backgrounds models that are based on the current detector status.

CHAPTER 2

Neutrino Physics, What We Know and How We Know It

2.1. First detection

The history of the neutrino begins with one of the most famous physicists of the 20th century, Wolfgang Pauli. Pauli hypothesized the existence of the neutrino in 1930 to solve a few mysteries surrounding the process of beta decay, namely to recover energy, momentum, and angular momentum conservation [1]. At the time of Pauli's speculation regarding the existence of the neutrino, there was no hope of detecting the neutrino given that it must be neutral and (according to Pauli) have mass on a similar scale to the electron.

Following Pauli's proposal, Fermi included the neutrino in his theory of beta decay in which a neutron could decay into a collection of three particles, a proton, neutrino, and an electron [2]. We now know that this is an effective theory of the full theory of weak interactions, which will be discussed shortly. The beta decay process discussed above can also happen in reverse, where an anti-neutrino interacts with a proton to create a neutron and a positron. Even in the middle of the 20th century there was some hope of detecting this signal since the positron and neutron create coincidence signals which could be measured using equipment available at the time.

The first experiment to detect the neutrino used this inverse beta decay process. This experiment, now known as the Cowan-Reines experiment after Clyde Cowan and Frank Reines, searched for the coincidence signal of the positron and the neutron. The positron annihilation will create two γ-rays totaling to 1 MeV of energy, and the capture of the emitted neutron, which will occur hundreds of microseconds after the positron annihilation, will create another γ-ray in short coincidence with the prompt positron annihilation. A large flux of $\bar{\nu}$s was needed to produce any noticeable signal in the detector since the expected cross-section was on the order of 10^{-42} barns. A source that can provide this large flux of anti-neutrinos are fission reactors, as the nuclear processes inside the reactor emits a very large flux of anti-neutrinos. Thus the Cowan-Reines experiment was conducted next to the Savannah River Plant, a nuclear reactor in South Carolina.

The experiment did indeed observe this signal and, furthermore, noticed this signal stopped when the reactor was shut down, constituting the first direct detection of the neutrino [3]. Frank Reines was later awarded the Nobel Prize for this discovery in 1995 [4]. Interestingly, this technique of using the coincidence between the prompt positron and the delayed neutron capture to detect $\bar{\nu}$ is still in use today.

2.2. Solar neutrinos and the solar neutrino problem

In addition to the creation of anti-neutrinos from nuclear processes in nuclear reactors, nuclear processes in the core of the Sun also create neutrinos. The search for these neutrinos created within the Sun, which we will call solar neutrinos, began with the Homestake experiment, which began taking data in 1967. The Homestake experiment was led by Ray Davis, and used the reaction

$$^{37}\mathrm{Cl} + \nu_e \rightarrow {}^{37}\mathrm{Ar} + e^- \tag{2.1}$$

to measure high energy solar neutrinos by counting the number of Argon nuclei created in a vat of C_2Cl_4, a commonly used cleaning detergent. The first results from the experiment resulted in far fewer neutrinos being detected than expected based on calculations done by John Bahcall [5, 6]. As the Homestake experiment took more data and Bahcall refined his calculations (which eventually became known as the Standard Solar Model (SSM)), the tension persisted with the measured flux of the solar neutrinos being about 1/3 of what was expected [7]. This tension grew into what is called the solar neutrino problem.

The high energy solar neutrinos detected in the Homestake experiment are produced primarily through the β-decay of ^{8}B

$$^{8}\mathrm{B} \rightarrow {}^{8}Be + e^+ + \nu_e \tag{2.2}$$

but there are other reactions in the Sun that can produce solar neutrinos. The interaction that creates the highest flux of solar neutrinos is the proton-proton fusion process, the spectrum of which is labeled as “pp” in Fig. 2.1.

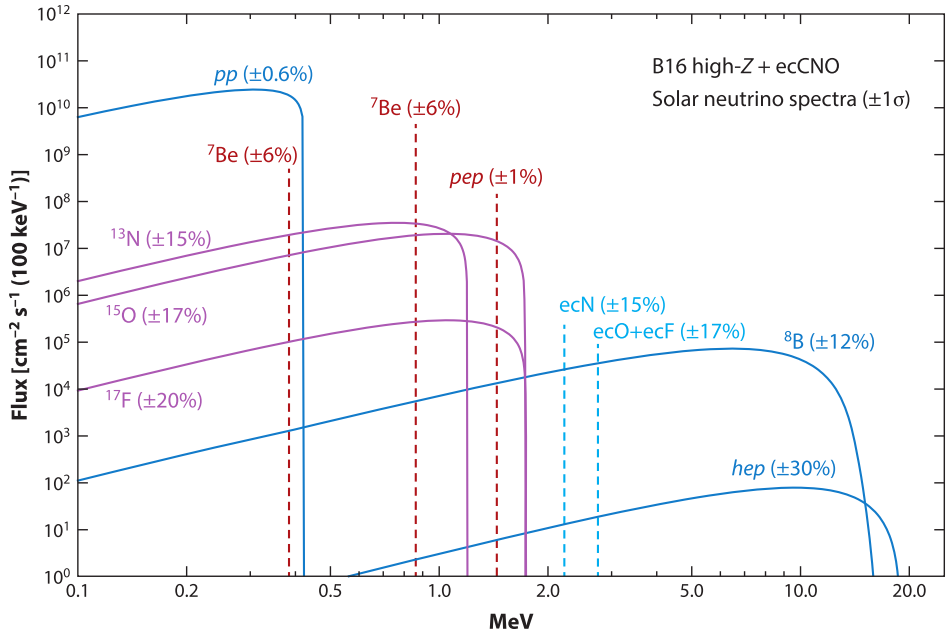


Figure 2.1: The fluxes of neutrinos from the various neutrino emitting processes in the Sun. Taken from [8]

Since the pp neutrinos are of a much higher flux than the ^{8}B solar neutrinos, they provided an excellent opportunity to either resolve or reinforce the solar neutrino problem. But given the low energy of the pp neutrinos, one cannot use the same approach as used in Eq. (2.1) since the energy threshold for such a decay is larger than the endpoint of the pp spectrum. Instead new experiments were designed to use a similar reaction on Gallium

$$^{71}\mathrm{Ga} + \nu_e \rightarrow {}^{71}\mathrm{Ge} + e^- \tag{2.3}$$

where the number of Germanium nuclei can be counted to measure the flux of pp neutrinos. This was done by multiple experiments in the 90s and early 2000s including GALLEX/GNO [9] and SAGE [10] but they measured a rate that was 50% of the prediction.

In addition to the measurements above, the Kamiokande-II detector also made a measurement of the ^{8}B neutrino flux using a new detection technique based upon the collection of photons called Cherekov radiation. Charged particles traveling faster than the speed of light in a medium (like

water) will generate bursts of light and the photons created in this process are called Cherenkov radiation. Particle physicists, equipped with the their ingenuity, realized that this process could be used to detect the scatter of electrons by neutrinos, i.e. the process

$$\nu_e + e^- \rightarrow \nu_e + e^-. \tag{2.4}$$

For ^{8}B solar neutrinos with MeVs of energy, the scattered electrons will move faster than the speed of light in water, leading to detectable bursts of light.

This new detector technology is a paradigm shift in neutrino detection for a few reasons. For one it allowed for realtime measurements of solar neutrinos interactions (since one did not need to stop taking data to acquire results like was done with the Homestake detector). Additionally, Cherenkov radiation is correlated with the direction of the moving particle. For for electrons scattered by neutrinos, the electron direction is correlated with the neutrino direction, which means that water Cherenkov detectors can also measure the direction in which the neutrino was traveling. This is especially useful for detecting solar neutrinos since the neutrinos should all be moving away from the Sun.

By the time Kamiokande-II began taking data, it was also believed that neutrinos come in three generations (or "flavors") paired with their respective charged leptons, ν_e, ν_μ, and ν_τ. The neutrino-electron scattering process to which the Kamiokande detector was sensitive can be caused by all three neutrino species, in contrast with the earlier detection techniques which were only sensitive to electron neutrinos. The Kamiokande detector did measure neutrinos that seemed to originate from the Sun based on the inferred directions of the neutrinos (see Fig. 2.2), but they measured a flux that was about 50% lower than expected.

At this point the solar neutrino problem had persisted for nearly 4 decades, and two main hypotheses had emerged during this time. One was that the true temperature at the core of the Sun was different than the theoretical predictions (which could explain the deficit of detected solar neutrinos), or that the neutrinos were transforming from one type to another (i.e. $\nu_e \rightarrow \nu_\mu$) along

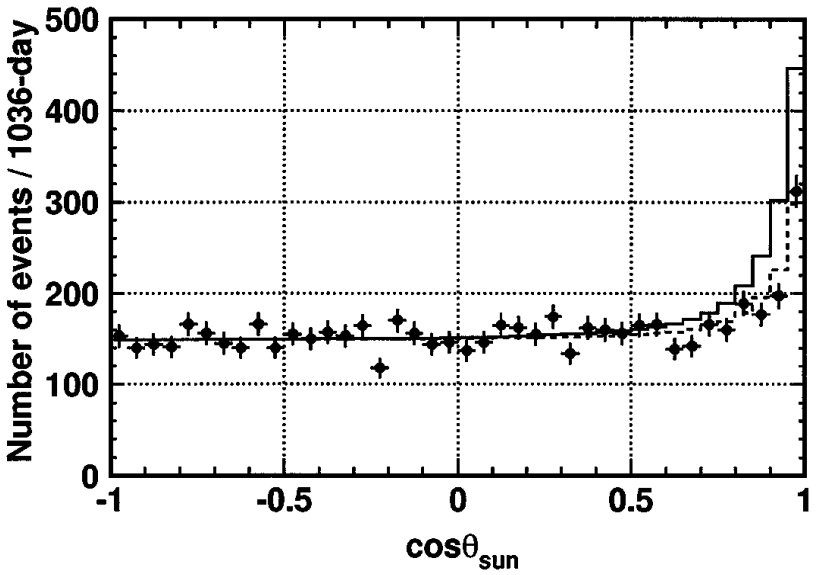


Figure 2.2: The angle between the direction of events in the Kamiokande-II detector and the direction to the Sun. The dashed line represent the expected signal based on the SSM, and the dashed histogram represents the fitted flux from the signal. From [11].

their journey from the core of the Sun to the Earth. The latter hypothesis is referred to as neutrino oscillations.

The resolution of the solar neutrino problem eventually came from the results of the Sudbury Neutrino Observatory (SNO) [12]. The approach taken by the SNO detector was specifically designed to directly test the theory of neutrino oscillations. To this end, the SNO detector used a heavy water target which was not only sensitive to the neutrino-electron scattering process seen in the Kamiokande detector, but the presence of deuterons allowed for additional interaction channels. The allowed interaction channels for SNO were:

$$\begin{aligned} \nu_e + {}^2\mathrm{H} &\rightarrow e^- + p + p \\ \nu_x + {}^2\mathrm{H} &\rightarrow \nu_x + p + n \\ \nu_x + e^- &\rightarrow \nu_x + e^- \end{aligned} \tag{2.5}$$

where the ν_x in the last two equations means that any neutrino flavor can participate. The second interaction channel is equally sensitive to all flavors and the third interaction channel is roughly 6x

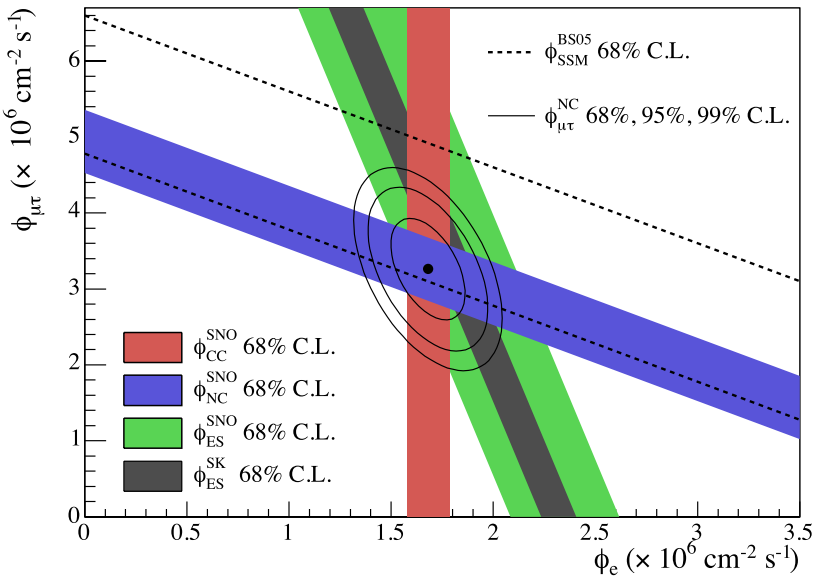


Figure 2.3: The measured fluxes of ν_e and $\nu_{\mu/\tau}$ neutrinos in SNO+ alongside the measured ν_e flux by Super-K and the prediction by the standard solar model (SSN). It can be seen that SNO can place limits on the $\nu_{\mu/\tau}$ flux by comparing the rates in the different interaction channels, while Super-K has no sensitivity to the $\nu_{\mu/\tau}$ flux. Taken from [14].

more likely to occur for ν_e relative to ν_μ and ν_τ. By comparing the rates in the three interaction channels, the SNO detector is able to directly determine if the missing ν_e solar neutrinos are oscillating into $\nu_{\mu/\tau}$ neutrinos.

SNO took data in three separate phases, with each additional phase improving the sensitivity to the second channel. The final results from SNO showed a total neutrino flux of $(5.25 \pm 0.16(\text{stat.})^{+0.11}_{-0.13}(\text{syst.})) \times 10^6 \text{cm}^{-2}\text{s}^{-1}$, which in agreement with the calculations of Bahcall and others (fluxes are shown in Fig. 2.3) [13]. This is now seen as the resolution of the solar neutrino problem. Following SNO's resolution of the solar neutrino problem, Ray Davis was awarded a half of the Nobel Prize in 2002 for the first detection of solar neutrinos.

A key confirmation of neutrino oscillations came from the successor to the Kamiokande experiment, Super Kamiokande or Super-K [15], which measured the oscillations of atmospheric neutrinos by comparing the downward and upward flux of electron neutrino and muon neutrino interactions. We could write an equally long discussion of Super-Kamiokande's measurement, but oscillations of atmospheric neutrinos are not as relevant to this dissertation. Arthur McDonald and Takaaki

Kajita were co-awarded the Nobel Prize in Physics in 2015 as spokespersons of SNO and Super-Kamiokande respectively, for the resolution of the solar neutrino problem and confirmation of neutrino oscillations [16].

While the solar neutrino problem is now considered to be solved, there still some specifics regarding the exact way in which the oscillations occur that are not well understood experimentally. A deeper understanding of how solar neutrinos oscillate as they move through the Sun could potentially provide evidence for new physics. Before we can discuss these details though, we must dive into the underlying theory used to describe neutrino oscillations.

2.3. Theoretical descriptions of neutrinos

Let us recall qualitatively what we know based on the measurements discussed in the previous section. Clearly, neutrinos are able to oscillate from one flavor to another as they propagate. But perhaps a subtle point is that the oscillations all seem to depend on the energy of the neutrinos as well, since the reduction in flux for the *pp* solar neutrinos was different than the reduction in flux for the ^{8}B solar neutrinos. Furthermore, we see oscillations over very different distances, one on the scale of the solar system (the Sun to the Earth), and the other on the scale of the Earth itself (atmospheric neutrinos). Clearly an interesting theoretical puzzle is afoot, and thankfully theorists have figured out the details for us.

2.3.1. Neutrinos in the Standard Model

First we will benefit from understanding the quantum field theoretic description of neutrinos and their interactions through the weak force. This chapter follows the section on neutrino mass and mixing from [14], with the relevant pieces for this work included here. We will start with the Standard Model (SM), considering only the charge current (CC) and neutral current (NC) processes in which the neutrino participates

$$\begin{aligned} -\mathcal{L}_{\text{CC}} &= \frac{g}{\sqrt{2}} \sum_{\ell} \bar{\nu}_{L\ell}\, \gamma^{\mu}\, \ell_L\, W_{\mu}^{\dagger} + \text{h.c.} \\ -\mathcal{L}_{\text{NC}} &= \frac{g}{2\cos\theta_W} \sum_{\ell} \bar{\nu}_{L\ell}\, \gamma^{\mu}\, \nu_{L\ell}\, Z_{\mu}^{0}. \end{aligned} \tag{2.6}$$

g is the coupling constant associated with the weak interactions, and θ_W is the Weinberg angle. The L subscripts in Eq. (2.6) signify that only left handed (or right handed anti-) particles participate in the weak interactions and the l subscript runs over the lepton generations. The force carriers of the weak force are the $W^{\pm}$ bosons and the Z boson, which are the fields W_μ and Z^0_μ. The left handed charged leptons and neutrinos come in pairs of lepton doublets, $L_{L_l} = (\nu_l, l)$ for $l \in (e, \mu, \tau)$, which encodes the neutrino and lepton flavors that are "paired" to one another through CC interactions.

The resulting Feynman diagrams that can be drawn are:



Figure 2.4: *Left:* Charged-current vertex. *Right:* Neutral-current vertex.

One can see that at charged current vertices the neutrino is transformed into its partner based on the weak doublet. At this point we can already understand the different interaction channels in SNO based on the allowed interactions in the standard model.

The first channel in Eq. (2.5) is a charged current interaction since a neutron is transformed into a proton, which requires the exchange of a charged W boson. Only the electron neutrino can participate in this interaction since the neutrinos are not high energy enough to produce a muon or a tau. The second listed interaction is mediated only through the neutral current since there is no exchange of charge or change of particle flavor between the initial and final states. The last channel is the most interesting — it can occur through both a NC process where the neutrino and electron exchange a Z boson *or* as a CC interaction where the incoming electron neutrino converts to an electron and emits a W^- which then converts another electron into an electron neutrino. Since Feynman diagrams interfere with one another, this results in a non-trivial change to the cross section of electron-neutrino elastic scattering for electron neutrinos relative to muon and tau

neutrinos. This is why the electron neutrinos participate in the this channel at 6x the rate of the other two flavors of neutrinos.

While not particularly useful at this moment, it is worth pointing out that in the SM masses for the charged leptons are generated by a Yukawa coupling of the scalar Higgs doublet with the right and left handed components of the charged fermions. Since only left handed particles interact through the weak force and neutrinos only interact through the weak force (ignoring gravity since we are talking about the SM), the SM only includes left handed (right handed anti-) neutrinos. This means that no Yukawa coupling terms of the same form can be written for the SM neutrinos, so neutrinos are massless in the Standard Model.

This presents an issue though, as the accepted description of neutrino oscillations requires massive neutrinos. So the Standard Model must be extended to include massive neutrinos.

2.3.2. Massive neutrinos

As was just discussed, adding massive neutrinos into the standard model requires the inclusion of some right handed neutrino fields. Since right handed particles do not participate in the weak interaction, these right handed neutrinos would not interact with any other particles, which is why they are called sterile neutrinos. Potential mass terms one can use to extend the SM are

$$-\mathcal{L}_{M_\nu} = M_{Dij}\,\bar{\nu}_{si}\,\nu_{Lj} + \frac{1}{2}M_{Nij}\,\bar{\nu}_{si}\,\nu^c_{sj} + \text{h.c.} \tag{2.7}$$

which defines two distinct methods for generating the neutrino mass, with the inclusion of m sterile neutrinos where the ν_ss are the sterile neutrino fields. ν^c is the charge conjugated field. M_D is a $m \times 3$ complex matrix and M_N is a symmetric $m \times m$ matrix.

The first term is called a Dirac mass term, which can be generated from electroweak symmetry breaking in the same was as the other leptons, i.e. through a Yukawa interaction with the Higgs field. If we assume there are 3 sterile neutrinos, we can identify each sterile neutrino combined with an active neutrino field ν_{Li} as a full four-spinor neutrino field, analogous to the four-spinor fields for the charged leptons. Rewriting the neutrino mass Lagrangian in terms of the eigenstates of the

mass Dirac mass term we have

$$-\mathcal{L}_{M_\nu} = \sum_{k=1}^{3} m_k \, \bar{\nu}_{Dk} \, \nu_{Dk}, \tag{2.8}$$

$$\nu_{Dk} = \left(V^{\nu\dagger} \, \vec{\nu}_L\right)_k + \left(V_R^{\nu\dagger} \, \vec{\nu}_s\right)_k . \tag{2.9}$$

Adding neutrino masses in this way to the SM has some downsides. For one, the addition of the right handed neutrino fields is an ad hoc addition to allow for us to write down a Dirac mass term which has no experimental motivation. Secondly, the masses of the neutrinos are orders of magnitude lighter than the next lightest lepton. A Dirac mass term does not explain why the neutrinos are so light, since the Yukawa coupling must be tuned to be very small to give neutrinos such small masses. A more elegant way to include mass into the SM is through the Majorana mass term.

The second term in Eq. (2.7) is the Majorana mass term. It turns out that one can rewrite Eq. (2.7) in terms of the mass eigenstates of the matrix $M_\nu = \left(\begin{smallmatrix} 0 & M_D^T \\ M_D & M_N \end{smallmatrix}\right)$ resulting in

$$-\mathcal{L}_{M_\nu} = \frac{1}{2} \sum_{k=1}^{3+m} m_k \bar{\nu}_{Mk} \nu_{Mk} \tag{2.10}$$

where each $\nu_{Mk} = \nu_{mass,k} + \nu^c_{mass,k}$. $\nu_{mass,k}$ is the eigenstate of M_ν with eigenvalue m_k. We can see that by definition $\nu_{Mk} = \nu^c_{Mk}$, which means these states obey the Majorana condition [17], named after Ettore Majorana. A consequence of the allowing for a Majorana mass term is that it allows for lepton number violating processes to occur. We can think of setting the entries of M_N to zero as requiring that lepton number be conserved.

If the mass eigenvalues of M_N are much higher than the scale of electroweak symmetry breaking, the diagonalization of M_ν will provide three light neutrinos and m heavy neutrinos. The light neutrino masses are are proportional to M_N^{-1} and the heavy neutrino masses are proportional to M_N, which can provide an explanation of why the masses of the active neutrinos are so light. The mechanism described above tells us that if neutrinos are Majorana particles, the heavy right handed

sterile neutrinos essentially force the active left handed neutrinos to have very small masses. This is called the *see-saw mechanism*. Both SO(10) Grand Unified Theories and left-right symmetric models are well known examples which involve see-saw mechanisms for the mass generation of the neutrinos [18, 19].

Now that we know we can extend the standard model to generate massive neutrinos, we can discuss neutrino oscillations.

2.3.3. Neutrino oscillations in vacuum

Neutrino oscillations can be understood as a quantum mechanical phenomenon. If we consider what was shown previously in our discussion on neutrino masses, we can see that in general the mass basis eigenstates and the interaction basis (or flavor basis) eigenstates are related through a unitary transformation

$$|\nu_\alpha\rangle = \sum_{i=1}^{n} U^*_{\alpha i}\,|\nu_i\rangle. \tag{2.11}$$

where i runs over the number of mass states, which for our purposes can always be assumed to be 3. If we are interested in the probability of measuring the flavor of the neutrino after some time t during which it is propagating freely, we need to calculate

$$P_{\alpha\beta} = |\langle\nu_\beta|\nu_\alpha(t)\rangle|^2 = \left|\sum_{i=1}^{n}\sum_{j=1}^{n} U^*_{\alpha i}\,U_{\beta j}\,\langle\nu_j|\nu_i(t)\rangle\right|^2. \tag{2.12}$$

Treating the neutrinos as a plane wave and assuming we are in the relativistic regime ($E \approx p + \frac{m_i^2}{2E}$), we know the neutrino state evolves like $e^{-iE_i t}|\nu_i(0)\rangle$. Equipped with this fact and the fact that the mass eigenstates are orthogonal, we arrive at

$$\begin{aligned} P_{\alpha\beta} = \delta_{\alpha\beta} &- 4\sum_{i>j}^{n} \mathrm{Re}\big[U^*_{\alpha i}\,U_{\beta i}\,U_{\alpha j}\,U^*_{\beta j}\big]\,\sin^2\left(\frac{\Delta m^2_{ij}L}{4E}\right) \\ &+ 2\sum_{i>j}^{n} \mathrm{Im}\big[U^*_{\alpha i}\,U_{\beta i}\,U_{\alpha j}\,U^*_{\beta j}\big]\,\sin\left(2\frac{\Delta m^2_{ij}L}{4E}\right), \end{aligned} \tag{2.13}$$

with $\Delta m^2_{ij} = m^2_i - m^2_j$ and L is the distance traveled by the neutrino, which is $\approx ct$. The transformation $U^*_{\alpha i}$ is the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [20].

The PMNS matrix for 3 massive neutrino states can take two different forms depending on whether the neutrinos are Dirac or Majorana. If they are Majorana the PMNS matrix must also contain Majorana phases and takes the following form:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\rm CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\rm CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\eta_1} & 0 & 0 \\ 0 & e^{i\eta_2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.14}$$

with $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$. For the angles $\theta_{ij} \in [0, \pi/2]$ and the CP violating and Majorana phases $\delta_{CP}, \eta_i \in [0, 2\pi]$. In the case of Dirac neutrinos, the Majorana phases carry no physical significance, allowing for the PMNS matrix to be written as

$$U = \begin{pmatrix} c_{12}\, c_{13} & s_{12}\, c_{13} & s_{13}\, e^{-i\delta_{\rm CP}} \\ -s_{12}\, c_{23} - c_{12}\, s_{13}\, s_{23}\, e^{i\delta_{\rm CP}} & c_{12}\, c_{23} - s_{12}\, s_{13}\, s_{23}\, e^{i\delta_{\rm CP}} & c_{13}\, s_{23} \\ s_{12}\, s_{23} - c_{12}\, s_{13}\, c_{23}\, e^{i\delta_{\rm CP}} & -c_{12}\, s_{23} - s_{12}\, s_{13}\, c_{23}\, e^{i\delta_{\rm CP}} & c_{13}\, c_{23} \end{pmatrix} \tag{2.15}$$

Returning to Eq. (2.13), we can see that the wavelength of neutrino oscillations are governed by the ratio of L and E and the PMNS matrix encodes the overall mixing between any two neutrino states. This means that different neutrino sources at different baselines are sensitive to different parameters of the PMNS matrix, which is summarized in Sec. 2.3.3 (recreated from Tbl. 14.6 of [14]).

Experiment	Dominant	Important
Solar Experiments	θ_{12}	Δm^2_{21}, θ_{13}
Reactor LBL (KamLAND)	Δm^2_{21}	θ_{12}, θ_{13}
Reactor MBL (Daya-Bay, Reno, D-Chooz)	θ_{13}, $\lvert\Delta m^2_{31,32}\rvert$	
Atmospheric Experiments (IC-DC,SK)	θ_{23}, $\lvert\Delta m^2_{31,32}\rvert$, sign($\Delta m^2_{31,32}$)	θ_{13}, δ_{CP}
Accel LBL $\nu_\mu, \bar{\nu}_\mu$ Disapp. (T2K, NOνA, MINOS, K2K)	$\lvert\Delta m^2_{31,32}\rvert$, θ_{23}	
Accel LBL $\nu_e, \bar{\nu}_e$ App. (T2K, NOνA, MINOS)	δ_{CP}, θ_{23}, sign($\Delta m^2_{31,32}$)	θ_{13}

We can see in the table that solar neutrino experiments are only sensitive to one mass splitting and θ_{13} is small relative to θ_{12} — this allows us to model solar neutrino oscillations in the two neutrino regime where there is only one $\Delta m^2 = \Delta m^2_{21}$ and mixing angle $\theta = \theta_{12}$:

$$P_{\alpha\beta} = \delta_{\alpha\beta} - (2\delta_{\alpha\beta} - 1)\sin^2(2\theta)\sin\left(2\frac{\Delta m^2 L}{4E}\right) \tag{2.16}$$

with θ usually called the solar mixing angle. The current best fit values for the solar oscillation parameters are $\theta_{12} = 33.68^{+0.73}_{-0.70}$ and $\Delta m^2_{21} = 7.49^{+0.19}_{-0.19} \times 10^{-5}\text{eV}^2$ [21].

Something worth briefly mentioning is that in this two neutrino oscillation limit, one can see that we are not sensitive to the sign of the mass splitting, Δm^2. It turns out that if one considers oscillations in the presence of matter the sign of the mass splitting becomes physically consequential, which is something we will see shortly. This is why future baseline experiments like DUNE [22] and Hyper-K [23] use beams which travel through large sections of Earth's crust.

2.3.4. Neutrino oscillations in matter

We can think of neutrino oscillations in matter as a perturbation to the free neutrino Hamiltonian used in the previous section. Skipping the details on calculation, intuitively we can see that only electron neutrinos will pick up a change in their potential energy due to CC interactions since there are no muons or taus in normal matter. The contribution to the potential energy is

$$V_{CC} = \sqrt{2} G_F n_e \tag{2.17}$$

where G_F is Fermi's constant and n_e is the electron density in the material. We can also use a similar line of reasoning to argue that there is no observable effect on the oscillations produced by the NC interactions in matter since they will affect each neutrino flavor in the same way. This means that the change in Hamiltonian for the neutrinos by the NC interaction only contributes a global phase to the neutrino oscillations, something to which we are not sensitive. The interaction Hamiltonian will have the form $V = \text{diag}(V_e(x), 0, 0)$ with $V_e(x) = \sqrt{2}G_F n_e(x)$ where $n_e(x)$ is still the electron density, we are just now allowing it to vary throughout the material.

Neutrinos interact in the flavor basis, but propagate in the mass basis which means we must use something equivalent the PMNS matrix to understand how matter affects neutrino propagation. For example, if we consider specifically solar neutrinos the two relevant neutrino states are ν_e and $\nu_{\mu/\tau}$ (since only electron neutrinos experience a different potential) with two mass eigenstates ν_1 and ν_2. We can relate the instantaneous mass eigenstates ν_1' and ν_2' of the Hamiltonian including the matter effects to the flavor basis using

$$\nu_\alpha = \widetilde{U}_{\alpha i}(x)\nu_i' \tag{2.18}$$

where $\tilde{U}$ is new 2x2 mixing matrix. The mixing angle of this new 2x2 mixing matrix is given by

$$\tan(2\theta_m) = \frac{\Delta m^2 \sin(2\theta)}{\Delta m^2 \cos(2\theta) - 2EV_e} \tag{2.19}$$

We can now see why matter effects can help one resolve the sign of the mass splitting, since θ_m is not invariant under the transformation $\Delta m^2 \to -\Delta m^2$.

There is also resonance condition when $2EV_e = \Delta m^2 \cos(2\theta)$, meaning the effective mixing between the two instantaneous mass eigenstates becomes maximal (i.e. $45°$). Furthermore, if $2EV_e >> \cos(2\theta)$ then the mixing matrix is such that the relationship between mass states and flavor states in vacuum is essentially swapped. Specifically, since ν_e are mostly ν_1 in vacuum, then in dense enough matter the ν_e asymptotically approaches being entirely in the ν_2' instantaneous mass eigenstate since θ_m becomes $\pi/2$. This effect where matter can greatly affect the propagation of neutrinos

is often referred to as the Mikheyev-Smirnov-Wolfenstein (MSW) effect and it plays a key role in understanding solar neutrino oscillations [24, 25].

2.3.5. MSW effect for solar neutrinos

The density of the Sun decreases from the core to the surface, meaning we can think of $n_e(R)$ for the Sun as a monotonically decreasing function. If we recall the resonance condition on Eq. (2.19), we can think of three different scenarios given the potential experienced by the neutrinos produced in the core of the sun, $A_0 = 2EG_F n_{e,0}$ where $n_{e,0} = n_e(R = 0)$. It's also helpful to define the potential at some distance R from the center of the Sun, $A_R = 2EG_F n_e(R)$.

The first scenario is if A_0 is much less than the resonance condition $A_R = \Delta m^2 \cos 2\theta$, which means that the neutrinos will never experience a potential that satisfies the resonance condition on their journey through the Sun. In this case, the propagation is approximately the same as that experienced in vacuum, so the survival probability of the solar electron neutrinos at the surface of the Earth is

$$P_{ee} = 1 - \frac{1}{2}\sin^2 2\theta > \frac{1}{2}. \tag{2.20}$$

The next scenario is when A_0 is much larger than the resonance condition, which in turn means that the resonance condition A_R will be met as the neutrinos travel to the surface of the Sun. Considering Eq. (2.19) , we can also see that in this scenario the ν_es at the moment of their creation will be almost entirely created in the ν_2' state as θ_m at the production point will be nearly $\pi/2$.

Another key feature of the MSW effect for solar neutrinos is that the resulting predictions depend on whether or not the evolution of the instantaneous mass eigenstates are adiabatic or non-adiabatic from the core to the surface of the Sun. The parameters that define solar neutrino oscillations are such that this process is adiabatic for all ranges of solar neutrino energies, but it is still interesting to consider why we need the process to be adiabatic. Let's assume the maximally non-adiabatic scenario, where the neutrinos are created in the ν_2' mass eigenstate in the core of the Sun and then the potential immediately turns off such that the neutrinos will now propagate as vacuum mass

eigenstates. In this scenario, the ν_2' and the ν_2 states are never mapped onto one another, meaning that the probability of a neutrino in the ν_2' state exiting the Sun to be in the ν_2 state is some probability less than 1.

On the other hand, when the process is adiabatic, the instantaneous mass eigenstates are continuously "rotated" onto the vacuum mass eigenstates as the matter potential changes through the interior of the Sun. By the time the neutrinos are propagating according to the vacuum Hamiltonian, the ν_2' and ν_2 states will be essentially identical, and the probability for the neutrinos to be in the ν_2 mass eigenstate is 1. This means the adiabatic case is much more effective at converting neutrinos created as ν_es into the ν_2 eigenstate.

Loosely speaking, this is analogous to photons passing through polarizers. If all the photons are polarized in one direction and we pass them through a polarizer at 45° relative to the initial polarization, the survival fraction will be only 1/2. But if we use an infinite number of polarizers all infinitesimally rotated relative to the previous, we can make the transition probability essentially 1.

In the case where $A_0 >> A_R$ and the neutrino evolution is adiabatic the survival probability is

$$P_{ee} = \frac{1}{2}[1 + \cos 2\theta_{m,0} \cos 2\theta] = \sin^2 \theta \tag{2.21}$$

which allows for P_{ee} to be much smaller than 1/2.

The last scenario is essentially in between these two scenarios, where the neutrinos are not always created in the ν_2' instantaneous mass eigenstate, which means that P_{ee} will end up being somewhere in between the other two limiting scenarios.

This now fully explains the results discussed in Sec. 2.2. The detectors that measured only the ν_e component of the higher energy ^{8}B solar neutrinos all measured fluxes roughly 1/3 of what was expected relative to the standard solar model (SSM). Experiments, like Kamiokande, which measured elastic scatters of the ^{8}B solar neutrinos were still sensitive to the other neutrino flavors,

but less so than the electron neutrinos, making them measure values slightly higher than 1/3. Experiments measuring the lower energy *pp* neutrinos measured fluxes closer to 1/2 the expectation, since the low energy neutrinos do not experience the matter effects. A large solar mixing angle θ_{12} alongside the MSW effect perfectly explains these results. This is the full solution to the solar neutrino problem, typically called the MSW effect with a large solar neutrino mixing angle, or the MSW-LMA solution.

A plot illustrating the different regimes discussed here can be seen in Fig. 2.5.

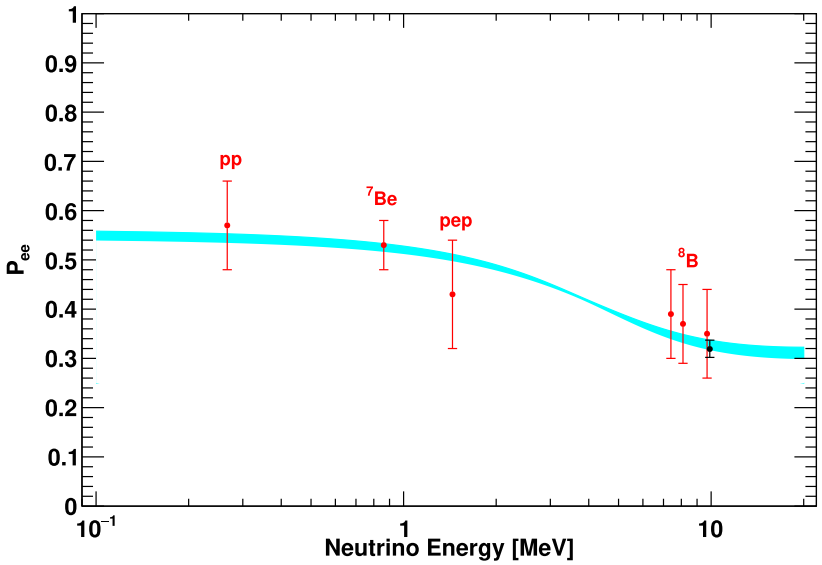


Figure 2.5: This caption is taken directly from the reference [14], Figure 14.3. "Electron neutrino survival probability as a function of neutrino energy. The points represent, from left to right, the Borexino pp, 7Be, pep, and 8B data (red points) and the SNO+SK 8B data (black point). The three Borexino 8B data points correspond, from left to right, to the low- energy (LE) range, LE+HE range, and the high-energy (HE) range. The electron neutrino survival probabilities from experimental points are determined using a high metallicity SSM from [26]. The error bars represent the $\pm 1\sigma$ experimental + theoretical uncertainties. The curve corresponds to the $\pm 1\sigma$ prediction of the MSW-LMA solution using the parameter values given in [27]. This figure is provided by A. Ianni."

This work aims to make a measurement of the MSW transition region (roughly speaking for neutrino energies between 2 and 8 MeV), ideally making a lower energy ^{8}B solar neutrino measurement than Borexino.

MSW transition region and sensitivities to new physics

Measuring this "MSW transition region" is important in and of itself as a confirmation of the MSW effect in solar neutrino oscillations. While the general consensus in the neutrino physics community is that the MSW-LMA solution is the entire story of neutrino oscillations, the MSW transition region has not been measured with the required resolution to definitively show that is the case. We have only made accurate measurements of the vacuum dominated and matter dominated oscillation regimes, which are the two edges of the x-axis in Fig. 2.5.

Aside from confirming the MSW-LMA solution, a measurement of the electron neutrino survival probability for an average neutrino energy of ∼7 MeV also provides limits on models that predict different shapes for the MSW transition region. This MSW transition region can be very sensitive to any models that couple to the active neutrinos since they will necessarily change the Hamiltonian of the neutrino, slightly shifting around the details of the resonance condition implied by Eq. (2.19). Following [28], we will discuss some generic models which do predict changes to the MSW transition region.

First we will discuss a generic way of including non-standard contributions to the neutrino Lagrangian. It is suggested in [29] that these non-standard contributions can be included as low energy four fermion operator. The inclusion of this operator transforms the two neutrino Hamiltonian for solar neutrino oscillations (in the flavor basis) into

$$\mathcal{H} = \frac{\Delta m^2_{21}}{4E} \begin{pmatrix} -\cos 2\theta_{12} & \sin 2\theta_{12} \\ \sin 2\theta_{12} & \cos 2\theta_{12} \end{pmatrix} + \sqrt{2} G_F n_e \begin{pmatrix} \cos\theta_{13} & \epsilon_1 \\ \epsilon_1 & \epsilon_2 \end{pmatrix}. \tag{2.22}$$

The effects of varying ϵ_1 and ϵ_2 can be seen in 2.6.

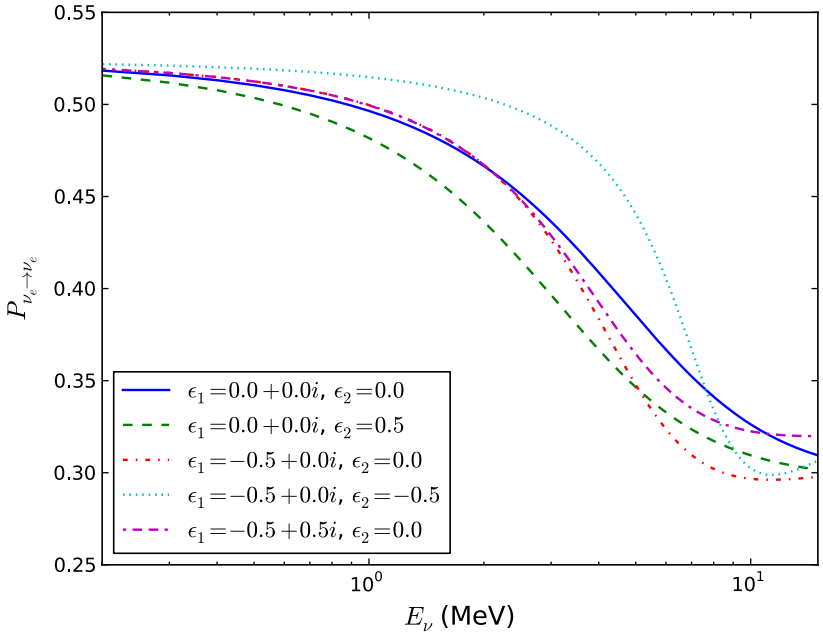


Figure 2.6: The MSW transition region for some range of NSI parameters ϵ_1 and ϵ_2. Taken from [28]

More specific models are discussed in [28], including mass varying neutrinos, which are theorized to try and resolve the coincidence that the energy density of dark energy and matter are similar in the present day universe, and the inclusion of long range leptonic forces into the Standard Model.

2.3.6. Neutrinoless double beta decay

As was discussed in Sec. 2.3.2, there is an ambiguity regarding how one can add massive neutrinos into the standard model depending on if neutrinos are Dirac or Majorana fermions. In the case where neutrinos are Majorana fermions, they are then by definition their own antiparticles. This allows for lepton number violating process like neutrinoless double beta decay ($0\nu\beta\beta$) to occur, where a double beta decay occurs and emits no neutrinos:

$$(A, Z) \to (A, Z + 2) + e^- + e^- \tag{2.23}$$

Observation of this $0\nu\beta\beta$ process would definitely prove that the neutrinos are indeed Majorana fermions, and that the small masses of the active neutrinos might be due to something like the

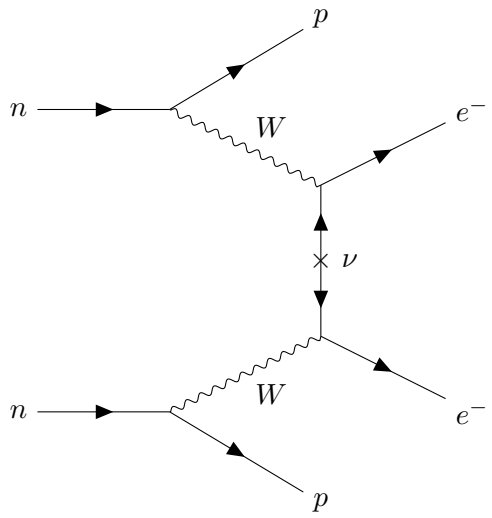


Figure 2.7: $0\nu\beta\beta$ Feynman diagram.

see-saw effect discussed previously. A Feynam diagram of the $0\nu\beta\beta$ process can be seen in Fig. 2.7

Neutrino experiments measure the half life of the $0\nu\beta\beta$ decay for a given nucleus. We can relate the half life to more fundamental physics quantities using

$$T_{1/2}^{-1} = G_{0\nu} \left| M_{0\nu} \right| \left(\frac{m_{\beta\beta}}{m_e} \right)^2 \tag{2.24}$$

where $G_{0\nu}$ is the phase space integral taking into account the final atomic state, $M_{0\nu}$ is the nuclear matrix element corresponding to the transition, m_e is the electron mass, and $m_{\beta\beta}$ is called the *Majorana mass* of the electron neutrino.

The fundamentally interesting quantity in neutrinoless double beta decay from a neutrino physics perspective is the Majorana mass, since the other terms in Eq. (2.24) depend on the nuclear properties of the nucleus being observed. The Majorana mass is understood in terms of the PMNS matrix (specifically the Majorana PMNS matrix): $m_{\beta\beta} = \left| \sum_i m_i U_{ei}^2 \right|$. Experiments will typically use 2.24 to convert from half life limits to $m_{\beta\beta}$ values. It is also worth mentioning that there is much complexity involved in calculating the matrix elements used to do this conversion leading to wide ranges in the estimates of their values. Thus $0\nu\beta\beta$ experiments will typically quote a range of limits on $m_{\beta\beta}$ while specifying the range of matrix elements used in the conversion.

Another important aspect of the direct detection of neutrinoless double beta decay is that it would provide the first direct evidence of a lepton number violating process.

2.4. Neutrino experiments

Now that we have discussed the history of neutrino observations and how that has informed our theoretical description of neutrinos, it is important to discuss the current experimental landscape of neutrino physics. This work is only concerned with solar neutrinos and neutrinoless double beta decay, so we will focus our discussion on experiments sensitive to those signals.

2.4.1. Solar neutrinos

In the present day, solar neutrino experiments typically follow a similar approach to SNO; a large vat of liquid surrounded by PMTs that will glow when solar neutrinos interact with particles inside the detector. While heavy water has not been deployed again since the de-comissioning of the SNO detector, success has been found both in standard water-Cherenkov detectors and scintillation detectors.

Super-K

On the water-Cherenkov detector end, Super-K has made measurements of the ^{8}B solar neutrino flux using elastic scatters. Of course, as was discussed in 2.2, since Super-K uses water rather than heavy water as its target it has no way to differentiate between electron neutrinos and other neutrino flavors. SNO still remains the best model independent solar neutrino flux measurement, but Super-K has a more accurate measurement of the flux derived just from the elastic scattering signal, which is $\Phi_{\mathrm{SK}} = (2.336 \pm 0.011(\mathrm{stat.}) \pm 0.043(\mathrm{syst.})) \times 10^6 \mathrm{cm}^{-2}\mathrm{s}^{-1}$ which assumes pure electron neutrino flavor component, without neutrino oscillation [30]. Super-K's measurements of the ^{8}B flux over their different phases can be seen in Fig. 2.8.

Super-K has also looked for another matter induced effect on solar neutrinos, called the day-night asymmetry, which results in the regeneration of electron neutrinos as solar neutrinos pass through the Earth at night. This would be measured as an increase in the solar neutrino flux measured by Super-K during the night relative to during the day. So far, Super-K has not seen a significant

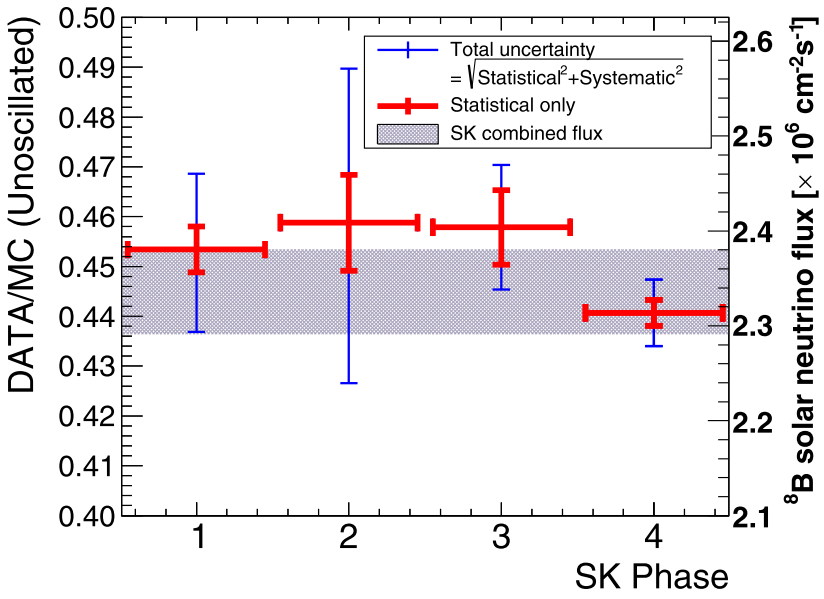


Figure 2.8: From [30]. Shows the ^{8}B flux across the different SK phases. The gray band shows the combined value. Note that the DATA/MC is higher than the P_{ee} since they assume a pure ν_e flux, which includes the elastic scatters from the other neutrino flavors as well.

signal of the day-night asymmetry [30].

Borexino

Borexino was a scintillator detector built to observe solar neutrinos. Borexino had very low backgrounds and good energy resolution, allowing for it to measure the *pp*, *pep*, CNO, ^{7}Be, and of course the ^{8}B neutrinos[31, 32]. The relevant analysis for the work presented in this dissertation is their ^{8}B measurement, which is what we will discuss here.

Borexino measured the ^{8}B flux in two distinct energy regions, high energy region 1 (HER I) and high energy region 2 (HER II).[1] HER I ranged from 1650-2950 photoelectrons (p.e.) and HER II range from 2950-8500 p.e. The conversion from p.e. to electron energy is essentially linear up to some arbitrary scaling (see Sec. 3.2), which for Borexino is approximately 500 p.e. per MeV. A picture showing the two energy regions can be found in Fig. 2.9

[1]The other solar neutrino sources Borexino measured were all lower energy, hence the name "high energy region".

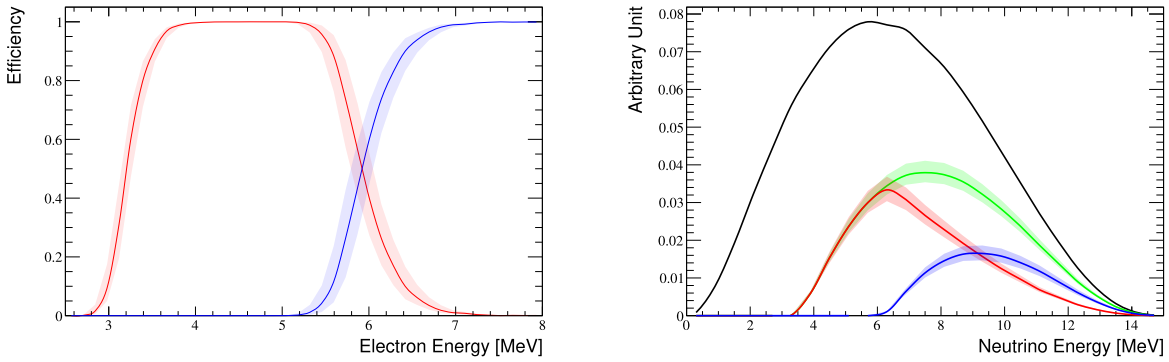


Figure 2.9: *Left:* The detection efficiency over the two different energy regions (HER I and HER II). The detection efficiency remains 1 until the 8500 p.e. upper edge in HER II. *Right:* The neutrino energies corresponding to HER I (red), HER II (blue), and the total HER (green). Both figures from [32].

The reason for measuring the flux in two distinct regions is to try and probe the MSW transition region by selecting a low energy sample of solar neutrino events. Even though HER I corresponds to something like 3 MeV to 6 MeV, the corresponding neutrino energies are smeared out by the kinematics of the elastic scattering of the electron by the incident neutrino. This effect can be seen on the right of Fig. 2.9. The corresponding average electron energy for HER I is 8.0 MeV and for HER II it is 9.9 MeV. Taking the SSM flux as an input and calculating the corresponding electron neutrino survival probability, Borexino calculated $P_{ee}(^8B_{\text{HER}}, 8.8\text{ MeV}) = 0.37 \pm 0.08$, $P_{ee}(^8B_{\text{HER-I}}, 8.0\text{ MeV}) = 0.39 \pm 0.09$, and $P_{ee}(^8B_{\text{HER-II}}, 9.9\text{ MeV}) = 0.35 \pm 0.09$. These values are the values included in Fig. 2.5. Note that the average neutrino energies quoted here are different than the ones in Fig. 2.5 because the values were changed between publications, which is not reflected in the plot from the PDG. The values quoted here are the most recent, and all future P_{ee} plots will use the new Borexino average neutrino energy values.

The measurement presented in this dissertation is functionally identical to the measurements done by Borexino above, but we will be going down to slightly lower energies than those in Borexino's HER I.

2.4.2. Neutrinoless double beta decay

Neutrinoless double beta has a diverse field of detectors and detector technologies. This is because the experimental task at hand welcomes a diverse range of approaches each with a different strategy. Each approach roughly falls somewhere on the axis of maximizing isotope vs. minimizing background counts. Of course, every neutrinoless double beta decay experiment tries to minimize background counts, but strategies which try to maximize isotope are willing to accept some increase in the background counts as long as the increase in expected signal is enough to increase the sensitivity of the experiment. On the end of minimizing background counts you have detectors like CUORE [33] and Legend-200 [34], and on the end of maximizing isotope you have large scintillation detectors like KamLAND-Zen[35] or SNO+.

This dissertation won't give a complete overview of all $0\nu\beta\beta$ detectors, but we will mention some of the current best results.

Kamland-Zen

The current best limit on the neutrinoless double beta decay half life comes from KamLAND-Zen, using a liquid scintillator doped with ^{136}Xe. KamLAND-Zen employed an approach very similar to the planned approach of SNO+, putting their isotope loaded scintillator in a "balloon" surrounded by an outer scintillator region, all of which is surrounded by 1325 17-inch and 554 20-inch photomultiplier tubes (PMTs) for calorimetry.

KamLAND-Zen, and other scintillator based $0\nu\beta\beta$ experiments, will not have the keV scale resolution the approaches aimed at minimizing background counts employ. Instead, KamLAND-Zen benefits from a large amount of isotope, which increases the exposure and therefore the number of $0\nu\beta\beta$ events expected for a given value of $T_{1/2}$, allowing it to tolerate higher background counts. The results of the full KamLAND-Zen dataset can be found in Fig. 2.10.

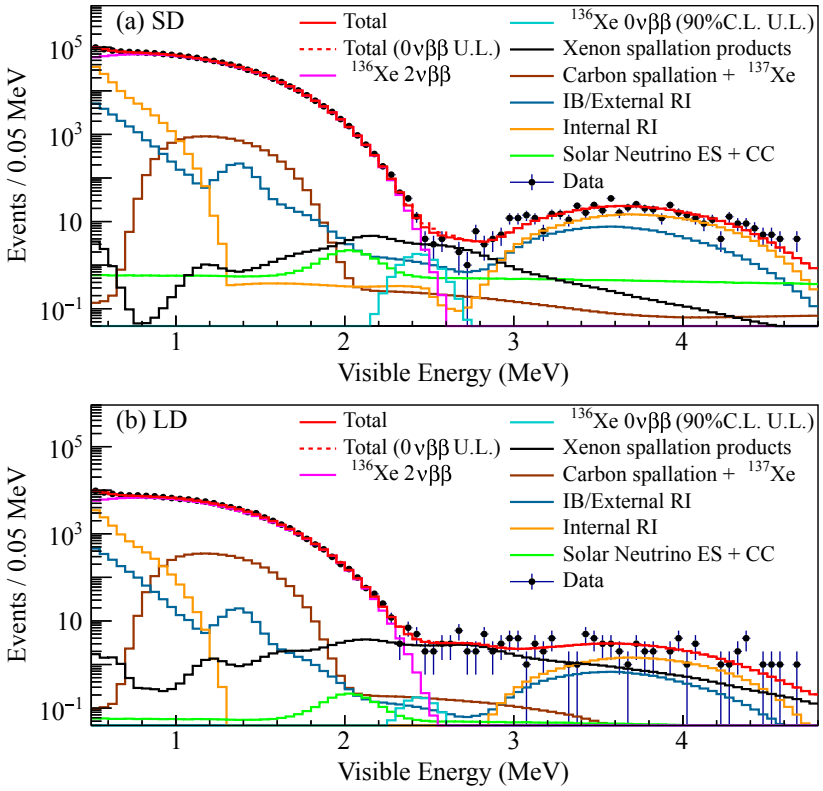


Figure 2.10: The energy spectrum of the full KamLAND-Zen dataset, after applying some backgrounds removing cuts. *Top:* The singles data (SD), which is supposed to be a dataset with a lower leakage of cosmogenic events. *Bottom:* The long-lived data, the remaining data aside from the singles data. From [35].

Using their full dataset, which corresponds to an exposure of 2.1 ton yr, the KamLAND-Zen collaboration set a limit of $T_{1/2} > 3.8{\times}10^{26}$ yr (90% C.L) which corresponds to an $m_{\beta\beta}$ of (28 − 122) meV (90% C.L.).

Legend-200

As was mentioned before, the Legend-200 detector employs the background counts minimizing strategy. This is accomplished through the use of extremely sensitive calorimeters which have energy resolutions down to the keV scale (while scintillator detectors have energy resolutions on the scale of 100s of keV). This is accomplished through the use of high-purity germanium (HPGe) detectors which collect the charge created from particles moving within the detector. The Ge in the HPGe detectors is isotopically enriched ^{76}Ge to 86%–92%, which is a double beta decay isotope.

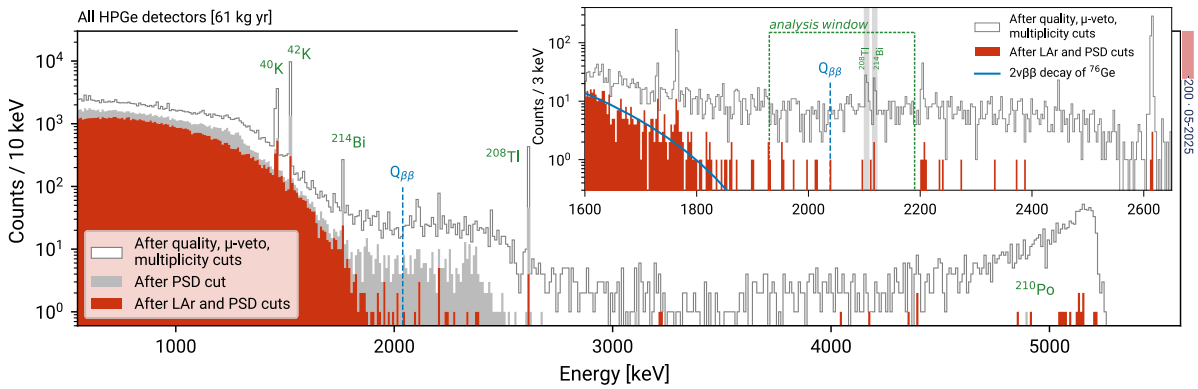


Figure 2.11: The energy spectrum of the first LEGEND-200 dataset. From [34]. The much finer energy resolution clearly be seen relative to Fig. 2.10.

Using 61 kg y of exposure, the half life limit set by the Legend-200 0.5×10^{26} yr (at 90% C.L.). The corresponding dataset can be seen in Fig. 2.11. Combining this result with the results of the MAJORANA DEMONSTRATOR [36] and GERDA [37], the limit is 1.9×10^{26} yr (at 90% C.L.), which corresponds to an $m_{\beta\beta}$ limit of $<$ 75–200 meV using nuclear matrix elements with values 2.35–6.34.

Comparisons of current and future results

A helpful way to understand the experimental landscape as a whole is to see how the half life limits and resulting $m_{\beta\beta}$ values of the different experiments compare when employing the same matrix elements in all calculations. This can be seen in Fig. 2.12. Clearly the choices of matrix element matter significantly. The predicted limits for SNO+ will be discussed more in Ch. 8.

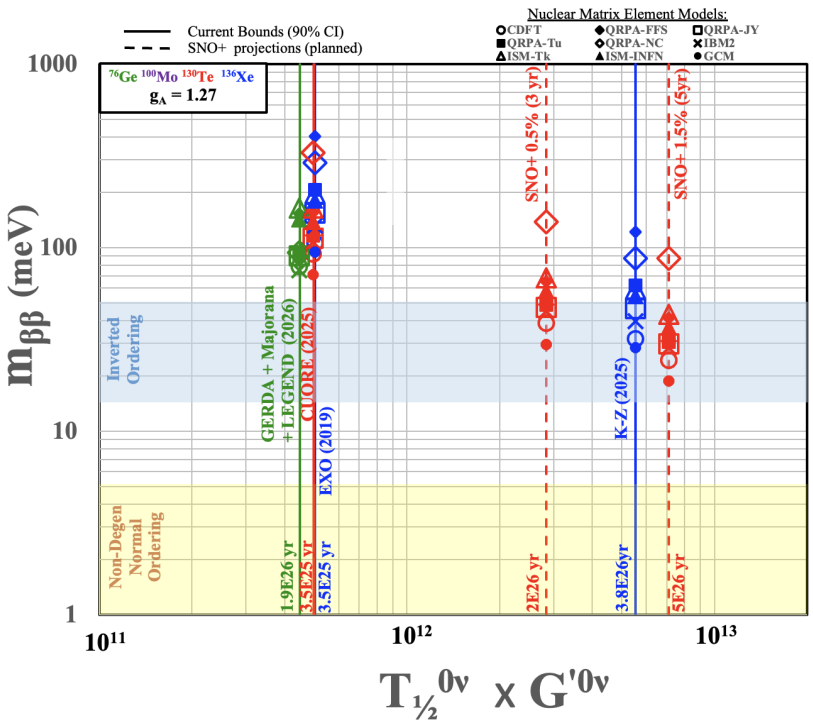


Figure 2.12: The phase space scaled half-life limits of various experiments and the corresponding $m_{\beta\beta}$ limits, calculated for various matrix elements. Figure is courtesy of S. Biller [38].

CHAPTER 3

The SNO+ Experiment

The SNO+ experiment is the successor to the Nobel prize winning Sudbury Neutrino Observatory, featuring much of the same infrastructure along with numerous electronics upgrades. Located 2km underground in the Vale Creighton mine, SNO+ operates as one of the deepest science experiments in the world.

Perhaps the most frequently asked question following a description of the location of the SNO+ detector is "why!?". This chapter will begin with an attempt to answer that question with some of the motivations that force neutrino physicists into large underground caverns in Sec. 3.1. Following that will be a discussion of the specifics of the SNO+ detector design in Sec. 3.2, then Sec. 3.3 contains a brief digression into liquid scintillators , and then an enumeration of the various phases over which SNO+ will take data will be discussed in Sec. 3.4. Finally, we will discuss the custom data acquisition system (DAQ) used by the SNO+ experiment in Sec. 3.5.

3.1. Neutrino observatories, in general

As was discussed in Ch. 2, neutrinos are exceedingly difficult to detect. This is due to the fact that neutrinos only interact through the weak force, making their cross sections much smaller than other interactions driven through the electromagnetic or strong forces. This means neutrino observatories must be very large in order to detect MeV scale neutrinos at any appreciable rate.

If we take the ^{8}B neutrino flux measured by Borexino in [32] for example, we can see that the measured rate of solar neutrino interactions is ~ 0.22 counts per day per 100 tons. The error on a rate measurement will roughly scale with the square root of the counts, so to make a measurement with 5% error in 5 years one needs roughly 400 solar neutrino events in 5 years, which would require a detector of at least 100 tons. In reality, experiments also have to contend with systematics in the detector response and radiological backgrounds, so more neutrino interactions are needed to provide the same sensitivity. This also brings us to another important consideration in neutrino

detector design, signal contamination from radiological backgrounds.

^{8}B solar neutrinos are the dominant solar neutrino species at MeV scale energies, consequently neutrino observatories usually look for their interactions within detectors at energies on a similar scale. Many natural occurring radioactive isotopes produce particles with similar energies and it is difficult to distinguish between these radioactive decays and solar neutrino interactions in a typical neutrino detector. Neutrino detectors must therefore minimize the presence of radioactive materials in order to make sensitive solar neutrino rate measurements.

In addition to natural background radiation, signals from neutrino interactions can also be hidden by interactions produced from cosmic-ray muons. Cosmic ray interactions in the upper atmosphere produce high energy muons that are constantly raining down on the surface of the Earth, creating another potential background in neutrino experiments. The issue with cosmic muons is two-fold. These muons are typically much higher energies than the solar neutrinos we want to detect, meaning interactions in detectors from muons often hide signals from interactions with neutrinos. But in addition to this effect, muons can also create short lived isotopes through spallation reactions with the stable materials within a detector, effectively adding to the background radiation. While air does not effectively attenuate these cosmic muons, the rock in the Earth's crust does, leading to neutrino detectors built deep underground.

Depending on the goals of a given neutrino experiment, different concessions must be made balancing size, cost, cleanliness, etc. For neutrino experiments interested in making measurements of solar neutrinos, the detector must be on the scale of 100s of tons since the rate of interactions cannot be increased by any other means. As a consequence, solar neutrino experiments must take great care making sure that they keep their detectors as clean, deep, and precise as possible in order to preserve their signal integrity. This brings us to the SNO+ detector.

3.2. The SNO+ detector

Roughly speaking, this section will describe the SNO+ detector moving from the "outside" inwards, starting with the components surrounding the detector and ending with the bulk material within

the acrylic vessel A more detailed account of the information in this section can be found in [39].

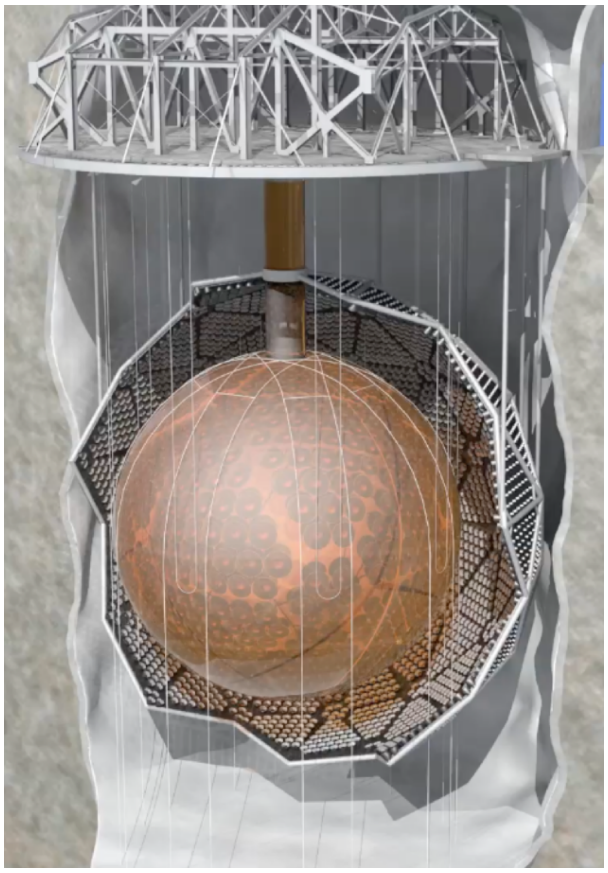


Figure 3.1: A rendering of the SNO+ detector.

As can be seen in Fig. 3.1, the SNO+ detector within a large cavity filled with water. Perhaps the largest advantage of the location of SNO+ is its depth. With a depth of 2070 m, or alternatively an overburden of 5890 meters water equivalent, the muon flux is only 0.27 muons per square meter per day. This corresponds to a rate of only ∼3 muons per hour over the entire detector. The low muon rate means that cosmogenics are not an issue in essentially every physics analysis, and any radioactive isotopes or neutrons produced by muons can be easily dealt with just by using a 20 second veto after events tagged as muons. The cut will again be discussed in Ch. 6.

In order to have somewhere to place the electronics and access the acrylic vessel, a large platform is built over the detector covering the entire cavity. This platform is usually referred to as the "deck". In addition to the deck, there is an additional sub-room on the deck that encompasses the part

of the acrylic vessel that pokes out above the deck. This smaller room is referred to as the "deck clean room", as it is there to try and provide an additional layer of shielding between the inside of the detector and the external environment.

As mentioned before, the cavity surrounding the PMT support structure (PSUP) is typically filled with ultra-pure water to both provide buoyant support for the PSUP and acrylic vessel (AV), and also shield from neutrons and gammas produced in the rock surrounding the cavity. The water level typically sits a few feet under the deck.

The PMT support structure is the spherical surface that provides the structure that holds the 9362 inward facing PMTs, as well as the 91 outward looking PMTs (OWLs) which are used to tag through-going cosmics. Unless the OWLs are explicitly mentioned, references to the PMTs implicitly refer to only the inward looking PMTs. The inward looking PMTs all are attached to Winston cone style concentrators to increase the effective coverage of the detector. The PMT support structure is not water tight, so water can freely flow between the cavity and the interior of the support structure — which is by design as the water buffer between the PMT support structure and the acrylic vessel will attenuate gammas produced from the PMT glass and the PSUP itself, a potential major background in the $0\nu\beta\beta$ phase. The faces of the inward facing PMTs are $\sim$8.35 meters away from the center of the AV. A picture taken in the region between the PSUP and the AV can be seen in Fig. 3.2

The AV itself has a 6 meter radius and a shell thickness of 5.5 cm. It is made of ultra violet transmitting (UVT) acrylic which increases the transmission of light through the AV surface, maximizing the number of photons that reach the PMTs. One of the major upgrades from SNO to SNO+ was the addition of the hold-down rope net, which can be seen in the schematic as the criss-crossing ropes that sit at the top of the spherical part of the AV. As the SNO experiment filled the AV with heavy water, the AV needed to be supported from above as the heavy water weighed more than the light water surrounding the AV. SNO+ on the other hand will primarily take data with the AV filled with liquid scintillator, meaning that the AV will want to float and will require support from the bottom since scintillator is lighter than water. The hold-down ropes apply this tension

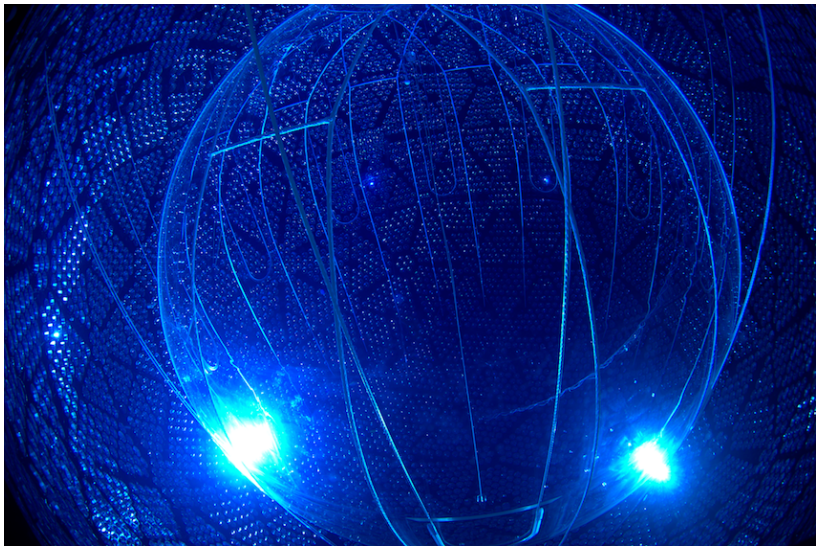


Figure 3.2: A picture taken from between the PSUP and the AV during the water phase. The hexagonal shapes that can be seen through the AV are the PMTs and concentrators.

that keeps the AV in place.

The innermost piece of the SNO+ detector is the target, or the bulk material that sits within the AV. Due to the phased approach of the SNO+ experiment, the target will change multiple times throughout the span of the SNO+ experiment, as will be discussed more fully in Sec. 3.4.

The PMTs used in the SNO+ detector are the same Hamamatsu R1408 PMTs used in the SNO detector. These are 8" hemispherical PMTs with a TTS of 1.6 ns, which is extracted from the standard deviation of a Gaussian fit [40]. A schematic of the PMT assembly including the concentrator can be seen in Fig. 3.3.

3.2.1. Chemical processing plants

Slightly distinct from the detector itself, the SNOLAB caverns surrounding SNO+ are filled with supporting machinery, much of which is designed to handle the loading of various substances into and out of the AV and surrounding cavity. SNO+ uses the same water plant as SNO, but new scintillator and tellurium processing plants were added to handle the additional materials that will go into SNO+. The scintillator plant is currently operating and the tellurium plants are at the final stages of commissioning.

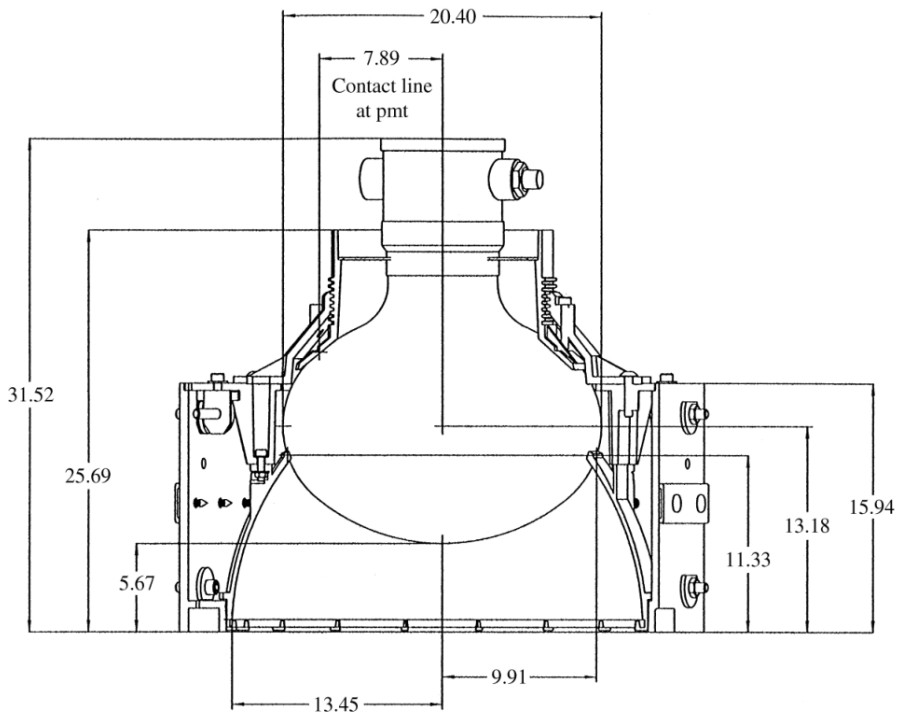


Figure 3.3: A schematic of the PMT and concentrator assembly in the SNO and SNO+ detectors [12]. Dimensions are in cm.

3.3. Liquid scintillators

As has already been mentioned, the SNO+ experiment is a liquid scintillator detector in contrast to SNO, which was a water-Cherenkov detector. Given the main physics goal of SNO+ being the observation of $0\nu\beta\beta$, using a high light yield target is a requirement of the physics objectives making liquid scintillator the obvious choice. Liquid scintillators are an attractive target to neutrino physicists due to their high light yield and radio-purity, making them especially appropriate for MeV-scale neutrino experiments.

The microscopic physics that produces light in liquid scintillators is complicated, but generally revolves around molecules that form rings, and these rings are shaped such that the resulting electron and vibrational energy levels on the rings are on the correct scale to produce photons in the blue to UV part of the spectrum [41]. An example of the ring-shaped phenyl groups present in linear alkylbenzene (LAB) can be seen in the left of Fig. 3.4.

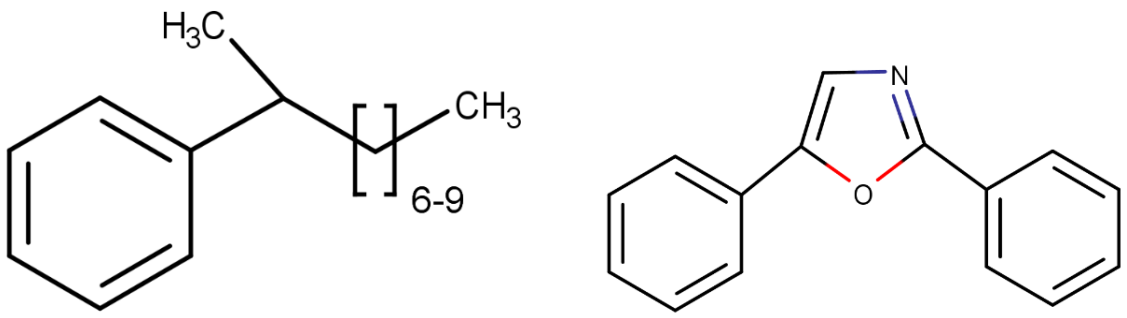


Figure 3.4: *Left:* The molecular form of LAB, which is the liquid scintillator used in the SNO+ experiment. *Right:* The molecular form of PPO, which is the primary fluor used in the SNO+ experiment

In addition to the primary scintillator, it is common to also use some number of additional fluors and wavelength shifters to try and both shift the light to a more transparent region of the primary scintillators absorption spectrum, and match the efficiency response of the photo-detectors. An example of one of the fluors used in the SNO+ experiment can be seen on the right of Fig. 3.4 This combination of primary scintillator, fluors, and wavelength shifters is usually referred to as a "cocktail".

The current cocktail used in the SNO+ experiment is LAB with 2.2 g/L PPO with an additional 2.2 mg/L bis-MSB, although data was taken with many other cocktails during the many phases of SNO+, which will be discussed in Sec. 3.4. The expected light yield for this cocktail was somewhere around 12000 photons/MeV, but the actual measured light yield after filling is lower (see Sec. 3.4.2 for a brief discussion). More information on the SNO+ scintillator can be found in [42].

The light emission from liquid scintillators also has a characteristic decay time since the excited states of the molecular structure take some time to deexcite and emit photons. Scintillators have multiple states that can be excited, therefore there are in general multiple decay constants that are all summed when one looks at the scintillator emission time profile. This is typically modeled as a weighted sum of exponential decays, optionally moderated by a rise time which accounts for the process of energy transference between the primary scintillator and secondary fluors. An example

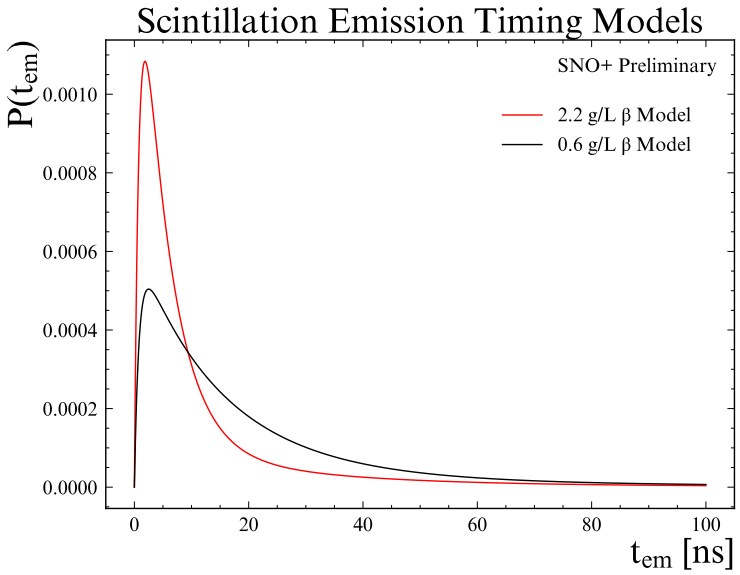


Figure 3.5: A comparison of the calibrated timing models used for 0.6 g/L PPO and 2.2 g/L PPO

of two different scintillators and the two different timing responses can be seen in Fig. 3.5.

3.3.1. Quenching

Another important property of liquid scintillators worth mentioning is the phenomenon of "quenching". One of the most useful properties of scintillators is their roughly linear response to particle energy. This means that relating the number of photons produced by the scintillator to the kinetic energy of the incident particle is simple and easy to extrapolate between calibration points.

While this linear relationship is *mostly* true for electrons in liquid scintillators, heavier particle like protons and α's have non-negligible quenching effects. This means they produce less light per unit energy deposited into the scintillator than electrons and the linear relationship between kinetic energy and photon production is no longer a valid approximation. This phenomenon is captured in an empirical sense by Birks' law [43]

$$\frac{dL}{dx} = S\frac{\frac{dE}{dx}}{1 + k_B\frac{dE}{dx}} \tag{3.1}$$

which relates the total number of produced photons L to the scintillation efficiency S (in units of

photons per MeV), the energy loss per unit distance of the particle $\frac{dE}{dx}$, and a particle and material dependent number k_B which is referred to as Birks' constant.

Eq. (3.1) applies to photon production in liquid scintillators in general, but for electrons k_B is very small and therefore there is a nearly linear relationship between the photons produced per unit track length (L) and the energy of the particle. But for large values of k_B, which is the case for heavier particles like α's, the linear relationship is broken — and this is the meaning of the colloquial statement that "α's are quenched".

3.3.2. Te loading

Novel techniques to load LABPPO with tellurium (Te) have also been developed for the Te-phase of SNO+ [44]. Stability has been shown over a wide range of concentrations and the effects of the loading process on light yield of the scintillator have also been studied. The liquid scintillator in SNO+ will eventually be loaded with Te using these techniques.

3.3.3. Cherenkov radiation in liquid scintillator detectors

A byproduct of the high light yield of scintillators is that while the Cherenkov light is still produced within the scintillator, the intensity of the scintillation light is enough to completely saturate the signal from the Cherenkov photons. Indeed, if we consider the Frank-Tamm formula

$$\frac{\partial^2 E}{\partial x \partial \omega} = \frac{q^2}{4\pi}\mu(\omega)\omega\left(1 - \frac{c^2}{v^2 n^2(\omega)}\right) \tag{3.2}$$

where dE is the energy emitted per unit length traveled by the particle dx per unit of frequency $d\omega$, v is the speed of the particle, and $\mu(\omega)$ and $n(\omega)$ are the permeability and index of refraction of the medium, respectively, we can see that photon production increases with increases in n — meaning that there is actually *more* Cherenkov light produced in LAB than there is in water. Whenever one shows Eq. (3.2), they are obliged to also mention that this only holds for charged particles where the speed of the particle is greater than the speed of light in the surrounding medium. This causes a "sonic boom" like effect, leading to coherent wavefronts of electromagnetic radiation, which is what we call Cherenkov radiation.

This trade-off between operating in a low light yield Cherenkov photon rich environment vs a high light yield scintillation photon rich environment has led to much discussion in recent years over how one may be able to combine the two techniques into one detector, a so called "hybrid Cherenkov scintillation" detector. We have already seen scintillation detectors utilize some of the information contained in the Cherenkov photons [45] [46], providing a tantalizing view of what could be done in a detector specifically designed to use both Cherenkov and scintillation light together. An example of one such detector is the Eos experiment[47], which is an experiment designed to serve as a test-bed for technologies which could be used in hybrid detectors.

3.4. Phases of the SNO+ experiment

In contrast to other neutrinoless double beta decay experiments, the SNO+ collaboration has elected to take a phased approach to loading the detector with tellurium. By phased approach, we mean that as the target of the detector changes from water to scintillator to Te-loaded scintillator, we stop and take data in each configuration to verify that we understand the performance of and backgrounds in the detector. This section will discuss the two phases that have already occurred, the water and scintillator phase, and the upcoming final phase, the Te-phase.

3.4.1. Water phase

Immediately following the hardware and electronics upgrades defining the transition from SNO to SNO+, the acrylic vessel was filled with 905 tons of ultra-pure water (UPW). This phase of the acrylic vessel being filled with UPW is referred to as the *water phase*. Data was collected during the water phase from May 2017 to June 2019. Filling first with ultra-pure water allowed for direct comparisons to data taken in SNO since the detector still operated as a water-Cherenkov detector, making it much more straightforward to understand the impacts of the upgrades and any impacts of the prolonged downtime.

To this end, multiple calibration sources were deployed in the SNO+ detector during the water phase. This includes a ^{16}N source and an Americium-Beryllium (AmBe) source, providing a tagged 6 MeV gamma and a 4.4 MeV gamma + neutron respectively, as well as a number of variable intensity laserball and LED sources. The laser sources were used to calibrate the water attenuation,

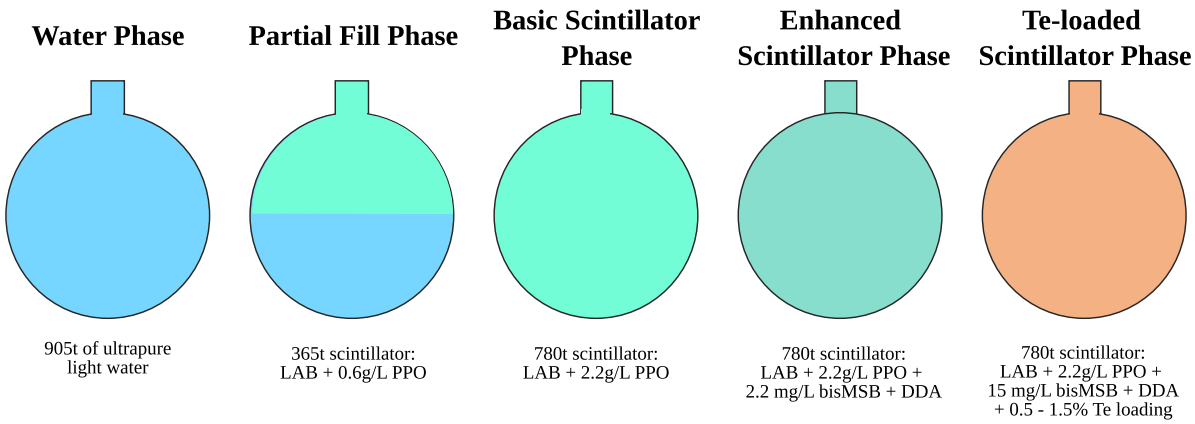


Figure 3.6: The phases of the SNO+ detector. Adapted from A. Inacio [54]

.

the effective attenuation of the acrylic vessel, and the angular response of the PMTs over a range of different wavelengths [48]. The ^{16}N was used to calibrate the energy, position, and direction reconstruction [49]. The AmBe source was used to measure the neutron capture cross section and the neutron detection efficiency in the detector [50].

In addition to the deployment of calibration sources, multiple physics results were obtained during the water phase. This includes a search for invisible nucleon decay [51], which set world leading limits in multiple decay channels, very low background measurements of the ^{8}B neutrino flux [52], and the first ever detection of reactor antineutrinos in a water Cherenkov detector [53]. A helpful figure for remembering the different phases can be found in Fig. 3.6.

3.4.2. Scintillator phase

Following the water phase, in 2019 the SNO+ detector then began to fill with a LAB + 0.6 g/L PPO cocktail, where 0.6 g/L means 0.6 grams of PPO per liter of LAB. This filling was interrupted by the COVID pandemic in 2020 when there was ∼365 tons of scintillator in the detector, meaning that the detector was stuck in a partially filled state with a water-LAB interface ∼75 cm above the AV equator. This detector configuration is referred to as the *partial fill* phase.

During the partial fill phase, several months of stable data was recorded between March and September of 2020, which was subsequently used in multiple physics analyses. This included the first

measurement of backgrounds relevant to the $0\nu\beta\beta$ phase, since many of the relevant backgrounds could not be observed during the water phase. A limited sensitivity measurement of Δm_{21} was also made during the partial fill phase [55].

The scintillator fill concluded in 2021 with 780 tons of LAB loaded into the detector. Data was taken with the 0.6 g/L PPO cocktail from April to June of 2021. Then in June of 2021, additional PPO was added to the scintillator cocktail bringing the PPO concentration up to 2.2 g/L. Data was then collected with the 2.2 g/L cocktail from April 2022 to June 2023. This configuration of 0.6 g/L or 2.2 g/L LABPPO without any water in the AV is referred to as the *basic scintillator phase*, and in this work references to the basic scintillator phase will always mean the 2.2 g/L scintillator phase unless 0.6 g/L is explicitly mentioned.

The addition of 1,4-bis(2-methylstyryl)benzene (bisMSB) as a secondary fluor occurred from July to December of 2023, the resulting cocktail being LAB + 2.2 g/L PPO + 2.2 mg/L bisMSB. Prior to the addition of the bisMSB, 6.5 mg/L of Butylated Hydroxytoluene (BHT), a stabilizing agent, was also added into the scintillator as a anti-yellowing agent. As was expected, the BHT did not have any effects on the optics. At the time this thesis is being prepared, which is August 2026, this is the cocktail that is currently in the AV. Data has only been collected up until February of 2026 though, since the detector was shut down due to the discovery of a leak in the AV. Discussions regarding the plan for the repair of the leak in the AV are ongoing. This phase with bisMSB added into the LABPPO cocktail is referred to as the *enhanced scintillator phase* or the *bisMSB phase*.

The light yield in the 0.6 g/L scintillator phase was ∼300 hits per MeV, for both the partial and fully filled configurations. The addition of PPO to bring the concentration from 0.6 to 2.2 g/L PPO was supposed to substantially increase the light yield of the detector, but we believe that due to overheating of the added PPO the light yield slightly decreased to ∼275 hits per MeV. The addition of bisMSB was also supposed to increase the light yield, and indeed it did provide an approximate 50% boost to the light yield, giving a final hits per MeV of ∼390.

Physics results have been extracted from the 0.6 g/L, 2.2 g/L, and bisMSB scintillator phases. A

combination of partial fill data and 0.6 g/L scintillator phase data were used to demonstrate the first evidence of event by event direction reconstruction in a high light yield scintillator [46]. Long baseline measurements of Δm_{21} were also made during these phases, and measurements of the solar oscillation parameters are also ongoing.

The data analyzed in this thesis is all from the 2.2 g/L scintillator phase, thus much more time will be spent discussing the calibrations carried out during this phase and results of the low energy ^{8}B neutrino flux during this phase will also be presented.

3.4.3. Tellurium phase

The next phase of the SNO+ experiment is the tellurium phase, which will begin with the loading of ∼3.9 tons of natural Te or 0.5% Te fraction relative to the LAB, which translates to a loading of ∼1.3 tons of ^{130}Te. R&D on the Te loading process has shown stability up to higher loading concentrations, therefore the current plan is to try and load to higher Te concentrations as soon as possible following the loading to 0.5%. The goal final loading concentration is 1.5%. This phase will be referred to as the *Te phase*, where 0.5% or 1.5% will be specified when necessary.

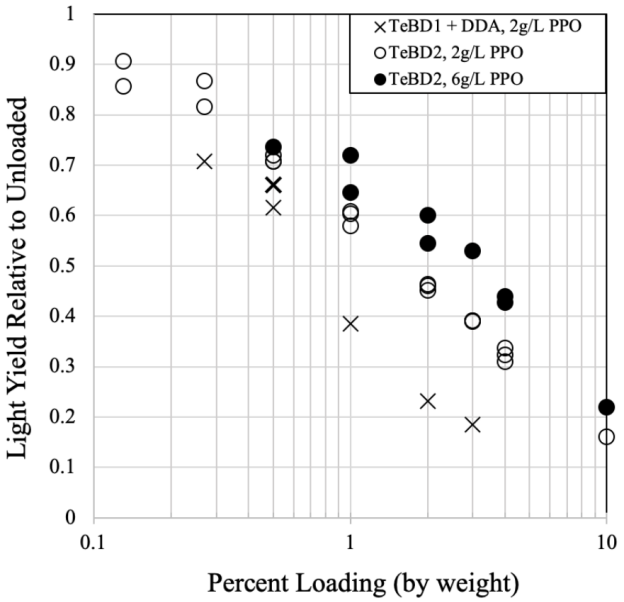


Figure 3.7: The change in scintillator light yield as a function of Te concentration. From [44].

Adding Te to the scintillator cocktail will decrease the light yield, which can be seen in Fig. 3.7

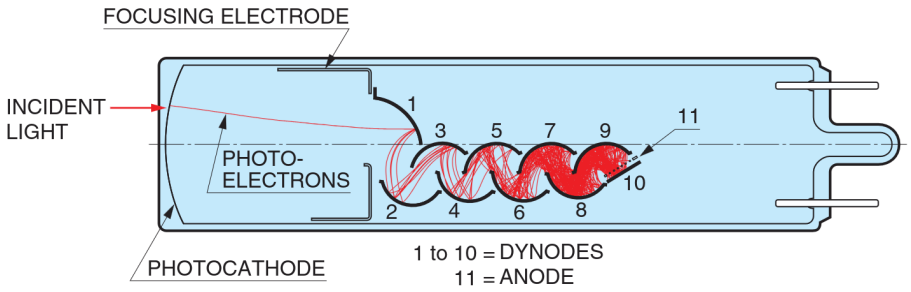


Figure 3.8: A schematic of a dynode chain like the ones found in the SNO+ PMT. From [56]

which is taken from [44]. Taking our current light yield and factoring in the reduction in light yield expected in the 0.5% and 1.5% Te phases, we expect to see around 260 hits per MeV in the 0.5% loading and 230 hits per MeV in the 1.5% loading.

This thesis will also present predictions of the sensitivity limits of the SNO+ detector in the Te phase.

3.5. SNO+ Electronics

Perhaps the best way to understand the SNO+ electronics is to travel along the path that converts a single photon into the "hit-level" information used in data analysis.

Assuming that a given photon is incident upon a PMT in the SNO+ detector and creates a photo-electron, the kV scale electric field inside the PMT will accelerate the photo-electron towards the dynode chain contained in the neck of the PMT. This dynode chain consists of many metal plates held at decreasing voltages, creating strong electric fields between one another. The initial photo-electron will eventually make contact with the first dynode in the chain, ejecting many electrons. These ejected electrons will then be accelerated towards the next dynode in chain, individually ejecting many electrons when each impacts the second dynode. This process is repeated for each dynode in the dynode chain ($\mathcal{O}(10)$), converting a single electron into around $\sim 10^7$, which are then finally absorbed by the final dynode in the chain. A schematic of a dynode chain can be seen in Fig. 3.8.

This number of electrons creates a measurable current on the scale of 100s of μA, which we can then

process to extract information regarding the photon that created the initial photo-electron, namely the start time and the charge of the pulse. PMTs have an inherent timing resolution, so if two photons arrive within a short period of time (short being on the order of ns), then the individual pulses created from each photo-electron are essentially just added together, therefore the charge of the pulse maps to the number of photons detected by the PMT.

At this point we are beginning to delve into the SNO+ electronics chain — note that a potentially helpful schematic showing the relationship between all the different components can be seen in Fig. 3.9.

The pulses from a set of 32 PMTs are all routed through a PMT interface card (PMTIC) and then to a set of 4 "daughterboards", which are mounted through surface pin connectors onto Front End Cards (FECs). The PMTICs are also responsible for supplying HV to the 32 connected PMTs, and contain 4 relays allowing for power to be turned off to groups of 8 PMTs. Each daughterboard processes the information from 8 PMTs in the analog domain using custom made ASICs. There are ASICs for both discrimination and integration, called the SNOD and SNOINT.

The SNOD is used to determine if the signal coming from a PMT crossed a per channel threshold, which implies that the PMT detected a photon. Upon crossing threshold, the SNOINT begins integrating the pulse using two independent gain paths, one being low gain and the other being high gain. The high gain path integrates the pulse over two different time ranges, $\sim$60 ns and $\sim$400 ns, respectively referred to as QHS (Q (charge) High-gain Short) and QHL (Q (charge) High-gain Long). The low gain path has one modifiable integration window and is referred to as QLX (Q (charge) Low-gain X). In principle, this allows for SNO+ to measure PMT charge over a wide dynamic range, and also provides some sensitivity to the time at which individual photons arrived.

The threshold crossing of the PMT pulse also begins what is referred to as the "TAC (time to amplitude converter) ramp" — a linear voltage ramp maintained by a custom "CMOS" chip on the daughterboards. In addition to handling the TAC ramp, the CMOS chip also stores the charge information recorded by the SNOINT in analog memory cells. The purpose of the TAC ramp will

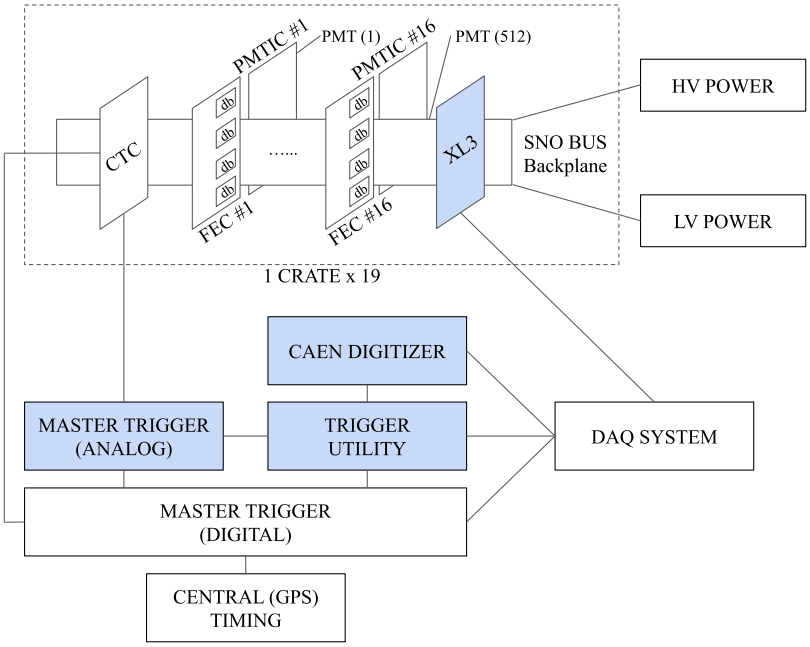


Figure 3.9: A schematic of the SNO+ electronics with components added during the upgrade from SNO to SNO+ in blue. Taken from T. Zummo [57]. Some of the components not involved during standard data taking will not be discussed in this section.

become clear shortly.

Even in pure darkness, PMTs will create pulses indistinguishable from those created by photoelectrons due to electrons spontaneously ejecting from the first dynode in the dynode chain. This noise is referred to as the "dark rate", and is typically on the order of kHz for standard PMTs. In order to try and differentiate between this constant noise and physics events within the detector, SNO+ employs a triggering system that only records data when many PMTs all record hits within a small time window.

At this point, we have discussed the local process of processing the hit information from a single PMT. But given that trigger decisions are driven by the global behavior of all the PMTs in the SNO+ detector, the per-PMT threshold crossing information must be passed along to some central authority to decide when data should be saved. Once the SNOD measures a PMT pulse that does cross the hit threshold, the CMOS chip will immediately begin to send some trigger information to the FEC. This includes two square waves indicating that the PMT was hit, one 20 ns wide and

another 100 ns wide referred to as N20 and N100, respectively. Two shaped copies of the PMT pulse are also created along low and high gain paths, referred to as ESUML and ESUMH, respectively. The FEC then sends these signals along a backplane to the crate trigger card (CTC). Speaking coarsely, the CTC will sum the PMT hit information from the 16 FEC cards within its "crate" (hence the name *crate* trigger card), adding together the N20, N100, ESUML and ESUMH from 512 PMTs. The SNO+ detector as a whole contains 19 of these crates in total.

After summing together the trigger information from its crate, the CTC then passes along these summed trigger signals to 7 MTCAs[2] (master trigger card - analog). The MTCAs sum the trigger signals from all 19 crates over three different gain paths (low, medium, and high), with one MTCA responsible for each trigger signal: N20, N100, ESUML, ESUMH, OWLN, OWLELOW, and OWLEHI. The OWL prefixed triggers are specific triggers for the OWL PMTs, or the outward looking PMTs. These PMTs face towards the water surrounding the SNO+ PMT support structure (PSUP), allowing them to collect Cherenkov photons produced in the buffer water which is useful for tagging muons. Each CTC is equipped with a dedicated path for the OWL PMTs, of which only some are populated. The N100, ESUML, and ESUMH from the dedicated OWL paths on the CTCs are then passed into a dedicated MTCA. These are the OWLN, OWLELOW, and OWLEHI respectively. Each MTCA compares its three analog sums against a configurable threshold, and upon a threshold cross it reports what is called a "raw trigger" to the MTCD (master trigger card - digital)

Specifically, the MTCA contains comparators that make the threshold crossing decisions, and then the outputs from the comparators are fed into a Xilinx CollRunner-II CPLD. Typically the CPLD just send the raw triggers on the output of the gain path that crossed threshold, but more complex triggering logic can be implemented on the CPLD, as will be discussed in Sec. 3.5.1

The MTCD is what makes the final decision on whether or not to emit a global trigger (GT). In principle, this decision can be based upon the raw triggers from any of the MTCAs, as well as a

[2]This board's true name is the MTCA+ as it features some modifications relative to the MTCA used in the SNO detector, but the "+" will always be implied in any reference to the MTCA in this work.

small number of separate trigger inputs that the MTCD supports. But during typical physics data taking, this decision to emit a GT is based upon a raw trigger from the low gain path N100. When the MTCD does indeed decide to emit a GT, the GT signal is sent from the MTCD to the CTCs in every crate, which then passes this signal along the backplane to each of its 16 FECs. The MTCD will also save information regarding which raw trigger caused the MTCD to trigger, a 24 bit global trigger identifier (GTID), and the time of the GT. The emission of the GT is the defining feature of an event in the SNO+ detector. The MTCD will also emit another signal called the lockout (LO) signal, which prohibits the MTCD from emitting another GT for 420 ns. This defines the length of the trigger window in the SNO+ detector, which is 420 ns.

When the FECs receive the GT from the MTCD, they will ask the daughterboards if they have any hit information from their respective PMTs. If they do, the FEC will then begin buffering this information into SODIMM cards present on each FEC. But since the timing of the trigger relative to when a PMT in the detector recorded a hit is arbitrary, one must be able to figure out when a PMT was hit relative to the emission of the GT.

Now we return to the purpose of the TAC ramp — since the ramp is linear it essentially acts as an analog counter proportional to the amount of time since the PMT was hit. Small TAC values indicate that the PMT was hit recently relative to when the GT was emitted, and large TAC values indicate the opposite. In terms of how the TAC corresponds to the trigger window, hits with large TAC values are early in the trigger window and events with small TAC values are late in the trigger window. For each hit PMT, the FEC will digitize and buffer the TAC, QHL, QHS, QLX, and GTID reported by the channel as well as some meta information about which PMT the information came from. Upon receiving a GT, each FEC will also increment its own GTID counter allowing for data in different crates to be mapped back to a GT. It is possible for FEC GTIDs and the MTCDs GTID to become out of sync, thus every 2^{16} events a signal is sent to the channels to synchronize the GTIDs across the entire detector.

Now an electronics minded individual might be wondering at this point what happens if a channel is hit, but there is no GTID forcing the readout of the channel for a very long time. Of course,

the analog buffers on the CMOS chip cannot store the charge and TAC information forever, and furthermore we care less and less about hits as the hit time grows further away from the trigger time. To account for these kinds of situations, the channels automatically stop the TAC ramp and clear all hit information ∼400 ns after a channel is hit.

Now the PMT level information from each of the FECs needs to be read out and bundled into a SNO+ event. The readout of the FECs is handled by the XL3, which asynchronously retrieves any buffered data from the FECs and then sends it over ethernet cables to servers that interface with all the XL3s and the MTCD. These servers are also responsible for retrieving trigger information from the MTCD, and then passing along this information to any client that requests it. The notable client of this server is the Event Builder, which is responsible for compiling all of the disparate and asynchronous information obtained from the FECs and MTCD/As into a sensible SNO+ event. This grouping of hits into events is done using the GTIDs reported by each channel when its hit information was read out. Perhaps the most commonly used phrase in the SNO+ collaboration is the term "nhit" which is just the number of hits grouped into each event.

During typical running data is acquired continuously, but we group hour long segments of data taking into "runs". Runs are given unique increasing IDs, which are typically incremented by one, but sometimes more in cases where the SNO+ experiment is switching between phases (Sec. 3.4). Run by run information regarding the detector state as well as any unusual detector conditions (like people working on the electronics or walking around on the deck) are also recorded. The run level distinction is mostly for bookkeeping, and on the processing side data is written out into individual 1GB files stored in a custom format, each of which is called a "subrun". During the most recent period of SNO+ data taking a run typically consisted of ∼15 subruns.

3.5.1. Deadtime and re-triggering

As was mentioned in the previous section, the MTCA through the CPLD can implement more complicated triggering logic. A dangerous class of background events for the SNO+ experiment are what we refer to as in-window or pileup events, or events where two pulses of light are generated within or nearly within one 420 ns trigger window. These events can have energies similar to the

$0\nu\beta\beta$ signal, potentially leading to mis-identification of some of these pileup events as $0\nu\beta\beta$ events. The main background producing these pileup events is ^{212}Bi^{212}Po , but some ^{214}Bi^{214}Po events will also have short enough time differences between the ^{214}Bi and ^{214}Po decays to end up being pileup events (discussions of the ^{214}Bi^{214}Po and ^{212}Bi^{212}Po backgrounds can be found in Sec. 4.3.1 and Sec. 5.1). Examples of the pileup vs non pileup events can be seen in Fig. 3.10

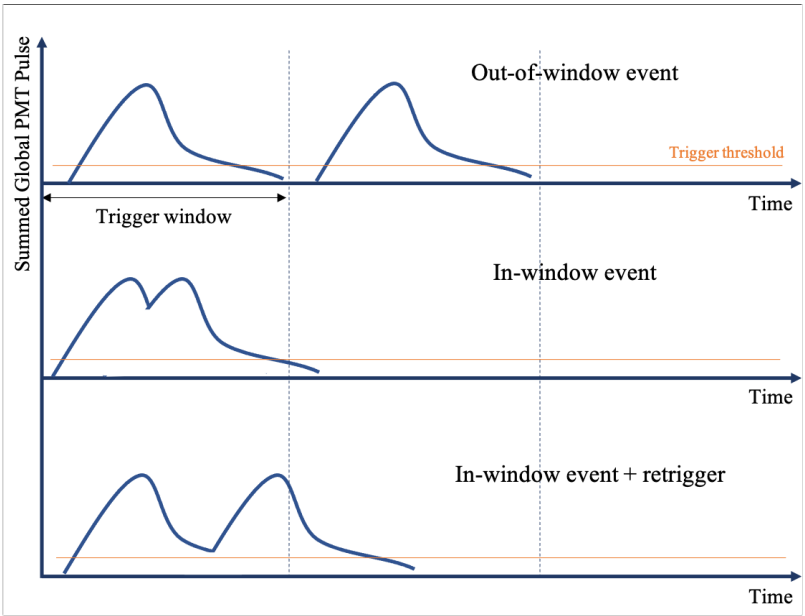


Figure 3.10: A cartoon representing different classes of pileup events based on the summed PMT pulses. Taken from A. Sofia [54]

Part of the reason pileup events present a dangerous background for the SNO+ experiment is because of the behavior of the trigger system. A common feature of a trigger based detector is that of deadtime, or a period of time where event or hit level information is unable to be read out by the electronics. In SNO+ there are two sources of deadtime, trigger deadtime and hit deadtime.

Trigger deadtime is due to the lockout (LO) signal, which is what prevents the MTCD from sending another GT for 420 ns after a trigger has occurred. The issue is that the GT takes around 200 ns to reach all the FECs in the detector, meaning that the hits that caused a GT to be generated will fall around 200 ns into the trigger window. Since the trigger window is 420 ns long, if another pulse of light is generated within the detector ~300 ns after the first trigger, then some of the hits from that pulse of light might not be read out since they are too late to fall in the trigger window

that already occurred, but too early to cause another trigger once LO has ended.

Hit deadtime on the other hand is due to the implementation details of how hits are mapped to individual events on the daughterboards. Once a hit occurs on a given channel, another bit called GTValid is set to high to inform the FEC that a channel has data that is ready to be read out. The GTValid must be shorter than the time over which the TAC ramp occurs, since a valid TAC must be available any time hit information can be read out. But GTValid must also be shorter than the LO signal since we do not want the same hit to be read out in two separate events. These requirements mean that there is always some delay between when LO ends and the TAC curve is reset, which implies that there is a small period of time in which a channel could be hit when the detector is ready to trigger, but the hit is not recorded due to the TAC curve not being reset yet.

To try and mitigate the effects of deadtime, a technique called "retriggering" was developed. Retriggering essentially forces the detector to emit a trigger as soon as LO ends to try and readout any hits that the detector might have otherwise not triggered on due to the deadtime effects described above. The conditions under which the detector will retrigger include what are called "automatic" retriggers, which are generated when events over a certain number of hits occurs in the detector. The other type of retriggers important for this work are LO retriggers and DGT retriggers. LO retriggers are generated if one of the outputs of the comparators on the MTCA is high at the end of LO, and DGT retriggers are generated if comparator is high during LO but after a delayed global trigger (DGT), or delayed copy of the trigger signal, has already gone from high to low.

Typically low level nhits cuts are applied on these retrigger windows to reject any events that have abnormally large numbers of hits in the retrigger window, since most physics analyses always expect clean retrigger windows following any physics events.

CHAPTER 4

Simulation and Reconstruction in SNO+

The main focus of this chapter is to describe how reconstruction algorithms can take the low level information recorded by the PMTs in the SNO+ detector and provide estimates of interesting quantities like interaction vertex times, positions, directions, energies, etc. An invaluable tool for the creation and evaluation of these algorithms is Monte Carlo (MC) simulation. Therefore before our discussion of the reconstruction techniques employed by the SNO+ detector, we will discuss the simulation framework developed and used by the SNO+ collaboration in Sec. 4.1.

The SNO+ collaboration has also developed a number of algorithms generally useful for every analysis. This not only includes algorithms used to reconstruct quantities computed in essentially every scintillator detector, like the position (Sec. 4.2.1) and energy (Sec. 4.2.2) of events, but also more SNO+ specific techniques which are referred to generally as classifiers (Sec. 4.2.3). Additional calibrations and corrections applied to the simulation and the data are discussed in Sec. 4.3.

4.1. Monte Carlo simulation of the SNO+ detector

Analysis of the data collected by the SNO+ experiment depends heavily on the use of Monte Carlo (MC) simulations, as well as general IO and analysis code. The simulation and analysis library used by the SNO+ experiment is called SNO+ RAT, which is a heavily modified version of the original RAT [58] software package created by S. Siebert for the the Braidwood Collaboration.

RAT is essentially a wrapper around the Geant4 particle physics simulation library [59], adding in a useful database system, overhauling the macro interfaces provided by Geant4, as well as overwritting some of the default physics behaviors of Geant4. The database system, which is called RATDB, allows for runtime fetching of information pertinent to almost every facet of the SNO+ detector: run state information, optics and physics tables, detector geometry information, etc. The databases can be overwritten locally or synchronized against remote databases, allowing for simulations produced on any machine with an internet connection to generate accurate simulations

for specific data runs. The changes to the macro interface supplied by Geant4 allow for changes to be made to the database within macros themselves. The overhaul to the macro interface also allows for geometries to be created within `geo` files, which are text files with line by line instructions for constructing a detector geometry from simpler volumes, which is often much preferable to writing C++ code to generate geometries like one does when using Geant4 directly.

Overwriting physics processes is another important functions of RAT since we often want to modify the behaviors of the processes that are relevant to our specific experiment. SNO+ RAT notably implements modified scintillation and Cherenkov processes as well a custom positiron annhilation process to correctly handle the formation of ortho-positronium bound states.

RAT also implements a number of custom event producers (which are also sometimes called generators) and processors. Examples of important producers include the electron gun, the Decay0 generator which simulates a number of complicated radioactive decays, and the AmBe generator, which simulates events generated by the Americium Berillium calibratation source used by the SNO+ experiment. The producers essentially act as an interface between the RAT user and Geant4 by generating vertices according to the RAT macro, and then passing them off to Geant4 for the actual physics simulation of those vertices. Typically, the relevant result of the Geant4 simulation is the detected photons, which are then passed back to RAT and ran through a custom, and very thorough, simulation of the SNO+ DAQ system. The simulation of the DAQ outputs information that mimics the real data that is produced by the detector. Of course, low level information regarding the vertices and other information handled by Geant4 can also be retrieved through RAT as well.

Following the simulations of the DAQ system, the output event information can then be passed to the RAT processors, which convert the hit-level information collected by the PMTs into values regarding the position, energy, etc. of the event. Since the simulation of the DAQ and the real DAQ spit out event information the same format, the data can be processed essentially the same fashion as the MC, except event information is read from the data files rather than being supplied by Geant4 simulations. The specific algorithms that go from the hit time information to more high

level quantities like position and energy will be discussed in the following section (Sec. 4.2).

4.1.1. Run by run productions

One of the benefits of the RATDB is that simulations can be generated according to the run conditions of the very runs one is processing in an analysis. Since many analyzers all use the same sets of runs, the SNO+ collaboration will run large scale "productions" which generate simulated datasets of the expected backgrounds in the detector. More information on these expected backgrounds can be found in Sec. 5.1

An important part of the detector state in regards to reconstruction performance is the number and distribution of online PMTs in the detector. The RATDB does keep a record of exactly what PMTs are on and off in a given run, and therefore these run-by-run productions should in principle recreate the effects of PMTs being turned off during data taking.

4.2. Reconstruction

Before we dive into the specifics of each reconstruction algorithm, it is worth pointing out some of the general assumptions made by the reconstruction algorithms used by SNO+. While MeV scale electrons will typically produce photons along their entire cm to few-cm length track as they stop, the reconstruction algorithms all either implicitly or explicitly assume that photons were created at a single point, unless otherwise stated. While this is wrong, we are not very sensitive to the specifics of the electron track because of the imperfect timing of the PMTs, optical effects within the detector, and the multiple-scattering that occurs as the electron propagates.

Another key assumption is that the position and energy reconstruction algorithms all work under the assumption that they are reconstructing a single electron vertex. Since heavier particles like alphas and protons have a much different ratio mapping emitted photons to particle kinetic energy, the energy reconstruction will report energies much lower than their true energy. Another way to phrase this is that we always report the energy of an event in terms of the electron equivalent energy, or the energy an required by an electron to create that number of photons. This effect of heavier particles creating fewer photons per unit of deposited energy is referred to as *quenching*

(Sec. 3.3). The single electron vertex assumption also fails when we reconstruct gammas, as a gamma can Compton scatter multiple electrons centimeters apart from one another as it moves through the detector.

It is important to understand that while we carry out the reconstruction under incorrect hypotheses, we are not completely insensitive to things like the event's true track length and particle type. Depending on the validity of the single vertex electron hypothesis, the performance of the position reconstruction changes and this provides a handle of how "electron"-like an event looks — and we can use this information in our analyses. This idea will be discussed more in depth towards the end of this chapter.

4.2.1. Position and time reconstruction

The first quantity reconstructed for a triggered event during SNO+ processing is the time and position (t, x, y, z), as almost every other reconstruction algorithm requires a position as an input. There are many different approaches one can use to extract a position from an event in a scintillator detector. Position reconstruction typically makes use of the time of flight differences of photons detected by different PMTs, as well as the charge deposited on each of the PMTs. In the limit that a detector is saturated with PMT hits, meaning that each PMT is collecting many photons, the arrival times of individual photons become difficult to measure, and charge based position reconstruction techniques that take the charge weighted average of the hit PMT positions can provide satisfactory results.

The SNO+ detector on the other hand mostly operates in the opposite limit, where hit PMTs collect mostly one photon per event and rarely see two. This means SNO+ can leverage the timing information of the individual photons to compute the position. In this section we will actually discuss two distinct approaches which are used in tandem in the SNO+ fitting pipeline. The first algorithm is called the "quadfitter" and the second is referred to as the "position likelihood fitter".

Quadfitter

The general idea behind quadfitter is that in the limit all the photons are emitted at the same time and take straight line paths from the particle's position to the PMTs, we have the following

$$r_\mu - v_{i\mu} = d_{i\mu} \;\text{ with }\; d_i{}^\mu d_{i\mu} = 0 \tag{4.1}$$

where i indexes over all the N hit PMTs in an event, r_μ is the true vertex position 4-vector, $v_{i\,\mu}$ is the hit PMT 4-vector. The second equality is a direct consequence of the zero length paths of light-like intervals in Minkowski space-time. Now since all the photons are not actually emitted at the same time due to the extent of the particle track, optical effects during propogration, and effects from the PMT timing uncertainty, $d_i{}^\mu d_{i\mu} = 0$ does not hold exactly. We can also note that $d_i{}^\mu d_{i\mu} = 0$ is really a coupled system of 4 equations for each hit PMT i, thus we cannot expect to find a solution in general since it will typically be highly over-constrained if we have more than 4 PMT hits.

The quadfitter algorithm avoids this issue by randomly selecting a subset of 4 PMT hits from the total set of hit PMTs. Since this subset consists of 4 $d_{i\mu}$ intervals, we have a fully determined system of equations and can analytically compute the event vertex that results in zero-valued space time intervals. Given infinite compute capability, one could do this for all $\binom{N}{4}$ combinations of PMTs but SNO+ frequently records events with 100s to 1000s of hits, and $\binom{500}{N}$ is already $\sim$2 billion combinations. Instead, the vertex will be typically computed for $\min(10000, \binom{N}{4})$ combinations of PMTs, and the vertex from each of these combinations is recorded. Each of this individual vertices is called a "quad-point" and the collection of them all together is the "quad-cloud". The median of the values in each dimension of the vertices is taken as the reconstructed vertex. The variance of the values in each dimension can also be used as a rough estimate of the uncertainty in the reconstructed vertex.

This reconstruction algorithm does not perform as well as the position-likelihood fitter we will discuss in the next section, but the utility of quadfitter comes from the fact that it is robust,

minimization free, and requires minimal changes when switching between materials with different optical properties. For these reasons, we run quadfitter to compute the seed of the more complicated position-likelihood fitter.

Position-likelihood fitter

As the name suggests, the position-likelihood fitter takes a likelihood based approach to position reconstruction. For a discussion on maximal likelihood estimation, see Sec. A.1. We want to evaluate a likelihood function of the following form

$$-\ln(\mathcal{L}) = \sum_i \ln(p(x_i)) \tag{4.2}$$

where the $p(x_i)$ term is the probability of a photon hitting a given PMT at a given time, represented with the 4-vector $x = (x_0, x_1, x_2, x_3) = (t, x, y, z)$, and i runs over all the hit PMTs in an event. In principle, $p(x_i)$ also has a marginal term $p(x_i|...)$ where the $|...$ term represents all the different processes that could affect $p(x_i)$. This includes the optics of the scintillator, the energy of the event, the specific run conditions during which the event occured — but crucially for our purposes, the term also includes where in the detector the event occured, the true vertex r. Therefore we can hypothesize different vertices, r', and determine which maximizes the likelihood, or minimizes the negative log likelihood.

As was stated in the last paragraph though, there are many factors that affect the p_i. Rather than trying to analytically account for these effects we employ a few tricks. First we use our in-depth Monte Carlo (MC) model to try and predict the shape of the likelihood function through a process we refer to as "coordination". For the position-likelihood fitter, this step roughly involves measuring and calibrating the optical model, and then simulating hundreds of thousands of electron events at different energies and locations in the detector to build up the likelihood function.

The next trick is to then define the likelihood function in terms of the residual times, or the path length corrected hit times, rather than the raw PMT hit times. Specifically, the hit time residuals

are defined as

$$t_r = t_{\text{hit}} - t_{\text{fit}} - t_{\text{tof}} \tag{4.3}$$

where t_{hit} is when the photon hit the PMT, t_{fit} is the fitted time from the position reconstruction algorithm, and t_{tof} is the time of flight of the photon. This path length correction is a difficult process in general, since photons cross multiple boundaries on their way to the PMTs and can also be absorbed and reemitted or scattered during propagation through the bulk material. While we cannot correct for effects of absorption and reemission as they are stochastic processes, we can and do correct for the refraction effects as photons pass through the material interfaces. Furthermore, the coordination step should account for the effects of absorption and reemission in a statistical fashion in the likelihood function.

Defining the likelihood in terms of the hit-time residuals allows us to ignore the locations of the PMTs in the detector since the path length correction accounts for the photon time of flight. Thus we can write our likelihood function in a simpler way,

$$-\ln(\mathcal{L}) = \sum_i \ln\left[p(t_{r\,i}|r')\right] \tag{4.4}$$

which depends only on the hypothesized position and the ith time residual $t_{r\,i}$. Now that we have a calculable version of the likelihood, the position fitting process becomes simple. We can just use any minimization algorithm to minimize the negative log likelihood, where at each evaluation we re-calculate the hit time residuals given the hypothesized vertex. Currently the Powell algorithm is used to do the minimization.

While using the hit-time residuals mostly corrects for the position dependence of the hit time residual distribution, we do still see residual effects that deform the distribution both in different radial shells and different energy regimes. The more straightforward effect is the one observed in the different radial shells, which is due to dispersion, or the wavelength dependence of speed of light in water. Since longer wavelengths travel faster than shorter wavelengths and Cherenkov and scintillation processes both produce photons in a range of wavelengths, events near the center of

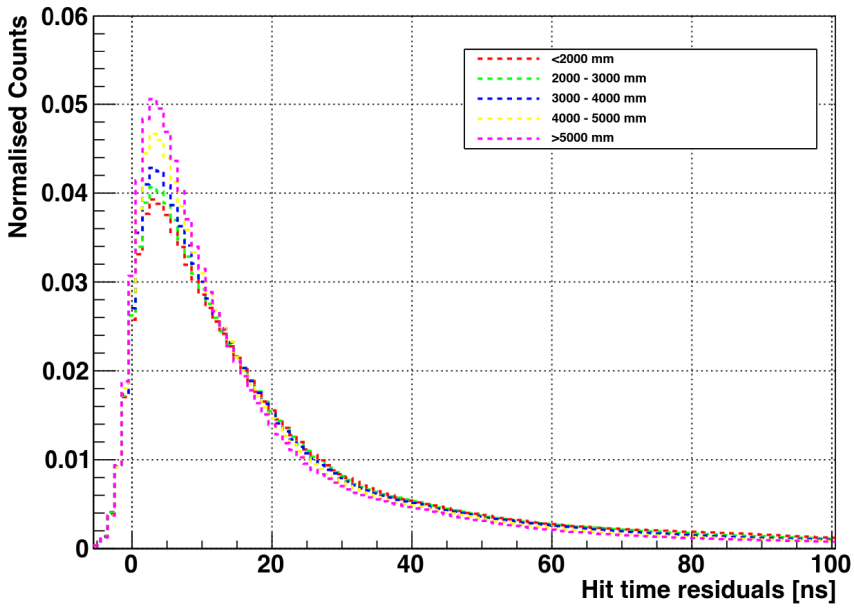


Figure 4.1: A plot showing how the hit time residuals change in different radial shells. A similar effect is seen 1-10 MeV energy range as well. Plot is courtesy of Will Parker [60].

the detector tend to have broader hit time residual distributions as the photons disperse as they travel through the bulk material. This effect can be seen in Fig. 4.1. One of the ways we try to correct for this effect is by generating the hit time residual distributions in different radial shells, and the one used during fitting is chosen based on the seed fed into the position-likelihood fitter.

As was mentioned above, we also observe that the hit time residual distributions become biased towards earlier times as event energies go up, which we believe to be due to the fact that we cannot measure the arrival time of any photons incident on PMTs that have already recorded a hit. Since higher energy events produce more photons, they are more likely to have multi-hit PMTs which will bias the hit time residual distributions towards earlier times as we only measure the arrival time of the first photon. This is effect is corrected for by changing the "effective" speed of light used when calculating the hit time residuals based on the number of hits in an event, since decreasing the speed of light introduces a bias in the hit time residuals that offsets this energy dependent bias. Both the dispersion and the multi-hit PMT effects manifest as radial biases in the position-likelihood reconstruction, which can be mostly corrected for using the approaches above.

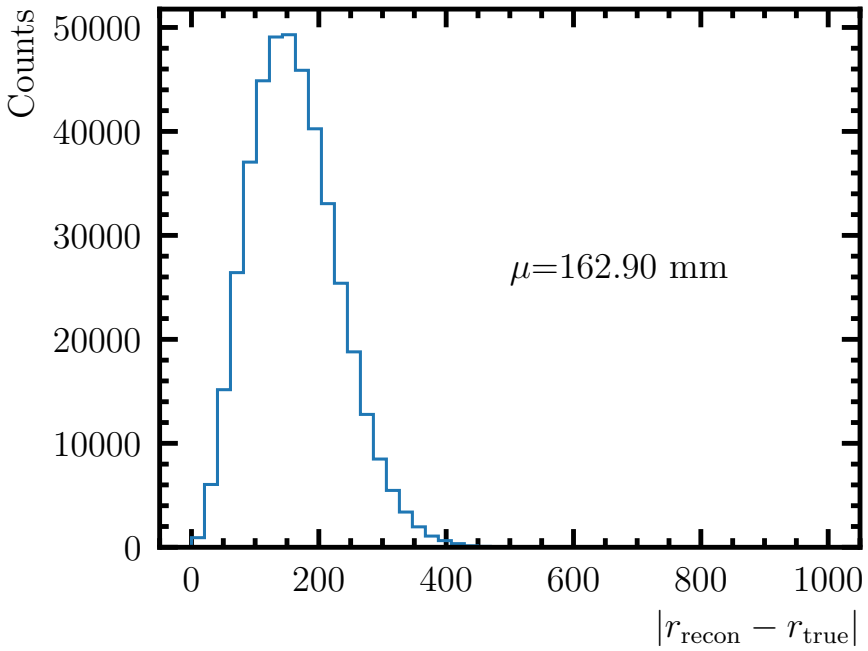


Figure 4.2: A plot showing the length between reconstructed and true positions of simulated $0\nu\beta\beta$ events (E∼2.5 MeV) in 2.2g/L LABPPO. Most events reconstruct within a few 10s of cm of their true position.

The remaining radial biases are on the order of $\sim 1\%$.

The performance of the position-likelihood fitter when used to reconstruct $0\nu\beta\beta$ events in 2.2 g/L LABPPO can be seen in Fig. 4.2.

4.2.2. Energy fitting

In it's simplest form, energy reconstruction in scintillator detectors can be done by taking the number of detected photons in an event and equating that to an electron equivalent energy using simulation or calibration data. Furthermore, the conversion between collected photons and electron energy is essentially linear, making this conversion easy to extrapolate.

This approach can be directly improved upon by also accounting for detector and optical effects by using simulation to come up with a mapping that takes into account how the location of an event in the detector correlates with the number of detected photons. This is essentially the energy reconstruction approach employed by SNO+, accomplished through an algorithm called the

"EnergyRThetaFuntional".

In SNO+, there is one additional complication due to the poor charge response of the PMTs used in the detector. The charge response forces analyzers to rely on the number of hit PMTs in the event rather than the number of collected photons, which means that a PMT that sees one photon and a PMT that sees 10 photons are treated identically in the sense that they both contribute one "hit" to the energy reconstruction. Fortunately, events with energies less than 10s of MeV are in the regime that most PMTs see only 1 photon, meaning this is a small and correctable effect.

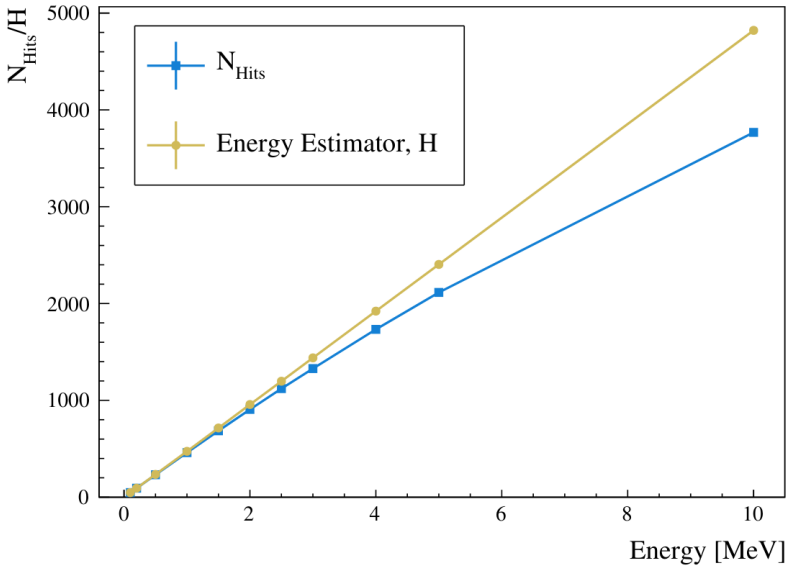


Figure 4.3: The H parameter (Sec. 4.2.2) is able to effectively reproduce an estimator that is linear in energy. From [61].

The EnergyRThetaFuntional accounts for this multi-hit effect by splitting the detector into S regions of approximately equal PMT solid angle, and defines a parameter that is linear in energy H:

$$H = -\sum_i N_{S\,i} \log\left(1 - \frac{N_{H\,i}}{N_{S\,i}}\right) \tag{4.5}$$

where $N_{S\,i}$ and $N_{H\,i}$ are the total number of PMTs and hit PMTs in segment i, respectively. H being a linear estimator of the energy is due to the fact that the number of photons collected by a PMT in any of the segments will roughly follow a Poisson distribution, which is centered around a

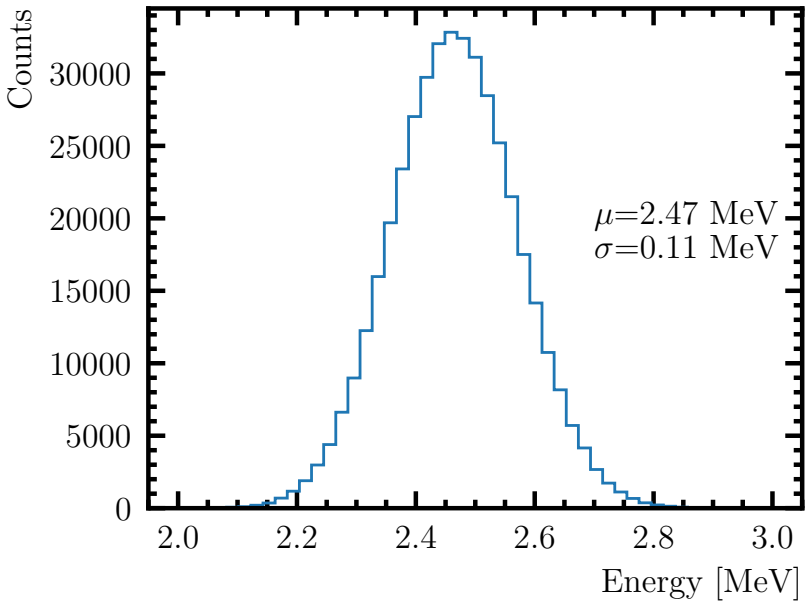


Figure 4.4: Reconstructed energies of $0\nu\beta\beta$ events in the 2.2 g/L LABPPO optics.

shared mean for all the PMTs in a given segment. So this H parameter essentially tries to account for our ignorance about the number of multi-hit PMTs based on the number of hits recorded by the rest of the PMTs in any of the given segments.

The next step is to then try and account for how the position affects the relationship between detected photons and true event energy. This is accomplished in effectively the approach described earlier in this section, where simulation is used to build a mapping from position to collected photons to account for the detector and optical effects. The EnergyRThetaFuntional uses one additional trick though; a piecewise function constructed from two continuous functions fit to the position dependence predicted by the MC is used to apply this correction. Using continuous functions rather than a simple lookup table ensures that no issues with discontinuities will bias the reconstruction results (aside from at the point where the two functions are stitched together). A piecewise function is used because the relationship between position and collected photons takes a different qualitative form past a radius of about 5.2 meters, so a different function must be used for this mapping at high radii. The mapping function is parameterized in r and θ as there is an azimuthal dependence to the detectors response due to the neck of the AV.

The performance of the energy reconstruction when applied to $0\nu\beta\beta$ events in the 2.2 g/L LABPPO optics can be seen in Fig. 4.4. A much more detailed evaluation of the energy bias and resolution will occur in the following sections on the calibration of the detector response.

4.2.3. Classifiers

In addition to the reconstructed quantities above, the SNO+ processing pipeline will also apply a number of "classifiers" to events to try and distinguish noise from physics events, and additionally distinguish between different classes of physics event types.

In time ratio (ITR)

Perhaps the classifier with the highest ratio of utility to simplicity, the in time ratio (ITR) simply calculates the ratio of hit time residuals in a -2 to 5 ns window relative to the total number of hits in an event. By definition, this value will range somewhere between 0 and 1, but is typically much closer to 0.25. This classifier is very effective at removing noise events, as noise events often have a much flatter hit time residual distribution than real physics.

Aside from noise events, any type of event does not fall into the single vertex electron hypothesis that the position reconstruction assumes will have an ITR value that deviates from what we'd expect to see from electrons. This means that ITR can distinguish between pure β events from β/γ or γ events since the β/γ and γ events are not well modeled by the single vertex electron hypothesis.

Furthermore, some of the most dangerous background in the SNO+ detector are what we refer to as in-window or pileup events, or events that consist of more than one decay in the 420 ns trigger window. The main source of these pileup events is the short lived ^{212}Po isotope which has a half life of 299 ns. In the cases where the ^{212}Po decay is more than a few ns after the ^{212}Bi decay, the double peaked nature of the hit time residual distribution results in a distinct ITR distribution for this class of events. A plot comparing the ITR distributions of different event types can be seen in Fig. 4.5.

Similar to energy, extensive ITR calibrations are done to not only to improve the data-MC agree-

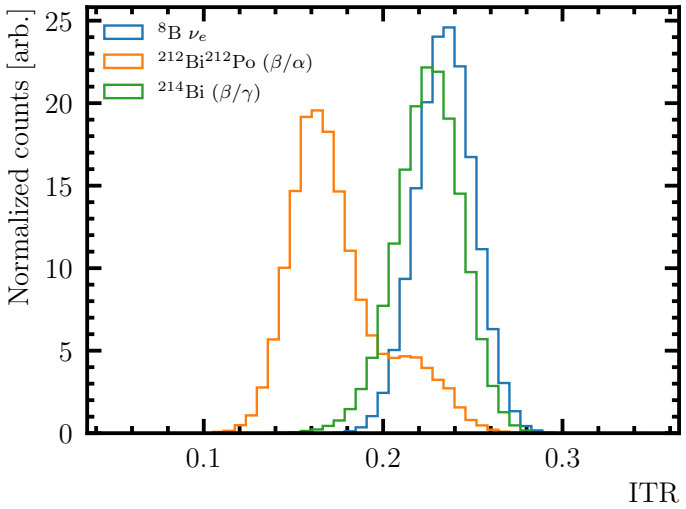


Figure 4.5: A plot comparing the expected ITR distributions of three distinct event classes. The ^{212}Bi^{212}Po are mostly in window β/α coincidences, giving them a distinct double peaked hit time residual distribution, resulting in a very different ITR distribution. ^{214}Bi events on the other hand are β decays that result in γ cascades that which occur on pico-second time scales relative to the initial β emission. Since the γ's will Compton scatter electrons 10s of cm away from the primary β emission, these events result in hit time residual distributions that are slightly broader than those from ^{8}B ν_e induced electron scatters, which are comprised of only one pure β.

ment, but to also limit and understand any potential systematic effects arising from ITR mismodeling. These calibrations will be discussed in the next section.

β_{14}

Another useful classifier is β_{14} which measures the isotropy of an event. Events during the scintillator phase should be very close to isotropic, and any events which are not isotropic are typically not physics events. Thus β_{14} is a very useful classifier when wanting to remove unwanted non-physics events from a dataset.

posFOM

The "position figure of merit" or posFOM is the depth of the likelihood space of the position-likelihood fitter at the reconstructed vertex. This is an effective cut as it essentially quantifies "how well" a specific event was modeled by the single electron vertex hypothesis, and therefore the hit time residual PDF. Since the position-likelihood fitter uses every single "cleaned" hit during reconstruction, this number needs to be scaled by the number of cleaned hits in an event to remain energy

independent. Therefore cuts on posFOM are typically implemented using posFOM/nhitsCleaned.

4.3. Calibrations and corrections

The inclusion of any additional backgrounds in the SNO+ detector can directly impact the sensitivity of the $0\nu\beta\beta$ measurement. For this reason, the collaboration has been extremely cautious regarding the deployment of internal sources as this is a process that will increase the backgrounds rate of the detector, we can only mitigate the extent to which the backgrounds will increase.

Because of this abundance of caution, no internal calibration source were deployed during the basic scintillator phase — the first internal calibration source deployment occurred during the enhanced scintillator phase. But this does not mean there is no way to calibrate the detector response during the basic scintillator phase, instead of relying on deployed calibration sources we can use tagged internal sources to calibrate the detector.

4.3.1. ^{214}Bi^{214}Po coincidence tagging

The internal background used for calibrations in this work is ^{214}Bi, which can be efficiently tagged in coincidence with the decay of its daughter isotope, ^{214}Po. The half life of ^{214}Po is 164.3 μseconds, which is already short relative to the trigger rate of the detector ($\sim$2kHz) and by utilizing coincidence cuts to select pairs of events that specifically look like ^{214}Bi^{214}Po pairs we can obtain very pure sample of these decays.

Typically these cuts involve selecting events such that there is a prompt event in the appropriate energy range for ^{214}Bi, that is close in both time and position to a delayed event within the appropriate energy range for ^{214}Po. It is important to recall that the ^{214}Po decay is an α decay, and the conversion of kinetic energy to photons in the scintillator is "quenched" (see Sec. 3.3.1), So even though the ^{214}Po α is emitted with 7.83 MeV of kinetic energy, the observed energy corresponds to an electron with closer to 0.9 MeV of energy. Energy distributions as well as timing and position separation distributions for prompt and delayed candidates can be seen in Fig. 4.6.

While we can obtain a pure sample of these events, as is corroborated by the agreement in the Δt plot in Fig. 4.6, there are still some difficulties presented by using tagged internal backgrounds

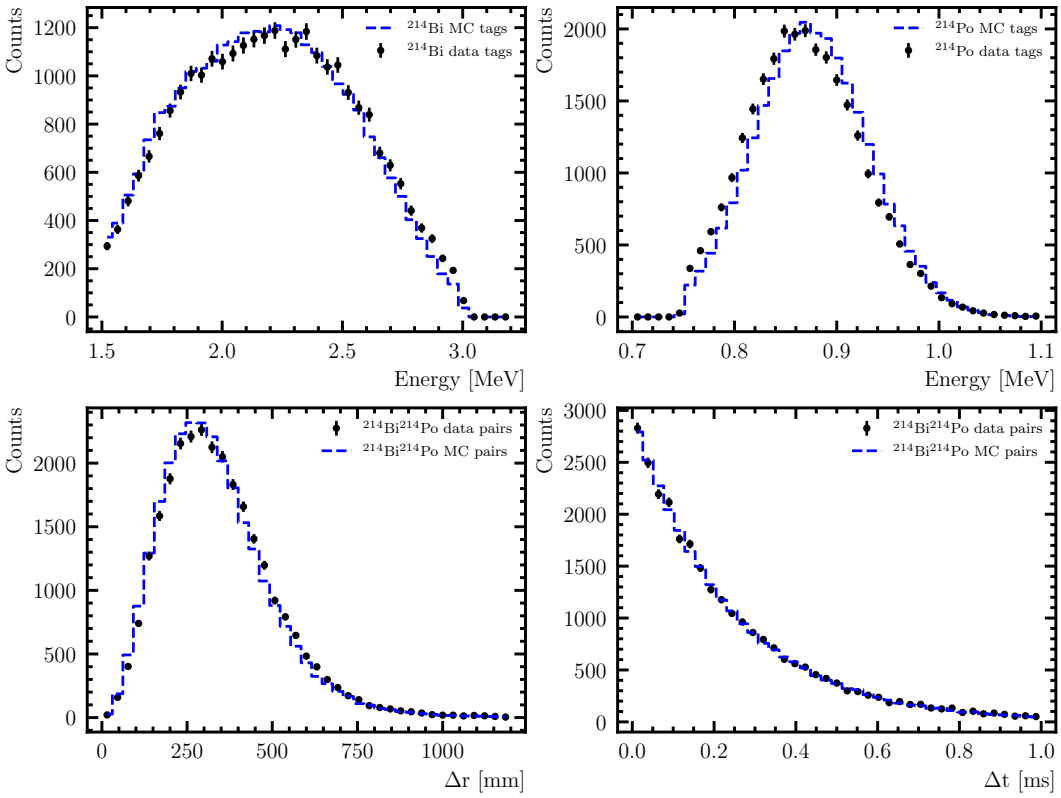


Figure 4.6: Various distributions of tagged ^{214}Bi^{214}Po pairs from the basic scintillator phase. Candidates from tagging on both data and simulation are shown. Δt and Δr are the separation between the events in position and time, respectively. There are small but clear systematic disagreements between the energy distributions for both the prompt and delayed events, which the calibration procedures discussed in this section are intended to fix. The tagging criteria used to select these events can be found in Tbl. 4.1, aside from a more aggressive r cut that only accepted events out to 4.5 m.

for calibration. Since so much effort is put into keeping the detector radiological "clean", we are typically statistically limited in the constraints we can place on systematics. Furthermore, we do not know any truth level information regarding the backgrounds unlike how one would know the position of a deployed source.

But what we do know is that for coincident backgrounds like the ^{214}Bi^{214}Po pairs, the two decays occur in the same location of the detector since the momentum passed to the nucleus is very small for β decays, so the difference between the Bi and Po locations is a good measure of our position reconstruction performance.

As can be seen in Fig. 4.6, the out of the box simulations are not in perfect agreement with data. The following section will discuss the techniques used in to try and minimize the systematic disagreement between data and simulation using this sample of tagged ^{214}Bi^{214}Po pairs.

4.3.2. Scintillator timing calibrations

Perhaps the most important part of the model calibrated with tagged ^{214}Bi^{214}Po decays is the scintillator timing response. Since photon emission from scintillators is driven by the excitation and deexcitation of different electronic and vibrational modes, there is also some inherent time delay between the charged particle exciting the scintillator and the scintillator emitting photons. This time delay is also very difficult to know a priori since the half lives of these excited states are difficult to calculate. Furthermore, since most scintillator cocktails use secondary fluors in addition to the primary scintillator, the interactions between the primary scintillator and secondary fluors and how that relates to photon emission can be very hard to model microscopically. This means that similar scintillator cocktails will often have slightly different emission time profiles, and therefore experiments must manually calibrate this time profile to for their own scintillator.

To carry out this calibration of the scintillator timing response, we take a pure sample of ^{214}Bi events from data and compare their hit time residual distributions to that of an MC generated sample of ^{214}Bi events. This process can then be repeated with ^{214}Po events to tune the α timing of the scintillator. In general, the MC and data distributions will not agree, so a different combination of parameters describing the scintillator emission time profile will be used to generate a new set of simulations, and the hit time residuals generated from this new set of parameters will then be compared to the hit time residuals found in data. This process is repeated over and over again until we find a combination of parameters that provides the best agreement between the data and MC. Fig. 4.7 shows the results of this calibration process.

The function used to describe the emission time profile of the scintillator is

$$f(t) = \sum_i \frac{A_i}{\tau_i - \tau_r} \left(\tau_i (1 - e^{-t/\tau_i}) + \tau_r (e^{-\tau_i/\tau_r} - 1) \right) \tag{4.6}$$

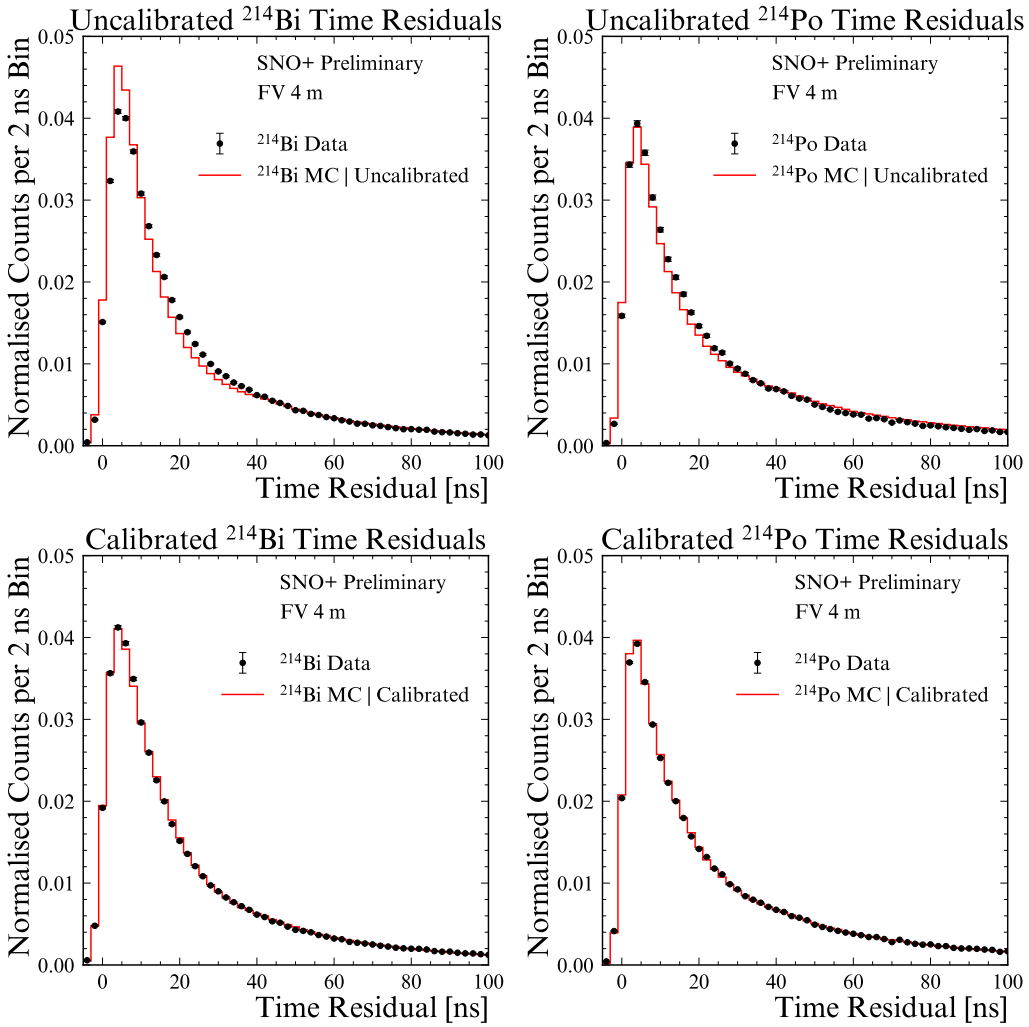


Figure 4.7: The hit time residual distributions of internal ^{214}Bi and ^{214}Po events before and after tuning the scintillator emission time profile.

where A and τ are vectors of amplitudes and decay times respectively, and τ_r is the risetime of the scintillator. Any number of elements can be placed into the vectors A and τ, but typically only 3 or 4 parameters are needed to effectively model the measured waveforms.

4.3.3. Energy and ITR corrections

Of course, the fitters used by the SNO+ experiment are not perfect, and reconstructed quantities will often have measurable biases between the data and MC. This is another area where we can rely on the ^{214}Bi events to help understand the detector response.

The cuts used to select the ^{214}Bi^{214}Po pairs used in this analysis are shown in Tbl. 4.1. As can be seen in the table, the events are split into two different groups based on run number to create separate "calibration" and "testing" datasets.

Cut	Prompt Event	Delayed Event
Valid Fit	True	True
Data Cleaning	Pass	Any
Radius	r $\leq$ 5.5 m	r $\leq$ 5.5 m
Energy	1.2 $\leq$ E $\leq$ 5.0 MeV	0.75 $\leq$ E $\leq$ 1.2 MeV
^{214}BiPo Tag	True (prompt)	True (delayed)
Hotspot Tags	False	False
Atmospheric $\bar{\nu}$ Tag	False	False
Δt	1 μs $\leq \Delta t \leq$ 1 ms	
Δr	$\Delta r \leq$ 1 m	

Calibration Runs	Testing Runs
300000 - 300701	$\geq$ 300702

Table 4.1: *Top:* The cuts applied to the ^{214}Bi^{214}Po pairs. The details of each of these cuts will be discussed in a later section. *Bottom:* The run ranges used for calibration and testing applied to the ^{214}Bi^{214}Po pairs

Once we have a selection of pure ^{214}Bi candidates in both data and MC (which is acquired by only taking the prompt events in ^{214}Bi^{214}Po pairs), we can calculate the shift or scale correction required to try and correct for the radial and z bias present in some parameters. Note that we are not correcting the data to the MC or vice versa, we are rather correcting the data and MC both to some arbitrary value in order to account for radial and z biases in a parameter. The two parameters for which it is important that we carry out this procedure are energy and ITR, as we will be explicitly fitting in these dimensions in this work. We first build and apply the correction in the radial dimension, but this leaves a residual bias in the z-dimension. Thus we apply a similar correction in z following the radial correction.

To build the correction map we slice the detector in radial or z slices, and then fit the mean of the resulting distribution with a Gaussian. Since both energy and ITR are roughly normally distributed for the ^{214}Bi events, this parameterization is acceptable for our requirements. The means of the

energy and ITR distributions before and after being corrected can be seen in Fig. 4.8.

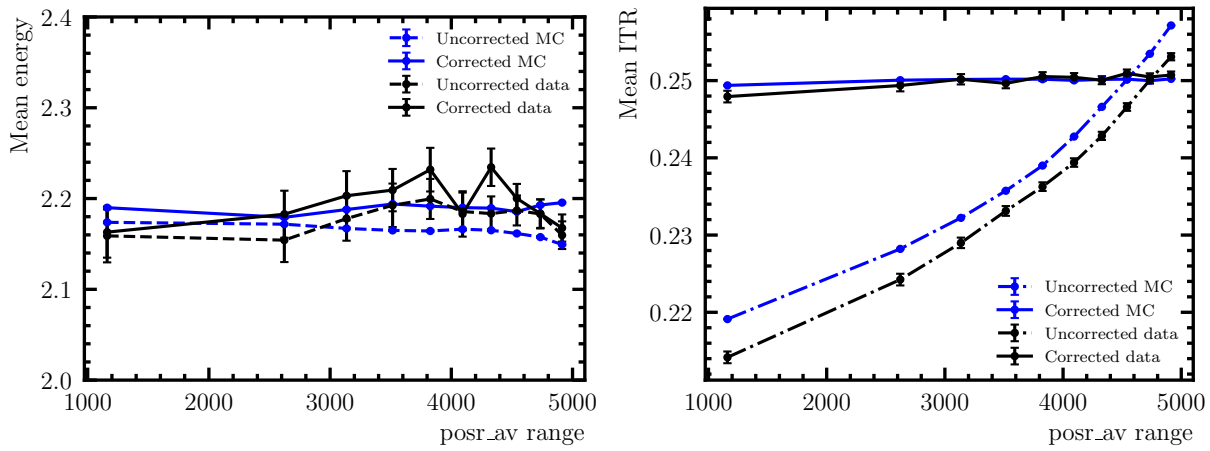


Figure 4.8: Mean energy (left) and ITR (right) distributions of tagged ^{214}Bi events before and applying the data driven corrections.

The energy correction is applied mostly to remain consistent with other solar neutrino analyses which go out to higher radii where the energy reconstruction picks up more of a bias, so the smaller effect of the energy correction is expected. The impact of the ITR correction, however, is immediately apparent.

Note that the mean value the ITR distributions are corrected towards is completely arbitrary, a value of 0.25 was chosen as it is close to the average uncorrected ITR for events with radii under 5 meters. The binning used in the radial shells is uniform in r^3 in order to provide equal volume bins. Direct comparisons between data and simulation in energy and ITR after applying these corrections will be presented in a later section.

4.3.4. Acrylic vessel offset correction

Another correction that must be applied to the data is the AV offset correction. As the material inside the AV has changed over the different phases of the SNO+ experiment, the buoyant force acting on the AV has changed as well. This has caused the AV to move around as the tension in the hold-down and hold-up ropes change, especially in the z-direction.

This is in contrast to the PSUP, which has remained in the same position over the entire course

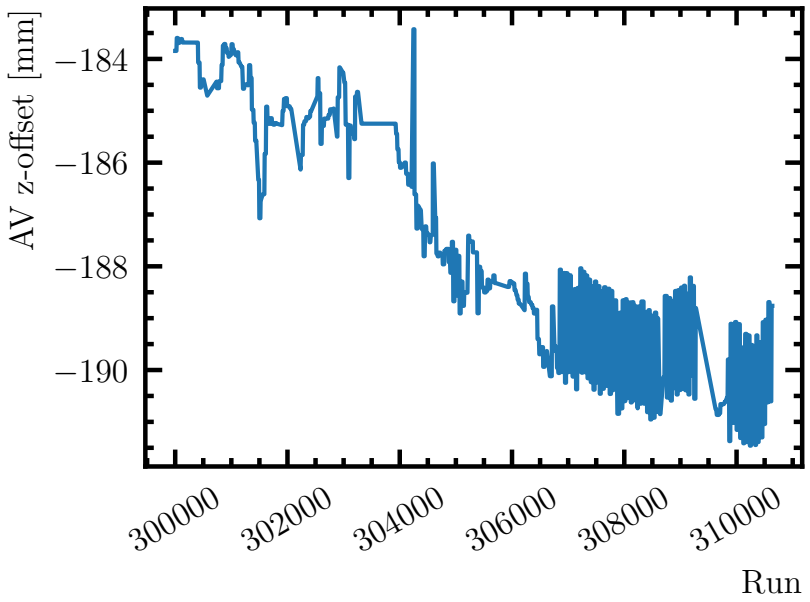


Figure 4.9: AV z-offset over the basic scintillator phase.

of the SNO+ experiement. Since all of the observations we make of the events inside the detector come from the PMTs, the positions we reconstruct are all in terms of the PMT coordinates. The physics we are interested in, on the other hand, is more usefully parameterized relative to the AV coordinates. Thus we are required to correct for the relative offset between the PSUP and the AV, and as we mentioned above this is something that can change from phase to phase.

Thankfully, we do monitor this effect using a sensor that tracks the position of the AV relative to the reference frame of the PSUP. These measurements of the AV offset are then applied to the reconstructed positions of events on a run by run basis. Fig. 4.9 shows the AV z-offset for each run in the basic scintillator phase. The offsets in the x and y directions are negligible. While this thesis attempts to always refer to the AV corrected coordinates as x, y, z, and r since we never use the PMT relative coordinates, internally the this corrected radius is referred to using a "_av" subscript which may appear in some places.

CHAPTER 5

Processing and Fitting of the Basic Scintillator Phase Dataset

As was discussed in Sec. 3.4, one of the benefits of the phased approach taken by the SNO+ experiment is it allows for the opportunity to understand the detector at intermittent stages before the Te phase. Importantly, the addition of Te to the detector will introduce a large $2\nu\beta\beta$ background that will dominate the low energy end of the energy spectrum. This means that the pre-Te addition scintillator phase is our only opportunity to try and understand the backgrounds below that will be swamped by the $2\nu\beta\beta$ events.

This can be seen more clearly in Fig. 5.1, which shows the "singles data" from the entire Kamland-Zen dataset [62]. The dominant background from 1 MeV onwards is the $2\nu\beta\beta$ decay of ^{136}Xe, which is their $0\nu\beta\beta$ isotope. Far below the $2\nu\beta\beta$ rate, one can find other backgrounds from many sources, notably internal and external radioactivity and spallation products from cosmic muons. Many of these backgrounds do not only sit under the $2\nu\beta\beta$ spectrum, they also make some of the dominant contributions to the $0\nu\beta\beta$ region of interest (ROI).

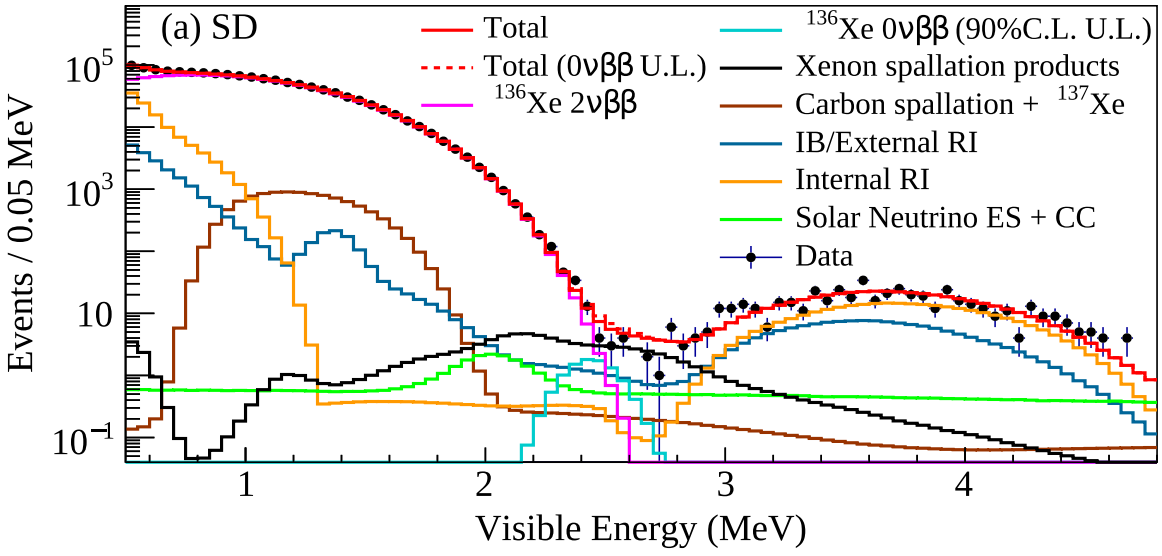


Figure 5.1: The backgrounds spectrum of events from the "singles data" part of the full Kamland-Zen dataset. From [62].

Now this begs the question — how can you say anything about the rate of the $0\nu\beta\beta$ events if you have other events creating signals right where one expects to see the $0\nu\beta\beta$ signal? Of course, this isn't a hopeless predicament, one can use strategies like coincidence tagging (see Sec. 4.3.1) to constrain the rates of any backgrounds connected through a decay chain, and likelihood fits use the distributions of the different backgrounds in parameters like energy and radius to differentiate them from the $0\nu\beta\beta$ signal. But one could always worry that perhaps there is some background lurking under the $2\nu\beta\beta$ distribution of which we are not aware, affecting the fitted rate on the $0\nu\beta\beta$ signal in some way we do not expect. The only way to try and address these worries would be to somehow peer below the $2\nu\beta\beta$ background, allowing one to explicity check if the backgrounds model in the ROI does indeed generalize to regions outside of the ROI.

While it's not possible to differentiate the $2\nu\beta\beta$ events from all other backgrounds on an event by event basis, the phased approach taken by the SNO+ experiment is one way to try and effectively turn "on" and "off" the $2\nu\beta\beta$ background. By taking data before and after adding isotope to the same scintillator, one can approximately measure what the backgrounds model looks like with and without the $2\nu\beta\beta$ backgrounds. This is not a perfect approach since loading the isotope into the detector is not a simple process — the optics of the scintillator change during Te loading and backgrounds will be introduced, but this staged loading process does provide some additional information that the un-staged process does not.

This chapter will discuss the expected backgrounds within the SNO+ detector in Sec. 5.1, and in Sec. 5.2 we will discuss how one can use coincidence and event level cuts to remove large quantities of these unwanted backgrounds. Additionally, in Sec. 5.3 we will introduce the fitting strategy used in this work, both the form of the likelihood function itself and the code used to do the fits.

5.1. Backgrounds in liquid scintillator detectors

Even before we have filled the SNO+ detector with scintillator, we already can make an educated guess regarding what backgrounds we can expect to see. While we go to great lengths to minimize the backgrounds inside the detector, some contamination is unavoidable. Indeed, when the liquid scintillator is purchased it will already be full of radioactive impurities which are reduced through

cleaning procedures before it is put into the acrylic vessel. More information on the scintillator cleaning can be found in [42].

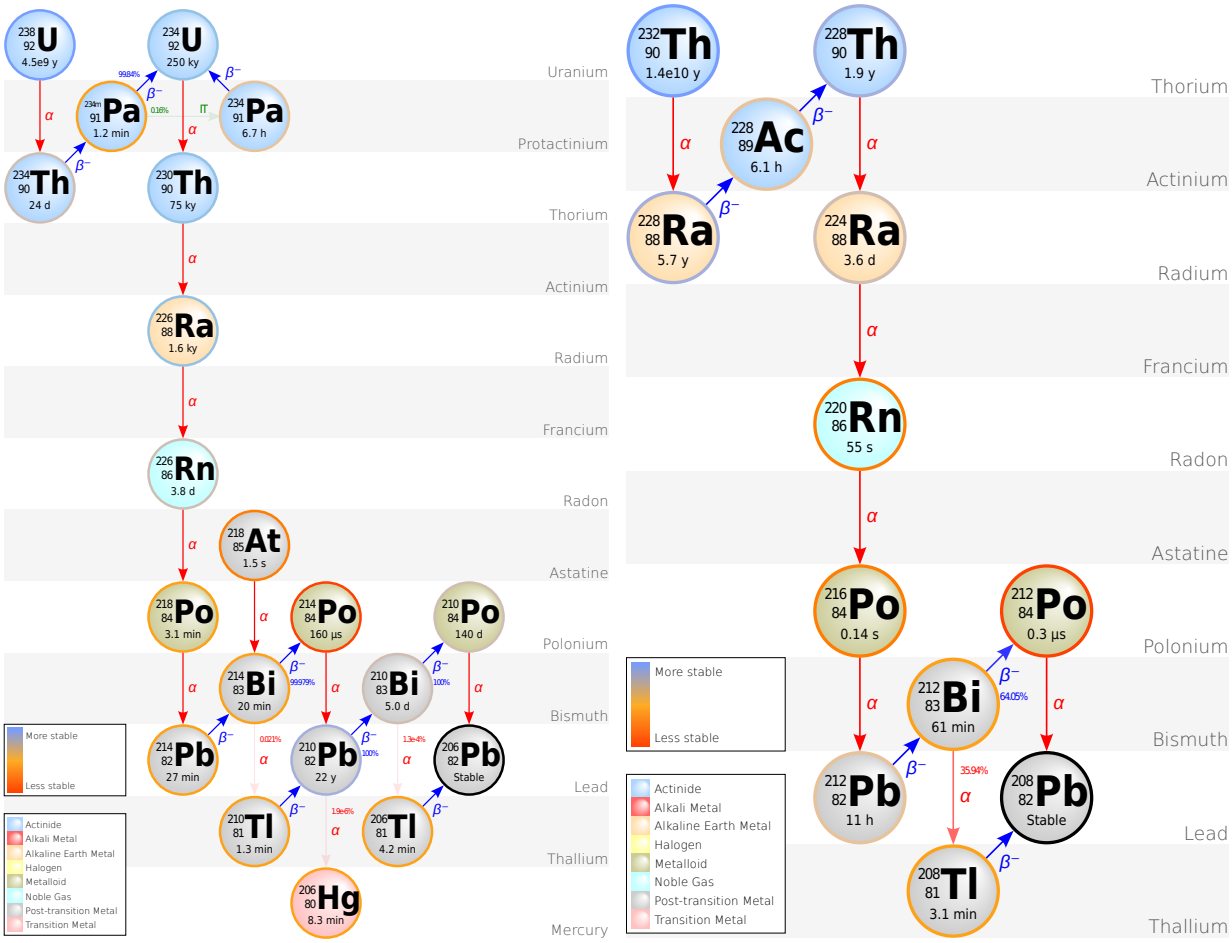


Figure 5.2: *Left:* Uranium-238 decay chain. *Right:* Thorium-232 decay chain.

Isotope	Mode	Half-Life	Q-value (MeV)
^{238}U			
^{234m}Pa	β^-	1.17 min	2.28
^{214}Pb	β^-	26.8 min	1.024
^{214}Bi	β^-	19.9 min	3.272
^{214}Po	α	164.3 μs	9.25
^{232}Th			
^{228}Ac	β^-	6.15 hr	2.127
^{212}Bi $\rightarrow$ ^{212}Po (64.06%)	β^-	60.55 min	2.252
^{212}Po	α	299 ns	8.954
^{212}Bi $\rightarrow$ ^{208}Tl (35.94%)	α	60.55 min	6.207
^{208}Tl	β^-	3.053 min	4.999

Table 5.1: The ^{238}U and ^{232}Th decay chain isotopes, decay modes, half-lives, and Q-values for decays relevant to this work.

The important radioactive isotopes for the work presented here are those in the ^{232}Th and ^{238}U decay chains. Since both of these isotopes are naturally occurring, they inevitably make their way into the scintillator and due to the long lived nature of both ^{232}Th and ^{238}U, they produce essentially a constant rate of backgrounds through their respective decay chains. Both decay chains can be seen in Fig. 5.2, and the isotopes contained within each chain are the dominant contributors to the "IB/External RI" and "Internal RI" backgrounds in Fig. 5.1. The chains are complicated, but fortunately there are only a handful of isotopes that end up producing decays over 1 MeV of observable energy in the detector.

In the ^{238}U decay chain, the important isotopes are ^{234m}Pa with a Q-value of 2.28 MeV which produces a β with only very rare γ emissions, ^{214}Bi which has a Q-value of 3.27 MeV and produces a β with intense γ emissions, and ^{214}Po which decays via a 9.25 MeV α (which is quenched!). ^{210}Tl is another seemingly dangerous background which has a Q-value of 5.48 MeV and decays via the emission of a β with intense γ emissions, but given its tiny branching ratio from ^{214}Bi (0.02%) we expect almost no contamination from this background in our detector, which is why it is omitted from the table.

In the ^{232}Th decay chain, ^{212}Bi with a Q-value of 2.25 MeV can pileup with ^{212}Po, due to the

short half life of the Po, reaching energies up to ∼3 MeV. These events are mostly removed by the classifiers like ITR, but some occur in such quick coincidence that they cannot be removed, creating one of the dominant backgrounds in the Te-phase. ^{208}Tl is another important background which sits above the $0\nu\beta\beta$ ROI that presents a major background for solar neutrino analyses under 5 MeV.

5.1.1. Externals

In addition to the radioactive isotopes inside the scintillator itself, another important class of background are the *externals*, or events where a β/γ decay occurs outside the acrylic vessel but one of the γ-rays travels into the detector and Compton scatters an electron within the AV. While the γ-rays are typically attenuated on the scale of tens of centimeters in the water and scintillator, the background rates in the AV, PSUP, and external water are high enough that the rare occurrence of a γ-ray traveling meters before Compton scattering is not negligible. These events are a very dangerous class of backgrounds for both the $0\nu\beta\beta$ and low energy ^{8}B solar ν analyses.

Fortunately, the external backgrounds do create a very particular signal in the SNO+ detector, namely in their radial distributions. Even though we cannot say where an individual external γ will Compton scatter an electron, we do know that when we sum over many external events the radial distribution of these events should follow an exponential distribution. This is why including the radial position of an event as a dimension in the likelihood greatly improves our sensitivity, it allows us to explicitly fit out the contributions of the external backgrounds.

The external backgrounds considered in this work include ^{214}Bi and ^{208}Tl (as these are the only isotopes with individual γ-rays over 2 MeV) in the AV, external water, hold-down ropes, and the PSUP. The ^{214}Bi and ^{208}Tl externals from the PSUP are simulated according to an analytic calculation of the expected rate in a shell closer to the AV than the PSUP in order to reduce the number of simulated events required to generate an appreciable signal inside the AV. To make this difference clear, we refer to the external simulations of events from the PSUP as the PMT beta gammas, or PMT$_{\beta/\gamma}$. Simulation is used to predict what the signal from these external backgrounds should look like in the SNO+ detector, in the same fashion as what is done for the other internal

backgrounds. A plot showing the different external backgrounds can be seen in Fig. 5.3.

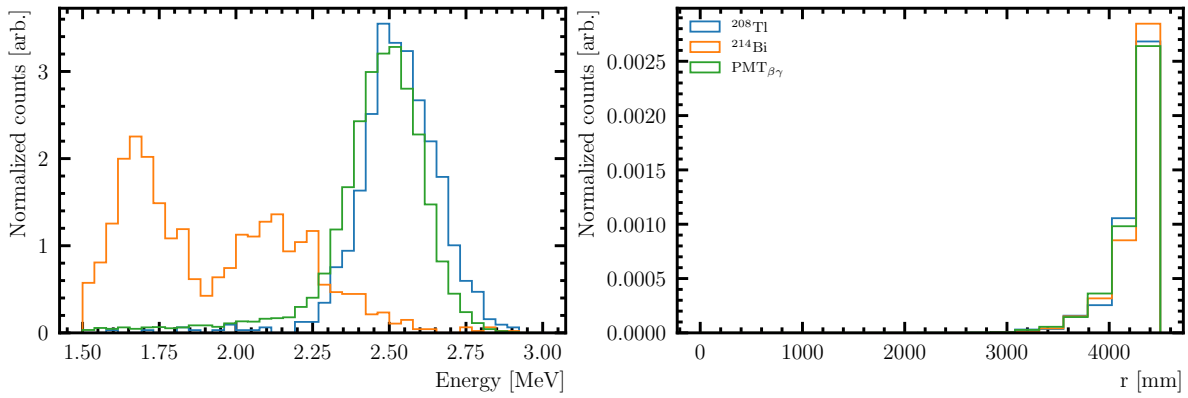


Figure 5.3: Results of the externals simulation for the external backgrounds considered in this analysis.

Note that in the plot we are showing the simulations of the externals on the AV and the PSUP. The simulations of the externals from the hold-down ropes and external water look practically identical to those from the AV.

5.1.2. Background classes

All the potential background sources discussed in this section can be grouped into different "classes" depending on what type of particle or particles are responsible for generating most of the collected light. The light detected in a physics event is created either from electrons, protons, or alphas. Gammas, neutrons, and neutrinos are all neutral particles so they do not produce scintillation light directly as they propagate through the detector. But of course, we are still able to detect these neutral particles through their interactions with charged particles in the scintillator.

For neutrons, protons recoiling off of the neutron as it thermalizes will often produce detectable amounts of light and then the neutron will capture on hydrogen (∼99% of the time) or carbon (∼1% of the time) which then de-excite by emitting 2.2 MeV or 4.95 MeV γ, respectively. As mentioned before γ-rays also do not produce scintillation light but at the MeV scale energies which we are interested in, the γ-rays will Compton scatter electrons which do produce scintillation light. Neutrinos create light within the detector in much the same way — neutral current or weak current

interactions will scatter electrons in the detector, creating detectable bursts of light. Additionally, $\bar{\nu}$ can interact weakly with protons in the detector to create neutrons and positrons, where the neutron can be detected through the process above, and the positron creates light as it propagates since it is charged, and then also produces two γ-rays with total energy $\sim$1 MeV when it annihilates with an electron in the scintillator.

The radioactive isotopes in the SNO+ detector typically emit βs, γs, a combination of βs and γs, or αs. For the analyses presented in this work there are no pure α backgrounds as they are all quenched and produce electron equivalent energies less than 1.5 MeV. As was briefly mentioned in Sec. 4.2.3, we do have the ability to statistically differentiate the signal classes from one another due to the differences in both their hit time residual distributions and energy distributions. Fig. 5.4 shows a few of the relevant backgrounds in the SNO+ backgrounds model.

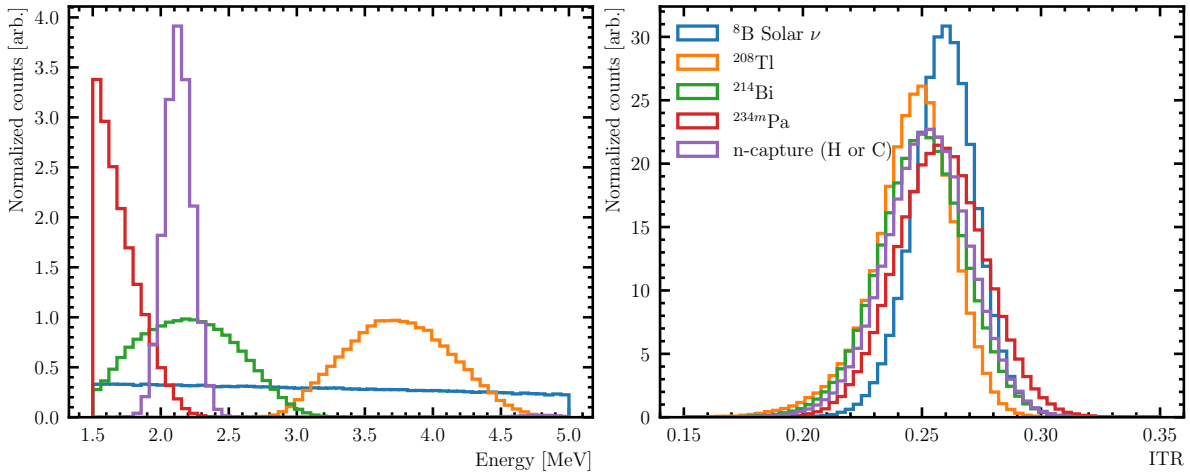


Figure 5.4: Energy (left) and ITR (right) distributions of some relevant backgrounds to this work.

The ^{234m}Pa and ^{8}B are both pure (or nearly pure for the ^{234m}Pa) β signals, but the ^{8}B events fall in a much broader energy range than the ^{234m}Pa events making the ITR distributions slightly narrower while the means remain close to one another. The reason for this is that ITR is slightly correlated with energy, making higher energy events have slightly narrower ITR distributions. The ^{234m}Pa is also a β decay with a Q-value of 2.28 MeV, and the lower energy limit of the histogram in Fig. 5.4 is 1.5 MeV which is why we only see what looks like an exponentially decaying tail.

^{208}Tl and ^{214}Bi on the other hand are β/γ events, meaning the radioactive decay consists of a β emission followed by very prompt (∼picosecond) γ de-excitations. As was discussed in Sec. 4.2.1, since photons are produced by both the electron emitted from the decay, as well as by the Compton scattered electrons from the de-excitation γs, the hit time residuals of these β/γ events are typically more spread out leading to ITR distributions biased towards lower values. This is also not a binary effect — depending on the relative intensities of the initial β and the de-excitation γ-rays the ITR distributions will shift around. The multi-particle nature of the β/γ decays also leads to interesting energy distributions, which look like spread out γ peaks. To understand these energy distributions one must remember that the γ-rays produced in the SNO+ detector by internal backgrounds are essentially all entirely contained at the FVs we are interested in, so the β and γ energies add together in a single event. The resulting shapes of the β/γ decays are therefore something like "smeared" γ distributions, where the constant energy of the γ defines the mean of the resulting distribution and the Q-value of the β defines how much smearing there will be around that mean.

The last type of event shown in the plots are the pure γ events, which are the events in the histograms labeled as "n-captures (H or C)". γ-rays emitted in a particular decay will always have the exact same energy due to energy conservation laws, so the resulting energy distributions are just Gaussians centered at the true value, where the width of the Gaussian comes from the energy resolution of the detector. Interestingly the ITR distribution of the pure γ events sits in between the pure β and β/γ histograms. The reason for this is that the γ itself does not emit any scintillation photons, so a γ can look identical to a β if it converts all its energy in the first Compton scatter. But the distributions are not identical because the γ can and does sometimes Compton scatter more than 1 electron as it travels through the detector, making it look more like a β/γ event.

The ability to use energy and ITR to distinguish between different types of backgrounds classes will be an extremely important piece of this analysis.

5.2. Cuts and coincidence tagging

Many event level cuts and coincidence tags are applied to the data to remove unwanted backgrounds. For both the $0\nu\beta\beta$ analysis and the low energy ^{8}B analysis, we expect no coincident events with our

signals, meaning that coincidence cuts are an effective way for us to maximize our signal to noise ratio. Therefore, distinct from the ^{214}Bi^{214}Po tagging discussed previously, the tagging cuts used here will all be very loose to try and remove as many coincident backgrounds as possible rather than select pure sample of the coincident events themselves.

5.2.1. Event level cuts

First we will discuss event level cuts. These are cuts that can be applied without any reference to other events in the detector, like cuts on energy, ITR, or other quantities.

Cut	Cut value
Data Cleaning	Pass
Valid Fit	True
Radius	r < 4.5 m
Energy	1.5 < E < 5.0 MeV
ITR	0.15 < ITR < 0.35
β_{14}	-0.1 < β_{14} < 0.04
necknhits	necknhits == 0
posFOM/nhitsCleaned	11.5 < posFOM/nhitsCleaned

Table 5.2: Typical cuts used to select events in the scintillator phases.

Table 5.2 presents the cuts used to select the events that are included in the likelihood fits presented in the final results of this chapter. Sometimes it is helpful to change the values of the cuts to look at different selections of events, and it will always be made clear when this occurs. Also note that these cuts are not the cuts that maximize the sensitivity of any signal. Since likelihood fits, which fiducialize automatically, are used to extract the rate measurements we just need to be sure that our cuts include the regions where we have the most sensitivity.

The data cleaning cut mentioned in Tbl. 5.2 is responsible for removing muons as well as events following within 20 seconds after a tagged muon. Additionally, this cut removes many of the instrumental backgrounds present in the SNO+ detector, including things like flashers.

5.2.2. General coincidence tag

The general coincidence tag is a cut that is essentially just a generalization of the cut used in Sec. 4.3.1, but with very loose prompt and delayed cuts to try to remove any events that have some kind of coincident signature. Most of the ∼2 kHz of events we record in the SNO+ detector are low energy ^{14}C decays or ^{210}Po decays on the AV surface. This means that for any given event in the energy range given in Tbl. 5.2, we should expect essentially no events within a window of a few ms and within about a meter of the event, unless it is some kind of coincident background. Aside from removing ^{214}Bi^{214}Po events, this general coincidence tag will also remove reactor $\bar{\nu}$ events which have a coincidence between the prompt positron annihilation and the delayed neutron capture.

Cut	Delayed Event
Valid Fit	Any
Data Cleaning	Any
Radius	r $\leq$ 5.5 m
Energy	0.75 $\leq$ E $\leq$ 16 MeV
Δt	$\Delta t \leq$ 2.5 ms
Δr	$\Delta r \leq$ 1.5 m

Table 5.3: Coincidence criteria for the general coincidence tag.

The cuts used to select the coincident pairs can be found in Tbl. 5.3. Note that the table only shows the cuts applied to the delayed event since we try to tag every event that passes the event level cuts shown previously, which defines what constitutes a prompt event.

5.2.3. Atmospheric $\bar{\nu}$ tag

While the atmospheric $\bar{\nu}$ rate is expected to be very low in SNO+ and the energies deposited in the detector by their interactions are far above our upper bound energy cut (7 MeV), each atmospheric interaction can spawn multiple neutrons whose capture on Hydrogen creates a 2.2 MeV γ, which can enter our ROI. Therefore it is necessary for us to remove these events, which can easily be done with the coincidence tagging criteria presented in Tbl. 5.4.

Cut	Prompt Event
Nhits	$3000 \leq$ Nhits
Δt	$\Delta t \leq 5$ ms

Table 5.4: Coincidence criteria to remove followers of atmospheric $\bar{\nu}$ events.

Note that this cut is applied backwards in time from a given event that passes the event level cuts. This is because this cut is for situations where a high-energy atmospheric interaction creates neutrons or other products which then cause interactions that do enter our ROI, such as neutron captures on Hydrogen. Again, the table only shows the cuts applied to the prompt event since we try to tag every event that passes the event level cuts shown previously, which defines what constitutes a delayed event.

5.2.4. Hot-spot cuts

Multiple hot-spots were identified during the scintillator phase, where a hot-spot is a location in the detector that appears to be intermittently creating intense flashes of light. These events can be tagged based on their location, the particular shape of their hit time residuals, and the distinct topology of the illuminated PMTs. While the events corresponding to the flashes themselves do not pass the event level cuts we apply, they can often be followed by many subsequent "ringdown" events caused from reflections as well as electronics noise due to the intensity of the flash. Since the hot-spots are always reconstructed to the same locations in the detector, we can be sure to remove all the ringdown events by vetoing a small window of time after hot-spot events. The coincidence cuts which accomplish this are listed in Tbl. 5.5.

Cut	Prompt Event
Nhits	$3000 \leq$ Nhits
x	$500 \leq$ x ≤ 900
y	$-900 \leq$ y ≤ -500
z	$5500 \leq$ z
Δt	$\Delta t \leq 1$ s

Cut	Prompt Event
Nhits	$3000 \leq$ Nhits
x	$0 \leq$ x ≤ 500
y	$-1250 \leq$ y ≤ -800
z	$5500 \leq$ z
Δt	$\Delta t \leq 1$ s

Table 5.5: Coincidence criteria to remove followers after flashes from the hotspots.

5.2.5. Dataset livetime and cut efficiencies

As has been mentioned, the dataset used in this analysis is the basic scintillator phase dataset of SNO+. The livetime of the dataset is 237.4 days. The expected signal efficiencies for all the background included in this analysis based on the cuts discussed in this chapter can be seen in 5.6.

Background	Efficiency
Ac228	0.127
Alphan Lab 13c Delayed	0.380
B8 Solar Nue High	0.125
B8 Solar Nue Low	0.0607
B8 Solar Nue Mid	0.048
Bi214 Av	3.25×10^{-5}
Bi214 Exwater Shell	5.75×10^{-6}
Bi214 Hdropes Shell	2.45×10^{-5}
Bipo212	0.0124
Bipo214	0.00064
C10 Scint	0.377
C11	0.06
Cno O15 Solar Nue	0.0001
Pa234m	0.033
Pb211	9.07×10^{-7}
Pep Solar Nue	4.40×10^{-7}
Pmt Betagammas	0.0014
Tl207	1.52×10^{-5}
Tl208	0.37
Tl208 Av	0.00015
Tl208 Exwater Shell	3.27×10^{-5}
Tl208 Hdropes Shell	0.00014

Table 5.6: Cut efficiencies for the various signals expected in the SNO+ dataset, calculated using the MC. It can be seen that externals and ^{214}Bi and ^{212}Bi events are effectively removed by the cuts relative to signals like the ^{8}B solar neutrinos.

5.2.6. Data tagging and processing software

Perhaps in a slight deviation from other large collaborations, a rite of passage for SNO+ analyzers is to develop their own event processing and tagging code. I am no exception.

A codebase to carry out the event tagging was developed, with special care taken to efficiently process the datasets in parallel and to keep track of the event level and coincidence cuts in some

convenient way. The code is written in C++, which is partially due to performance considerations, but also due to the fact that the SNO+ datasets are all stored as ROOT ntuples [63], which makes it less convenient to use languages other than C++ or python. Aside from having to link against the monolith that is the ROOT ecosystem, another goal during the development of the processing code was to minimize the number of dependencies in an effort to maximize simplicity. As of now, the code only has two dependencies, ROOT and yaml-cpp [64].

Early on in my experience as a graduate student, it became clear that most of the bugs in my software stemmed from poor bookkeeping. Even more frightening, poor bookkeeping could lead to *hidden bugs*, meaning bugs which do not prevent code from compiling or running, but rather manifest down the line during higher level analysis. Bugs of this type are unacceptable, especially when one considers that the perfect bookkeeping device is the very thing that is executing the code, the computer.

Thus, a great deal of effort was spent trying to write the processing code in such a way that all the *state* is separate from the code and localized as much as possible, making it much easier to reason about what the code is doing. This main way in which this is accomplished is all the information that defines the files to be processed, the cuts to be applied, the information extracted from the events, etc. are all defined within configuration YAML files which are supplied to the executable at runtime. This means adding or changing a coincidence cut, writing out additional information from events, or adding in a new set of simulation files to process can all be done without editing the code at all.

Much effort went into profiling the code to make it as fast as possible, and it was discovered that with a naive approach where every single thread read from a file and processed events, the majority of the running time of executable was spent on waiting on read requests over the NFS mounted data drives. Furthermore, the default behavior of the ROOT `TTree::GetEntry` methods were causing frequent read requests over the NFS mounts. Based on the documentation of the `TTree` class, it appeared that there should be a way to force the first call to `GetEntry` to cache the entire `TTree` in RAM rather than going back over NFS to load more entries, but none of the options seemed to

guarantee this behavior.

Instead, the best approach seemed to be using low level IO calls to force a full file to be read into memory with only a single syscall, and then use the `TMemFile` class to read the raw buffer as a file rather than interfacing with the standard `TFile` class. The optimal setup ended up being a combination of three threads which continuously reading the contents of ROOT files using the technique discussed above, and then using the rest of the cluster's capacity (∼57 additional threads) to do the actual processing of the events in the files.

Since SNO+ processes the data and simulation such that the structure of the resulting output files are identical, the analysis code discussed in this section is ran over both the simulation and data. Only particular simulations are set-up to generate coincident events, namely the ^{214}Bi and ^{212}Bi event generators, and these coincident background simulations are all ran through the analysis code and the coincidence tagging logic in the same way as the data.

The non-coincident background simulations are also ran through the analysis code to trim the simulation results since we only care about events over ∼1 MeV of energy within the 5.5 m FV. This makes the disk space taken up by the production datasets smaller and easier to work with. Another important use of the background simulations is to estimate our efficiencies and sacrifices of the backgrounds, including effects of the coincidence tagging logic. This means that the trimming functionality of the analysis code mentioned above must also take into account the effects of the trimming, which is all handled through meta information tracking that is generated by the analysis code and consumed by the fitter that will be discussed imminently.

The codebase is hosted on GitHub in a private repository, but access can be granted to any interested individuals.

5.3. The likelihood function

Following the application of the coincidence tags and event level cuts, we are left with some collection of events, created by internal radioactive decays, external radioactive decays, solar ν interactions, reactor $\bar{\nu}$ interactions, etc. Essentially both the low energy ^{8}B ν flux measurements and the

$0\nu\beta\beta$ lifetime measurement can be made by counting the events which we believe are attributable to these signals, but how can we tell if an event is a solar event rather than an internal radioactive decay?

This is accomplished using the likelihood function, see Sec. A.1 for an in depth discussion and introduction. The general idea behind the likelihood function is that if one knows the types of events produced by each of the expected signals, one can calculate the combinations of these signals that is most likely to have generated the events contained in the data. By "type of event", we mean the defining features of a given event, like energy, position, ITR, etc. and the different signals are the various processes that produce events in the detector, like radioactive decays and neutrino interactions. Constructing the likelihood function requires that we know the features from each of the signal classes we want to consider; for this we can use simulation (Sec. 4.1).

When we construct the likelihood function, we want the dimensions of $\mathcal{L}$, conversely the dimensions of the data points we feed into the likelihood x_i, to be the parameters in which we have the sensitivity to differentiate the backgrounds from the signals we are most interested in. Typically these dimensions are energy and radius — energy is how we are able distinguish most radiological backgrounds from more interesting signals and radius allows us to differentiate events originating from within the detector to those originating from regions outside the inner surface of the AV. When two signals are hard to differentiate using energy alone, and they are also different particle types (i.e. a pure β vs a β/γ), we can either fit in or cut on a classifier like ITR to try and increase our signal to noise ratio.

5.3.1. Deriving the likelihood function

To state precisely how the different signals factor into our likelihood function, we first define joint probability distribution functions (PDFs) for all the signals based on our results from simulations, which we can call p_j where j runs over all the signals. Each p_j is a function in the dimensions we want to fit in, energy, radius, and at times ITR as well. We then calculate the resulting combined distribution by adding together the individual PDFs along with arbitrary scalings for each of the

signals. We can write this in terms of the signal PDFs p_j and the signal normalizations N_j as

$$\nu' = \sum_j N_j p_j \tag{5.1}$$

where ν' is the combined distribution.

In this work, we always are using a binned likelihood approach, which means PDFs are discrete weights at different bin centers. Note that T' defined above is the counts per bin width. For reasons we will understand in a second, it is also useful to define a quantity in terms of the raw counts per bin T as

$$\nu_i = \sum_j N_j p_{ji} w_i \tag{5.2}$$

where w_i is the volume of bin i and ν_i means the contents of the histogram ν at bin i. Now we can think of ν as a prediction of the counts in each of the bins, which can be interpreted as a mean value λ of a Poisson distribution. This makes computing the likelihood simple. Recall that by definition the likelihood is the probability of observing the data x given the parameters of our model θ and some PDF. Since we know the per bin counts are Poisson distributed we can write

$$\mathcal{L}(x|\theta) = \prod_i \frac{\nu_i^{c_i}}{c_i!} e^{-\nu_i} \tag{5.3}$$

where c_i is the data counts in bin i and $\theta = (N_1, N_2, ..., N_k)$ is the different normalizations of the k signals. Taking the log of this function and dropping the terms which are constant we finally have

$$\ln\left[\mathcal{L}(x|\theta)\right] = \sum_i -\nu_i + c_i \ln(\nu_i). \tag{5.4}$$

5.3.2. Priors and constraints

In likelihood fitting, we often have a general idea of the values of some of the fit parameters prior to factoring in any of the data being fit. We can see a natural way to include this information if we take a Bayesian approach (see Eq. (A.4)); we know that the posterior is proportional to the

product of the likelihood and the prior

$$p(\theta) \sim \mathcal{L}(x|\theta)\pi(\theta). \tag{5.5}$$

Any knowledge regarding the value some parameter should take can be encoded into the prior. In this work, the priors are always Gaussian terms, where the mean and sigma of the Gaussian encode the value that we think is most likely and how confident we are in that most likely value. If we consider the log of the prior, for Gaussian terms on each parameters we have

$$-\ln\left[\pi_i(\theta_i)\right] = \frac{(\theta_i - \mu_i)^2}{2\sigma_i^2} \tag{5.6}$$

where i runs over the fit parameters with constraints, and μ_i and σ_i are the mean and standard deviations of the fit parameter.

Given our Gaussian prior terms, we know the priors are separable, i.e. that $\pi(\theta) = \prod_i \pi_i(\theta_i)$, or in other words all the priors are uncorrelated. Therefore to calculate the posterior, $p(\theta|x)$, the prior term can be added to the log-likelihood:

$$\ln(p(x)) = \ln(\mathcal{L}(x|\theta)) + \ln(\pi(\theta)). \tag{5.7}$$

This is essentially consistent with approaches which include an auxiliary term into their likelihood which comes from another likelihood function $\mathcal{L}(y|\theta')$ where y is a separate dataset and $\theta' \subseteq \theta$ is a subset of the parameters that the dataset y constrains.

The parameters for which we are interested in including these constraints includes both the background normalizations and the systematic parameters.

5.3.3. Systematic parameters and constraints

Another common situation in likelihood fitting is working with PDFs which do not perfectly capture the data we are attempting to describe. In those circumstances, it is convenient to float systematic

effects in the likelihood, or fit parameters distinct from the normalizations in that they encode an effect that changes the shapes of the PDFs themselves. These are usually effects like shifts, Gaussian convolutions (which are typically called smears), and scales.

Systematic parameters do not necessarily appear directly in the likelihood function, instead the PDFs are implicitly functions of the systematic parameters since they change the value of the likelihood function for a specific point θ through their effects on the PDFs. But it often is helpful to include a term into the likelihood that does explicitly depend on the systematic parameters to prevent them from transforming the PDFs in ways we know are not realistic. A natural way to include a term of this nature is to use priors.

To inform the choice of priors, we can use tagged ^{214}Bi events (Sec. 4.3.1). While we do apply correction in energy and ITR as described in Sec. 4.3.3, there are still small global differences in the distributions seen in data and predicted by MC. The systematics we elect to include for the energy dimension include a scale and smear, and for ITR we include a shift and a smear. The details of how the shift, scale, and smear systematics are applied will be discussed in section Sec. 5.4.3, but for now one can just think of the scales and shifts as a scale and shift of the datapoints being loaded into the histograms. The Gaussian smear can be thought of as a convolution between the PDF and a Gaussian centered at zero with the standard deviation set to the systematic value.

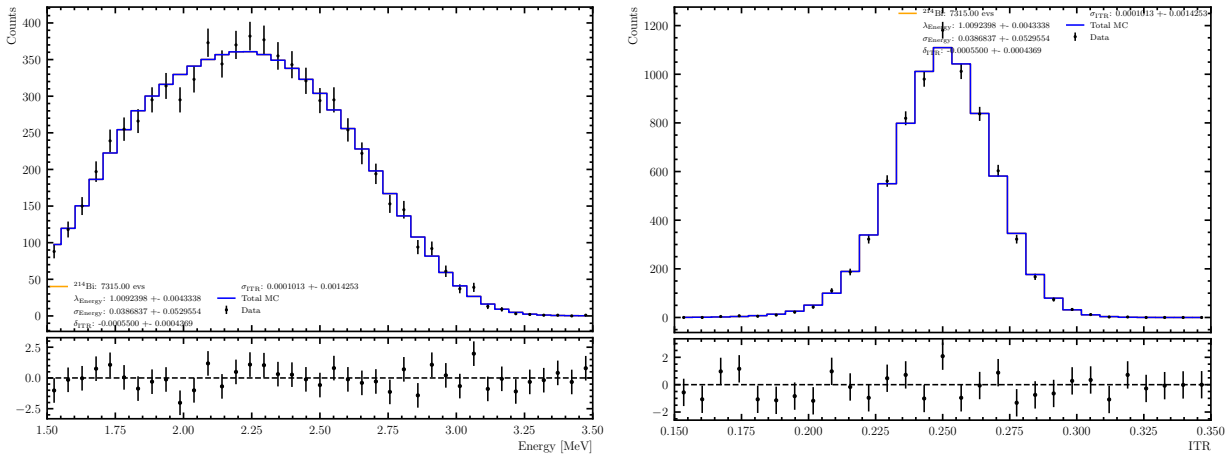


Figure 5.5: *Left:* Energy marginal of systematic constraint fits. *Right:* ITR marginal of systematic constraint fits.

The fits used to find the constraints also employ the same likelihood approach previously discussed in this section, except there are no constraints applied to any of the parameters and there is no parameter floated for the number of ^{214}Bi events since we are assuming the events in the dataset are only from that signal class. During fitting, the systematics are then applied to the PDFs of the ^{214}Bi events and the shape of the likelihood space gives us the mean and standard deviation to use for the priors on the systematics. The results of the fit can be seen in Fig. 5.5 and the resulting posterior profile curves (Eq. (40.41) [14]) can be seen in Fig. 5.6.

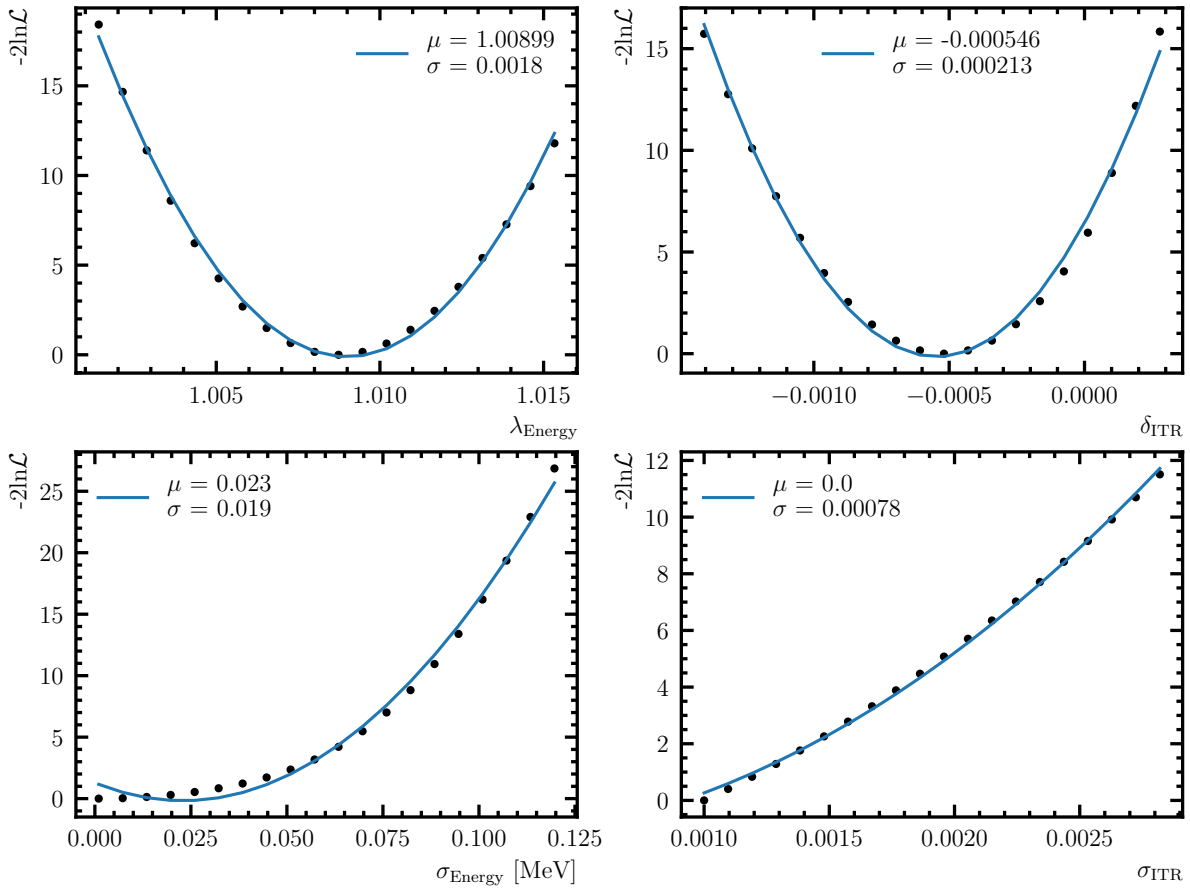


Figure 5.6: -2ln$\mathcal{L}$ profile curves for the systematic effects considered in this analysis. λ represents a scale, δ represents a shift, and σ represents a smear. The numbers in the legend are the mean and standard deviation derived from a quadratic fit to the likelihood profile points.

The resulting mean and standard deviation extracted from the quadratic fit in the log-likelihood space is used to set the Gaussian prior on each of the systematic parameters. While each systematic

is clearly not a perfect Gaussian, we consider this effect to be a "systematic on a systematic", or a higher order effect to which we are not very sensitive. Note that in general the errors in the legends of plots may not agree with the errors extracted from nested sampling posteriors or posterior profiles. The reason for this is the errors in the plot legends come from numerical estimates of the Hessian at the minimum in the log likelihood space, which allows us to approximate the correlation matrix. This is just an estimate though, and the estimates from the posterior profiles or the posteriors themselves will be more accurate, especially when the posterior is not Gaussian.

5.3.4. Background rate constraints

As we saw in Sec. 5.1, many of the radiological backgrounds present in the SNO+ detector are pieces of the same decay chains. For decay chains where each of the intermediate isotopes are short lived, where short lived is relative to the timescale of the datasets ($\mathcal{O}$(weeks)), the rate of one part of the chain informs the rate on all the other parts of the chain. Since the coincidence ^{214}Bi^{214}Po and ^{212}Bi^{212}Po backgrounds can be tagged to provide a pure sample, using the number of tags and our tagging efficiency we can obtain an accurate estimate of the total number of these nuclei in the detector. This estimate of the concentration of the coincident isotopes can then be used as the concentration of the entire radioactive chain itself since all backgrounds must be in secular equilibrium.

Fig. 5.7 shows the g/g concentration of ^{238}U and ^{232}Th throughout the basic scintillator phase based on the rate of tagged coincidence events from each chain. We can calculate the average g/g concentrations for the entire time range by averaging all the datapoints in Fig. 5.7, which gives a ^{238}U g/g of 6.20$\pm$0.13 and a ^{232}Th g/g of 7.53$\pm$0.72. Using this g/g concentration, we can predict the number of decays we expect from all the backgrounds belonging to each of these chains, which are the isotopes listed in Tbl. 5.1.

In addition to using the g/g concentrations to constrain certain rates, cosmogenic backgrounds can be constrained based on the generation of muon induced spallation backgrounds in other similar experiments. For example, Borexino [66] is not nearly as deep as SNO+ and used the same primary scintillator, meaning that we can use the measured production of cosmogenic backgrounds

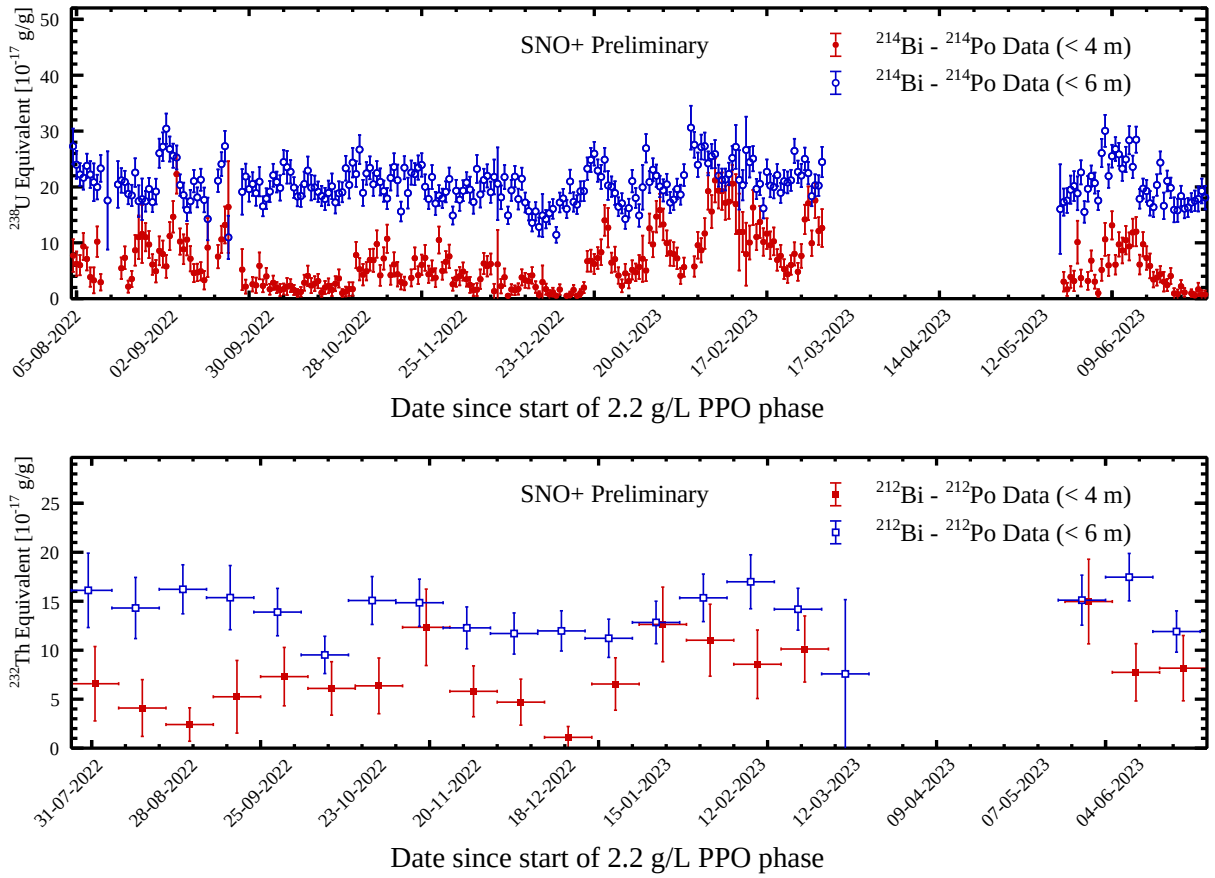


Figure 5.7: *Top:* The g/g concentration of ^{238}U based on the rate of tagged ^{214}Bi^{214}Po events. *Bottom:* The same but for ^{232}Th based on the rate of tagged ^{212}Bi^{212}Po events. Both images courtesy of G. Milton [65].

in Borexino, scaled down by some factor to account for our increased depth, as our expectation for the cosmogenic backgrounds in SNO+. Similarly, we can use measurements from other experiments to constrain the rates of solar ν events of all kinds, aside from situations in which we are interested in making measurements of the solar ν flux ourselves.

The final place from which we constrain rates in our detector is from analyses that measure signal rates in correlated but distinct regions of the detector. For example, the rate of external events increase exponentially as we move out towards the edges of the AV and the PSUP. During the water phase, a measurement was made of the external backgrounds using various cuts to select external events of different types [57]. The results of these measurements are used as input constraints of

the external backgrounds in the fits. In addition to these prior constraints, we also explicitly fit in the radial dimensions to use the high r regions where the externals dominate to constrain low r regions where we expect the externals to be almost entirely attenuated.

Finally, putting all this information together results in Tbl. 5.7, which shows the expected rates and the error on that rate from all the relevant signals.

Background	Rate [counts per year]	Uncertainty
Ac228	7,521.00	$\sqrt{N}$
Alphan Lab 13c Delayed	22.30	$\sqrt{N}$
B8 Solar Nux	1,151.75	$\sqrt{N}$
Bipo212	7,521.00	$\sqrt{N}$
Bipo214	18,979.00	$\sqrt{N}$
C11	895.00	$\sqrt{N}$
C10 Scint	18.00	$\sqrt{N}$
Pa234m	18,979.00	∞
Tl208	2,707.00	$\sqrt{N}$
Tl207	136,018.00	$\sqrt{N}$
Pb211	136,018.00	$\sqrt{N}$
Pep Solar Nue	7,800.00	$\sqrt{N}$
Bi214 Av	2.7×10^6	8.2×10^6
Bi214 Exwater Shell	2.0×10^7	∞
Bi214 Hdropes Shell	5.7×10^5	1.7×10^6
Tl208 Av	3.2×10^5	9.6×10^5
Tl208 Exwater Shell	8.6×10^5	∞
Tl208 Hdropes Shell	3.3×10^5	9.9×10^5
Pmt Betagammas	7.6×10^4	8.4×10^4

Table 5.7: Full volume pre-cut rates per year for the expected backgrounds in the basic scintillator phase. Columns where the uncertainties are left blank signifies that the uncertainty used in the fit is the Poisson error of the predicted rate after applying all cuts and a livetime scaling, or in other words the error on the rate shown here is small relative to the post-cuts Poisson error. For columns with uncertainties, the error is propagated into the uncertainty on the background normalization fit parameter. ^{234m}Pa has an infinite error since the rate might not be in equilibrium with ^{214}Bi tagging results because of long lived isotopes seperating the bottom and top of the ^{238}U chain.

5.4. Fitter

snopfit, the program used to carry out the likelihood fits shown in this thesis, is a completely custom codebase written for this particular analysis exclusively by the author. Everything is written in the Zig programming language, which is a systems programming language intended to be a modern

version of C. Zig provides much of the hardware level control provided by the C programming language, with many quality of life changes that significantly improve the programmer's experience relative to a language like C.

5.4.1. Why Zig?

When deciding what language to use to write the fitting program used in this analysis, the languages considered included Python, C, C++, Rust, and of course, Zig. Zig was deemed to be the best option for a number of reasons.

For one, since Zig is a systems level programming language, it can be as fast or even faster than C in some scenarios. This is distinct from Python, which is often much slower than lower level languages due to its high level syntax. While one can program Python in a way that will result in performance closer to C's (like using numpy and vectorized operations), it forces you to program in a certain way that might not always be convenient. Additionally, Python being a dynamically typed language can make large projects very complex as one must mentally track which variables are allowed to be which types, something that statically typed languages will do automatically at the cost of convenience. Python does support type hints now, but it is still no-where near as informative as a truly statically typed language. It is worth mentioning that the snopfit package uses Python and Matplotlib [67] for all of its postfit processing and plotting code, which has been a perfect fit.

While Zig is as fast as C, Zig tries to also provide some tools to make the language more memory safe, making it easier to avoid and find bugs that either leak memory or cause crashes at runtime. These tools include things like the `DebugAllocator`, which serves as an interface to request heap memory from the host system like C's `malloc`, but it will also track allocations and deallocations so that it can notify the user of mismanaged memory — and this exists in the standard library. Another important feature of the Zig programming language is its first tier support for C interoperability. Zig can compile and link against any C library since the Zig compiler can also function as a C compiler and transpiler to Zig code. This makes pulling C libraries into one's Zig project usually something that takes only a few minutes.

Another benefit of Zig over C is `comptime` or compile time support, which allows one to write Zig code that is evaluated at compile time rather than runtime. Zig's compile time supports generics, which means that types can be left undetermined in certain situations, and the Zig compiler will automatically deduce the correct type signatures to use for variables based on context in the code.

A simple example where this is helpful is for implementing a function that takes the maximum of an array. In Zig, one can only write this function once using generics, and then any type that supports ordered comparisons can use this function. In C one would need to explicitly write this function for every type, or use some macro tricks to force the compiler to generate functions of the correct type signatures during compilation.

One might point out that the other two options mentioned, C++ and Rust, both have all the benefits that we have mentioned so far regarding Zig. While that is true, Zig also has something that C++ and Rust do not have; simplicity. Zig is a programming language that prides itself on its simplicity, one of the main features of the programming language is that there is "no hidden control flow", meaning that the only way for a function to be executed is for the function to actually be called within the code. Rust and C++ on the other hand have things like operator overloading which by definition results in function calls without explicitly calling a function. For example, a line of code that sets two structs equal to one another in Rust and C++ might do many different things since the `=` operator can be overloaded — Zig intentionally does not support this.

Another benefit of no hidden control flow is that allocations should never be "hidden". `Allocator`s take center stage in the Zig programming language, and in the standard library any function that can allocate memory will require an `Allocator` to be passed in as a function argument, making the allocation clear and also hinting at ownership semantics. Furthermore, the zig standard library supports higher level `Allocator`s like `ArenaAllocator`s, which makes pre-caching and overwriting of buffers that are required for calculations occurring repeatedly much easier to keep track of. The use of `ArenaAllocator`s in snopfit allowed for calculations of the likelihood to be allocation free, even when including things like systematic effects.

Why does this discussion on hidden control flow and hidden allocations matter to a physicist though? Well for one, these conventions make it much easier to reason about the code. One never has to wonder about what overridden class' method is being called at runtime, which itself might also be an overridden operator, the execution will always follow the exact step by step instructions the code provides. This makes debugging much easier and faster. Furthermore, allocations in general are very slow as they require a system call to request more memory from the host system, something that takes orders of magnitude more time than operations like addition or subtraction. Often times allocations in hot loops can be one of the main causes of performance issues, and this is something I had programmed for years without knowing. Of course, one can write languages with hidden allocations in a way that will minimize allocations, but Zig pushes all of that complexity to the forefront of the programming language forcing one to think about these considerations during development.

Perhaps the best way to explain the benefits of programming in Zig is that it requires an understanding of computer hardware itself, whereas programming in other languages like C++ or Rust require an understanding of the semantics and abstractions of that specific language on top of also needing to understand of the underlying hardware.

5.4.2. snopfit dependencies

Similar to the analysis package discussed in Sec. 5.2.6, a development goal of snopfit is to try and minimize the number of dependencies required by the codebase. For one, this makes deploying and updating the software much easier since the number of versions one needs to keep track of is minimized. This also helps keep the complexity of the codebase small, since in principle all the code in snopfit should have a clear purpose. Furthermore, minimizing dependencies requires one to maintain a deep level of understanding of the analysis being performed since there is no abstraction that removes complexity.

Of course, there is a cost benefit analysis to be done when deciding whether or not to pull in a dependency. As such, snopfit does depend on GSL [68], zig-yaml [69], tinyexpr [70], and the C HDF5 library [71]. GSL handles all the matrix math needed for calculating inverses and also

functions for sampling distributions like the error function. zig-yaml handles the parsing of the YAML files used to configure snopfit at runtime. tinyexpr evaluates mathematical expressions allowing the configuration files to include cuts which themselves include arithmetic operations. Finally, the HDF5 C library is used for all the file IO, which is possible as SNO+'s ROOT ntuples are all converted from ROOT files to HDF5 files using a helper script. Technically we also pull in `llfit_zig` [72] repository as a dependency which handles the likelihood calculation, but this repository is also a solo project of mine so I do not consider it an external dependency.

5.4.3. Application of systematics

The most naive approach for handling the shift, scale, and smear systematics applied to the PDFs during the fitting process would be to apply the systematic effects directly to the points loaded into the PDF. The issue with this approach is two-fold, for one it can be slow for PDFs that load many datapoints into many bins, since each iteration of the fitter requires re-binning of the PDFs. Perhaps more importantly, though, is if we applied the smear systematic by adding normally distributed random values to the points used to construct the PDFs, the likelihood function becomes inherently random itself since each evaluation will lead to slightly different PDFs. This is undesirable as deterministic likelihood functions are often easier to work with, one reason being due to our reliance on minimizers to find the maximum of the likelihood function

The following paragraphs explain the specific approach used in this work to apply the systematic transformations. Note that references to the PDFs or histograms in following discussion imply 1D histograms or PDFs. Since we do fit in N-dimensions in this work we need to also apply these systematic effects to N-d PDFs, but this can be accomplished by simply applying the 1D transformations discussed below to every 1D slice in a given dimension of an N-d PDF. This is a well defined procedure since we do not consider correlated scale, shift, or smear systematics, which means that the same transformation is applied to every single 1D slice.

Rather than adding random values to the datapoints themselves, the smear is instead defined as a convolution directly with the binned histograms where the kernel is a Gaussian centered at zero with standard deviation set to the fit parameter representing the smear. Furthermore, during the

Gaussian convolution a finely binned interpolated histogram is used to represent the PDF, and then neighboring bins are then recombined to map the convolved PDF back to a PDF that uses the original coarser binning. The reason for this re-binning is it allows for convolutions to be applied that have much smaller standard deviations than the bin sizes, which is an important use case in our fits since the smearing systematics can be small on the scale of our bin widths.

Applying the smear systematic as a convolution makes it slightly more difficult to apply the shift and scale systematics though, as it is convenient to maintain a consistent approach in how the systematic effects are applied and there is no obvious way to scale or shift a histogram by sub-bin amounts. The fitting program developed for this analysis instead linearly interpolates along the bin centers and bin heights of the PDFs to implement these systematics. Specifically, the domain of the linear interpolation, which is the bin centers, is transformed according to the desired shift and scale, and then the interpolation is evaluated at the transformed bin centers, which essentially shifts and scales the binned PDFs in a continuous fashion.

Fig. 5.8 shows the posterior profiles of the systematic parameters in fits to simulated ^{214}Bi events who have had their energy and ITR distributions scaled, shifted, and smeared relative to the MC truth. The transformations applied to the datapoints in the figure are the naive versions of how the scale, shift, and smear systematics are applied in the fit. Namely the energies of the points are all multiplied by 1.1 to simulate a global energy scaling, the ITRs of all the points are shifted by 0.001, and both the ITR and energies of the ^{214}Bi events are smeared by adding samples from a normal distribution centered at zero with standard deviations of 0.01 and 0.03, respectively. As can be seen, the fitter is very close to picking out identical values to those used in the naive transformations, with some small biases present.

While having the fitter pick out the exact values used in the naive transformations would be nice, the only requirement is that the fitter must be able to recreate the dataset in the PDFs through applications of the sysematics. In fact, there is no reason one would expect to perfectly fit out the values used in the transformations discussed in the last paragraph given that the systematics operate on the binned PDFs, which is completely distinct from the operations applied to the ITR

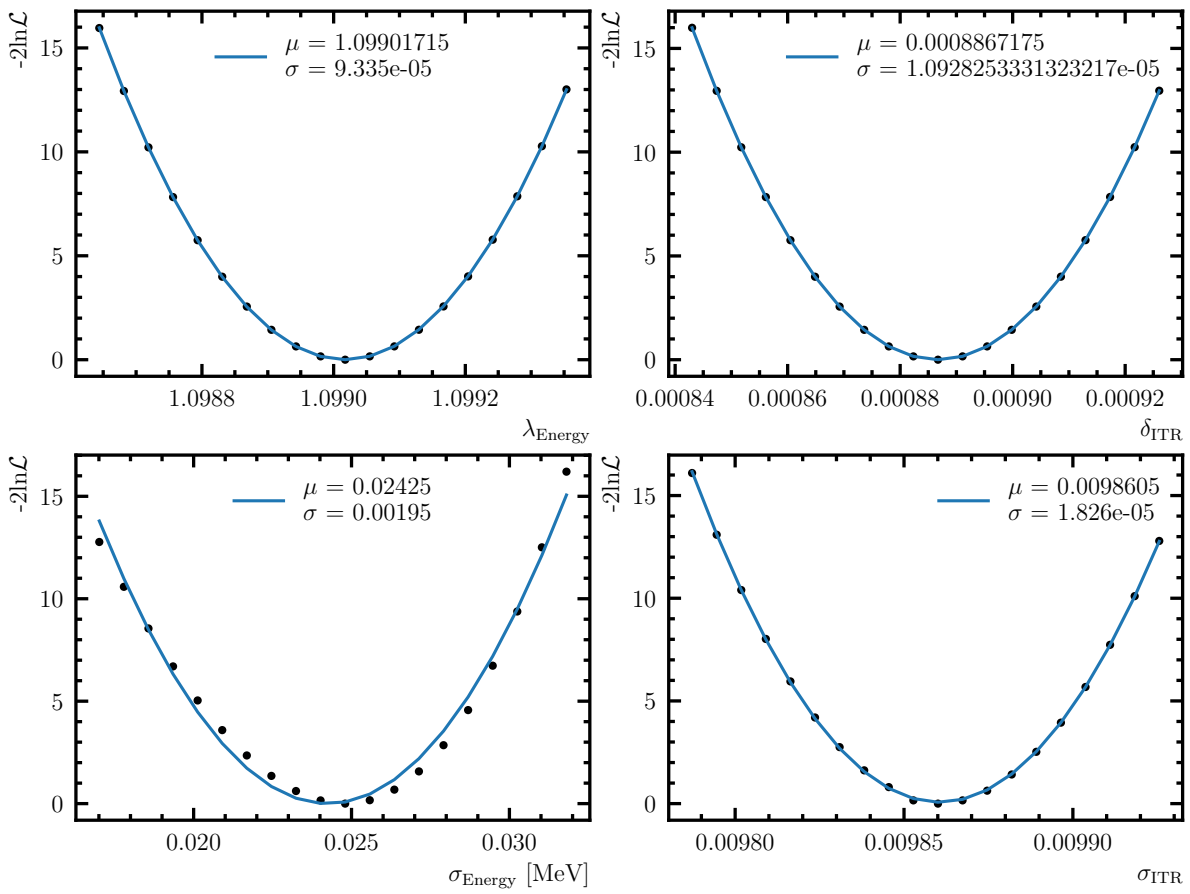


Figure 5.8: Posterior profiles of the systematic parameters when fitting to a manually transformed version of the PDFs used in the fit. The values used in the transformations are $\lambda_{\text{Energy}} = 1.1$, $\delta_{\text{ITR}} = 0.001$, $\sigma_{\text{Energy}} = 0.03$ MeV, and $\sigma_{\text{ITR}} = 0.01$.

and energy coordinates themselves to make the validation dataset. It can be clearly seen that this can be done in Fig. 5.9 as we see the residuals are very small once the systematics are applied at the values obtained from the MLE.

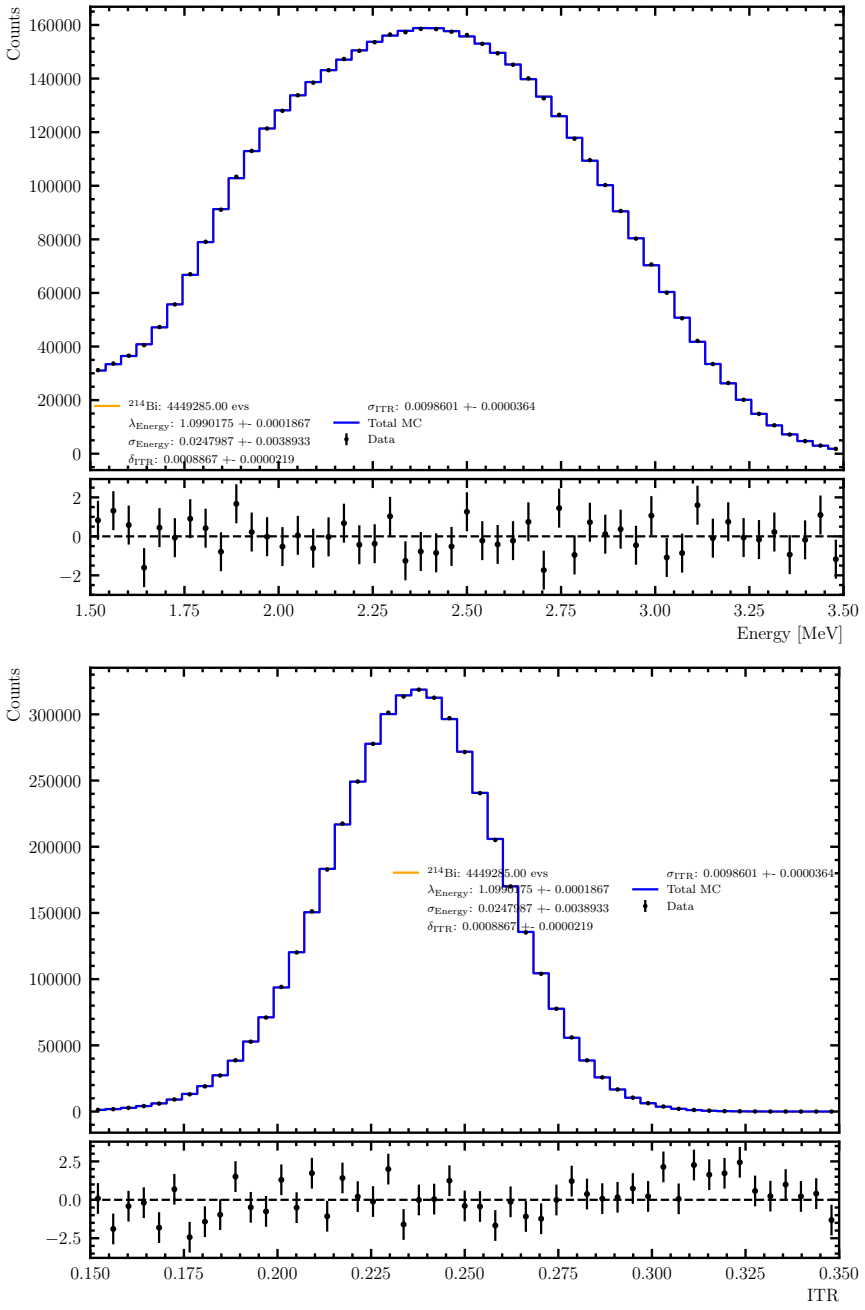


Figure 5.9: *Top:* Energy marginal of systematic validation fits. *Bottom:* ITR marginal of systematic validation fits.

Furthermore, the scale of the biases present are still only percent level effects, and percent level effects on top of systematic effects are typically ignored since they should have a very small impact

on the results.

It is also worth noting that snopfit supports what are called "buffer bins" or bins that do not affect the likelihood calculation but probability is allowed to flow from the buffer bins into other bins during the application of systematics. This is important for situations where large smears, shifts, or scales are being applied, since not including these bins can lead to artifacts at the edges of the PDFs. Maybe the clearest example of this effect is Gaussian smearing near the edges of a PDF. Since the Gaussian is symmetric, one can think of a convolution with a Gaussian kernel as removing probability from a bin and replacing it with equally weighted probability from the bins above and below it. If we are smearing the first bin of a PDF though, we can see that only the bins in the forward direction can contribute probability to the first bin, leading to an effect that biases bins near the edges of a PDF to lower probabilities. The buffer bins fix issues like this.

5.4.4. Runtime fitter configuration

Given the inherent complexity involved in this analysis, it can quickly become difficult to manage all the files, cuts, corrections, etc. that which are inputs into the fitter. To try and manage all of this state, snopfit supports runtime configuration of the fitter through YAML configuration files. An example of the YAML configuration file that defines the fits used to extract the systematic constraints from tagged ^{214}Bi events can be seen in Appendix D.

As can be seen in the example YAML file, snopfit supports defining the dimensions, files, corrections, systematics, cuts, and more to be used in a fit. This means that changing cuts, adding backgrounds to a fit, and adding a dimension to the fit can all occur without editing the snopfit source code. This also makes managing different fit scenarios very simple, each fit scenario can exist in a separate YAML file and one can specify whatever configuration file they would like to use to define the fit.

The cut string processing uses custom written code that converts the string into an abstract syntax tree (AST). The tree is walked to find all the event fields that need to be loaded from input files to evaluate the cut string. Additionally, the cut strings also support arithmetic combinations of parameters, for example `posFOM/nhitsCleaned > 11.5`. During the parsing of the cut string, the

code keeps track of which nodes of the AST will require the evaluation of arithmetic expressions, and then passes those nodes through tinyexpr when the cuts are being applied.

Another nice functionality provided by snopfit is the ability to define the corrections to apply the simulations within the configuration files. The corrections files are just json files where each entry in the json file defines a separate correction. Using tinyexpr also allows the way the correction is applied to be defined at runtime, so corrections can all use the same interface regardless of the specifics of how it is applied to the events. Perhaps the largest benefit, though, is that the corrections are applied every time files are loaded into the analysis which avoids the need to pre-process files through scripts that apply the correction, cutting out an error prone analysis step. This is how both the energy and ITR corrections are applied to the MC and data in this analysis.

5.4.5. Tracking cut efficiencies

One additional advantage of including the cuts in the configuration files is that we can keep track of all the signal efficiencies within the fitting framework. As was mentioned in Sec. 5.2.6, the coincidence tagging software also applies cuts to trim the data and MC, and retains some meta information regarding the number of events that were trimmed in order to allow for unbiased efficiency calculations. snopfit takes in this meta information, as well as number of remaining events after applying all of the cuts in the configuration file, to calculate what the expected efficiency is for a given signal. This is used to not only back out the full volume rates of the signals after the fitting, but it also allows us to scale the predicted full volume pre-cuts rates used as prior constraints based on the expected efficiency given the cuts and fiducial volume of a specific fit.

CHAPTER 6

Constructing a Backgrounds Model

6.1. The backgrounds model

Finally we can discuss the *actually* interesting part, fits to the basic scintillator phase dataset.

In this section, we may at times switch between 3.3 meter and 4.5 meter fiducial volumes (FV). The 3.3 m FV is simpler to understand since the detector response is more consistent throughout the entire region (Sec. 4.2.1) and the background rates above 2.0 MeV are much lower due to the larger buffer from external events. This also means we can fit in just energy and ITR when looking in the 3.3 m FV. The 4.5 m FV on the other hand maximizes our sensitivity since it increases the efficiency for selecting signal events while explicitly fitting out the external component. The final results of this chapter will all use fits in the 4.5 m FV in energy, ITR, and radius as these will maximize the sensitivity to both the $0\nu\beta\beta$ signal and the ^{8}B signal.

Now let us look at our first fit to the basic scintillator phase dataset. Fig. 6.1 shows the energy and ITR marginals of a 2D fit in the 3.3 meter FV. The data points pass all the cuts discussed in Sec. 5.2, which means we do not expect any non-physics backgrounds present in events fitted here. Furthermore, the data points should consist of only the signals we expect, which is discussed in Sec. 5.1, at rates consistent with the values in Sec. 5.3.4. First let us discuss the things we do understand, we can see ^{234m}Pa dominating the low energy part of the spectrum. While the rate of ^{234m}Pa being fit here is higher than we'd expect based on the coincidence tagging within the ^{238}U chain, there is a long lived isotope between the ^{214}Bi^{214}Po part of the chain and the ^{234m}Pa part of the chain, meaning that they can indeed be out of equilibrium which does appear to be the case. We can also see the small ^{208}Tl background between 3 and 5 MeV. Furthermore, we see low backgrounds between 2.5 and 3 MeV, which is both expected and required for success during the Te-phase. Plots focused on the solar neutrino spectrum will be shown in log-scale which will make it much easier to interpret the region around 2.5 MeV, which we will save for a future section.

Now for the part we don't understand. It can be clearly seen that something is missing both in the 1.5 to 2.5 MeV region of the energy spectrum and from 0.20 to 0.25 in ITR.

It is hard to see how this could be a systematic effect since we see good agreement between the data and MC for the tagged ^{214}Bi events, and the optimal values for the systematic parameters in the fit in Fig. 6.1 are already many sigma away from the expected values. Indeed, while it may seem like residuals in the fit marginals are not that large, one can see in Tbl. 6.1 that the optimal values of many of the systematics are in severe tension with their respective prior constraint, summarized in the last column.

Systematic	Prior	Fit result	Disagreement
λ_{Energy}	1.001 ± 0.004	1.011	2.5σ
σ_{Energy}	0.05 ± 0.03	0.22	5.6σ
δ_{ITR}	-0.0010 ± 0.0004	-0.0026	4.0σ
σ_{ITR}	0.0001 ± 0.0001	0.0001	0.0σ

Table 6.1: Comparison of systematic priors and their optimal values from the likelihood fit shown in Fig. 6.1. Large tension between the prior constraints and the fitted values can be seen, summarized in the last column. Note that the constraint values listed in the table here differ from those discussed in the previous chapter since this fit uses prior constraints obtained from ^{214}Bi events in the 3.3 m FV. With that said, the differences between the priors in the two FVs are small.

We believe this means we are missing the contribution of some kind of signal in our model, and the fitter is varying the systematics to try and resolve this issue, but there is no combination of the systematics that is consistent with the prior constraints and provides good agreement between the data and the model. Perhaps a clearer way to show this is to hold the systematics at their expected values (mean of the prior Gaussian), and then do the fit just floating the signal normalizations. This fit can be seen in Fig. 6.2, which shows the disagreement between the data and the model much more clearly.

This disagreement in both the energy and ITR spectrum has been a mystery for SNO+ for a few years now, and much effort has gone into trying to understand what could be causing this effect. In the following sections we will refer to this anomaly as the "low energy excess", where low means lower energy than the $0\nu\beta\beta$ and the low energy ^{8}B solar ν ROIs, and excess just refers to the

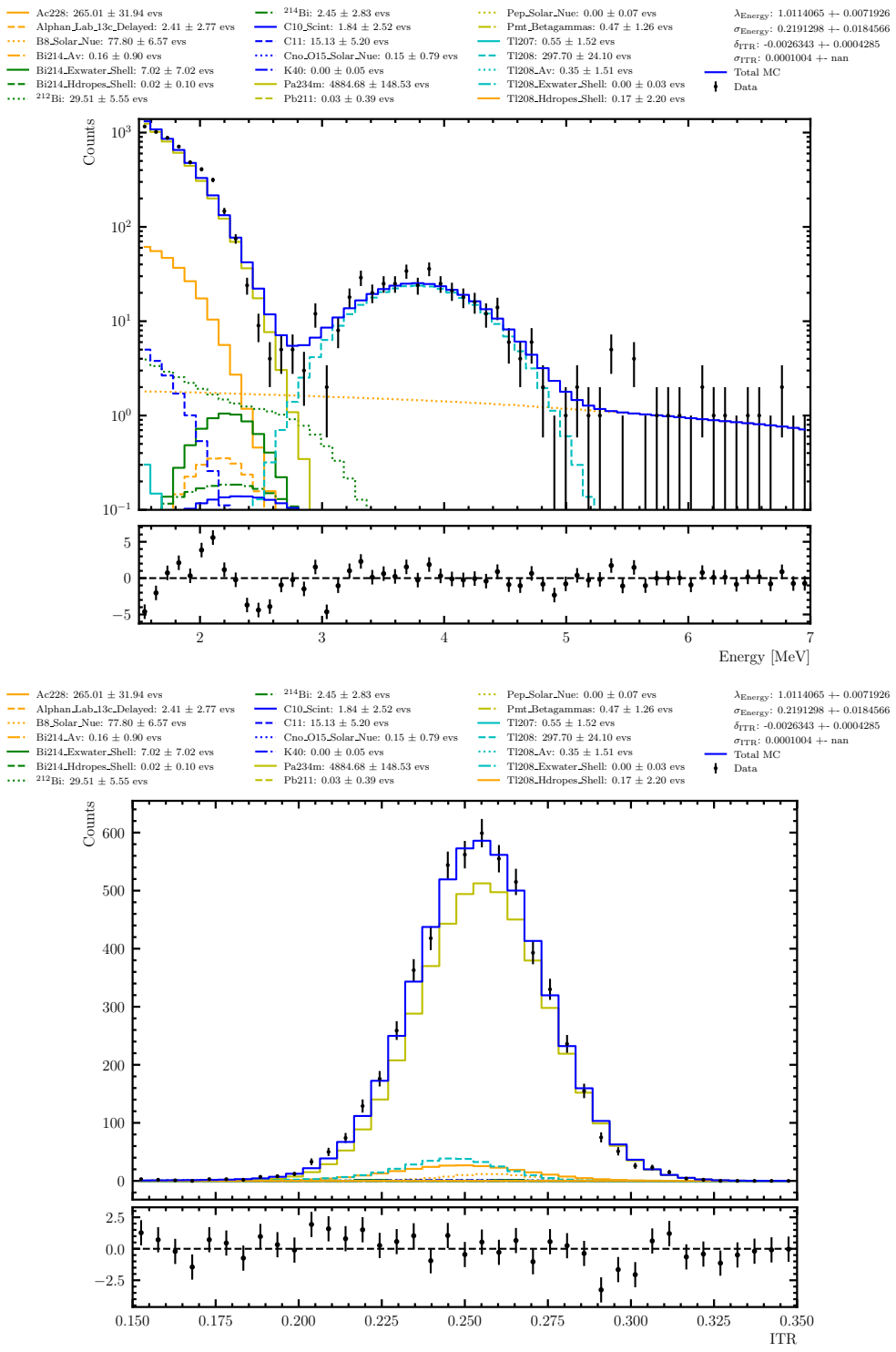


Figure 6.1: *Top:* Energy marginal of fit to SNO+ data between 1.5 and 7.0 MeV. *Bottom:* ITR marginal.

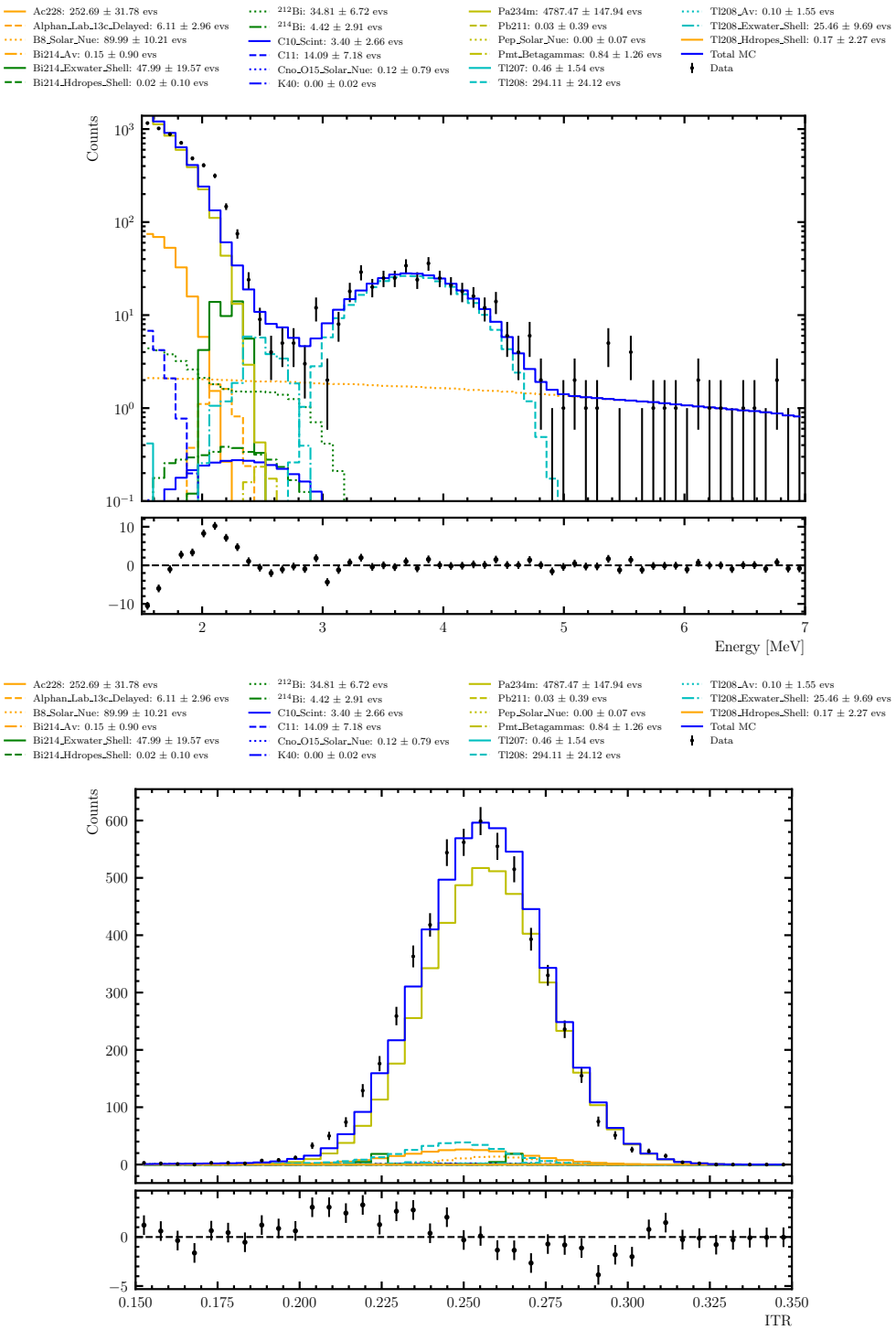


Figure 6.2: *Top:* Energy marginal of the same fit as shown in Fig. 6.1 with the systematic parameters held constant at their optimal values from the systematic constraint fits. *Bottom:* ITR marginal.

fact that we see more events in this region than can be explained by the presence of only ^{234m}Pa. Specifically, the events in the low energy excess will be events from 1.5 to 2.5 MeV. Sometimes references will be made to a 2.2 MeV peak, this stems from a period of time when we believed the background to consist of only a 2.2 MeV gamma peak, which is now considered to be only part of the explanation. When selecting events thought to belong to the 2.2 MeV peak, we typically select events with energies ranging from 2.0 to 2.5 MeV.

Many different ideas have been put forth regarding what could explain this low energy excess. Four general explanations for these events will be discussed in the following sections. The current leading theory is that there is both a γ and $\beta\gamma$ source in the detector not included in our backgrounds model. The rest of this chapter will be spent trying to motivate and develop this theory.

6.2. Low energy excess investigations

6.2.1. General coincidence tagging studies

Since coincidence tagging is one of our most powerful tools for removing and constraining backgrounds from the data taken in the SNO+ detector, many scans were done on the events comprising the low energy excess to look for some kind of coincidence signal that we may have missed. All of these studies have turned up empty-handed as it looks like there are no correlations between these events and other events. Fig. 6.3 shows the results of looking for general coincidences between low energy excess events and any other event within 1 meter and 1 second both into the past and future. Accidental coincidences will both appear flat in time and Δr^3, which is indeed what we see in the tagging results.

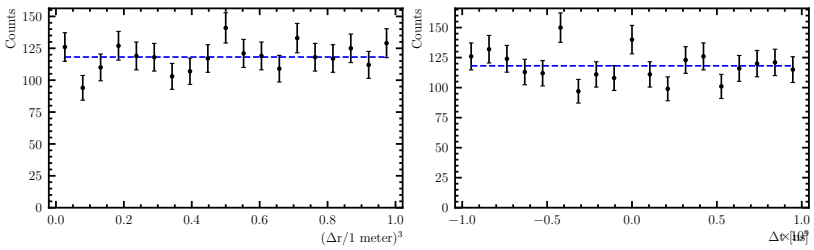


Figure 6.3: A search 1 second into the past and into the future for any coincidence signals with Δr less than a meter. There is no clear preference for low Δr or Δt values, which is what we would expect for some kind of coincidence signature — rather the results seem entirely consistent with accidental coincidences from high rate backgrounds like ^{210}Po and ^{14}C.

6.2.2. Instrumentals and mis-reconstructions

It has also been thought that these events may be some form of instrumental backgrounds, either from noise in the electronics or light produced inside the detector from sources other than physics events. SNO+ applies a myriad of data cleaning cuts to remove instrumental backgrounds, all of which are included in the analysis mask cut that is applied to the events shown in the fits above. Furthermore, these kinds of events typically have hit time residual distributions and hit topologies that differ drastically from physics events.

A similar class of events to instrumentals are mis-reconstructions, or events that are truly physics events but due to some unknown effect the events are reconstructed far away from their true positions. Just like instrumentals, this is a large hypothesis space and it is hard to exhaust every potential way in which a mis-reconstruction could occur. The best we can do is see if the low energy excess events have some distinct feature relative to a known class of background events.

The low energy excess events have been hand scanned to try and see if these events look distinct from standard physics events. It's worth mentioning that during these studies, we did find classes of events that we *do* want to remove from our datasets, but the vast majority of these events all "look" just like standard SNO+ physics events. For example, we discovered a class of "pileup" events where two AV surface events will occur in the same trigger window on opposite sides of the AV, which can lead to mis-reconstructions that place these events close to the center of the AV. These events can pass many of our typical classifiers, but they are very efficiently removed by β_{14}

which was not being applied before this class of event was discovered. An event view of a standard event and a pileup event can be seen in Fig. 6.4.

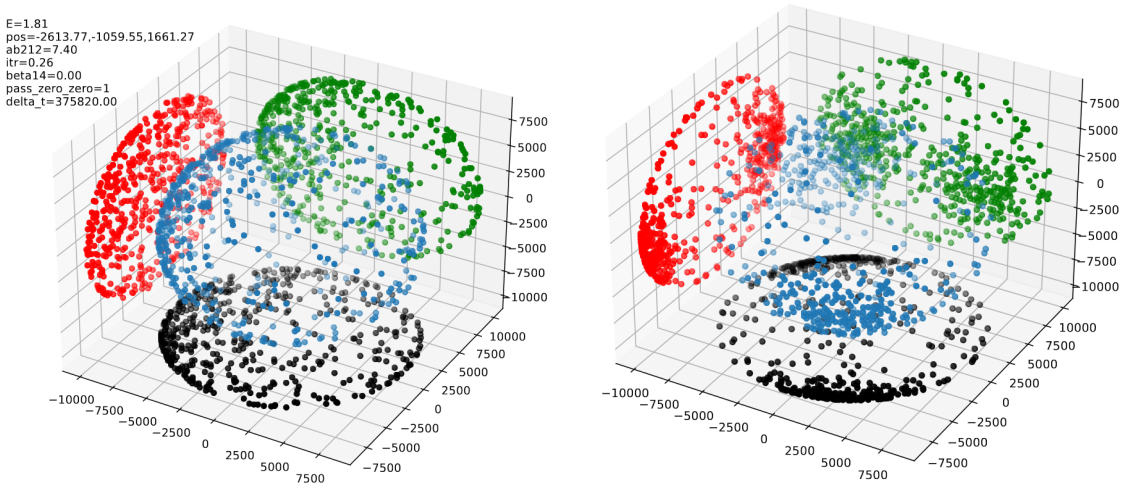


Figure 6.4: A comparison between a "normal" physics event and a "pileup" physics event. The event display shown here is custom. The black, green, and red points represent the x, y, and z projections of hit PMTs in an event, and the blue points are the true 3D locations of the PMTs. The pileup event (right) clearly has a distinct hit topology, which is captured by the β_{14} classifier.

As was said in the previous paragraph though, the pileup events accounted for just a small number of the low energy excess events and the rest of the events in the excess appear to look convincingly like other physics events. Fig. 6.5 shows the summed hit time residuals of the excess events relative to the summed hit time residuals of tagged ^{214}Bi events, and it can be seen that they agree well with one another. Therefore any instrumental backgrounds or mis-reconstructions would have to be created in such a way that they look convincingly like other physics events, which is very unlikely.

6.2.3. Neutron captures

The coincidence tagging studies also directly test one of the first hypothesized explanations of the disagreement between the backgrounds model and the data, which is a 2.2 MeV γ emitted from neutron captures on hydrogen. As one can see in Fig. 6.2, the residuals in the energy marginal resemble the energy distribution created by a 2.2 MeV γ. The ITR distribution of the pure γ events may help improve the agreement between the data and the MC in the ITR dimension as well. A fit with an additional neutron capture background freely floated can be seen in Fig. 6.6, which shows that the neutron capture peak does indeed help fill in some of the excess events. Furthermore,

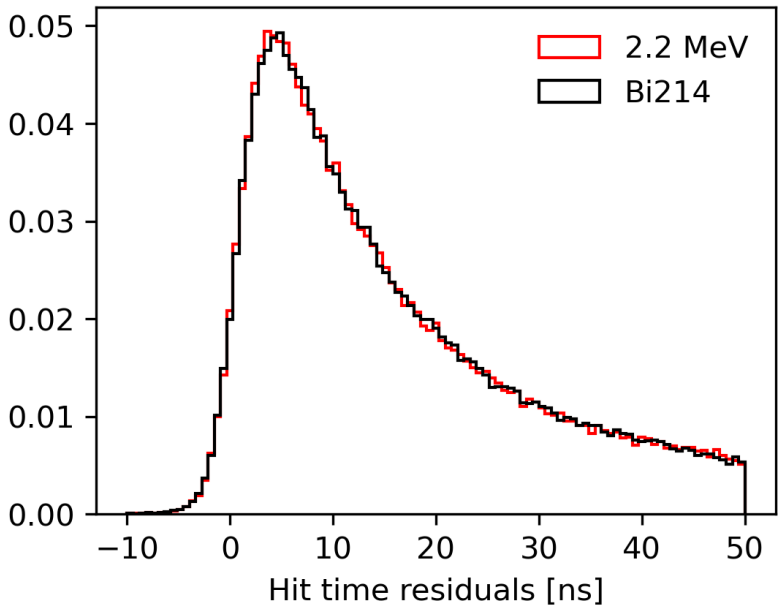


Figure 6.5: The summed hit time residuals of the excess events (red histogram labeled "2.2 MeV") relative to the summed hit time residuals of tagged ^{214}Bi events. The two agree with one another, suggesting that the low energy excess events appear to be physics events.

the systematics are much closer to their expected values in this fit relative to the fits without the neutron capture peak floated. Note that the fit suggests something like 600 neutron captures in this fit in the 3.3 m FV, which is many more neutron captures than we expect as will be discussed below.

Since neutron captures on hydrogren seem to be a plausible explanation, we must consider all the ways in which neutrons can be created within the detector. One process which produces neutrons has already been discussed, neutrons created in $\bar{\nu}$ interactions, specifically from both reactor and atmospheric neutrinos.

Reactor neutrino interactions, which are weak current interactions (see Fig. 6.7), will always produce one neutron since the exchange of a W boson transforms a proton into a neutron. Unfortunately, the reactor antineutrino rate is well known and also directly measured by SNO+ in the reactor antineutrino analysis, which has shown rates consistent with expectations. This rate only produces around 11 antineutrino interactions, or equivalently 11 neutrons, in this dataset [73].

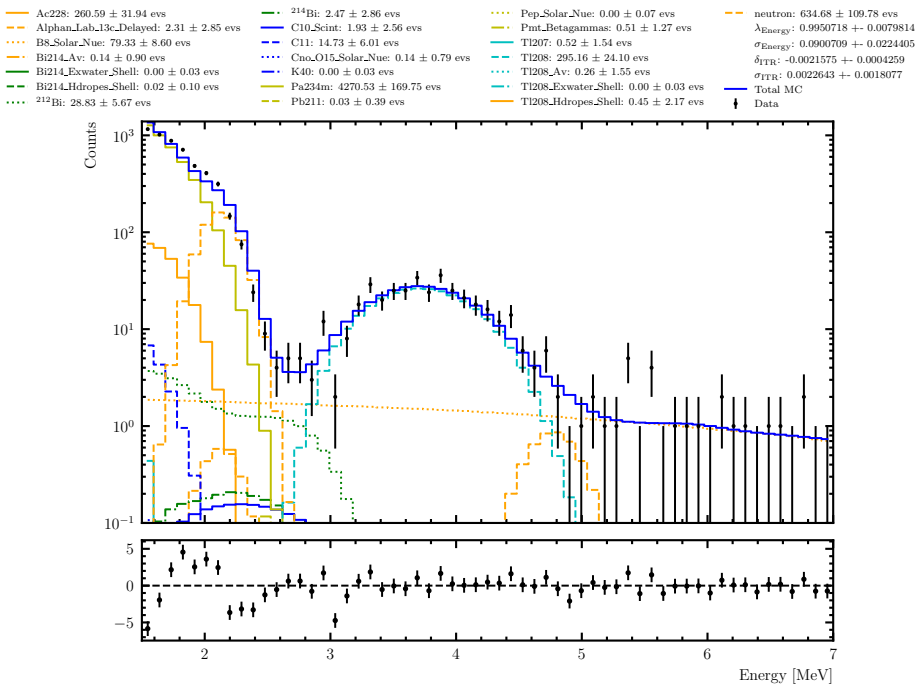


Figure 6.6: A fit with an additional free neutron capture component.

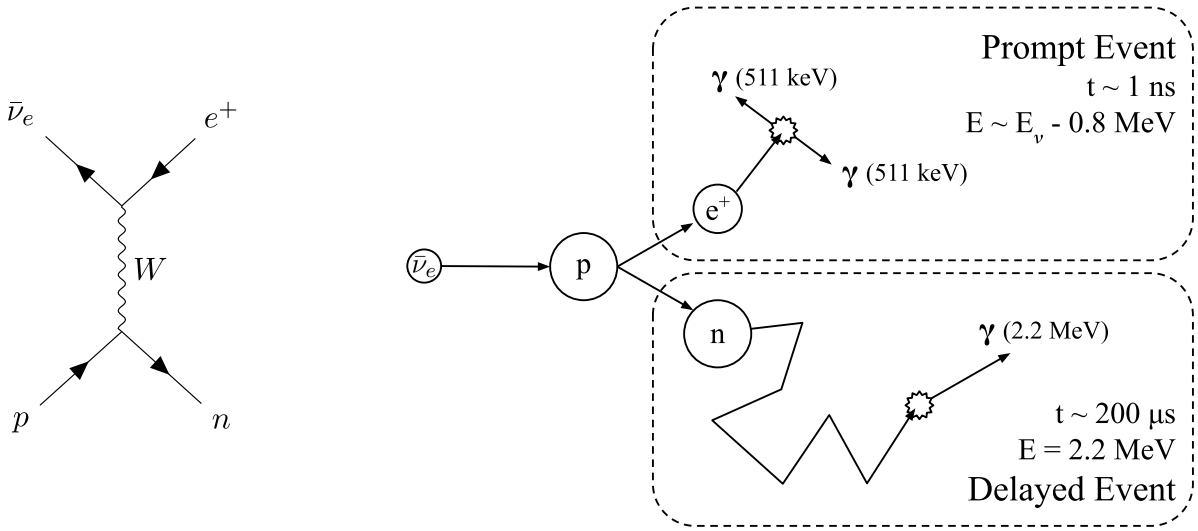


Figure 6.7: Schematic representing reactor antineutrino interactions in SNO+. From [57].

Furthermore, these neutron producing interactions will necessarily also produce positrons which are guaranteed to annihilate producing a prompt event of at least ∼1 MeV of energy before the neutron capture. The neutron will capture in a few hundred microseconds, emitting the 2.2 MeV gamma that we think might be what is the low energy excess. This coincidence between a >1

MeV event and a 2.2 MeV event means that the reactor antineutrino signal is very easy to tag, and should indeed be tagged by our general coincidence tag described in Sec. 5.2.

The atmospheric neutrinos on other hand can interact in ways that produce many neutrons in the detector, potentially being a source of neutrons in the detector. But GENIE [74] simulations of the SNO+ detector have been performed which show that the expected antineutrino contribution to this analysis is negligible [75].

Another source of neutrons are (α, n) interactions, where an α produced in a radioactive decay can interact with a nucleus in such a way that a neutron is emitted. The highest rate α emitter in the SNO+ detector is ^{210}Po, with a rate orders of magnitude larger than the next source of αs which is ^{214}Po. The production of neutrons from ^{210}Po was calculated in [76], and the rate was found to be 5.3μHz, with roughly a 40% uncertainty to account for uncertainties in the cross-section. Even with this large uncertainty, the number of α events expected in the entire basic scintillator phase dataset is only 68 over the entire detector volume, so one would need to increase the (α, n) rate by over an order of magnitude in order for it to explain this low energy excess.

Furthermore, two of the (α, n) processes produce high energy prompt signals: one involves the production of an excited state of ^{12}C which emits a 4.4 MeV de-excitation γ, and the other involves an excited state of ^{16}O which emits either a γ or a positron-electron pair, depositing about 6 MeV of energy into the detector in either case. The third channel of the (α, n) interactions involves the neutron scattering protons as it thermalizes, and the recoiling protons generate a continuous energy distribution of events with mean energy of around 1 MeV. All three of these processes are followed by the neutron capturing on hydrogen or carbon in the scintillator, and thus a 2.2 MeV γ-ray.

All three of these chains are tagged by the cuts used in the general coincidence tagging, the two chains that produce high energy prompt signals are tagged with nearly 100% efficiency while the prompt events from proton recoils should be tagged at about $\sim$70% efficiency. This means that even if the rate of the (α, n) decays is off by some large amount, we would still tag most of these events due to their particular coincident signals and see large numbers of events that look like (α, n)

in our tagged event selection, which we do not.

Another source of neutrons is from cosmic muons as they travel through the scintillator, but the 20 second veto applied after muon events is orders of magnitude longer than the neutron half life in scintillator so this background is negligible. Isotopes produced by muons can also undergo $\beta + \mathrm{n}$ decays, but these isotopes are all short lived and therefore removed by the muon veto.

These are the main sources of neutrons in the SNO+ scintillator and they all produce neutrons at rates far below what is required to explain the low energy excess. One source of neutrons not discussed here that will be discussed later is fast neutrons produced on the outside of the AV which then travel into the AV before they thermalize. This will be discussed along with other potential explanations that originate outside the AV.

6.2.4. Contamination from external events

External events are also another potential source of unexpected backgrounds, either due to mis-reconstructions that place events outside of our FV cut within the FV, or due to higher external rates than expected leading to larger external rates in the 3.3 meter FV. Mis-reconstructions were discussed above, all of which applies here, but a higher rate of externals can be explicitly tested by fitting out to higher fiducial volumes and seeing if the externals can explain the low energy excess.

Fig. 6.8 shows an example of such a fit with the water external backgrounds all allowed to float freely, since the backgrounds in the water have been shown to fluctuate throughout the basic scintillator phase. The fit shown here is now in 3 dimensions, energy, ITR, and $(\mathrm{r}/\mathrm{r}_{AV})^3$. As can be seen in the fit, there is still a clear component missing from the backgrounds model, the radial dimension of the likelihood fit shows that the rate of the externals in the region where they dominate (above 4.2 meters or so), is not high enough to explain the low energy excess in the 3.3 meter FV.

Of course, there could still be some effect where externals are mis-reconstructed such that the radial distribution of external events exponentially decays up until a certain point around 4 meters, where it then turns into a flat line. This is the worst case scenario for SNO+, since it would mean that the external rate seems to be exactly what we'd expect in the region where it dominates, and then

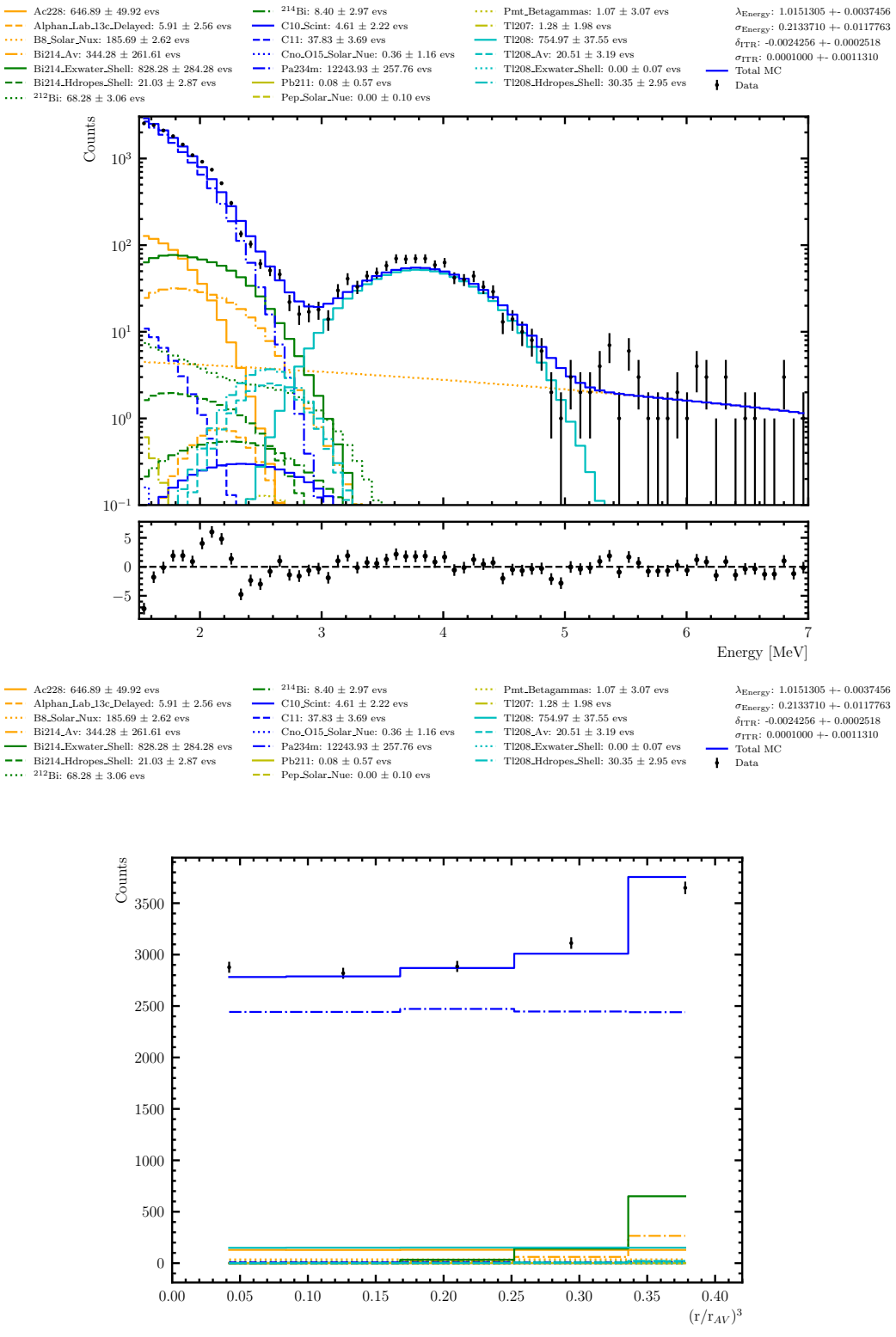


Figure 6.8: *Top:* Energy marginal of a 3D fit out to 4.5 meters in energy, ITR, and $(r/r_{AV})^3$. *Bottom:* $(r/r_{AV})^3$ marginal.

contributes orders of magnitude more in the regions where we believe it is negligible. One way we can simulate such an effect is by redoing the fits in the 3.3 m FV without using any constraints on the externals from outside the 3.3 m FV, and allow them to freely float to see if they can explain the low energy excess events. The fit shown in Fig. 6.9 shows such a fit using ^{214}Bi AV events as the flat-tailed externals (since one can see in Fig. 5.3 that the external ^{208}Tl events will not fit the low energy excess events), and as can be clearly seen even in this idealized scenario the externals look different from the low energy excess events. There are some caveats to the fit shown in Fig. 6.9 which are discussed in the figure caption.

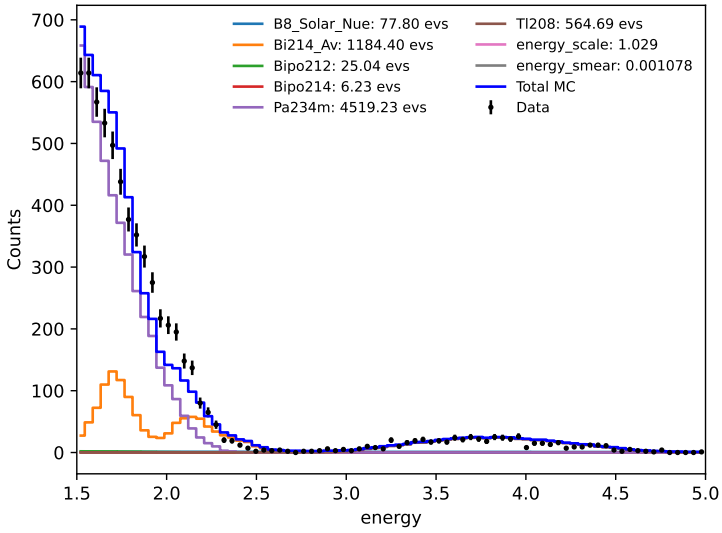


Figure 6.9: A fit with the ^{214}Bi AV events freely floated without using any radial information to constrain the rate. Since external events that make it all the way into the 3.3 m FV are rare, it is difficult to generate a large enough simulation dataset to have a high statistics sample of events close to the center of the AV. Instead, the external PDF used in this plot applies a radial cut of 4.5 meters to select the events to load into the PDF, while using the 3.3 meter cut for all the other backgrounds.

The variation in the external rates throughout the basic scintillator phase provides another way in which we can test how the external events are contributing to the low energy excess. This is shown in Fig. 6.10, where we compare the rate of events in the low energy excess region in the 3.3 and 5 meter FVs.

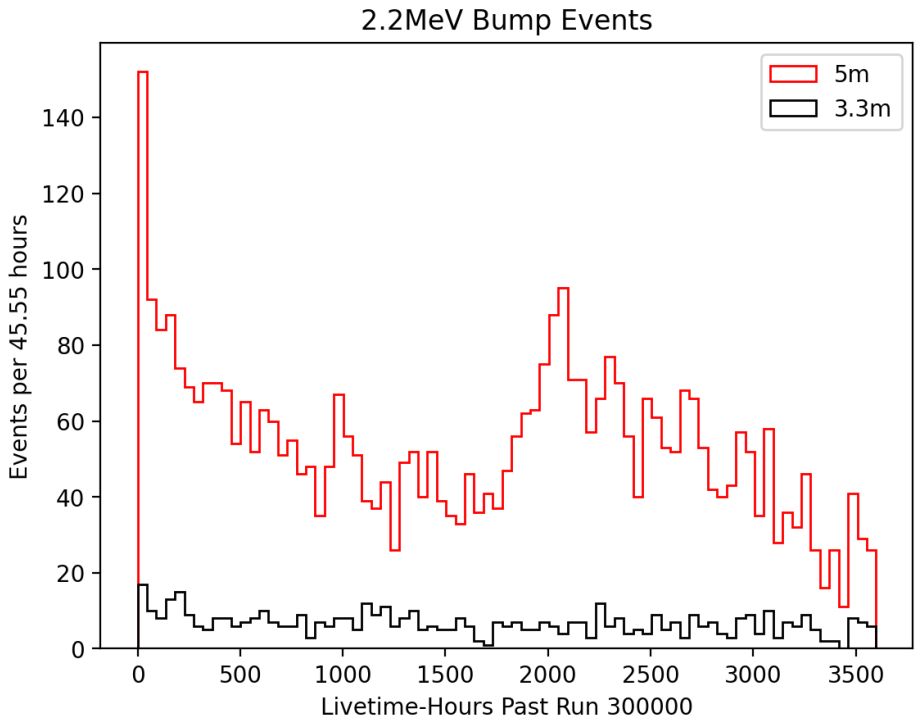


Figure 6.10: The rate of events in the low energy excess in two different fiducial volumes. The red histogram goes out to larger radius, and the fluctuation in rate of the external events can be clearly seen in the histogram. The black histogram does not show any correlation with the red histogram, which one would expect to see if the low energy excess events were caused by mis-reconstructed externals.

Since the externals become a dominant background at around 4 meters, in the 5 meter FV we can clearly see the fluctuation of the rate of events in the low energy excess region due to the rate fluctuations of the externals themselves. These fluctuations are also corroborated with other analyses focused on the external rates. On the other hand, we can clearly see that the rate in the low energy excess region in the 3.3 m FV remains relatively constant, even during times where the external rates increase by factors of 2 or 3. This also provides additional evidence that these events are not somehow connected to the external backgrounds.

6.2.5. Rare interaction or process

Perhaps the most exciting explanation is that we are seeing the signature of some rare interaction or process. Unfortunately, this is unlikely to be the case since other scintillator experiments like Borexino should have seen a similar signal if this was indeed some kind of rare neutrino interaction,

or some other process which led to the production of electron scatters.

Fig. 6.11 shows what the signal in the 3.3 m FV in SNO+ would look like if Borexino detected the same signal, weighted by the exposures of each experiment. The two red lines represent two rates, one where we consider the events between 2 and 2.5 MeV as a distinct signal, and one where we consider the rate of all events from 1.8 to 2.5 MeV together which is labeled as the "^{234m}Pa + 2.2 MeV" rate. The red lines in the plot are added by the author, and demonstrate what it would look like if one takes the rates seen in SNO+ and divides it equally among the bins from 1.8 to 2.5 MeV in the plot of the Borexino data. If we truly overlaid the shape of the spectra seen in SNO+, the red lines would be peaked at much higher values. This provides strong evidence that Borexino should have seen this signal if it was some kind of general physics signal, rather than a background specific to the SNO+ experiment.

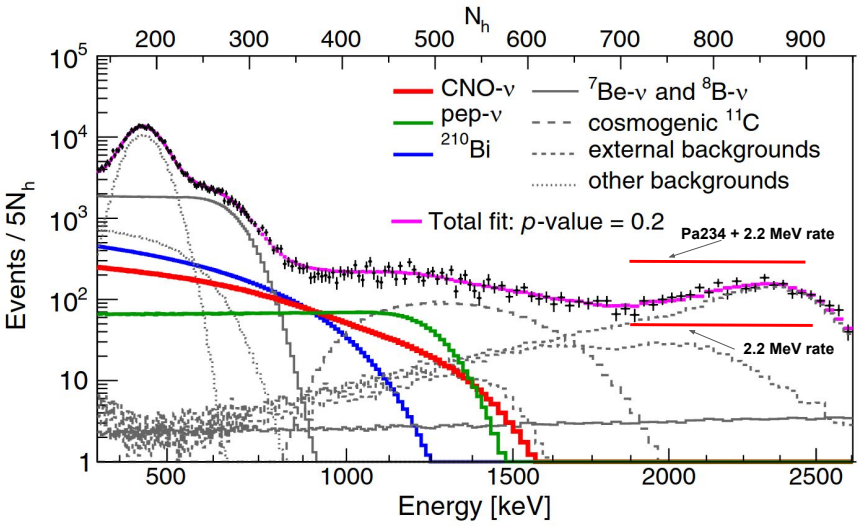


Figure 6.11: The energy spectrum seen by Borexino in their CNO results paper with red lines superimposed to represent the rates Borexino would see if the low energy excess was present in their dataset. From [45], with red lines representing rates seen in SNO+ artfully drawn by the author.

6.2.6. Rare isotopes

Another hypothesis is that there is some isotope in the detector that we have not considered yet that explains the low energy excess. Of course, there are many known isotopes and we have only simulated and included in our backgrounds model a miniscule fraction of all the isotopes that could

be in our detector, so perhaps we just missed one or two. This isotope would have to be rare in the sense that it's something not typically found in scintillator detectors since that would be something well known in the literature.

When we consider all the different criteria this isotope must meet, though, it becomes harder and harder to find any potential explanations. For one, the isotope must be long lived since we do not see any obvious evidence of the isotope decaying away throughout the basic scintillator phase. Since we do not observe any clear coincidence structure, the half life of any parent or daughter isotopes must be long enough that we would not see it in our coincidence tagging (something like seconds to minutes). And of course, the decay must be high enough energy to actually produce events in the low energy excess region, which means the endpoint of the decay must be around 2.5 MeV.

It turns out that this set of criteria drastically limits the number of potential isotopes. An internal SNO+ document created by V. Lozza goes through all the potential isotopes in detail, including information regarding their half lives, production modes, Q-values, decay modes, etc. Without going into the detail of each isotope, it is worth listing out the isotopes individually to give a sense of the number of isotopes that have been considered: ^{46}Sc, ^{51}Ti, ^{90}Y, ^{104}Rh, ^{111}Pd, ^{119}In, ^{136}Cs, ^{139}Ba, ^{141}La, ^{145}Ce, ^{194}Ir, ^{227}Fr, ^{58}Co, ^{102}Rh, ^{102m}Rh, ^{144}Pm, ^{150}Eu, ^{172}Lu, ^{194}Au, ^{207}Bi.

Many of the isotopes on this list are not long lived enough to provide the constant rate of low energy excess events we are seeing in the detector. To be thorough, we have also consider the production of these isotopes through neutron absorption or spallation from cosmic muons. Back of the envelope calculations have shown that one would need the number of atoms of certain isotopes to be close to the number of protons in the detector to produce these isotopes at a rate which maintains the rate of events in the low energy excess region.

Two isotopes did seem promising though, with appropriately energetic endpoints and long enough lifetimes to remain throughout the entire basic scintillator phase: ^{102}Rh and ^{102m}Rh.

6.2.7. ^{102}Rh and ^{102m}Rh as potential explanations of the low energy excess

^{102}Rh and ^{102m}Rh both fit many of the requirements discussed previously of a potential source of the events in the low energy excess. ^{102}Rh has a half life of 207 days and a Q-value of 2.32 MeV, making it a promising candidate. ^{102m}Rh also has a long half life, 3.7 years, and an endpoint that fits perfectly within the low energy excess window of energies, 2.46 MeV.

Including ^{102}Rh and ^{102m}Rh into the backgrounds model and fitting the dataset provides compelling evidence that these isotopes are what we see in the detector. This can be seen in Fig. 6.12 which shows decent agreement between the data and MC, and importantly the systematic parameters are much closer to the means of their prior distributions.

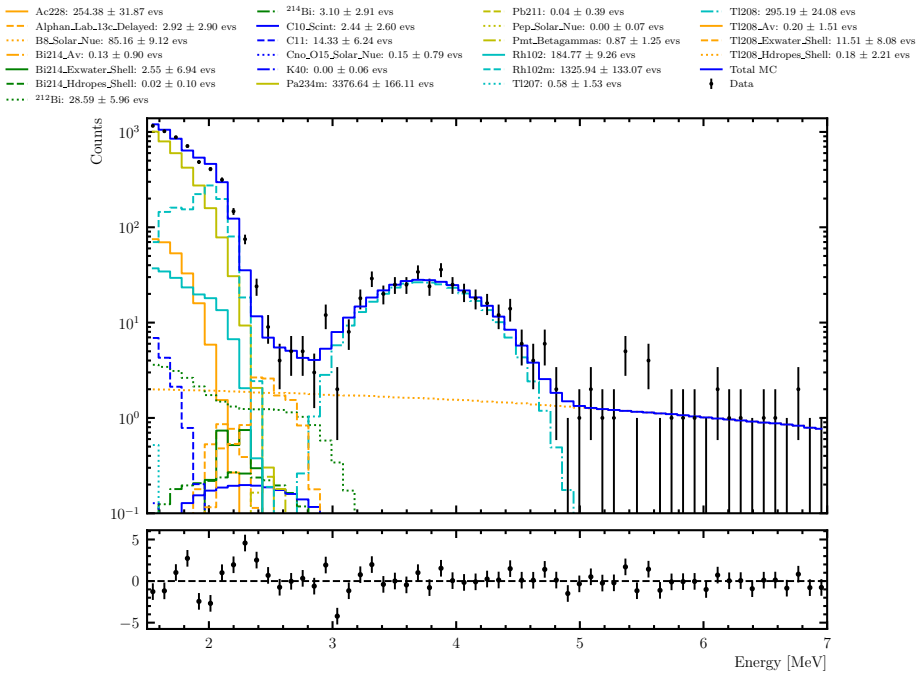


Figure 6.12: Energy marginal of 2D fit in the 3.3 m FV with ^{102}Rh and ^{102m}Rh included in the backgrounds model.

It is difficult to produce these isotopes in the quantities needed to explain the low energy excess, especially in a deep underground detector like SNO+. The production of both isotopes would need to come from the neutron activation of either ^{102}Pb or ^{103}Rh, and for this to take place inside the detector, given the number of neutrons produced in the detector, the concentration of either the

^{103}Rh or the Pb would need to be on order of the proton density, which is impossible of course.

But we can posit that the Rh isotopes are produced above ground where both muons and neutrons are far more abundant, and then the isotopes are somehow piped into the detector. This is a hard scenario to justify given all the effort that goes into keeping the scintillator as clean as possible, and given how rare Rh is in general. Thus we must have very good evidence to support the claim that isotopes of Rh are the source of the low energy excess, something even more convincing than the fits alone.

A new kind of fitter was developed to try and provide this evidence, which attempts to reconstruct events under a different hypothesis than the single-electron hypothesis. ^{102}Rh is a positron emitter, and positrons have a 36% chance to form a bound state between an electron and a positron called ortho-positronium before annihilating [77]. This bound state has a longer half life than the free positron in the scintillator, 2.7 ns. The 2.7 ns half life is long enough that we should have some sensitivity to fit out the time difference between the positron emission and then the annihilation, especially when we consider that the two processes should occur at the same vertex.

Multi-vertex fitting

As hinted at in the previous paragraphs, the idea of the multi-vertex fitter is that for the specific signal of positron annihilation, we can try and look for events where the positron forms the bound state ortho-positronium before annihilating. We have a nice advantage in that we know that the emission and annihilation vertices should be in the same spot for these events.

The multi-vertex fitter works in much the same way as the position-likelihood fitter discussed in Sec. 4.2.1, but instead of floating only one vertex we float two, but defined such that the two vertices differ only in their times. Furthermore, we always know the energy released in the positron annihilation since it is just $2m_e = 1.02$ MeV, which allows us to back-calculate the energy the positron must have been emitted with to create an event with the detected total energy.

Combining all these ideas we can define a likelihood where we have

$$p_T(x|r_1, r_2) = p(x|r_1) * \frac{N_{\text{ann}}}{N} + p(x|r_2) * \frac{N - N_{\text{ann}}}{N} \tag{6.1}$$

where $r_1 = (t_1, x, y, z)$ is the annihilation vertex and $r_2 = (t_2, x, y, z)$ is the positron emission vertex. $p(x, r_i)$ is the normal hit time residual PDF used in the standard position-likelihood fitting, N is the total number hits in the event and N_{ann} is the expected number of hits from the annihilation γs, which is just a constant calculated ahead of time. Again, the two vertices differ only in their times, and we expect $\Delta t = t_1 - t_2$ to be distributed according to an exponential decay function with a half life of 2.7 ns.

This fitter was tested both on ^{102}Rh simulations and reactor antineutrino events (which produce positrons) the results of which can be found in Fig. 6.13. As can be seen in the figure, the approach seems to work well based on both the technique applied to MC and a (small) sample of data which consists of the tagged reactor $\bar{\nu}$ events from the basic scintillator phase.

Unfortunately, when we apply this reconstruction algorithm on the events in the low energy excess, we do not see the same tail in the Δt distribution between the two vertices. These results can be seen in Fig. 6.13. This suggests that the extent to which the two Rh isotopes improved the agreement between data and MC was completely coincidental.

At this point, it feels clear that we must take a step back and try and think of a more powerful approach where we can test collections of possible explanations for this low energy excess rather than stepping through the hypothesis space isotope by isotope. This is a hard problem though, how can we even start to step through this space that is made up of a continuum of possible backgrounds models? This is what we will attempt to do in the next section.

6.3. A generic backgrounds model

As mentioned at the end of the previous section, a more powerful approach than trying to find potential isotopes that explain the low energy excess is instead trying to come up with a generic approach that can continuously search through the space of all possible backgrounds models. At

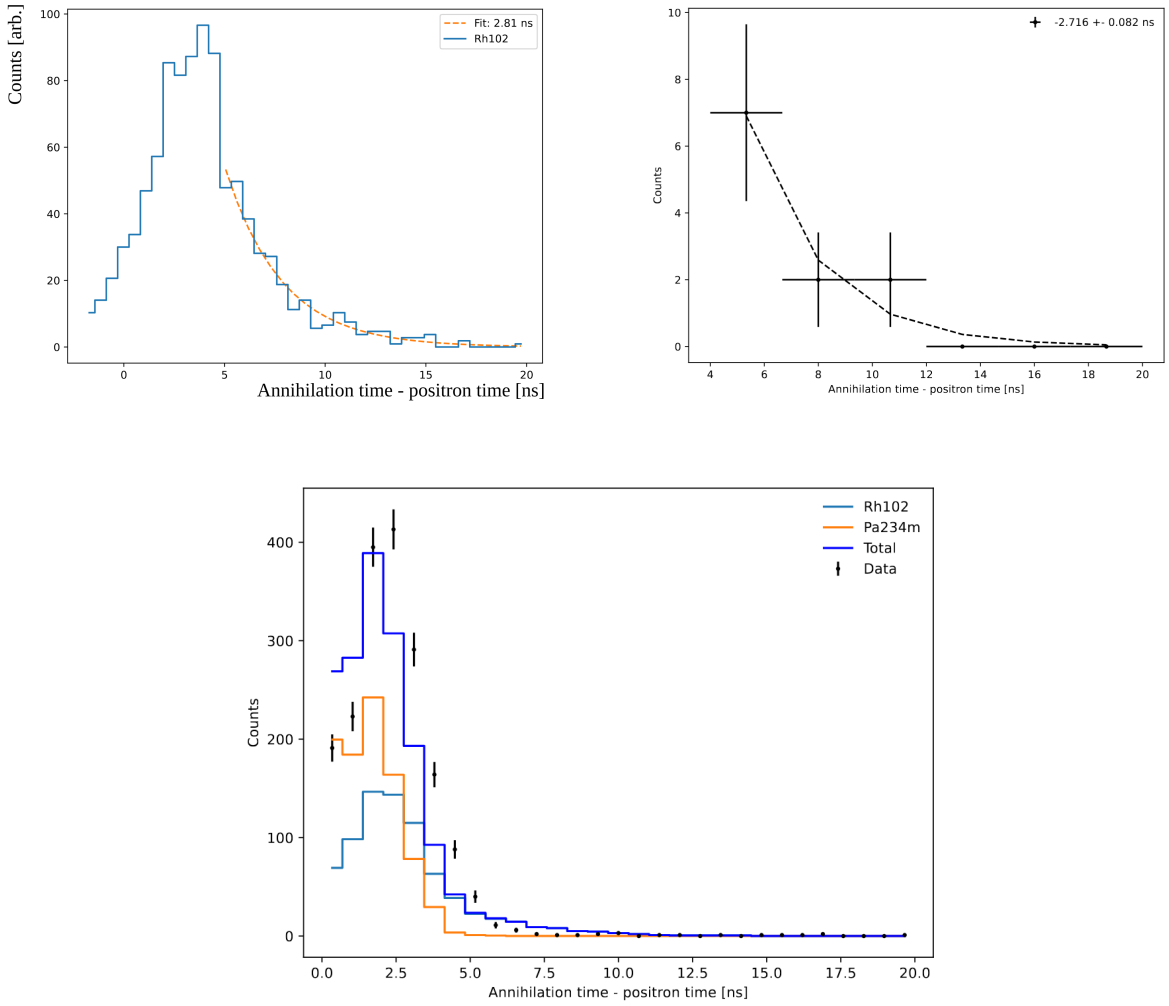


Figure 6.13: *Top left:* The time separation between the two vertices in the multi-vertex fitter from simulations of ^{102m}Rh decays. *Top right:* The time separation between the two vertices in the multi-vertex fitter when applied to tagged reactor $\bar{\nu}$ events. An exponential fit to a histogram of the time differences agrees nicely with the expectation based on the ortho-positronium half life. *Bottom:* The results of applying the multi-vertex fitter on the low energy excess events. There is no clear signal of the formation of ortho-positronium.

first glance, this feels like a hopelessly under constrained problem, one where it is impossible to tell if any progress has been made towards creating a model with explanatory power.

If we think at the coarsest level about what could cause the low energy excess events at this point though, we can make some real progress towards trying to parameterize this anomalous

background. For example, the end of the last section showed that it is unlikely these events are positron emissions, since we should see the half life of the orthopositronium in the multi-vertex fitter. We also know that these events cannot be explained by some kind of α emission since the αs are quenched in the scintillator, and there is no way to generate αs of a high enough energy to explain these events. For similar reasons, it's also very unlikely that these events could be the result of protons moving through the scintillator, and both protons and αs would have hit time residual distributions distinct from ^{214}Bi events, which goes against what we see in Fig. 6.5.

At this point, we can conclude that these events must be caused by some kind of β, γ, or β/γ emitter that we haven't been able to name yet. To first order, the β and γ emitters are both easy to parameterize by just a single number, the endpoint for the β and just the energy for the γ. The β/γ on the other hand can be parameterized with two variables, the total energy shared between the β and the γ and the fraction of the energy that goes to the β endpoint. Equipped with these parameterizations, we can think of the possible models that can explain the low energy excess as collections of L β, M γ, and N β/γ emitters where L, M, N are all just integers.

Occam's razor tells us that we should prefer the simplest models, which means that we should order the space of all models based on the number of parameters a given model has. Each β and γ add a single free parameter to the model, the endpoint, whereas a β/γ emitter adds two free parameters, the endpoint and the beta fraction (the fraction of energy that goes to the Q-value of the β). This means the first 4 models we should consider are: 1 β, 1 γ, 1 β and 1 γ, 1 β/γ. This process can be continued for as long as one desires. Each β, γ, or β/γ will be referred to as an "empirical background" to communicate the fact that these signals are data driven in the sense that the datapoints inform us about the shape of the signal.

Of course, we will need something a little more formal than Occam's razor to compare between each of the models. We will take a Bayesian approach and use techniques from the sub-field of Bayesian model comparison to inform our "belief" in one model relative to another.

6.3.1. Bayesian model comparison

The fits presented thus far have all consisted of a single model M which is described by some number of parameters $M(\theta_1, ..., \theta_j)$. In this analysis, we have found ourselves in a situation where there is some inherent uncertainty in what backgrounds model is the correct backgrounds model, and we need to compare models with different numbers of parameters, i.e. models with different numbers of βs, γs, and β/γs. Furthermore, we want to compare or combine predictions from the different models, for example when we measure the low energy ^{8}B neutrino flux. The quantity we would like to compute then is $p(M|x)$, or the probability of a model M given the data x. The following discussion follows the material of Appendix A.

As seen in the appendix, applying Bayes' theorem gives us

$$p(M|x) = \frac{p(x|M)\pi(M)}{p(x)} \tag{6.2}$$

where now $p(x)$ is a sum over the different models we are considering, $p(x) = \sum_i p(x|M_i)\pi(M_i)$, and $\pi(M)$ is the prior we assign to a given model. We can see that for a given model, $p(x|M)$ is equivalent to the evidence for that model, or the integral over the entire likelihood space

$$Z_M = \int \mathcal{L}(x|\theta, M)\pi(\theta)d\theta = p(x|M), \tag{6.3}$$

therefore the posterior probability for a given model is directly proportional to the evidence. Rewriting $p(M|x)$ we now have

$$p(M|x) = \frac{Z_M\pi(M)}{p(x)}. \tag{6.4}$$

This framework also provides us with a way to marginalize over our uncertainty in the backgrounds model when different models share a given parameter θ_j. This can be represented as the weighted sum of the individual model posteriors of the parameter of interest. where the weights are the prior

weighted evidences of the different models:

$$p(\theta_j|x) = \sum_i \frac{p(\theta_j|x, M_i) Z_{Mi} \pi(M_i)}{p(x)} \tag{6.5}$$

It is important to note that the evidence calculation naturally contains an effect similar to Occam's razor, where more complicated models are penalized relative to simpler models with the same predictive power. This is encoded in the prior weighted measure of the integral, $\pi(\theta)d\theta$. For simplicity we will consider a scenario in which we have two competing models, one with two parameters and prior volume l^2 and one with three parameters with prior volume l^3, or in other words all the parameters encompass the same prior volume l. Let's also assume that the additional parameter in the second model provides no additional prediction power, so the likelihoods are identical between the two models.

One can see that in this simple scenario, if the evidence of the 2 parameter model is Z_2, the evidence of the 3 parameter model is $Z_3 = \frac{Z_2}{l}$, since the prior volume enclosed increases by a factor of l. Another way to interpret this result is the prior space became less dense as we added in the third parameter, and this decrease in density is reflected in the evidence. Any parameters added to the model will make the prior space less dense, but this can be in general offset by an increase in the value of the likelihood function if the parameter does indeed make the fit much better. This is a very powerful result of Bayesian model comparison, and a mathematical way of representing Occam's razor.

In the analysis presented in this thesis, we will consider many different background models, and as such we will make use of the results in this section to compare and combine their results. This means accurately calculating the evidence for the different models will play an integral role in this analysis.

6.3.2. Nested sampling

As the integral in Eq. (6.3) is evaluated over the entire parameter space θ, it will often be a $\sim$30 dimensional integral in this work. It has no direct analytical form given that we are using a binned

PDF approach, which means that we must rely on numerical techniques to calculate the integral.

Monto Carlo integration using the VEGAS algorithm [78] was initially used to try and evaluate the integral, but this approach frequently did not sample the space adequately enough to find all the high density regions of the parameter space, and taking enough samples to properly scan the entire space led to prohibitively long run times. Additionally, in lucky cases where the algorithm did find the high density regions in the parameter space, the algorithm would often over-sample just that part of the parameter space. This lead to results that could be systematically biased both high and low relative to the true integral.

Eventually it was decided to use the nested sampling algorithm instead [79], which is specifically designed for calculating evidences within the Bayesian framework. The nested sampling algorithm essentially samples the likelihood space according to the prior distributions of the parameters, and in doing so the multidimensional integral shown in Eq. (6.3) turns into a 1D Riemann sum over shells of constant likelihood with widths $\pi(\theta)d\theta$. A more detailed discussion of nested sampling can be found in Appendix B.

The results of nested sampling run are a collection of points that sample the likleihood space according to the prior, each weighted by the product of the volume of the prior space enclosed at that sample and the value of the likelihood at that sample. These weights are the individual summands of a Riemann sum that approximates the evidence. The weights and the samples can also be used to approximate the posterior of a given parameter:

$$p(\theta_i) = \sum_{j=0}^{\mathrm{M}} \frac{Z_j \theta_{ij}}{Z} \tag{6.6}$$

where θ_{ij} just means the i^{th} component of the j^{th} sample and M runs over the total number of samples. This is also shown in Eq. (B.5), we've just shown it here as well for convenience.

A nested sampling implementation developed by the author has been included in snopfit (Sec. 5.4). A more in depth discussion of nested sampling as well as the specifics of the algorithm implemented

for this work are discussed in Appendix B. It is worth mentioning here, though, that we use a "slice sampling" approach to generate the new samples. Both 1D and ND slice sampling approaches are used, 1D for the normalizations and ND for the systematic parameters since they are expensive to update and updating one requires recomputing the effects of all of the systematic parameters.

To verify the algorithm, nested sampling runs were done with likelihood functions and priors that result in analytically calculable evidences.

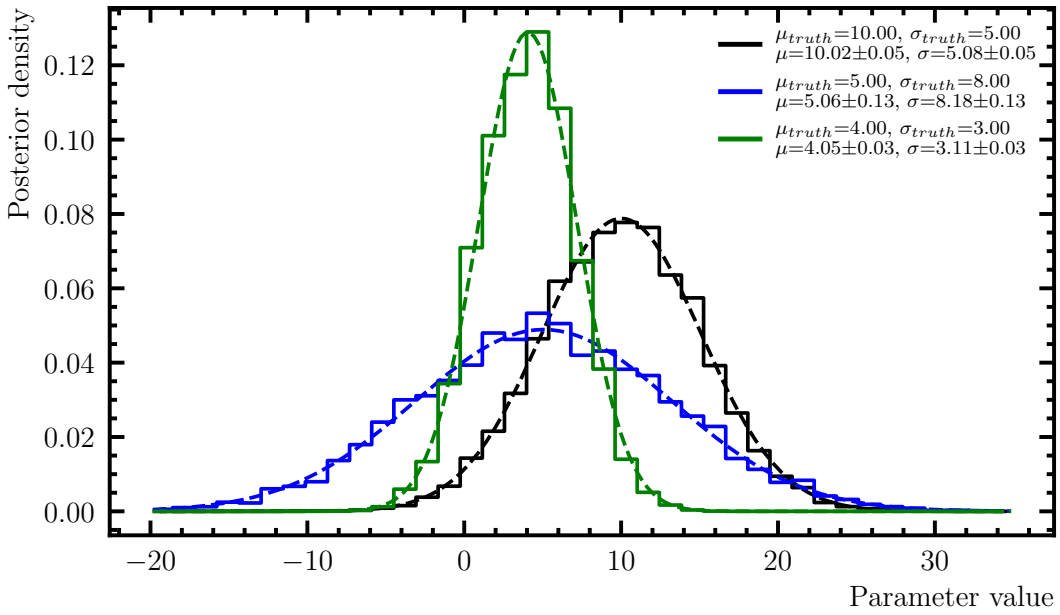


Figure 6.14: The results of a nested sampling run with three parameters. The priors for every parameter are uniform. The likelihood is the product of Gaussian functions, one for each parameter with the mean and standard deviation defined in the legend.

Figure Fig. 6.14 shows the results of a nested sampling run where the likelihood function is the product of Gaussians, with one Gaussian for every parameter. It can be seen in the figure that the posterior distributions are what one would expect given the parameters that define each of the Gaussian functions.

6.3.3. PDF generation and interpolation

One more difficulty in trying to fit arbitrary parameterizations of backgrounds is that since the parameters defining the different empirical backgrounds models (the endpoint and β fraction) are

continuous, we cannot generate simulations of the entire parameter space of models. Instead we must rely on some interpolation strategy to generate PDFs in between the points at which we do simulate the empirical backgrounds.

The simulations to generate the empirical backgrounds models use the RAT β decay generator to generate βs according to the β spectrum of the ^{214}Bi nucleus. This nucleus was chosen at random, and since the details of the nucleus only change the shape of the β decay spectrum by a small amount we are not sensitive to this choice. Of course, the endpoint of the β spectrum is set according to the empirical background PDF we want to generate. The γ empirical backgrounds are simpler — these simulations are ran using the `gun` generator, which "shoots" gammas throughout the detector at specified energies.

The β/γ simulations are ran using a combination of the two approaches above, where both a β and a γ are generated in the same event window, generating events that are distinct from either pure β and pure γ events, as was discussed in Sec. 5.1. As mentioned above, this class of empirical background is defined by two parameters, an endpoint and a beta fraction, where the beta fraction defines the ratio of β endpoint relative to the energy of the γ. Note that this means that a some of the β/γ events will have βs with arbitrarily low energies since the energy of the β in each individual event is still sampled form a β spectrum. One can also think of the β and γ empirical backgrounds as limiting cases of the β/γ backgrounds, where the β is the limit in which the beta fraction goes to 1, and the γ is the limit in which the beta fraction goes to 0. Fig. 6.15 shows how the energy and ITR [3] distributions of the β/γ events change as a function of the endpoint and beta fraction.

Now that we know how to construct the PDFs for a particular endpoint or endpoint and beta fraction, we can begin to try and attack the problem of interpolating between the different PDFs.

Linear interpolation

One potential solution to this problem is interpolation, or using simulation to generate the transformed distributions at discrete points and using interpolation to build representative PDFs between

[3] The plot shows the multisite distribution, not the ITR distribution, but these two parameters convey essentially the the same information.

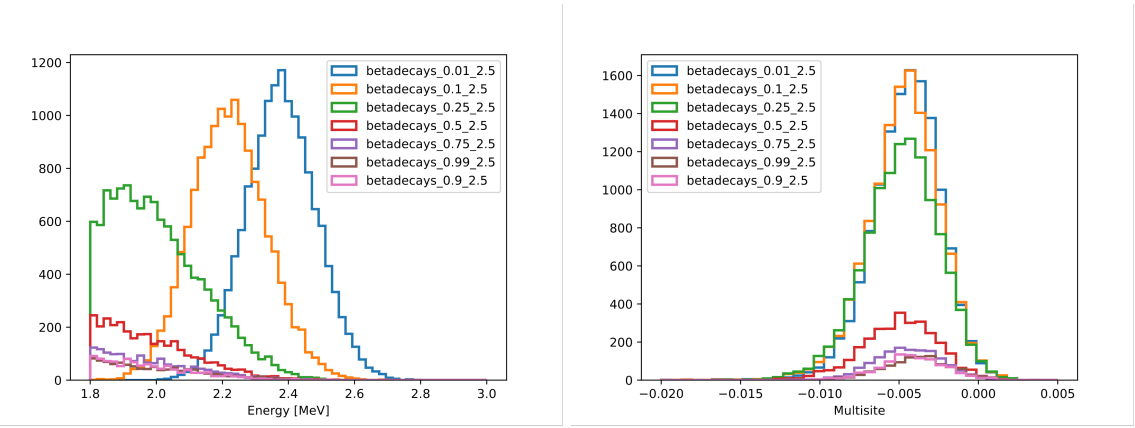


Figure 6.15: Energy and multisite distributions of β/γ backgrounds simulated with different endpoints and beta fractions. Note that multisite is essentially a more advanced version of ITR that essentially carries the same exact information.

the simulated points. The most obvious approach to do this interpolation is linear-interpolation which may work for certain types of re-weightings of the PDFs, but in general this will lead to "ghosting effects" which are almost always unwanted. An example of these ghosting effects can be seen in Fig. 6.16, where even though the two distributions we are interpolating between are unimodal, the interpolated distributions develop a bimodal behavior due to the difference in means of the two input distributions.

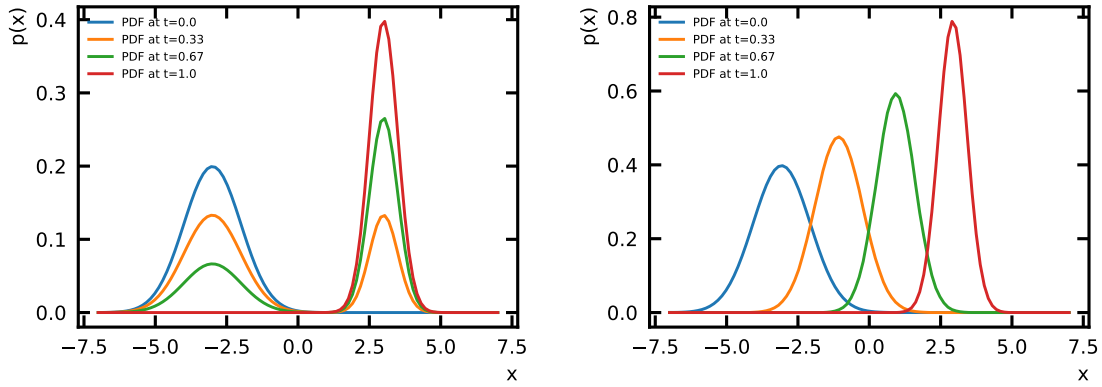


Figure 6.16: Comparison between linear and McCann interpolation of two Gaussians with different means and standard deviations at 4 different interpolation points. *Left*: Linear interpolation showing the "ghosting" effect that can occur when PDFs do not lie on top of one another. *Right*: McCann interpolation.

This ghosting effect makes linear interpolation unusable for applications where the probability mass

of the PDFs shift at the different interpolation points. We need something that somehow is aware of how probability "moves" as we move in the parameter space in which we are attempting to interpolate. This brings us to the field of optimal transport.

Optimal transport based interpolation

Optimal transport re-frames interpolation in terms of how probability is *transported* as one moves between PDFs. This new interpretation proves to be very useful, as one can define interpolation in terms of the paths probability masses will trace out as we vary the parameters in which we are interpolating. An introduction to optimal transport and a review of the basics of measure theory needed to understand some of the important results of optimal transport can be found in Appendix C.

The results from optimal transport that we are interested in is that the optimal mapping (or the mapping that minimizes the distance traveled by probability masses as they move from on distribution to the other), is a solvable problem in 1 dimension. Specifically, the map from one 1D distribution μ to another ν is given by Sec. C.2

$$x \to F_\nu^{-1}(F_\mu(x)). \tag{6.7}$$

where F_λ is the quantile function of the distribution λ, or the inverse cumulative distribution function (CDF). Furthermore, one can construct an intermediate map, or a map that interpolates between the two distributions, using the following Sec. C.2:

$$F_t^{-1} = (1-t)F_\mu^{-1} + tF_\nu^{-1} \tag{6.8}$$

where t varies from 0 to 1. This map is instead meant to map the uniform distribution from 0 to 1 onto the interpolated distribution, which is a more useful parameterization.

Now that we have built up some useful results by applying optimal transport in 1D, we can begin to layout the PDF interpolation algorithm. Imagine some space $V = \mathbb{R}^n$, and at every point in V we have some multi dimensional PDF $p_v(x)$. v in $p_v(x)$ represents the coordinate in v at which

$p_v(x)$ resides, which is in the space of V. x resides in the coordinate system on the PDF itself, some coordinate in $\mathbb{R}^c$ where c is the dimensionality of the PDF. We will call the space V the interpolation space, since it contains the points over which we would like to interpolate. Fig. 6.17 illustrates how one can think of the PDFs as being embedded within the interpolation space.

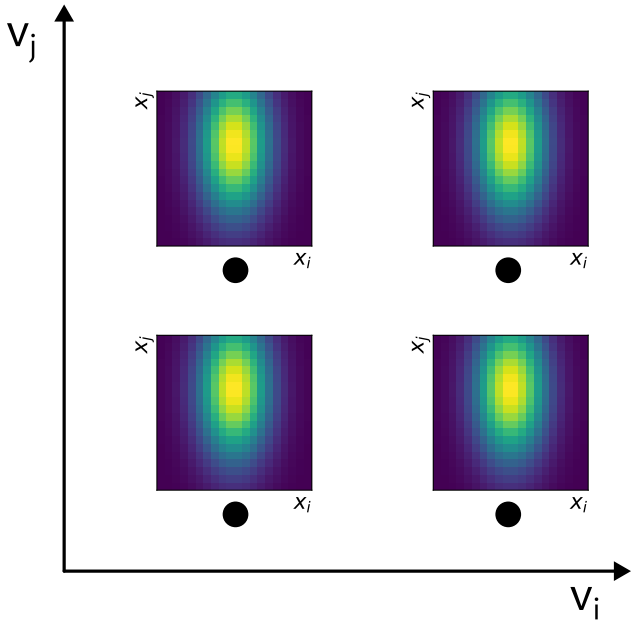


Figure 6.17: A representation of the interpolation space V and the coordinates on the PDFs x. Each black dot is a coordinate in V, and we associate each coordinate in v with a distinct PDF. In this figure both the interpolation space and the PDFs are 2-dimensional, but in general the dimensionalities of each can be different.

Of course, V is continuous so if there is no analytic way to describe how $p_v(x)$ changes as one explores the space V, then an interpolation strategy must be used. As was shown in Fig. 6.16, linear interpolation often creates unwanted artifacts in the PDFs, so an optimal transport based approach is much preferred. The issue with optimal transport is that finding the optimal transport maps for dimensions greater than 1 is difficult as no closed form solutions are known, so one must resort to numerical techniques.

The PDFs employed in this work are typically more than 1 dimension, so we must come up with an approach that is applicable for an arbitrary number of dimensions. Unfortunately, the techniques developed to calculate optimal transport maps in more than 1 dimension require minimization,

which is also undesirable since this would require a "meta"-minimization step during likelihood fitting [80]. Another way to try and attack the problem is to recall that for any arbitrary dimension PDF, we can always construct 1D distributions by marginalizing over every dimension aside from one. Therefore, if we develop a strategy that uses only on the marginals of PDFs we are free to use the results shown in Sec. C.2 for 1D distributions, and we can generalize to arbitrary dimensions in a very straightforward manner — we just repeat the process for each dimension of the PDF.

The inputs of our interpolation algorithm are pairs of input PDFs and points in the interpolation space V, as well as the point in V at which one wants to interpolate, which we will call the interpolation point. For simplicity, this discussion will restrict the input points of the PDFs to be distributed along a uniform grid, therefore for any interpolation point the nearest input points are those that form the n-dimensional hyper-rectangle that surrounds it (this work only deals with spaces where the dimensionality of V is less than or equal to 2, which means the "hyper-rectangles" are either line segments or rectangles.). This isn't a requirement of this technique in general, it can actually be used on an arbitrary set of points for the input points as long as they span all the points at which one may want to interpolate. Since the input points and PDFs do not change, we can assume that for each input PDF we also cache a CDF and quantile function (QF) to avoid recomputing these each time the algorithm is ran.

We begin the interpolation process by finding the hyper-rectangle that bounds the interpolation point. There are many solutions to this problem, any number of algorithms can be used to find these bounding points. Once we have these points, we start the per PDF dimension part of the algorithm. To keep track of which dimension of the PDF we are currently evaluating we will use $i_d = [0, c]$ which increments every time we move on to the next dimension. First we calculate the Wasserstein barycenter (W_b) the i_d marginals of the PDFs located at the points in the bounding box. The weights used for the vertices of the bounding box are identical to those used in linear and bilinear interpolation. For 1D interpolation with lower and upper bounds v_u and v_l the weights are $w_u = \frac{v_l}{v_u - v_l}$ and $w_l = \frac{v_u}{v_u - v_l}$. In 2D the situation is more complicated, but one effectively weights by the area of the rectangle formed by the interpolation point and the vertex opposite of

the point being weighted. Fig. 6.18 shows this graphically. The Wasserstein barycenter is then calculated according to Sec. C.2, where the weights w_i are defined as just described and the F_{ν_i} are the marginals of the PDFs at the bounding points for the current dimension.

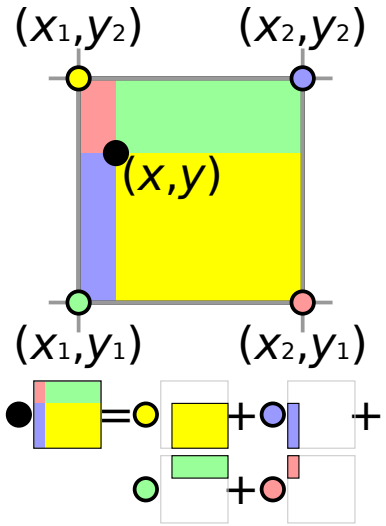


Figure 6.18: Weights assigned to vertices of the bounding rectangle in bilinear interpolation. From [81].

Once we have calculated W_b, we can calculate the map T_i from the marginals of each of the i bounding points to W_b using Sec. C.2. Now if we assume that the marginals transform in a similar way to how each 1D slice of the PDF in i_d transforms, applying the map to each 1D slice in the PDF will given an approximation to the true PDF at the interpolation point. Note that this doesn't result in a PDF whose 1D slices are all identical — the 1D slices all retain their prior shape, that shape has just been transformed in the way that the marginal transforms when optimally transported to the interpolation point. Since the same transformation is applied to every slice, we also avoid any issues that arise from other dimensions being affected by the interpolation procedure since probability moves between the bins in the 1D slices in an identical fashion from slice to slice. This process is then repeated for every dimension in the PDF. An illustration showing these steps visually can be seen in Fig. 6.19.

As the process described in the last paragraph is applied to the PDF at each bounding point, it will result multiple different interpolated PDFs at the interpolation point, one for each of the bounding

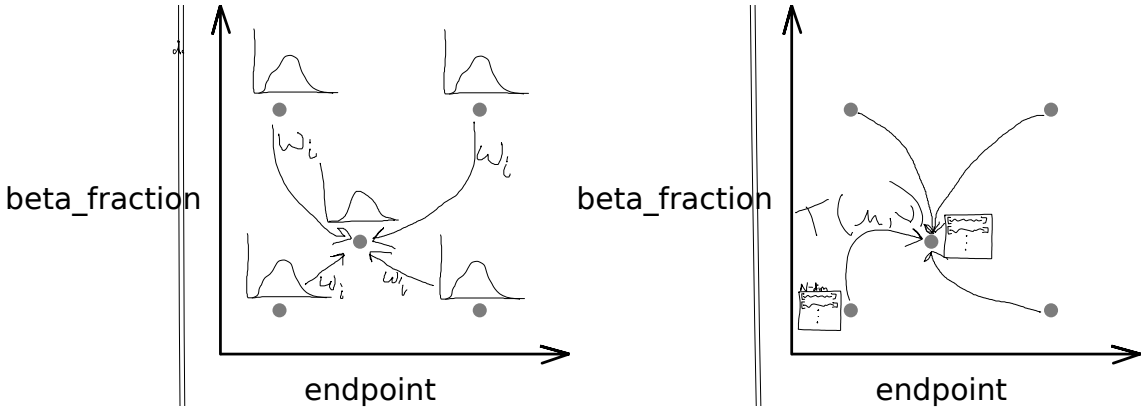


Figure 6.19: *Left*: Illustration of the W_b part of the interpolation algorithm. For each dimension in the PDF we calculate the marginals and then use the weighted QFs of the marginals to calculate W_b. *Right*: Illustration of the application of the map found in the previous step to each 1D slice in PDFs.

points. This is because the optimal transport map applied to the 1D slices will not in general recreate the same PDF from each of the bounding points. At this point, though, if the assumption that the mapped 1D slices will approximate the true PDF is indeed true, there should only be minor differences between the resulting PDFs. In this limit linear interpolation directly in the PDF space should not produce any artifacts, so the final PDF is constructed by averaging over each of the PDFs weighted by the same weights used when calculating the Wasserstein barycenters.

This specific interpolation algorithm appears to be novel, and is a invention of the author. Studies comparing it to more empirical approaches like the ones used in the ROOT framework are planned but have not been carried out yet.

Returning back to the problem at hand, using this technique to interpolate between two 2D PDFs for the β/γ empirical background can be seen in Fig. 6.20. As can be seen in the figure, the interpolated distribution matches the true distribution closely. Importantly, there does not seem to be a systematic difference between the two distributions.

6.3.4. Fits including empirical backgrounds

Now that we have all the required machinery, we can finally look at fits which float some number of these empirical backgrounds. This section does not intend to try and prove which model is optimal,

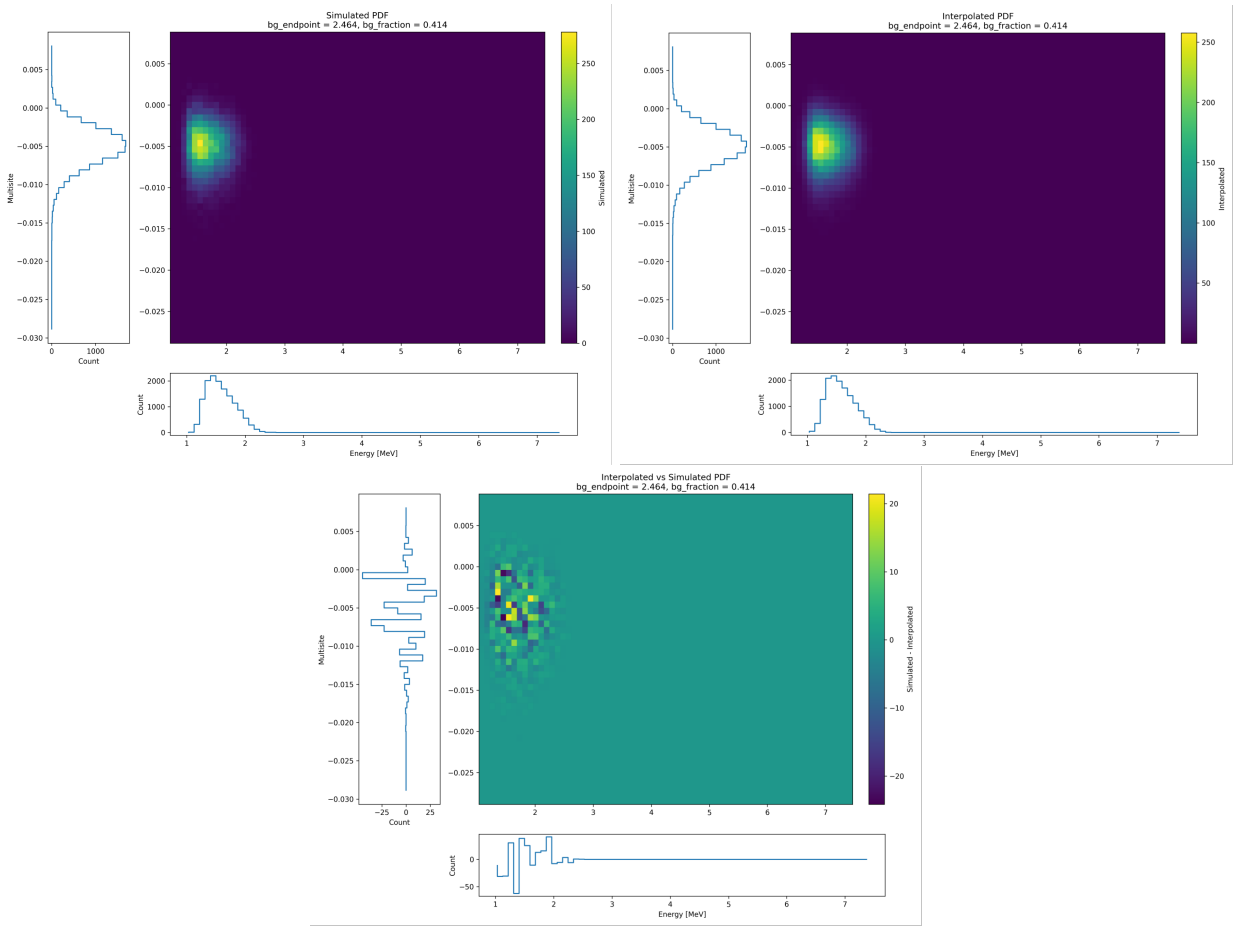


Figure 6.20: *Top:* An example of the simulated PDFs for a given empirical backgrounds (left) and the interpolated empirical background (right) with interpolation points at 2.25 and 2.5 MeV in energy and 0.40 and 0.425 in beta fraction. *Right:* The residuals between the two distributions on top.

which will be done in the next section, but rather we will show that a given model which includes empirical backgrounds does have some explanatory power. The way will do this is by fitting a model which includes an additional γ and β/γ to the basic scintillator phase dataset. This fit will provide us with an endpoint (or for γ-rays we typically just call this the energy) for the γ and an endpoint and a beta fraction for the β/γ. We can then simulate this particular model in the enhanced scintillator optics and fit the enhanced scintillator dataset with this particular model, to see if it indeed has any predictive power. It is important to note that we are not floating the shape of the empirical backgrounds in the fit to the enhanced scintillator phase, we are essentially holding the shape of the empirical backgrounds constant at their MLE values from the fit to the

basic scintillator phase dataset.

The fit to the basic scintillator phase dataset and the enhanced scintillator phase dataset can be seen in Fig. 6.21.

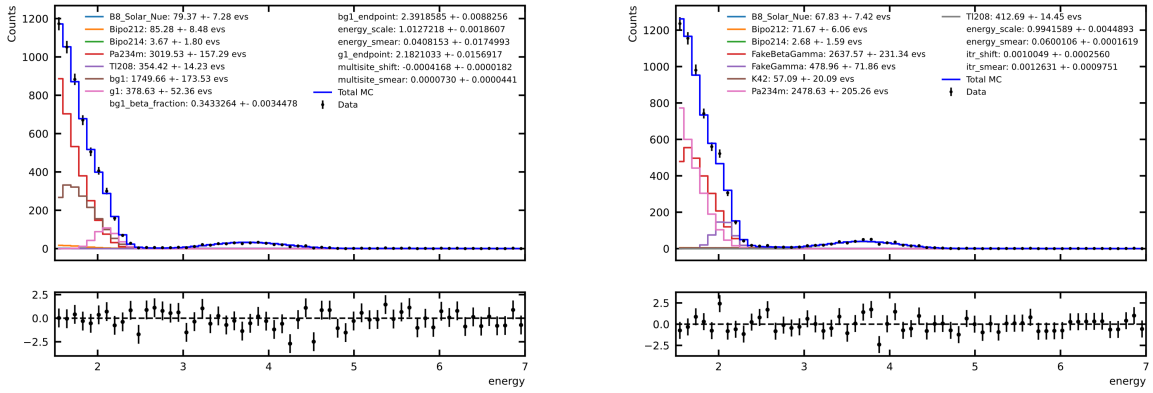


Figure 6.21: *Left:* A fit to the basic scintillator phase dataset in the 3.3 m FV where the shapes of the two empirical backgrounds are floated. The fit prefers a γ endpoint of 2.18 MeV and a β/γ endpoint of 2.39 MeV with a beta fraction of 0.34. *Right:* A fit to the enhanced scintillator phase dataset using empirical backgrounds simulated only at the optimal shape parameters found in the fit to the basic scinitllator phase. Good agreement between the data and MC can be seen in both fits, suggesting that the model predicts the empirical backgrounds present in the enhanced scintillator phase.

Note that this test is also more supporting evidence that the events in the low energy excess are indeed physics events. The change from the basic scintillator phase to the enhanced scintillator phase involves a substantial increase in the light yield, which is reflected in the empirical backgrounds since they are simulated using the optics model corresponding to each phase. It is hard to see how an instrumental effect would transform in the same way as real physics events when switching between the phases.

CHAPTER 7

A Measurement of the Low Energy ^{8}B Neutrino Survival Probability

Now we are finally at a point where we can do fits to extract a physics result from the basic scintillator phase. Given the complexity involved in this approach, it is worth reviewing how we plan to go from fits over many different backgrounds models to a measurement of the low energy ^{8}B solar neutrino flux. As a reminder, almost all the complexity of this analysis comes from the fact that we have some inherent uncertainty in our backgrounds model, which must be factored into any physics results extracted from this dataset. This can be accomplished by computing the evidence for multiple models and marginalizing over the posterior distributions of all the models under consideration.

Of course, the space of models is infinite since we can always add more and more empirical backgrounds into the model. Even worse, as we consider models with more and more free parameters (through the addition of empirical backgrounds), the agreement between a given model and the data will either improve or stay the same purely based on the fact that the models have gotten more complicated. The solution to this problem is of course the Bayesian model comparison approach discussed in Sec. 6.3.1. Specifically, since the prior space becomes less dense as we add parameters into the model, more complex models are "de-weighted" relative to simpler models, which can be offset by increase in the value of likelihood, or if the additional parameter provides better agreement between the data and the model.

In some sense, we can think of this prior density as a regularizer as we marginalize over all the different possible backgrounds models. Ideally at some point additional models will start to decrease in evidence such that they contribute negligible amounts to the posterior distributions of the shared parameters, and we can conclude that we can ignore backgrounds models with any additional complexity. This also provides some support for our approach of iterating over the different models in order of increasing numbers of free parameters — if at some point we start to see the evidence for more complex backgrounds models decreasing we know that this means we have reach a maximum

in the evidence. This is guaranteed since more complex backgrounds models are just powersets of the simpler models.

Now that we are equipped with a strategy to deal with the uncertainty in the backgrounds model for the basic scintillator phase, we can extract physics results from the dataset. This chapter will focus on extracting the ^{8}B solar neutrino survival probability using the PDF interpolation and likelihood fitting techniques discussed previously.

7.1. Defining the region of interest

First we must decide upon the appropriate region of interest (ROI) in which we will fit for the low energy ^{8}B neutrino rate. We want to push the region we are fitting over as low in average neutrino energy as possible, while still maintaining a reasonably accurate result. The approach employed in this work is to construct an Asimov dataset based on the background rates fit in the previous section, and then optimize the significance of the ^{8}B signal in the energy region between the low energy excess and the ^{208}Tl peak, with the significance defined as

$$P = \frac{S}{\sqrt{S+B}}. \tag{7.1}$$

S is the counts from the ^{8}B signal and B is the combined counts from all other backgrounds in the ROI.

The energy marginals of the Asimov dataset and corresponding fit can be seen in Fig. 7.1. This Asimov fit and resulting ROI optimization was done over energy, ITR, and rcubed, but the only dimension of the optimization that we need at this moment is energy since we must manually split the ^{8}B signal into the different energy regions when we do the fits to the real data. All we need to make sure of is that we span the region that maximizes the sensitivity in both ITR and radius, as the likelihood fit will then fiducialize for us. The resulting optimized energy region for the low energy ^{8}B solar neutrino rate measurement is 2.4-3.25 MeV, which we will call the *low energy region* (LER). Since the energy range of the fits shown in this work is 1.5 to 7.0 MeV, this also defines two other ^{8}B regions, one which we will call the *very-low energy region* (VLER) (1.5 to 2.4 MeV)

and the other we will call the *high energy region* (HER) (3.25 to 7.0 MeV).

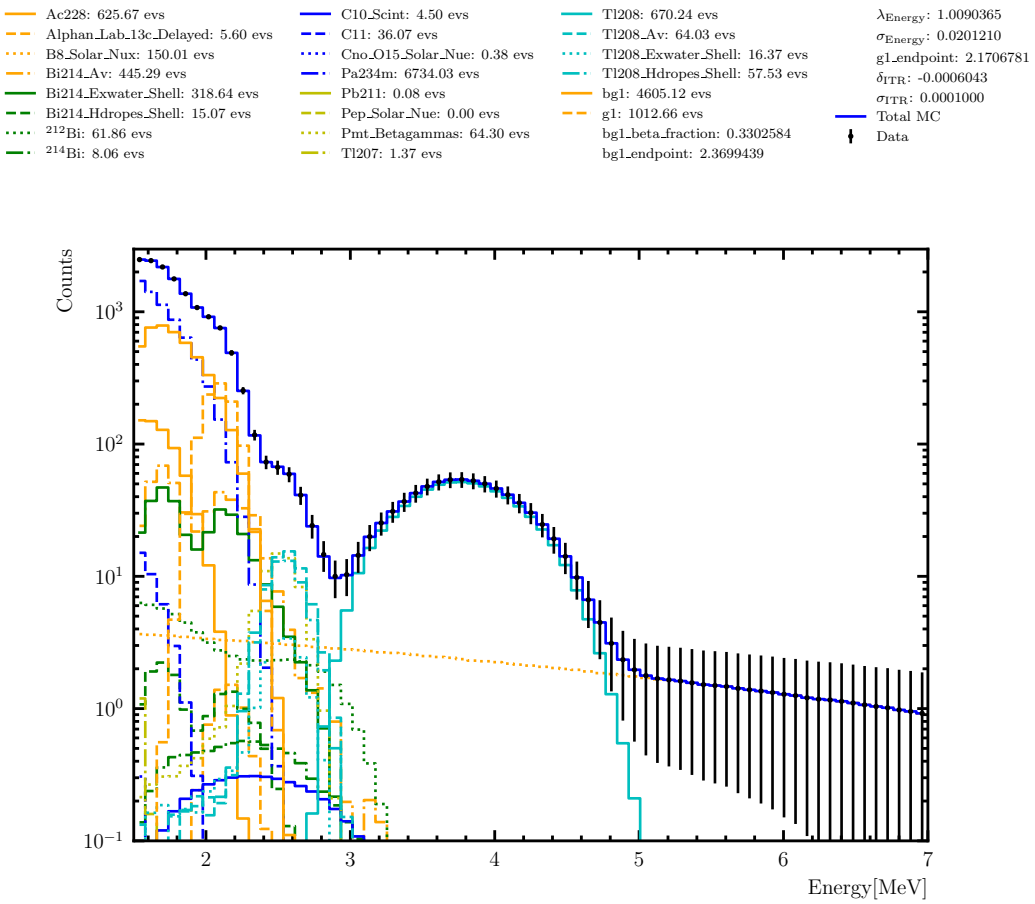


Figure 7.1: The energy marginal of the Asimov dataset and the resulting fit used to maximize the ^{8}B ROI.

The next step is relating the low energy region to the true neutrino energies it probes. This can be seen in Fig. 7.2. It can be seen in the plot that the low energy region corresponds to an average neutrino energy of 7.13 MeV. This factors in the effects of oscillations according to the MSW-LMA solution. The solar neutrino PDFs used in the fits also have neutrino oscillations applied, and combine the elastic scattering distributions of ν_e, ν_μ, and ν_τ interactions, scaled by the appropriate fractions.

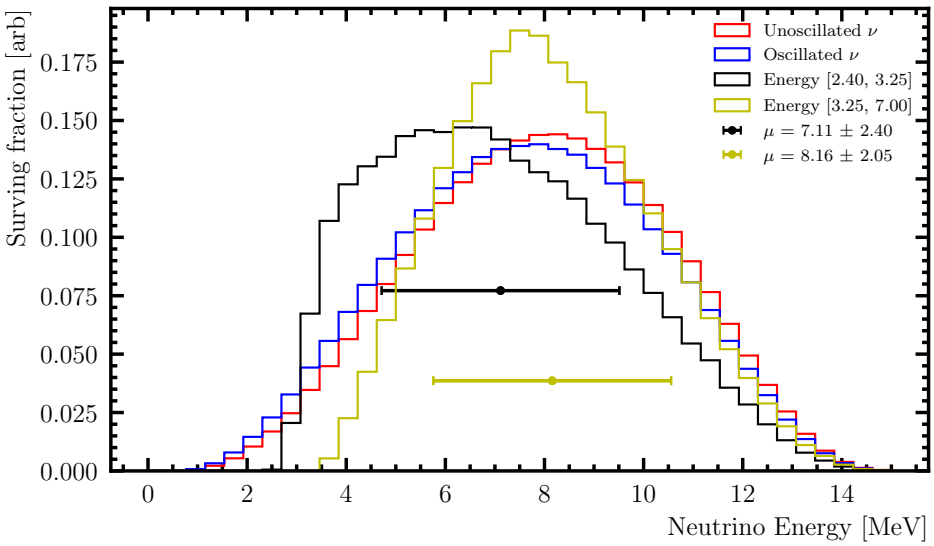


Figure 7.2: The true neutrino energies corresponding to the low energy region and high energy region in this work. The black point is the mean $\pm$ the numerical standard deviation of all the energies in the black histogram, and the same is true for the yellow point and the yellow histogram.

7.2. Model comparisons and marginalized results

For similar reasons that likelihood fits almost always work in the log-likelihood space, the nested sampling algorithm also estimates the log of the evidence, $\log Z$. The evidences of the various models considered in this work can be seen in Tbl. 7.1.

As is mentioned in the caption of the table, the nested sampling runs of the models which did not include a β/γ signal required the standard deviations on the priors to be increased in order to prevent numerical precision issues from popping up during the run. Specifically, the energy smear prior standard deviation increased by a factor of 5 and the ITR shift prior standard deviation increased by a factor of 100. In the limit that the maximum likelihood value of a paramter is far from the maximum prior value of a parameter, this *always* provide an upper bound since we are increasing the overlap between the prior and the likelihood. Therefore the log$Z$ values calculated for those models is overestimated by some unknown amount, but this is fine for reasons we will discuss in the next paragraph.

As can be seen in the table, the evidences clearly peak for a model that includes a γ and a β/γ,

Model content	log$Z$	Added free parameters
0β, 0γ, 0β/γ	183025.28	0
1β, 0γ, 0β/γ	183104.68	2
0β, 1γ, 0β/γ	183152.28	2
0β, 0γ, 1β/γ	183148.38	3
2β, 0γ, 0β/γ	183103.96	4
0β, 2γ, 0β/γ	183153.16	4
1β, 1γ, 0β/γ	183152.63	4
1β, 0γ, 1β/γ	183148.92	5
0β, 1γ, 1β/γ	183176.25	5
3β, 0γ, 0β/γ	183103.27	6
0β, 3γ, 0β/γ	183150.51	6
2β, 1γ, 0β/γ	183152.28	6
1β, 2γ, 0β/γ	183154.33	6
0β, 0γ, 2β/γ	183153.61	6

Table 7.1: The log of the evidence calculated for each of the models considered in this work. Note that the models lacking a β/γ background had to be ran with the systematic constraint on the energy smear increased by a factor of 5 and the constraint on the ITR shift increased by a factor of 100 for the slice sampling algorithm in those dimensions to not break down. This breaking down of the algorithm was due to the fitted value for those systematics being so far from MLE, leading to numerical instability since the probability of being that far away from the prior mean becomes vanishingly small.

and all other models have evidences at least $\sim$30 lower in log$Z$. Since the posteriors are weighted by the evidences, not the log of the evidence, the next best model relative to the 1 γ 1 β/γ model is the 2γ model, which has a log$Z$ of 183155.01. The contribution to the posterior of γ β/γ model is then $e^{183153.61-183176.25} = 1.5 \times 10^{-10}$. This means the posteriors of all the other models essentially provide no change to the posterior of the 1 γ 1 β/γ model, so the following results are all only shown for the 1 γ 1 β/γ model. This is why we don't care that we calculated upper bounds on the evidences for some of the models, even these upper bounds do not contribute to the marginalized posterior in any substantial way.

The marginals of the fit with the fit parameters at their MLE values can be seen in Fig. 7.3.

We can also use the nested sampling weights as an estimate of the posterior density (see Appendix B). The posterior density of the VLER, LER, and HER can be seen in 7.4.

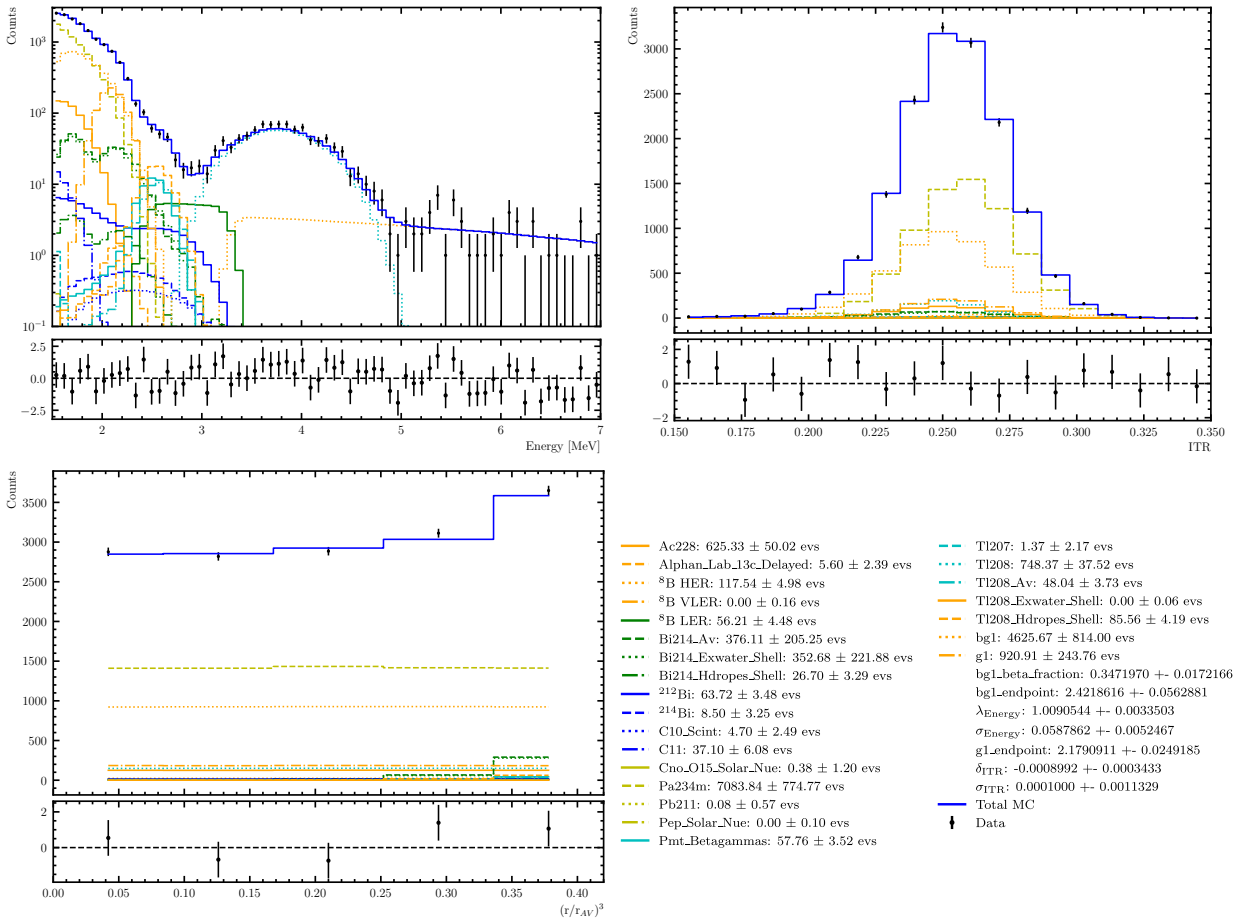


Figure 7.3: Marginals in each dimension of the 3D fit with an addition γ and β/γ empirical background. As can be seen in the legend, the resulting β/γ parameters are an endpoint of 2.42 MeV and a beta fraction of 0.35. The energy of the γ is fit to 2.18 MeV.

The posterior distribution for the VLER shows that we basically have no sensitivity to that part of the ^{8}B spectrum, which makes sense given that is where most of the substantial backgrounds in the analysis appear. The LER and the HER both produce well behaved posterior distributions. To extract a credible interval from the posterior distributions, we use Kernel Density Estimation (KDE) to construct a smooth distribution out of the weighted histogram, and then integrate the KDE distribution on either side of the mean until we reach 34% of the probability to find the 1-σ intervals. The resulting counts for the LER is $55.3^{+15.8}_{-13.6}$ and for the HER we have $119.9^{+16.6}_{-14.8}$.

Using an external value for the flux, i.e. that predicted by the B16-GS98 SSM, we can convert these values to a survival probability. This can be done using the fitted rate of elastic scatters in

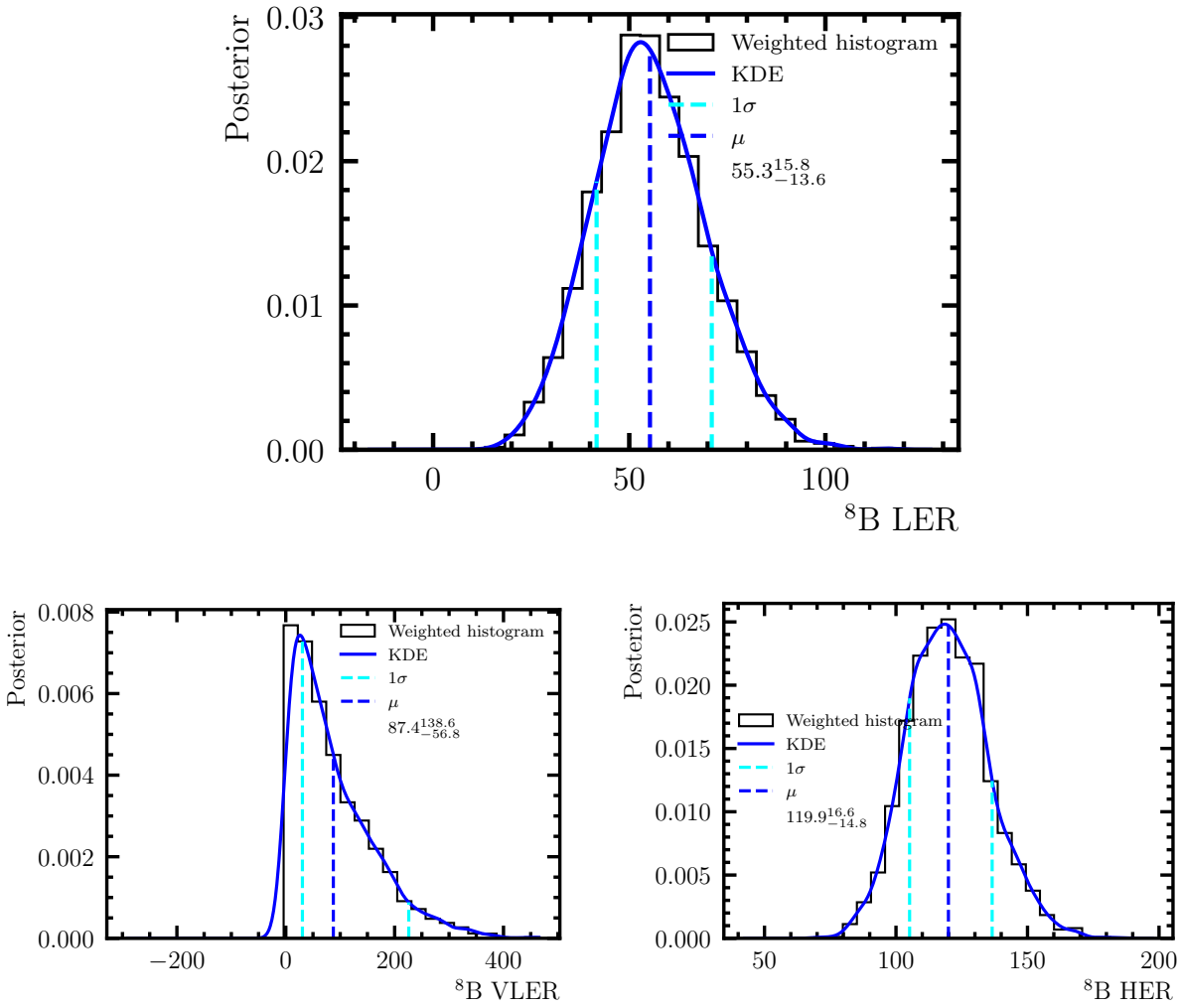


Figure 7.4: Posterior distributions of the three solar neutrino energy regions.

the LER and HER, which we will generically call R_{ES}, and comparing it to the expected number of scatters. Specifically, we want to compute

$$P_{ee} = \frac{R_{\mathrm{ES}} - R_{\mu/\tau}}{R_{\nu_e} - R_{\nu_\mu/\nu_\tau}}. \tag{7.2}$$

R_{ν_e} and $R_{\mu/\tau}$ can be undestood as the number of interactions expected given a pure beam of that

flavor of neutrino, which is given by

$$R_{\nu_e} = N_e \Phi(^8\mathrm{B}) \epsilon_e \int \sigma_{\nu_e e^-}(E_\nu) S_R(E_\nu) dE_\nu \tag{7.3}$$

and

$$R_{\nu_\mu/\nu_\tau} = N_e \Phi(^8\mathrm{B}) \epsilon_{\nu_\mu/\nu_\tau} \int \sigma_{\nu_\mu/\nu_\tau e^-}(E_\nu) S_R(E_\nu) dE_\nu \tag{7.4}$$

where ϵ_α is the signal efficiency for a given neutrino flavor, E_ν is the incident neutrino energy, and $S_R(E_\nu)$ is a unit normalized efficiency of the energy cuts applied for the energy region in which the flux is being measured. The flux $\Phi(^8\mathrm{B})$ is taken to be $5.03 \times 10^6 \mathrm{cm}^{-2}\mathrm{s}^{-1}$ [82], the number of electron targets N_e is taken to be 2.6262×10^{32} [83], and $\sigma_{\nu_\alpha e^-}$ is the neutrino electron scattering cross section for a given neutrino flavor.

Turning the crank and calculating a survival fraction for the LER we find $P_{ee}(\bar{E_\nu} = 7.11 \text{ MeV}) = 0.67^{+0.26}_{-0.22}$ and for the HER we have $P_{ee}(\bar{E_\nu} = 8.16 \text{ MeV}) = 0.30^{+0.07}_{-0.06}$ A plot showing the P_{ee} values measured here relative to the ones measured by Borexino can be seen in Fig. 7.5

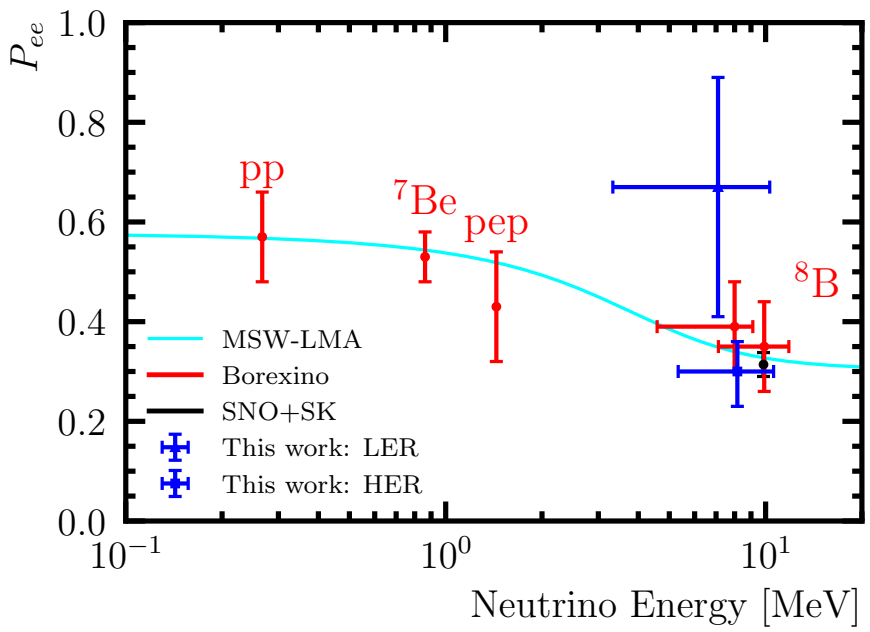


Figure 7.5: The survival probabilities measured by Borexino [32, 31], as well as the ones measured in this work.

Of course, the errors on the LER region make it hard to say anything definitive regarding the agreement between the MSW transition region and the measured value of P_{ee}. The measured survival probability of 0.67 is significantly higher than the expectation of ∼0.36, with a tension of around 1.4σ. While this may provide a hint of an unexpected high flux of low energy neutrinos in the ^{8}B spectrum, it is difficult to make any definitive statements without a more accurate measurement. Additionally, the measured and expected P_{ee} for the HER agree well.

SNO+ does have another dataset with roughly 3 times the livetime that will be added to this analysis, which will provide much more stringent constraints. It is also worth noting that this result probes a region of solar neutrino energies that has not been probed by any experiment to date. While the average neutrino energy detected in the LER is similar to that seen by Borexino and water Cherenkov detectors, as can be seen in Fig. 7.2, there is a population of neutrinos from ∼2-3 MeV measured in this fit that has not been measured before.

CHAPTER 8

Tellurium Phase Fits and Sensitivity Projections

While there is not yet Tellurium in the SNO+ detector, enough progress has been made towards the Te-phase (for descriptions of the different SNO+ phases see Sec. 3.4) that it is worthwhile to review the physics capabilities of the experiment. One notable improvement over previous sensitivity projections is that we can now fold the current optical model of the scintillator into our optical models used to generate simulations of the Te-phase datasets.

Unfortunately, the scintillator in the detector has a much lower light yield than was expected before the scintillator phase (3.4.2). This means that unless there are some unexpected effects that cause the light yield to decrease less aggressively during the Te loading, the previous simulations of the Te phase over-predict the light yield by a non-negligible fraction.

This chapter will present new results of fits to the Te-phase simulated datasets taking into account this decrease in light yield. These simulated datasets are called "productions" in SNO+ jargon. The fitting framework and approach used in this chapter are all identical to those discussed in previous chapters, except we will not need to utilize the nested sampling algorithm since we are no longer considering multiple models.

As a quick reminder, the region of interest (ROI) for the neutrinoless double beta decay analysis is very similar to the low energy ^{8}B solar neutrino analysis showed in the last chapter. The $0\nu\beta\beta$ endpoint for ^{130}Te is 2.53 MeV, which sits within the ROI for the ^{8}B analyis. This means much of the discussion from the last two chapters regarding backgrounds also applies here.

8.1. Expected background rates

The background rates used for the Te loading analyses can be seen in Tbl. 8.1. These rates are include the impacts of the measured rates currently in the detector, as well as the expected effects of the Te loading process, which is expected to increase the backgrounds by a substantial amount.

Background	Rate (per year)
B8 Solar Nue	1.15×10^{3}
Te130 2n2b	4.60×10^{6}
Bipo212	6.29×10^{4}
Bipo214	4.10×10^{5}
Tl210	8.61×10^{1}
Pmt Betagammas	6.20×10^{4}
Pa234m	6.70×10^{5}
Tl208 Hdropes Shell	3.26×10^{5}
I130	5.47×10^{0}
Co60	1.84×10^{0}
Tl208 Av	3.15×10^{5}
Ga68	2.03×10^{1}
Tl208 Exwater Shell	8.62×10^{5}
Na22	1.96×10^{0}
Alphan Telab 13c Delayed	1.00×10^{2}
Bi214 Av	2.69×10^{6}
Tl208 Hupropes Shell	3.50×10^{4}
I130m	6.50×10^{0}
Bi214 Hdropes Shell	5.71×10^{5}
Bi214 Hupropes Shell	6.10×10^{4}
Tl208	2.26×10^{4}
Bi214 Exwater Shell	2.03×10^{7}
Co56	7.50×10^{-1}
Sc46	5.30×10^{-1}
Ac228	6.29×10^{4}
Co58	3.01×10^{0}

Background	Rate (per year)
Te130 2n2b	1.37×10^{7}
B8 Solar Nue	1.15×10^{3}
Bipo212	4.80×10^{4}
Bipo214	4.47×10^{5}
Tl210	9.39×10^{1}
Pa234m	6.90×10^{5}
Pmt Betagammas	7.60×10^{4}
C10 Scint	1.77×10^{1}
I130	1.64×10^{1}
Co60	5.09×10^{0}
Ga68	4.43×10^{1}
Tl208 Hdropes Shell	3.26×10^{5}
Tl208 Av	3.15×10^{5}
Na22	5.10×10^{0}
Tl208 Exwater Shell	8.62×10^{5}
Alphan Ls 13c Delayed	6.80×10^{1}
Tl208	1.73×10^{4}
I130m	1.95×10^{1}
Y88	1.58×10^{-1}
Sb124	1.43×10^{-1}
Agm110	6.21×10^{-2}
Bi214 Exwater Shell	2.03×10^{7}
Rh106	5.32×10^{-3}
Bi214 Hdropes Shell	5.71×10^{5}
Ac228	4.80×10^{4}
Rh102m	8.81×10^{-1}
Sb126	1.49×10^{-4}
Ag110	8.40×10^{-4}
Rh102	5.53×10^{-1}
Bi214 Av	2.69×10^{6}
Sb126m	1.63×10^{-5}
K40	5.20×10^{4}
C11	8.85×10^{2}

Table 8.1: Expected backgrounds rates in the 0.5% (*left*) and 1.5% (*right*) phases.

8.2. Calibrations

Similar to what was shown in Sec. 4.3.3, the energy and ITR values are not quite independent of radius and ITR in both Te productions, requiring that we apply a correction similar to what was done with the basic scintillator phase data.

The results of the calibrations can be seen in Fig. 8.1.

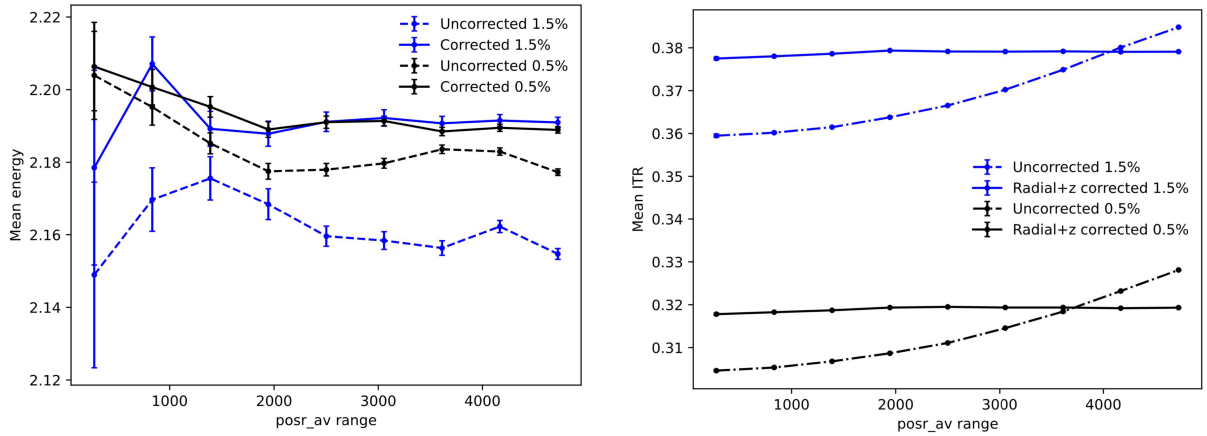


Figure 8.1: *Left:* The results of the energy correction applied to the tagged Te production ^{214}Bi events. *Right:* The results of the ITR correction applied to the same events.

Similar to what we saw in the basic scintillator phase, the correction in energy is small inside the 4.5 m fiducial volume while the correction in ITR is large. We can also see a stark difference in the ITR means for the 0.5% and 1.5% productions, which is expected since the decay time constants of the scintillator model used in the 1.5% production are shorter than they are for the 0.5% production. This is because we add the secondary fluor bisMSB as we add Te, which in turn makes the scintillator "faster"

8.3. Coincidence and classifier cuts

Since the ^{238}U and ^{232}Th rates go up in the Te-phase, much effort was spent on trying to optimize our classifier cuts to provide the highest sensitivity to the 0ν signal. ITR effectively removes one of the dominant backgrounds of in-window ^{214}Bi and ^{212}Bi events as they typically have distinct double peaked hit time residual distributions, biasing them towards lower ITR values.

The ITR cut used on each production dataset is determined by finding where the derivative of the receiver operating characteristic (ROC) meets the following condition:

$$\frac{d\epsilon_s}{d\epsilon_b} = \frac{\epsilon_s}{\epsilon_s + 2\epsilon_b} \tag{8.1}$$

where $\epsilon_s(c)$ and $\epsilon_b(c)$ are the signal and background efficiency as a function of the ITR cut c. The signal is only the 0ν signal and the background is assumed to just be the ^{212}Bi^{212}Po background since that is the dominant in window background. The ROC curves can be seen in Fig. 8.2. The optimal ITR cut removes ∼98% of the in-window ^{212}Bi^{212}Po background.

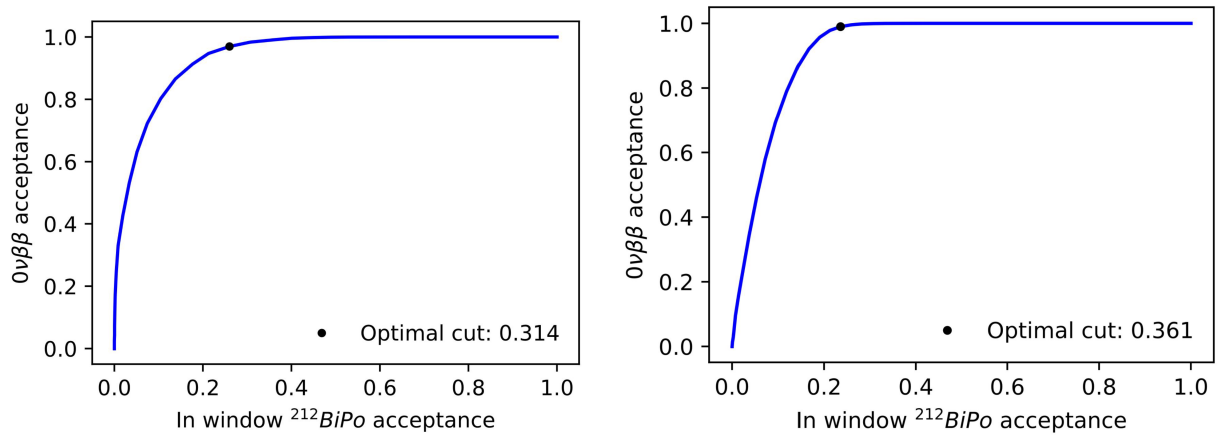


Figure 8.2: *Left:* The ROC curve of the ITR cut for the 0.5% Te production where the signal is just the 0ν events, and the background is only the in-window ^{212}Bi^{212}Po events. Also shown is the optimal ITR cut based on the derivative of the ROC curve. *Right:* The same but for the 1.5% production.

Coincidence cuts similar to the general coincidence cut discussed in Sec. 5.2 are also applied to remove the out of window coincidence backgrounds like ^{214}Bi^{214}Po .

8.4. Backgrounds in the region of interest

It is important to precisely understand the background contributions within the $0\nu\beta\beta$ ROI. We represent this visually in two distinct ways, one is a spectral plot of the backgrounds within a small window around the 0ν signal, and the other is a pie chart where the sizes of slices are defined by the backgrounds contributions in the 0ν ROI. Note that the cuts used to create the plots shown in this section are supposed to give some intuition on which backgrounds will be important in the Te-phase rather than optimize the significance of the result.

The spectral plots for the 0.5% and 1.5% productions can be seen in Fig. 8.3.

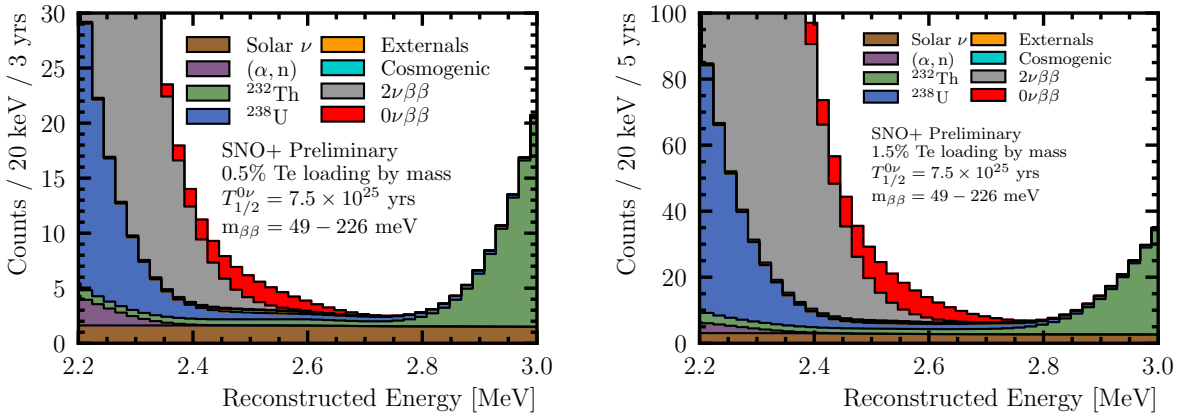


Figure 8.3: Plots showing the expected background contributions in a region near the 0ν ROI for 0.5% (*left*) and 1.5% (*right*) productions. Note the difference in y-axis scales.

Clearly the backgrounds are much higher in the Te-phase than we saw in similar plots from the basic scintillator phase. Furthermore, we can see the impact of the dominant Te-phase background, the $2\nu\beta\beta$ decay of ^{130}Te. This is especially clear when comparing the 1.5% and 0.5% plots directly. Not only does the $2\nu\beta\beta$ increase by a factor of 3, but the worsening of the resolution, which is directly due to the decrease in light yield caused by the addition of more Te, can clearly be seen in the much larger tail of the $2\nu\beta\beta$ distribution. It is also clear that this tail substantially overlaps with the $0\nu\beta\beta$ distribution. These conclusions are also made clear in the pie charts, shown Fig. 8.4.

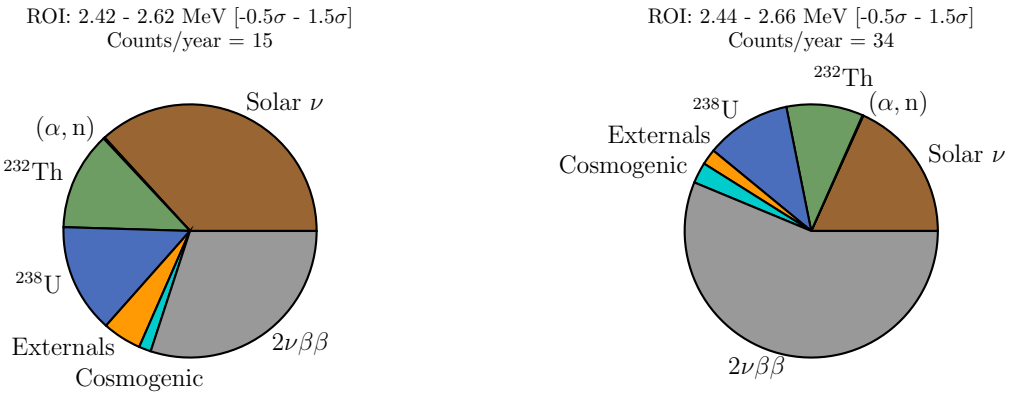


Figure 8.4: Pie charts showing the expected background contributions in the 0ν ROI for 0.5% (*left*) and 1.5% (*right*) productions. The ROI is somwhat arbitrarily defined as -0.5 σ to 1.5 σ around the $0\nu\beta\beta$ peak where σ is the energy resolution at the 0ν endpoint.

As is mentioned in the caption, the ROI used in the pie charts does not maximize the sensitivity to the $0\nu\beta\beta$ signal. In reality, as the loading increases the ROI that does maximize the sensitivity moves towards higher energies to avoid large contributions from the $2\nu\beta\beta$ signal. This is effectively the balancing act one must play when deciding the amount of isotope to load into an experiment like SNO+. The $2\nu\beta\beta$ rate increases linearly with the loading, but the energy resolution will eventually become so poor that it becomes impossible to differentiate the $0\nu\beta\beta$ events from the 2ν background. Thus there is an optimum somewhere as one loads more isotope; for SNO+ this sits somewhere roughly around 1.5% - 2.5% loading.

8.5. Half life and $m_{\beta\beta}$ limits

Of course, the interesting parts of this chapter are the expected sensitivities of the SNO+ detector. The fits in this chapter are 2D fits in energy and $(r/r_{AV})^3$, where we cut on ITR at the optimal value shown in Sec. 8.3 rather than included as a dimension in the fit. The fits go out to radius of 5 meters, whereas fits in the previous chapters only went out to 4.5 meters. The fits are all performed to Asimov datasets. Once the MLE values have been found, we walk the parameter space of the $0\nu\beta\beta$ signal and re-minimize the -ln $\mathcal{L}$ at each point to compute the log-likelihood profile ((Eq. (40.41) [14])).

The log-likleihood profile is then fit with a parabola, and the parameters of the parabolic fit are then converted into the standard deviation of a Gaussian. This standard deviation is then used to calculate the corresponding 90% C.L. Even though this is not completely rigorous, as we will see the profile curves are approximated very well as parabolas (and therefore Gaussians), meaning any systematic errors introduced in this statistical treatment will be small. The profile log-likelihood can be seen in Fig. 8.5.

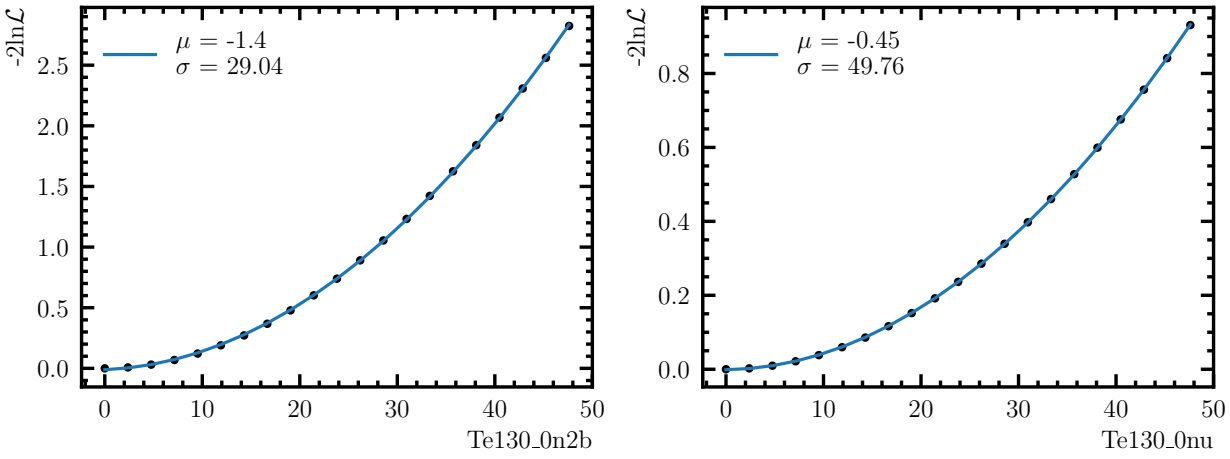


Figure 8.5: $-2\ln\mathcal{L}$ constructed by profiling the log-likelihood with the profile points shown in black. The blue curve is a parabolic fit, which is then converted to a normal distribution, the parameters of which are shown in the legend. These parameters are then used to analytically calculate the 90% C.L.

The resulting half life limits associated with the limits computed from the profiles shown in the figure are 1.7×10^{26} yrs for 3 years of livetime at 0.5%, and 4.3×10^{26} yrs for 5 years of livetime at 1.5% loading. Using a phase space factor of $G = 3.70 \times 10^{-14}\text{yrs}^{-1}$ and matrix elements ranging from 1.37 to 6.366 [84], the resulting $m_{\beta\beta}$ limits are 35 - 164 meV for the 0.5% dataset, and 20 - 95 meV for the 1.5% dataset. Therefore we can see it is unlikely for SNO+ to set a world leading limit at 0.5% since Kamland-Zen's limit is 3.8×10^{26} yrs, but at 1.5% loading SNO+ can set a world leading result in under 5 years of livetime.

It is worth recalling Fig. 2.12. As can be seen in the plot, SNO+ does gain a little bit on Kamland-Zen because the nuclear matrix elements are in our favor, but we will still need to load to 1.5% to

definitely pass Kamland-Zen's limit. This is reflected in SNO+'s current schedule, as we intend to load to 1.5% as soon as the Te to do so is acquired.

We should also mention that the limits shown in Fig. 2.12 are slightly more agressive than the ones found in this dissertation, where the disagreement is by about 20%. The reasons for this are known, the ROI external rates and $2\nu\beta\beta$ rates used in this analysis and the analysis used to calculate the limits shown in Fig. 2.12 disagree, and when this disagreement is manually removed then the half life limits calculated by the two approaches are identical. Work is ongoing to try and find the source of the disagreement.

CHAPTER 9

Conclusions

The main physics result of this dissertation is the measurement of the electron neutrino survival probability using the lowest energy population of ^{8}B solar neutrinos to date. This measurement required a large amount of work developing the backgrounds model at around 2 MeV, since a naive backgrounds model based on what is typically present in scintillator detectors did not adequately explain the backgrounds seen in the SNO+ detector.

Instead, an "empirical" backgrounds model was constructed based on the basic scintillator dataset, which consists of a generic β/γ emitter, with an endpoint of 2.42 MeV and the fraction of energy given to the β relative to the γ is 0.35, and another pure γ emitter with an energy of 2.18 MeV. The ability to float these empirical backgrounds during fits to the data required the creation of a novel new optimal transport PDF interpolation strategy, allowing for the smooth interpolation of complex N-dimensional PDFs.

Furthermore, a Bayesian approach was used to compare other empirical models and quantitatively show that this model is indeed the most well motivated based on the data and the relative complexities of the models considered. This required the use of the nested sampling algorithm, which provides estimates of the log of the evidence for each competing model, as well as estimates of the posteriors of all the parameters in the fits.

This empirical model was then used in likelihood fits to the 237.4 day basic scintillator phase dataset to measure the ^{8}B solar neutrino flux down to energies of 2.4 MeV, providing survival probabilities sensitive to never before probed ^{8}B solar neutrino populations. The resulting survival probabilities are $P_{ee}(\bar{E_\nu} = 7.11 \text{ MeV}) = 0.67^{+0.26}_{-0.22}$ for the low energy region (LER) and $P_{ee}(\bar{E_\nu} = 8.16 \text{ MeV}) = 0.30^{+0.07}_{-0.06}$ for the high energy region. $\bar{E_\nu}$ is the average solar neutrino energy in the each of the respective regions. The LER measurement is in tension with the MSW-LMA solution at 1.4σ. While the large errors on the survival probability in the LER make it difficult to make

decisive statement regarding the validity of the MSW solution in the transition region, including all additional data taken by the SNO+ detector will increase the livetime by a factor of 3 which should allow for a much more accurate measurement. The HER measurement does agree well with the MSW-LMA solution in the well established matter effects dominated part of the solution.

Following the solar neutrino fits, we also discussed the expected $0\nu\beta\beta$ sensitivities of the SNO+ detector during the Te phase. Shown in this dissertation were updates to the expected sensitivities using optical models and backgrounds models based on the current detector state. The predicted half life limits are 1.7×10^{26} yrs for 3 years of livetime at 0.5%, and 4.3×10^{26} yrs for 5 years of livetime at 1.5% loading. The resulting $m_{\beta\beta}$ limits are 35 - 164 meV for the 0.5% dataset, and 20 - 95 meV for the 1.5% dataset. At 1.5% loading SNO+ can set a world leading result in under 5 years of livetime.

In the near future, the inclusion of more data to the low energy ^{8}B rate fits will provide a much more stringent test on the MSW-LMA solution in the transition region. Novel techniques involving improved multisite classifiers and the utilization of directional information from the solar neutrino events to discriminate between backgrounds and the ^{8}B events are also being investigated to improve the results here. No comment is made on the sensitivity of these potential future analyses, but they can only improve our current understanding of the MSW transition region.

APPENDIX A

The Likelihood Function and Bayesian Inference

A.1. The likelihood function and maximal likelihood estimation

In principle, the idea behind the likelihood function is simple — first consider a random variable x as a collection of measurements $(x_1, x_2, ..., x_N)$ which are each distributed according to some probability distribution function (PDF) $f(x_j)$ for $x_j \in x$. Note that each x_j can in principle be some vector itself if each measurement consists of multiple variables, like a Cartesian x and y position. We can then calculate the probability $p(x)$ of observing some sequence of independent measurements of x as

$$p(x) = f(x_1)f(x_2)...f(x_N) = \prod_{j=1}^{N} f(x_j).$$

If $f(x)$ can be parameterized in terms of some collection of parameters $\theta = (\theta_1, \theta_2, ..., \theta_M)$, then we can also write

$$\mathcal{L}(x|\theta) = \prod_{j=1}^{N} f(x_j|\theta) \tag{A.1}$$

where we have replaced p with $\mathcal{L}$, as is typically used to represent the likelihood. We have also introduced the notation $f(x_j|\theta)$ which just means the function f evaluated at x_j given some value θ. We will often drop the (...|) or (|...) if it is not contextually important, especially when discussing $\mathcal{L}$ since we are usually interpreting $\mathcal{L}$ as a function of x, i.e. $\mathcal{L}(x|\theta) \rightarrow \mathcal{L}(\theta)$.

The idea behind MLE is to try and find the θ_i that maximizes the probability $p(\theta|x)$ for a given set of observations x. Of course, we know that a maximum in the likelihood indicates a zero-crossing in the derivative of the likelihood, so we can express this in terms of partial derivatives as

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = 0. \tag{A.2}$$

We will refer to the point that fulfills this conditions a $\hat{\theta}$. Finding $\hat{\theta}$ is the game of MLE.

In the complicated situations physicists typically find themselves in, $\mathcal{L}(\theta)$ will not be described by some analytic function. Often the PDF's used to calculate $\mathcal{L}$ will be generated through simulation, therefore there is no way to analytically calculate at what collection of the θ_i will satisfy the condition in A.2. Instead, we rely on numerical techniques to calculate the value of $\mathcal{L}(\theta)$ for different values of the θ_i and to find the points that fulfill A.2.

Rather than using A.1 directly, we often instead work in terms of $-\ln\mathcal{L}(\theta)$, which is called the *log-likelihood function*. This provides a few benefits, the major one being that if we recall that $\ln(xy) = \ln(x) + \ln(y)$ then we can see that A.1 becomes

$$-\ln\mathcal{L}(\theta) = \sum_{j=1}^{N} \ln f(x_j)$$

which is much easier to compute since addition is a cheaper operation than multiplication, and since N might be very large, this helps keep the numbers involved in the computation within double precision. The minus sign is a convention to ensure the maximal likelihood occurs at a minimum rather than a maximum. We also don't need to worry about the point of maximal likelihood being different than the point of minimal log-likelihood since the log function is monotonically increasing.

While $\mathcal{L}(x|\theta)$ is a useful quantity, another useful result we want to consider is $p(\theta|x)$, called the *posterior density*, or more plainly the probability of a given value of θ given the data x. If we recall Bayes' theorem:

$$p(A|B) = \frac{p(B|A)P(A)}{p(B)}$$

we can write $p(\theta|x)$ as

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}. \tag{A.3}$$

Immediately, we can recognize $p(x|\theta)$ as the likelihood function $\mathcal{L}$ but we have not yet encountered the other two quantities $p(\theta)$ and $p(x)$. Perhaps the easier quantity to interpret is $p(\theta)$, which is a measure of the belief we have in a particular value of the θ_i's before considering the measurements x. This term is what we typically refer to as a *prior*, often written as $\pi(\theta)$.

$p(x)$ on the other hand becomes more clear when we recall the fact that $p(\theta|x)$ must be normalized over the space θ, implying that $p(x) = \int \mathcal{L}(\theta|x)\pi(\theta)d\theta$. This term is the probability that we observed the set of measurements x given our prediction of the distribution of x for any given θ (the likelihood), while also factoring in our prior beliefs in the values of the θ. $p(x)$ is quite literally the probability of measuring x, given any possible form of our model and our prior beliefs regarding the model. People will often refer to $p(x)$ as the *evidence* or Z.[4]

We are finally at the point where we can write the posterior density in its most recognizable form:

$$p(\theta|x) = \frac{\mathcal{L}(x|\theta)\pi(\theta)}{Z} \tag{A.4}$$

A.2. Bayesian model comparison

The previous section considered a single model M which is described by some number of parameters $(\theta_1, ..., \theta_j)$ Now if we consider a situation where we have multiple competing models and we want to compare or combine predictions from the different models, we must try to compute $p(M|x)$, or the probability of a model M given the data x.

Applying Bayes' theorem again, we find

$$p(M|x) = \frac{p(x|M)\pi(M)}{p(x)} \tag{A.5}$$

where now $p(x)$ is a sum over the different models we are considering, $p(x) = \sum_i p(x|M_i)p(M_i)$, and $\pi(M)$ is the prior we assign to a given model. Although in the previous sections we stated $p(x) = Z$, there was an implied conditional term since this was the $p(x)$ for a given model, M. Being pedantic, we should write $Z_M = \int \mathcal{L}(x|\theta, M)\pi(\theta)d\theta = p(x|M)$, therefore the posterior probability for a given model is directly proportional to the evidence. Rewriting $p(M|x)$ we now have

$$p(M|x) = \frac{Z_M\pi(M)}{p(x)}. \tag{A.6}$$

[4]While this normalization is often treated as an afterthought, it will become important shortly.

If we want to compare the relative probabilities between two models, M_1 and M_2, we can use the Bayes factor B_{12}

$$B_{12} = \frac{p(M_1|x)}{p(M_2|x)} = \frac{\frac{p(x|M_1)\pi(M_1)}{p(x)}}{\frac{p(x|M_2)\pi(M_2)}{p(x)}} = \frac{Z_1\pi(M_1)}{Z_2\pi(M_2)} \tag{A.7}$$

and if the priors are the same then then $B_{12} = \frac{Z_1}{Z_2}$.

We can also think of "marginalizing" over the different models given a parameter θ_j that is shared between all the models under consideration. This can be represented as the weighted sum of the individual model posteriors of the parameter of interest. where the weights are the prior weighted evidences of the different models:

$$p(\theta_j|x) = \sum_i \frac{p(\theta_j|x, M_i)Z_{Mi}\pi(M_i)}{p(x)} \tag{A.8}$$

In the analysis presented in this thesis, we will consider many different background models, and as such we will make use of the results in this section to compare and combine their results. This means accurately calculating the evidence for the different models will play an integral role in this analysis.

APPENDIX B

Nested Sampling

B.1. Nested sampling

This discussion will closely follow Skillings' original paper introducing nested sampling [79], while also taking some inspiration from [85].

Recall the evidence Z from Eq. (A.4) which is the normalization constant of the posterior $p(\theta|x)$ and equal to $\int \mathcal{L}(x|\theta)\pi(\theta)d\theta$. For the measurement presented in this thesis, the number of parameters in a given model (which is the dimensionality of θ) is typically between 20-30. Therefore calculating the evidence requires one to evaluate an N-dimensional integral (where N is the number of parameters) over the entire parameter space allowed by the prior, $\pi(\theta)$. [5]

Obviously, calculating Z becomes prohibitively expensive for complicated models if one is not careful. Nested sampling sidesteps the dimensionality issue by converting the N-dimensional integral over θ into a 1-dimensional integral over some new measure. Of course, we must take care when selecting this new measure to keep the result of the 1D integration consistent with the N-D integration. Rather than trying to explicitly calculate a transformation to this new measure, the nested sampling approach instead builds up an approximation to this 1D integral through strategic sampling of the likelihood space.

First we define the cumulative prior mass enclosed within some boundary in the likelihood space

$$\int_{\mathcal{L}(\theta)>\lambda} \pi(\theta)d\theta = X(\lambda) \tag{B.1}$$

which is a function of the value of the likelihood, λ, at which we impose the boundary. λ will vary from 1 to 0, starting out covering the entire prior space when $\lambda = 0$ and covering none of the

[5] Typically, the direct calculation of the evidence is avoided, essentially by considering ratios of the likelihood between different values of θ to approximate the posterior distribution. Unfortunately, this analysis will require the consideration of many different models, and in order to compare between the different models we must explicitly calculate Z.

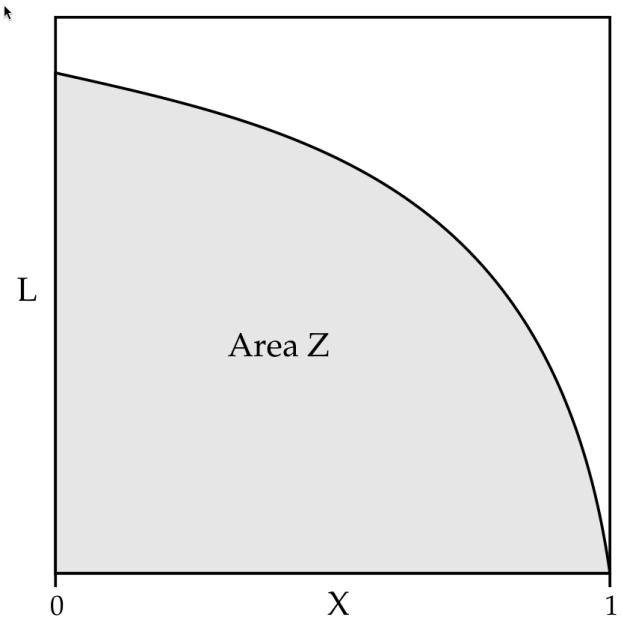


Figure B.1: Plot of the 1-dimensional function $f(X)$ and the integral Z. Taken from [79].

space when $\lambda = \max \mathcal{L}$. We can also define the inverse function that maps the cumulative prior mass X back to λ, $f(X)$ where $f(X(\lambda)) = \lambda$. Fig. B.1 schematically represents what this function f might look like. This allows us to write the volume contained in one small region of $f(X)$ as $f(X)\Delta X = \lambda\pi\Delta\theta = \Delta Z$. Going from discrete to differential quantities we find

$$Z = \int_0^1 f(X)dX. \tag{B.2}$$

While we have made a big step forward by converting the N-dimensional integral in θ into a 1-dimensional integral in X, evaluating Eq. (B.2) does require us to order X in terms of decreasing likelihood. In other words, we have essentially unraveled each of the dimensions of θ, using the likelihood to decide how to order the points, and any $\theta = (\theta_1, ..., \theta_j)$ will get mapped to some point X_k. If we assume that we have carried out this sorting operation, then approximating the integral in Eq. (B.2) becomes easy — we can compute a 1D Riemann sum using $Z \approx \sum_i \Delta X_i L_i = \sum_i (X_i - X_{i-1})L_i$.

Now we must try and come up with a way to decide on the X_i. First note that typically the bulk

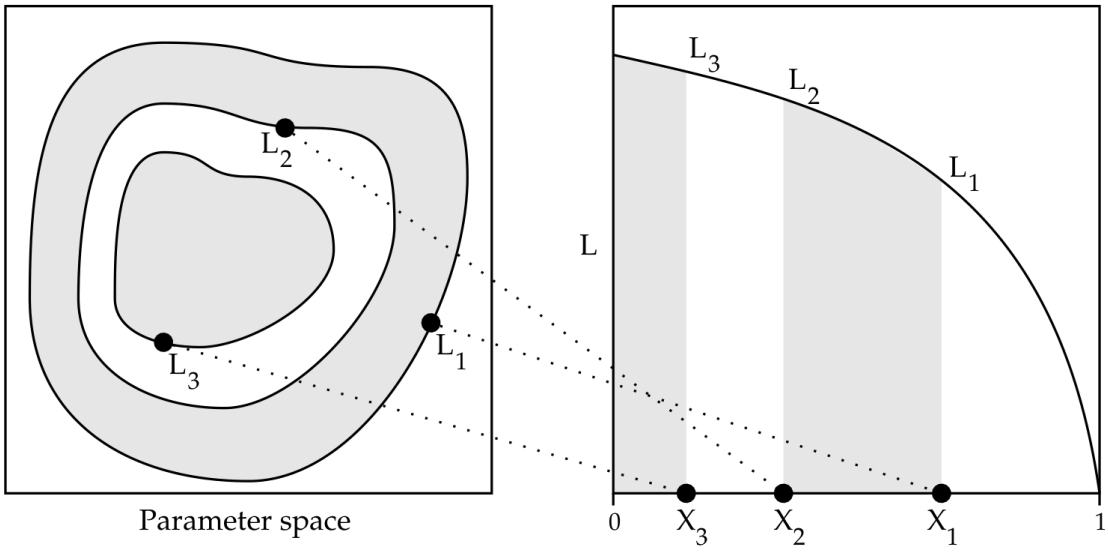


Figure B.2: The mapping of θ to X. From [79]

of the posterior mass lies in some small region of the prior space, so we would benefit by spacing the steps linearly in log$X$ rather than X, i.e.

$$X_1 = t_1,\ X_2 = t_1 t_2,\ ...,\ X_i = t_1 ... t_i,\ ...,\ X_k = t_1 ... t_k$$

where each t_i is between 0 and 1. One process for choosing a new point could be to randomly select a point X_i from the prior space such that $X_i < X_{i-1}$, and the relationship between subsequent points would be $X_i = X_{i-1}t$, with ts being uniformly distributed between 0 and 1, just as we wanted. But this still leaves the difficulty of going from values for the Xs to L_i. But here is the trick of nested sampling, if we instead sample in θ according to the prior $\pi(\theta)$ for points that satisfy the likelihood constraint $\mathcal{L}(\theta_i) > \mathcal{L}(\theta_{i-1})$, then we end up with random points distributed in the same way as if we did the sampling directly in X. By sampling in θ we avoid explicitly working in terms of the X coordinates, saving us from having to actually carry out the sorting of the likelihood points.

We still need to know the spacing ΔX_i at each step in order to calculate the Riemann sum. While we can't find the spacing for a specific choice of the Xs (since that would require us to carry out the sorting procedure), we do know statistically how each X is related to previous Xs, i.e. by some

factor $t_i \in \text{Uniform}(0, 1)$. If we consider a sampling technique where we sample N points distributed within the restricted prior volume and take the point with lowest likelihood as our sample, we now have the following relationship between X_i, t_i, and N

$$X_0 = 1, \qquad X_i = t_i X_{i-1}, \qquad p(t_i) = N t_i^{N-1} \tag{B.3}$$

where $p(t_i)$ is the probability distribution of the t_is, which comes from the fact that t_i is the maximum of N samples from $\text{Uniform}(0, 1)$.

Using $p(t_i)$, we can show that the log of the ratios t between samples when using this technique has mean and standard deviation

$$E[\log t] = -\frac{1}{N} \text{ and } \text{Var}[\log t] = \frac{1}{N}. \tag{B.4}$$

The intuition here is that as we have more samples to consider, the rate at which the prior volume shrinks between each sample decreases exponentially. Specifically, after i steps we can expect the prior mass to shrink by $\log X_i \approx -(i \pm \sqrt{i}/N)$. This statistical description of the prior mass shrinkage every time a new sample is drawn according to the technique outlined above allows us to approximate the weights of the Riemann sum. Thinking in terms of the θ's rather than the X's, each sample in X represents a contour of constant L_i with prior mass X_i, as is shown in Fig. B.2. One can try to sample the likelihood space directly to measure this shrinkage factor, but the approach taken in this work is always to just use the statistical relationship defined above. An important note is that when this sampling process is done iteratively, only the lowest likelihood sample in the N samples needs to be replaced, the other N-1 samples are already a uniform sample of the prior space that fulfill the likelihood constraint.

Now we can full define the nested sampling algorithm:

Algorithm 1 Top level nested sampling algorithm

N $\in \mathbb{N}$
samples $\leftarrow$ [N]
for j in N **do** ▷ Fill array with initial samples from prior
    Find point θ by sampling from $\pi(\theta)$
    samples[j] $\leftarrow \theta$
end for
$Z \leftarrow 0$
$X_0 \leftarrow 0$
terminate $\leftarrow false$
$i \leftarrow 1$
while not terminate **do**
    Find sample $\in$ samples with lowest likelihood and index l
    $\theta_L \leftarrow$ samples[l]
    $\mathcal{L}_L \leftarrow \mathcal{L}(\theta_L)$
    $X_i \leftarrow e^{-i/N}$
    $w_i \leftarrow X_i - X_{i-1}$
    $Z \leftarrow Z + \mathcal{L}_L w_i$
    $\theta \leftarrow$ getNewSample($\mathcal{L}_\mathrm{L}$) ▷ getNewSample must sample from $\pi(\theta)$, but constrained such that $\mathcal{L}(\theta) > \mathcal{L}_L$
    samples[l] $\leftarrow \theta$
    i $\leftarrow$i + 1
    terminate $\leftarrow$ shouldTerminate() ▷ Decide on whether or not algorithm should terminate
end while
for θ in samples **do**
    $Z \leftarrow Z + \frac{\mathcal{L}(\theta)}{N} X_i$ ▷ Fill in remaining space $0 < X < X_i$ based on average likelihood of surviving points.
end for

As can be seen above, the only parameter required by the nesting sampling algorithm is the number of samples considered at each iteration, N. This value is typically referred to as the number of *live points*.

While this algorithm provides a top level view of how nested sampling works, there is still some complexity that we have abstracted away, contained in the functions "shouldTerminate" and "getNewSample". First we will discuss shouldTerminate, which decides on when the nested sampling algorithm should stop. In [79], Skilling provides a useful heuristic based on the *information* $H = \int \log(\frac{dP}{dX}) dX$, where $dP = p(\theta)d\theta$ and $dX = \pi(\theta)d\theta$. Colloquially, the information represents how much additional posterior mass is being gained per additional prior mass. A natural

place to end the algorithm is when additional iterations are providing only modest improvements in the information, H. The amount of information we expect to gain per iteration also depends on the number of live points, so we want to terminate when $i > \alpha \mathrm{N} H$ where α is some arbitrary factor that is $\sim$1. In the nested sampling runs used in this analysis α is set at 2.

B.1.1. Slice sampling

Arguably the most challenging part of the nested sampling algorithm is the process of selecting new samples, which is taken care of by "getNewSample" in algorithm algorithm 1. Selecting new samples is difficult because as the dimensionality of the parameter space grows, the fraction of the parameter space with $\mathcal{L}(\theta) > \mathcal{L}_L$ shrinks exponentially. This means that effectively sampling the prior space becomes one of the bottlenecks in the computation of this algorithm, but perhaps even more importantly the validity of the nested sampling algorithm hinges on the new samples being pulled uniformly over the prior space $\pi(\theta)$.

In this work we employ a technique called slice sampling, introduced in [86], to replace the lowest likelihood points at each iteration in the nested sampling loop. Slice sampling efficiently and uniformly samples from a multidimensional probability density by updating the coordinates of a starting point one by one, avoiding the difficulties of rejection sampling many dimensions. Slice sampling can also be generalized to updating multiple coordinate dimensions at once, which is useful for minimizing the time spent updating dimensions that result in expensive likelihood calculations, like systematics that are applied as shape changes to the underlying model PDFs. We will first discuss the case where we update coordinates one by one and then move on to the multi-dimensional case.

Slice sampling requires that we start with a point $\hat{\theta} = (\hat{\theta}_0, \hat{\theta}_1, ...)$ which we will update to a new point. In our nested sampling loops it is easy to find the starting point $\hat{\theta}$ since the live samples are already are a uniform sample of the coordinate space constrained such that $\mathcal{L}(\theta) > \mathcal{L}_L$, the likelihood value of the point we are replacing, so we randomly set $\hat{\theta}$ to one of the live points. The next step is to then take the first dimension we would like to update and find the minimum and maximum values such that $\mathcal{L}(\hat{\theta}) = \mathcal{L}_L$, where we only vary $\hat{\theta}$ in the dimension that we are currently

updating. The minimum and maximum values then define the lower and upper bounds (L, R) over which we can sample $\pi(\theta)$[6] to find a new value for the coordinate. Then this process is repeated for every dimension.

There are a few options for finding the bounds (L, R) over which we sample the prior. The technique used in this work is referred to as the "stepping out procedure" in [86], which ,as the name suggests, just means that we step out by some width w from the ends of the region until we have a segment where $\mathcal{L}(\theta) < \mathcal{L}_L$ at both ends. One pitfall of using this approach is that when we draw a sample from $\pi_i(\theta_i)$ and update $\hat{\theta}$, there is no guarantee that $\mathcal{L}(\hat{\theta}) > \mathcal{L}_L$ since the bounds will always extend slightly past $\mathcal{L}(\theta) = \mathcal{L}_L$ due to the nature of the procedure. The part of (L, R) that does not satisfy the likelihood constraint may in fact be a substantial fraction of the entire region since the width we extend the different ends by should be large enough to quickly find the bounds of the region surrounding the initial starting point. Therefore to ensure that the new point is still within the likelihood constraint, we must calculate $\mathcal{L}(\hat{\theta})$ every time we update $\hat{\theta}$, and try another point along the dimension θ_i if the likelihood value of the point is not above the constraint.

The stepping out procedure also requires an input length that defines the step size w that we step out by at each point in the loop. At the outset of the nested sampling run, the only information we can use to decide on the value of w is the bounds of the parameter, since that guarantees that we will sample over the entire parameter space. As the fit progresses though, the likelihood constrained region of the parameter space that is accessible is typically much smaller than the bounds of the parameter. To monitor the shrinkage of the accessible regions of the parameter space we can keep track of the sampling bounds that resulted in acceptable parameter values and use that information to guide what the initial step size should be for subsequent parameter updates. In this work, we use a per parameter exponential moving average (EMA) to keep track of these widths.

Now that we have high and low bounds, we could find a new point through brute force — just re-sampling $\pi_i(\theta_i)$ from (L, R) until we find a point that satisfies the criteria. But we can minimize

[6]Typically $\pi(\theta)$ is separable, meaning that $\pi(\theta) = \prod_{i=0}^{N} \pi_i(\theta_i)$ where N is the dimensionality of θ. This means we can trivially sample the prior along a single dimension θ by sampling from a uniform distribution and pushing that through the quantile function of π_i.

the number of times we calculate the likelihood if we update the bounds over which we sample every time we fail to find a point that meets our criteria. The updating process used in this work is the same as what was outlined in [86], if the new value of θ_i is smaller than $\hat{\theta}_i$ we replace L with θ_i, and if the new value of θ_i is larger than $\hat{\theta}_i$ we replace R with θ_i. This allows us to refine the range over which we do the sampling in a way that will always converge to a point that satisfies the criteria.

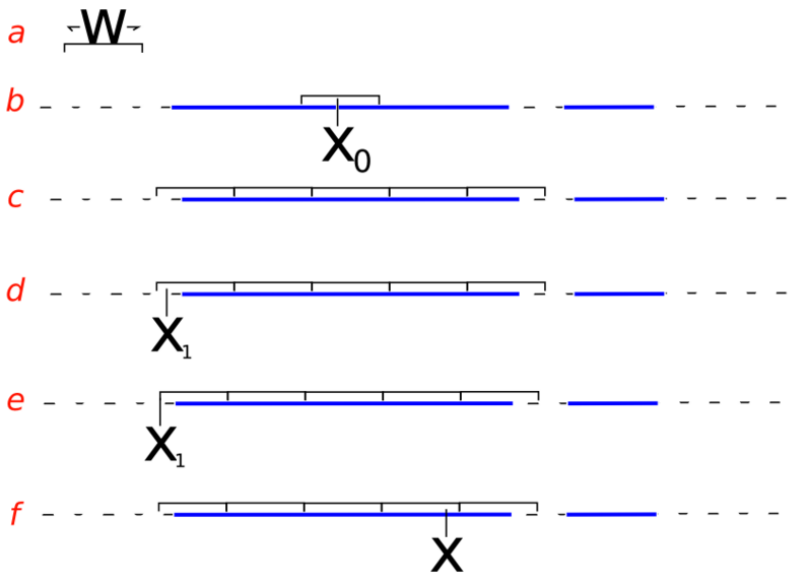


Figure B.3: A summary of slice sampling. Process begins with a step width w and a point x_0. There is some allowed region over which we would like to sample, represented in blue. c shows the result of the stepping out procedure where upper and lower bounds have been found where the allowed region ends. In d a random sample as has been taken over the range in black, resulting in a point that falls in the region between the true lower bound of the allowed region and the lower bound of the region found in the stepping out procedure. At this point we update the lower bound of the sampling region to be x_1, shown in step e. Step f shows a sample in the allowed region being selected, which results in the termination of the slice sampling loop. From [87].

A helpful image outlining this process can be seen in Fig. B.3.

B.1.2. N-dimensional slice sampling

The process is much the same when we want to update multiple dimensions at once. We will call the multiple dimensions being updated simultaneously the "expensive subspace", since typically we will group any parameter that is expensive to recalculate into this collection of dimensions.

The first step is to randomly choose a unit direction vector with the same dimensionality as the expensive subspace, and then repeat the process used in the 1D space along this random direction. To accomplish this, we take a vector where each element is an independent sample from a Gaussian distribution with mean and standard deviation 1, and then use the unit vector of this vector as the direction vector. It turns out that this process uniformly samples the surface of an (N+1) dimensional sphere [88].

Once we have the random direction, we essentially just repeat the 1 dimensional process along the line traced out by the direction vector through the expensive subspace. But the validity of the nested sampling approach requires that we correctly sample from prior distribution, which is more difficult since there is no trivial way to sample along this line while also sampling uniformly in $\pi(\theta)$. What we can do, though, is work in what is referred to as the unit-hypercube space, where we map each dimension θ_i onto a new dimension, which we will refer to as u_i, where the prior becomes uniform on the interval $u_i \in [0, 1)$. This can be accomplished by using the quantile function of the prior PDF to transform points in u-space back to θ space. Now that we are working in a transformed version of the expensive subspace where each parameter is uniformly distributed from 0 to 1, uniformly sampling along the line traced out through the space by the random vector will still provide unbiased sampling of the priors in θ once we transform back from u-space to θ space.

Another difficulty introduced by the multiple dimensions is that the initial step-size s at each iteration are not as straightforward to set. To see why this is the case, we can imagine a scenario in which the expensive subspace is 2 dimensional. Even though we are able to make the prior distributions uniform by working in u-space, this of course does not guarantee that the likelihood space is uniform. Therefore we can imagine a scenario where the $\hat{x}$ direction in the expensive subspace has a region where $(L)(\theta) > \mathcal{L}_L$ that is much smaller than the $\hat{y}$ direction. In this case, if the random direction vector is pointed completely along the $\hat{x}$ or $\hat{y}$ direction, the starting step width will be either much too large or much too small if previous direction vectors have been pointed mostly along the other dimension.

The way we can fix this is to keep track of the correlations between the direction vector and the

bounds that resulted in the acceptable sample, all of which is encoded in the covariance matrix, Σ, of the parameters that form the expensive subspace. The live points can be used to calculate Σ of the expensive parameters, which must be continuously updated as live points are replaced. In the nested sampling runs used in this work, Σ is updated every 10^{th} replacement to the live points. s is calculated each time the expensive parameters are updated using $s = d^T \Sigma d$.

While this process should provide new points that are uncorrelated from the current live points in the nested sampling run, other nested sampling libraries have found that slice sampling can select new points correlated with the seed points [85]. To try and mitigate this effect, the slice sampling algorithm can just be chained together where the output of one iteration can be used as the input to the next. If the slice sampling is indeed selecting correlated points, this process can help the points de-correlate over many iterations. In this work, somewhere around 10-20 slice sampling iterations are used per explore step in the nested sampling loop. This can be prohibitively slow if expensive parameters are updated every iteration of the slice sampling, so expensive parameters are instead updated every 5^{th} slice sampling iteration, and are always updated on the first iteration.

Even though the nested sampling algorithm was designed to calculate the evidence, the samples and weights of the removed live samples (which are referred to as dead samples) during the run also provide a great deal of information. These points are essentially prior-weighted samples of the likelihood space, which themselves can be weighted by the $Z_j = w_j \mathcal{L}_j$ of the j^{th} sample relative to the total evidence Z. Formally, we can say that the posterior of a given dimension $p(\theta_i)$ can be calculated using

$$p(\theta_i) = \sum_{j=0}^{\text{M}} \frac{Z_j \theta_{ij}}{Z} \tag{B.5}$$

where θ_{ij} just means the i^{th} component of the j^{th} sample and M runs over the total number of samples.

APPENDIX C

Measure Theory, Optimal Transport, and PDF Interpolation

C.1. Probability through measure theory

We physicists, especially myself, tend to take a laissez faire attitude towards our treatment of statistics. But when working with probabilities as pure mathematical objects, some rigor is necessary. Thus, before we dive into optimal transport, it will be helpful to provide some background information on the measure theoretic approach to probability theory. A more formal treatment can be found in [89], sections specifically useful for the discussion here are 2, 3, and 10-14. Often times rigor comes at a cost and can leave little room for the use of our intuition, so I will try to toe the line between formality and approachability in this section.

The foundational objects for probability theory are the sample space Ω, an event space $\mathcal{F}$, and a probability function P. Ω contains all the possible outcomes of a random process, such as heads (H) or tails (T) for flipping a coin or the subset $[0, 1.0] \subset \mathbb{R}$ when sampling the uniform distributions from 0 to 1. $\mathcal{F}$ contains subsets of Ω which are called events, which can be simple events like the coin landing on heads, or more complicated events like a dice roll resulting in an prime number. More precisely, $\mathcal{F}$ is a *σ-field* [7] , meaning that elements of $\mathcal{F}$ are measurable, and is required to contain Ω and be closed under the formation of complements or countable unions. The largest set $\mathcal{F}$ could be is the powerset of Ω, but it can be smaller. P is a function that maps elements of $\mathcal{F}$ to values from $[0, 1]$ which are probabilities, or the likelihood of the occurrence of a given event. These objects together form what is referred to as a probability space. Note that every subset of Ω need not belong in $\mathcal{F}$, which is why this formalism is useful — this helps correctly formalize the concept of probability when the sample space is continuous.

We must also formalize the idea of a random variable. Using our intuition, it is clear we want

[7] We are avoiding a lot of detail here regarding the construction of the field $\mathcal{F}$ such that it is a σ-field since that doesn't provide much utility for our purposes. Requiring that $\mathcal{F}$ is a σ-field avoids pathological examples that make it difficult to assign probabilities in situations where $\Omega \subseteq \mathbb{R}$. When we are working with sample spaces that are subsets of the real numbers, we will take $\mathcal{F}$ to be the Borel sets on Ω which is what we mean we say "measurable space".

a function that maps the potential outcomes of a random event, which is the sample space, to relevant values given the outcome. In the case of the coin flip, we have the sample space $\{ H, T \}$ and we can define random variable as the mapping that returns 1 in the case of heads and -1 in the case of tails. More formally, we would say that a random variable is a function that maps elements of the sample space Ω to another measurable space. A more general random variable is the average of 2 dice rolls — the sample space is $\{ \{ 1, 1 \}, \{ 1, 2 \}, \{ 2, 1 \} ..., \{ 12, 12 \} \}$ but E in this case would be a $\{ 1, 1.5, 2, ... \}$, all the possible averages. We can also see that the random variable must have probabilities for every possible element in it's image, so we require that $X^{-1}(o) \in \Omega$ for any $o \in E$. Without this requirement we could have outcomes of a random variable that are not in the domain for the probability function of our probability space, which is clearly not what we want.

In cases where the random variable X maps to a measurable space $E \subseteq \mathbb{R}$, we can see that X is a real valued function, $X : \Omega \to \mathbb{R}$. Typically the term "random variable" is restricted to only real valued functions, which is how we will use the term moving forward. A familiar example of a random variable is the sample of a uniform distribution from $[0, 1]$; in this case the sample space and measure space E are both $[0, 1]$. We can trivially extend this idea of a random variable to the concept of random vectors, where we define random vectors to just be collections of multiple random variables all on a shared probability space.

We will also make use of the concept of a measure, which is a set-function that maps elements of the event space $\mathcal{F}$ to some positive number $[0, \infty]$. Typically measures are written as μ or ν, and a measure acting on some set $A \in \mathcal{F}$ is $\mu(A) = [0, \infty]$. We also require the following relationship for measures applied on disjoint elements $A_1, A_2, A_3, ...$ of $\mathcal{F}$ where $\bigcup_i A_i \in \mathcal{F}$:

$$\mu(\bigcup_i U_i) = \sum_i \mu(U_i). \tag{C.1}$$

Probability measures, measures whose outputs represent probabilities, are naturally described in this formalism as measures where $\mu(\Omega) = 1$, which is a generalization of the idea that a PDF integrated over its domain must be 1.

A probability distribution of a random variable is then the probability measure μ on $\mathbb{R}$ defined by

$$\mu(A) = P(\{\omega : X(\omega) \in A\}),\ A \subseteq \mathbb{R}\,, \omega \in \Omega. \tag{C.2}$$

It is common to write $P(X)$, suppressing the technicality that X is a map from Ω to a measure space E. Another way to express this relationship is to say that μ is the law of X, or law$X \sim \mu$, which just means that the random variable X is distributed according to some measure μ. In a frequentist interpretation, if one evaluated X many times and made a histogram of the results they would find the results distributed according to μ. This generalization allows us to separate the probability of the outcomes of a random event, which is encoded in the measure, from the space over which we interpret that measure, which is the x when we write $p(x)$ for some one dimensional probability distribution. We can also relate this to the more familiar concept of a probability distribution function in the case of a continuous measure, specifically

$$\mu(A) = \int_A \mu(x)dx \tag{C.3}$$

where $\mu(x)$ is the probability distribution function.

We will also make use of the "push-through" of a measure μ through map T, $T_{\#}\mu[A] = \mu[T^{-1}(A)]$, where $T^{-1}(A)$ is the pre-image of A. This allows us to reinterpret a distribution over one space as distribution in another space by the map T. A simple example that can provide some intuition is that the push-through of some two-dimensional probability distribution $\pi(dx, dy)$, where $dx, dy \subset \mathbb{R}$, through the map p_x defined as $p_x(dx, dy) \to dx$ is the marginal in x of $\pi(dx, dy)$. For some $A \subseteq \mathbb{R}$ we have

$$p_{x\,\#}\pi[A] = \pi(p_x^{-1}[A]) = \int_A \int_{\mathbb{R}} \pi(x, y)dydx = \int_A \pi(x)dx \tag{C.4}$$

where $\pi(x) = \int_{\mathbb{R}} \pi(x, y)dy$ is the marginal density. As physicists we usually think of probability distributions being evaluated at some value, i.e. $\pi(x_0) = y$, but formally we mean is the probability mass contained in the density over some small region dx is $\pi(x_0)dx$, which is what $\int_A \pi(x)dx$ represents, and this is the marginal density of $\pi(x, y)$.

C.2. Optimal transport

The field of optimal transport provides a mathematical description of the interpolation of probability distributions, from an albeit slightly different perspective. Optimal transport attempts to provide solutions to a more general problem: given two distributions, what is the mapping from one distribution to the other that minimizes some cost? The problem can be stated more formally as [90]

$$C(\mu,\nu) = \inf_{\gamma\in\Gamma(\mu,\nu)} \int c(x,y)\gamma(x,y)dxdy \tag{C.5}$$

where $\mu(x) = \int \gamma(x,y)dy$, $\nu(y) = \int \gamma(x,y)dx$, $c(x,y)$ is some cost function for moving one unit of mass from x to y, and Γ is the set of all possible paths that map μ to ν, which is the set over which we want to minimize $c(x,y)$. Another way to think of γ is any function of x and y whose marginals are μ and ν. While the mathematicians working in the field of optimal transport work in more general topological spaces, for our purposes we can always assume $x, y \in \mathbb{R}^n$.

A convenient choice for the cost function is the distance function of the metric space over which μ and ν are defined, which we will call $d(x,y)$. This allows us to define a more useful concept, the *Wasserstein distance of order p*, W_p, defined as

$$W_p(\mu,\nu) = \left(\inf_{\gamma\in\Gamma(\mu,\nu)} \int d(x,y)^p\gamma(x,y)dxdy\right)^{1/p} = \left(\inf_{\pi\in\Pi(\mu,\nu)} \int d(x,y)^p\pi(x,y)d\pi\right)^{1/p} \tag{C.6}$$

where we have also combined the map into the measure using $\pi(x,y)d\pi = \gamma(x,y)dxdy$. When defined in such a way it can be shown that W_p defines a distance between μ and ν themselves, allowing us to promote the space of distributions to which μ and ν belong to a metric space.

The ability to think of probability distributions as belonging to a metric space is very useful, especially in the context of interpolation. One natural way to define our interpolation function is to think of the path we would want it to trace out in the metric space defined by W_p. The most obvious choice is to require the path traced out in the W_p space as we interpolate between μ and ν to be a geodesic in the metric space, or the path that minimizes the distance between μ and ν in the metric space.

Now let us setup the dynamic problem. We are interested in the paths probability masses follow as they are moved between the distributions μ_0 and μ_1. Since we are interested in how a distribution might be continuously deformed as we move along geodesics in this Wasserstein space, it is natural to extend our concept of the mapping $\pi(x, y)$ which connects the distributions to a more complicated object Π which instead is the collection of the *distance minimizing* paths that the different masses take during the transportation. $\pi(x, y)$ was agnostic regarding this information, it only encoded where probability mass started and ended. Π on the other hand encodes the information of the path taken as the mass is moved from one place to another. We can define the elements of Π as the curves $\gamma(t)$, where Π is the distribution of many of these curves and $t \in [0, 1]$ parametrizes the distance along the curve. In probability speak, $\gamma(t)$ is a random variable itself, and at a given time the law$(\gamma(t))$ provides us with the probability distribution μ_t, i.e. the interpolated probability distribution. In this context, we will limit ourselves to situations where the starting and ending distributions, $\mu_0(\chi)$ and $\mu_1(\chi)$ are both over the same metric space, (χ, d) where χ is a measurable space and d is the distance function of the metric. Again, we are always interested in the case $\chi = \mathbb{R}^n$.

Equipped with these new tools, we can make some statements about the curves $\gamma(t)$. We require law$(\gamma(0)) = \mu_0$ and law$(\gamma(1)) = \mu_1$ as these are the starting and ending conditions of the curves. We will also introduce $e_t(\gamma) := \gamma(t)$, a function at each t that maps any curve γ to a point in the metric space χ. We define a transference plan between μ_0 and μ_1 such that

$$\pi_{0,1} = (e_0, e_1)_{\#}\Pi. \tag{C.7}$$

Note that this object is entirely analogous to $\pi(x, y)$ (when working in $\mathbb{R}^n$), and this is consistent with the statements above that law$(\gamma(0)) = \mu_0$ and law$(\gamma(1)) = \mu_1$ since the marginals of π should be the two distributions we are interpolating between. Furthermore, the plan in eq. C.7 is *optimal* by construction since we defined Π to be the measure only over the distance minimizing curves between μ_0 and μ_1. And finally, given the optimal transference plan it can be shown that the following statements are equivalent (Corollary 7.22 [90]):

(i) μ_t is the law of $\gamma(t)$ where γ is a random (minimizing, constant-speed) geodesic such that $(\gamma(0), \gamma(1))$ is an optimal coupling

(ii) $(\mu_t)_{0 \leq t \leq 1}$ is a geodesic curve in the space $P_p(\chi)$.

The space over which the distributions are defined, χ and the metric of that space, d, are enough to define the paths that the probability masses follow as they move between the two distributions. They just follow the geodesics of the space χ!

This is very useful in interpolation for probability distributions over $\mathbb{R}^n$ since geodesics in flat space are just straight lines. This means the $\gamma(t)$'s are all just straight lines connecting the probability masses between the beginning and ending distributions. This is also distinct from linear interpolation, since we are not interpolating over distribtions themselves, but rather interpolating over the movement of probability between the two distributions which happens to follow straight line paths.

While finding the optimal transference even in flat space can be difficult, in 1 dimension over $\mathbb{R}$ we can derive some useful results. The following discussion is all assumed to be in 1D where the space over which distributions are defined is $\mathbb{R}$. First, define the cumulative distribution function (CDF) of some probability density $\mu(x)$ as

$$\int_{-\infty}^{x'} \mu(x)dx = F_\mu(x') \tag{C.8}$$

and the quantile function as F_μ^{-1}, such that $F_\mu^{-1}(F_\mu(x)) = x$. Put more simply, the quantile function takes in a probability and maps it back to the point x_0 where $F_\mu(x)$ reaches that probability. It can be shown that in 1D the Wasserstein p-distance is given by (Remark 2.30 [80])

$$W_p = \left(\int_0^1 \left| F_\mu^{-1}(u) - F_\nu^{-1}(u) \right|^p du \right)^{1/p} \tag{C.9}$$

or just the total difference between the quantile functions.

We can also see that we can construct a map from the measure $\mu(x)$ to $\nu(x)$ through combinations

of the quantile functions and CDFs. One way we can think of constructing this map is by first trying to find some way to map both μ and ν onto the same underlying distribution, like the uniform distribution from 0 to 1. First consider the push through of μ through F_μ for some arbitrary interval $A = [0, u]$ for $u \in [0, 1]$ [8]

$$F_{\mu\,\#}\mu[[0, u]] = \mu[F_\mu^{-1}([0, u])]. \tag{C.10}$$

Applying the definition of the quantile function we have

$$\mu[F_\mu^{-1}([0, u])] = \mu[\{x : F_\mu(x) \leq u\}] = u \tag{C.11}$$

where we have used the fact that $F_\mu^{-1}([0, u])$ is just the range of values where the CDF is less than u. As for the last equality, we know by definition that F_μ is just the summed probability of μ meaning $\mu[\{x : F_\mu(x) \leq u\}] = u$. Another way to see this is by considering that in $\mathbb{R}$ we can state that $\mu[A] = \int_A \mu(x)dx$ where $A \subseteq \mathbb{R}$. Since A is defined as the set $x : F_\mu(x) \leq u$ it is guaranteed that μ integrated over A is μ. And since u is arbitrary, this holds for any value between 0 and 1.

Finally, we can recognize that this transformed measure $\rho = F_{\mu\,\#}\mu$ is a measure where the probability mass is equal to the interval over which the measure is evaluated for any interval 0 to 1, which is the uniform distribution, implying $F_{\mu\,\#}\mu = \text{Uniform}(0, 1)$. This actually formalizes the frequently used trick in which one samples a distribution using uniform sampling evaluated on the quantile function (we just proved it in the reverse direction).

Of course, this same push through can be applied to ν, and the inverse of this transformation can also be used to instead transform Uniform$(0, 1)$ to ν. Putting this together we have $F_{\nu\,\#}^{-1}[F_{\mu\,\#}\mu] = (F_\nu^{-1})_\#\text{Uniform}(0, 1) = \nu$.

It turns out that this maps is indeed *the* optimal map from μ to ν [80]. We will not formally prove this, but it is clear this must be the case if one considers that for flat-space geodesics, any crossing paths of the probability masses would leave room for a more optimal map in which the paths do

[8] While we define F_μ as a function from scalar values to scalar values, we use it in the push through as a map from one measure to another, so over subsets of $\mathbb{R}$ in this instance. This is an abuse of notation, but since F_μ is non-decreasing due to it's construction, we don't need to worry too much about distinguishing between the two cases.

not cross. Since the CDF is non-decreasing, this push through of the CDF of one measure to the quantile function of another guarantees non-crossing paths, thus this must be the optimal map.

If we consider how a point in x moves during this mapping process we find that it moves according to the functional convolution of the CDF and the QF:

$$x \to F_\nu^{-1}(F_\mu(x)). \tag{C.12}$$

Recall that these distributions are all defined over $\mathbb{R}$, and geodesics in flat space are all straight lines. As was discussed above, this implies that the paths traced out by probability masses as they are interpolated from μ to ν along the dynamic transport map must also be straight lines. We can therefore write the optimal dynamic map as $T_t(x) = (1-t)x + tF_\nu^{-1}(F_\mu(x))$. Another way to write the dynamic map is using a flat underlying measure rather than one distributed according to μ. This can be accomplished through the transformation $x \to F_\mu^{-1}(x)$, giving a new map

$$T_t = (1-t)F_\mu^{-1} + tF_\nu^{-1}. \tag{C.13}$$

Now T_t represents the map that takes a uniform measure and maps it onto the interpolated PDF, but this is exactly what the quantile function does, thus we can write

$$F_t^{-1} = (1-t)F_\mu^{-1} + tF_\nu^{-1} \tag{C.14}$$

or that in 1D, one can interpolate between PDFs by interpolating between the quantile functions. This form of interpolation is often referred to as McCann's interpolation [91]. The right of fig. 6.16 shows this interpolation strategy being used to interpolate between two Gaussian distributions. It can be shown that interpolation along geodesics in the Wasserstein space is equivalent to interpolating in the variances and means of Gaussian distributions, which is an often used technique when working with normal distributions. [80].

Another useful concept in optimal transport is the "Wasserstein barycenter", the generalization of

the idea of a barycenter in Euclidean space, Given some set of N measures $\{\nu_1, \nu_2, ..., \nu_N\}$, the Wasserstein barycenter is the measure that minimizes the function

$$f(\mu) = \frac{1}{N} \sum_{N} W_p^p(\mu, \nu_i) \tag{C.15}$$

or in other words, the measure that minimizes the distance in Wasserstein space between itself and all the measures in a given set. As was true before, we can find a closed form solution for the calculation of the Wasserstein barycenter in 1D. The minimizing measure is the measure whose quantile function is just the average of the quantile functions of the other measures:

$$F_\mu^{-1} = \frac{1}{N} \sum_{N} F_{\nu_i}^{-1} \tag{C.16}$$

If we instead consider arbitrary rather than equal weights on each of the measures, we have

$$F_\mu^{-1} = \frac{1}{N} \sum_{N} w_i F_{\nu_i}^{-1} = W_b \tag{C.17}$$

which is referred to as the generalized Wasserstein barycenter (W_b).

APPENDIX D

Example YAML snopfit configuration file

```yaml
fit:
  name: ppo_bi214
  systematics:
    energy_scale:
      dimension: energy
      type: scale
      sigma: inf
      value: 1.00
      constraints: [0.90, 1.1]
      float: true
    energy_smear:
      dimension: energy
      type: smear
      sigma: inf
      value: 0.02
      constraints: [0.001, 0.5]
      float: true
    itr_shift:
      dimension: itr
      type: shift
      value: -0.003
      sigma: inf
      constraints: [-0.2, 0.2]
      float: true
    itr_smear:
      dimension: itr
      type: smear
      value: 0.001
      sigma: inf
      constraints: [0.0001, 0.10]
      float: true
  datasets:
    ppo:
      dimensions:
        energy:
          bin_info: [1.5, 3.5, 40]
          buffer_bins: [2, 2]
        itr:
          bin_info: [0.15, 0.35, 30]
          buffer_bins: [2, 2]
      load:
        cuts: "
```

```
    (fitValid == 1) and
    (itr > 0.15) and (itr < 0.35) and
    (runID >= 300702) and
    (runID <  310637) and
    (energy > 1.0) and (energy < 7.0) and
    (posr_av > 0) and (posr_av < 4500) and
    (analysis_mask == 1) and
    (bipo214_flag == 1) and
    (hotspot1_flag == 0) and
    (hotspot2_flag == 0) and
    (antinu_flag == 0) and
    (beta14 > -0.1) and (beta14 < 0.04) and
    (necknhits == 0) and
    (posFOM/nhitsCleaned > 11.5) and
    (bipo214_delta_r < 1000) and
    (bipo214_delta_t > 3000) and
    (bipo214_delta_t < 1000000)
    "
data:
  file: ""
  livetime_file: ""
  corrections_file: ""

mc:
  path: ""
  dataset_root: ""
  corrections_file: ""
  background_defaults:
    cuts: "
      (fitValid == 1) and
      (itr > 0.0) and (itr < 0.6) and
      (runID >= 300702) and
      (runID <  310637) and
      (energy > 1.0) and (energy < 7.0) and
      (posr_av > 0) and (posr_av < 4500)
      "
    efficiency_cuts: "
      (fitValid == 1) and
      (itr > 0.15) and (itr < 0.35) and
      (energy > 1.0) and (energy < 7.0) and
      (posr_av > 0) and (posr_av < 4500)
      "
    float: true
    scale: true
    constraints: [0, inf]
    rate: 1
    sigma: 0
    systematics:
```

```
91                ["energy_scale", "energy_smear", "itr_shift", "itr_smear"]
92            backgrounds:
93              Bipo214:
94                rate: nev
95                scale: false
96                sigma: inf
97                float: false
```

BIBLIOGRAPHY

[1] Wolfgang Pauli. "Pauli letter collection." Typed copy. URL: https://cds.cern.ch/record/83282.

[2] Fred L. Wilson. "Fermi's Theory of Beta Decay." In: *American Journal of Physics* 36.12 (Dec. 1968), pp. 1150–1160. ISSN: 0002-9505. DOI: 10.1119/1.1974382. eprint: https://pubs.aip.org/aapt/ajp/article-pdf/36/12/1150/11891052/1150_1_online.pdf. URL: https://doi.org/10.1119/1.1974382.

[3] C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation." In: *Science* 124.3212 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103. eprint: https://www.science.org/doi/pdf/10.1126/science.124.3212.103. URL: https://www.science.org/doi/abs/10.1126/science.124.3212.103.

[4] Nobel Prize Outreach. *The Nobel Prize in Physics 1995*. NobelPrize.org. Accessed 18 Aug 2026. 2026. URL: https://www.nobelprize.org/prizes/physics/1995/summary/.

[5] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. "Search for Neutrinos from the Sun." In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[6] John N. Bahcall. "Solar Neutrinos. I. Theoretical." In: *Phys. Rev. Lett.* 12 (11 Mar. 1964), pp. 300–302. DOI: 10.1103/PhysRevLett.12.300. URL: https://link.aps.org/doi/10.1103/PhysRevLett.12.300.

[7] Bruce T. Cleveland et al. "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector." In: *APJ* 496.1 (Mar. 1998), pp. 505–526. DOI: 10.1086/305343.

[8] Gabriel D. Orebi Gann et al. "The Future of Solar Neutrinos." In: *Annual Review of Nuclear and Particle Science* 71.Volume 71, 2021 (2021), pp. 491–528. ISSN: 1545-4134. DOI: https://doi.org/10.1146/annurev-nucl-011921-061243. URL: https://www.annualreviews.org/content/journals/10.1146/annurev-nucl-011921-061243.

[9] F. Kaether et al. "Reanalysis of the Gallex solar neutrino flux and source experiments." In: *Physics Letters B* 685.1 (2010), pp. 47–54. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2010.01.030. URL: https://www.sciencedirect.com/science/article/pii/S0370269310000729.

[10] J. N. Abdurashitov et al. "Measurement of the solar neutrino capture rate with gallium metal. III. Results for the 2002–2007 data-taking period." In: *Phys. Rev. C* 80 (1 July 2009), p. 015807. DOI: 10.1103/PhysRevC.80.015807. URL: https://link.aps.org/doi/10.1103/PhysRevC.80.015807.

[11] K. S. Hirata et al. "Observation of ^{8}B solar neutrinos in the Kamiokande-II detector." In: *Phys. Rev. Lett.* 63 (1 July 1989), pp. 16–19. DOI: 10.1103/PhysRevLett.63.16. URL: https://link.aps.org/doi/10.1103/PhysRevLett.63.16.

[12] J Boger et al. "The Sudbury Neutrino Observatory." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 449.1 (2000), pp. 172–207. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(99)01469-2. URL: https://www.sciencedirect.com/science/article/pii/S0168900299014692.

[13] B. Aharmim et al. “Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory.” In: *Phys. Rev. C* 88 (2 Aug. 2013), p. 025501. DOI: 10.1103/PhysRevC.88.025501. URL: https://link.aps.org/doi/10.1103/PhysRevC.88.025501.

[14] F. Takahashi et al. “Review of Particle Physics.” In: *Int. J. Mod. Phys. A* 41 (2026), p. 2630011. DOI: 10.1142/S0217751X26300115.

[15] T. Kajita, E. Kearns, and M. Shiozawa. “Establishing atmospheric neutrino oscillations with Super-Kamiokande.” In: *Nuclear Physics B* 908 (2016). Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015, pp. 14–29. ISSN: 0550-3213. DOI: https://doi.org/10.1016/j.nuclphysb.2016.04.017. URL: https://www.sciencedirect.com/science/article/pii/S0550321316300554.

[16] Nobel Prize Outreach. *Advanced Information*. NobelPrize.org. Accessed 18 Aug 2026. 2026. URL: https://www.nobelprize.org/prizes/physics/2015/advanced-information/.

[17] Ettore Majorana and Luciano Maiani. “A symmetric theory of electrons and positrons.” In: *Ettore Majorana Scientific Papers: On occasion of the centenary of his birth*. Ed. by Giuseppe Franco Bassani. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 201–233. ISBN: 978-3-540-48095-2. DOI: 10.1007/978-3-540-48095-2_10. URL: https://doi.org/10.1007/978-3-540-48095-2_10.

[18] Murray Gell-Mann, Pierre Ramond, and Richard Slansky. *Complex Spinors and Unified Theories*. 2013. arXiv: 1306.4669 [hep-th]. URL: https://arxiv.org/abs/1306.4669.

[19] Rabindra N. Mohapatra and Goran Senjanovi ć. “Neutrino Mass and Spontaneous Parity Nonconservation.” In: *Phys. Rev. Lett.* 44 (14 Apr. 1980), pp. 912–915. DOI: 10.1103/PhysRevLett.44.912. URL: https://link.aps.org/doi/10.1103/PhysRevLett.44.912.

[20] Z. Maki, M. Nakagawa, and S. Sakata. “Remarks on the Unified Model of Elementary Particles.” In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. DOI: 10.1143/PTP.28.870.

[21] Ivan Esteban et al. “NuFit-6.0: updated global analysis of three-flavor neutrino oscillations.” In: *Journal of High Energy Physics* 2024.12 (Dec. 2024). ISSN: 1029-8479. DOI: 10.1007/jhep12(2024)216. URL: http://dx.doi.org/10.1007/JHEP12(2024)216.

[22] B. Abi et al. *Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume I: Introduction to DUNE*. 2020. arXiv: 2002.02967 [physics.ins-det]. URL: https://arxiv.org/abs/2002.02967.

[23] Hyper-Kamiokande Proto-Collaboration et al. *Hyper-Kamiokande Design Report*. 2018. arXiv: 1805.04163 [physics.ins-det]. URL: https://arxiv.org/abs/1805.04163.

[24] S. P. Mikheyev and A. Yu. Smirnov. “Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos.” In: *Sov. J. Nucl. Phys.* 42 (1985), pp. 913–917.

[25] L. Wolfenstein. “Neutrino oscillations in matter.” In: *Phys. Rev. D* 17 (9 May 1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369. URL: https://link.aps.org/doi/10.1103/PhysRevD.17.2369.

[26] Núria Vinyoles et al. “A New Generation of Standard Solar Models.” In: *The Astrophysical Journal* 835.2 (Jan. 2017), p. 202. ISSN: 1538-4357. DOI: 10.3847/1538-4357/835/2/202. URL: http://dx.doi.org/10.3847/1538-4357/835/2/202.

[27] Johannes Bergström et al. "Updated determination of the solar neutrino fluxes from solar neutrino data." In: *Journal of High Energy Physics* 2016.3 (Mar. 2016). ISSN: 1029-8479. DOI: 10.1007/jhep03(2016)132. URL: http://dx.doi.org/10.1007/JHEP03(2016)132.

[28] R. Bonventre et al. "Nonstandard models, solar neutrinos, and large θ_{13}." In: *Phys. Rev. D* 88 (5 Sept. 2013), p. 053010. DOI: 10.1103/PhysRevD.88.053010. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.053010.

[29] Alexander Friedland, Cecilia Lunardini, and Carlos Peña-Garay. "Solar neutrinos as probes of neutrino–matter interactions." In: *Physics Letters B* 594.3-4 (Aug. 2004), pp. 347–354. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2004.05.047. URL: http://dx.doi.org/10.1016/j.physletb.2004.05.047.

[30] K. Abe et al. "Solar neutrino measurements using the full data period of Super-Kamiokande-IV." In: *Phys. Rev. D* 109 (9 May 2024), p. 092001. DOI: 10.1103/PhysRevD.109.092001. URL: https://link.aps.org/doi/10.1103/PhysRevD.109.092001.

[31] M. Agostini et al. "Simultaneous precision spectroscopy of *pp*, ^{7}Be, and *pep* solar neutrinos with Borexino Phase-II." In: *Phys. Rev. D* 100 (8 Oct. 2019), p. 082004. DOI: 10.1103/PhysRevD.100.082004. URL: https://link.aps.org/doi/10.1103/PhysRevD.100.082004.

[32] M. Agostini et al. "Improved measurement of ^{8}B solar neutrinos with 1.5 kt y of Borexino exposure." In: *Physical Review D* 101.6 (Mar. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.101.062001. URL: http://dx.doi.org/10.1103/PhysRevD.101.062001.

[33] C Arnaboldi et al. "CUORE: a cryogenic underground observatory for rare events." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 518.3 (Feb. 2004), pp. 775–798. ISSN: 0168-9002. DOI: 10.1016/j.nima.2003.07.067. URL: http://dx.doi.org/10.1016/j.nima.2003.07.067.

[34] H. Acharya et al. "First Results on the Search for Lepton Number Violating Neutrinoless Double-β Decay with the LEGEND-200 Experiment." In: *Phys. Rev. Lett.* 136 (2 Jan. 2026), p. 022701. DOI: 10.1103/25tk-nctn. URL: https://link.aps.org/doi/10.1103/25tk-nctn.

[35] S. Abe et al. "Search for Majorana Neutrinos with the Complete KamLAND-Zen Dataset." In: *Phys. Rev. Lett.* 135 (26 Dec. 2025), p. 262501. DOI: 10.1103/jkf6-48j8. URL: https://link.aps.org/doi/10.1103/jkf6-48j8.

[36] I. J. Arnquist et al. "Final Result of the Majorana Demonstrator's Search for Neutrinoless Double-β Decay in ^{76}Ge." In: *Phys. Rev. Lett.* 130 (6 Feb. 2023), p. 062501. DOI: 10.1103/PhysRevLett.130.062501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.130.062501.

[37] M. Agostini et al. "Final Results of GERDA on the Search for Neutrinoless Double-β Decay." In: *Phys. Rev. Lett.* 125 (25 Dec. 2020), p. 252502. DOI: 10.1103/PhysRevLett.125.252502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.252502.

[38] S. Biller. private communication. 2026.

[39] V. Albanese et al. "The SNO+ experiment." In: *Journal of Instrumentation* 16.08 (Aug. 2021), P08059. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/08/p08059. URL: http://dx.doi.org/10.1088/1748-0221/16/08/P08059.

[40] C.J Jillings et al. “The photomultiplier tube testing facility for the Sudbury Neutrino Observatory.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 373.3 (1996), pp. 421–429. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(96)00067-8. URL: https://www.sciencedirect.com/science/article/pii/0168900296000678.

[41] Stefan Schoppmann. “Review of Novel Approaches to Organic Liquid Scintillators in Neutrino Physics.” In: *Symmetry* 15.1 (Dec. 2022), p. 11. ISSN: 2073-8994. DOI: 10.3390/sym15010011. URL: http://dx.doi.org/10.3390/sym15010011.

[42] M.R. Anderson et al. “Development, characterisation, and deployment of the SNO+ liquid scintillator.” In: *Journal of Instrumentation* 16.05 (May 2021), P05009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/05/p05009. URL: http://dx.doi.org/10.1088/1748-0221/16/05/P05009.

[43] J B Birks. “Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations.” In: *Proceedings of the Physical Society. Section A* 64.10 (Oct. 1951), p. 874. DOI: 10.1088/0370-1298/64/10/303. URL: https://doi.org/10.1088/0370-1298/64/10/303.

[44] D.J. Auty et al. “A method to load tellurium in liquid scintillator for the study of neutrinoless double beta decay.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1051 (2023), p. 168204. ISSN: 0168-9002. DOI: 10.1016/j.nima.2023.168204. URL: http://dx.doi.org/10.1016/j.nima.2023.168204.

[45] D. Basilico et al. “Final results of Borexino on CNO solar neutrinos.” In: *Phys. Rev. D* 108 (10 Nov. 2023), p. 102005. DOI: 10.1103/PhysRevD.108.102005. URL: https://link.aps.org/doi/10.1103/PhysRevD.108.102005.

[46] A. Allega et al. “Event-by-event direction reconstruction of solar neutrinos in a high light-yield liquid scintillator.” In: *Physical Review D* 109.7 (Apr. 2024). ISSN: 2470-0029. DOI: 10.1103/physrevd.109.072002. URL: http://dx.doi.org/10.1103/PhysRevD.109.072002.

[47] T. Anderson et al. “Eos: conceptual design for a demonstrator of hybrid optical detector technology.” In: *Journal of Instrumentation* 18.02 (Feb. 2023), P02009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/18/02/p02009. URL: http://dx.doi.org/10.1088/1748-0221/18/02/P02009.

[48] SNO+ collaboration et al. “Optical calibration of the SNO+ detector in the water phase with deployed sources.” In: *Journal of Instrumentation* 16.10 (Oct. 2021), P10021. DOI: 10.1088/1748-0221/16/10/P10021. URL: https://doi.org/10.1088/1748-0221/16/10/P10021.

[49] M. Anderson et al. “Search for invisible modes of nucleon decay in water with the SNO+ detector.” In: *Phys. Rev. D* 99 (3 Feb. 2019), p. 032008. DOI: 10.1103/PhysRevD.99.032008. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.032008.

[50] M. R. Anderson et al. “Measurement of neutron-proton capture in the SNO+ water phase.” In: *Phys. Rev. C* 102 (1 July 2020), p. 014002. DOI: 10.1103/PhysRevC.102.014002. URL: https://link.aps.org/doi/10.1103/PhysRevC.102.014002.

[51] A. Allega et al. “Improved search for invisible modes of nucleon decay in water with the SNO + detector.” In: *Phys. Rev. D* 105 (11 June 2022), p. 112012. DOI: 10.1103/PhysRevD.105.112012. URL: https://link.aps.org/doi/10.1103/PhysRevD.105.112012.

[52] M. Anderson et al. “Measurement of the ^{8}B solar neutrino flux in SNO+ with very low backgrounds.” In: *Phys. Rev. D* 99 (1 Jan. 2019), p. 012012. DOI: 10.1103/PhysRevD.99.012012. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.012012.

[53] A. Allega et al. “Evidence of Antineutrinos from Distant Reactors Using Pure Water at SNO+.” In: *Phys. Rev. Lett.* 130 (9 Mar. 2023), p. 091801. DOI: 10.1103/PhysRevLett.130.091801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.130.091801.

[54] A.S. Inácio. “Data Analysis of the Water and Scintillator Phases of SNO+: from Solar Neutrino Measurements to Double Beta Decay Sensitivity Studies.” PhD thesis. Universidade de Lisboa, 2022.

[55] SNO+ Collaboration et al. *Initial measurement of reactor antineutrino oscillation at SNO+*. 2025. DOI: https://doi.org/10.1140/epjc/s10052-024-13687-5. arXiv: 2405.19700 [hep-ex]. URL: https://arxiv.org/abs/2405.19700.

[56] Hamamatsu Photonics K. K. *PHOTOMULTIPLIER TUBES Basics and Applications FOURTH EDITION*. 2017. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/PMT_handbook_v4E.pdf.

[57] A. Zummo. “Measurement of Reactor Antineutrino Oscillation with SNO+.” PhD thesis. University of Pennsylvania, 2024.

[58] S. Seibert et al. URL: https://rat.readthedocs.io/en/latest/.

[59] 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 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[60] “Position Reconstruction in ScintFitter.” SNO+ Internal Document 8607.

[61] T. Kroupová. “Improving the Sensitivity to Neutrinoless Double Beta Decay in SNO+.” PhD thesis. University of Oxford, 2020.

[62] S. Abe et al. “Search for Majorana Neutrinos with the Complete KamLAND-Zen Dataset.” In: *Physical Review Letters* 135.26 (Dec. 2025). ISSN: 1079-7114. DOI: 10.1103/jkf6-48j8. URL: http://dx.doi.org/10.1103/jkf6-48j8.

[63] R. Brun and F. Rademakers. “ROOT: An object oriented data analysis framework.” In: *Nucl. Instrum. Meth. A* 389 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X.

[64] jbeder. URL: https://github.com/jbeder/yaml-cpp.

[65] “BiPo214/212 plots for approval (Both PPO and bisMSB).” SNO+ Internal Document 9115.

[66] G. Alimonti et al. “The Borexino detector at the Laboratori Nazionali del Gran Sasso.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 600.3 (2009), pp. 568–593. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2008.11.076. URL: https://www.sciencedirect.com/science/article/pii/S016890020801601X.

[67] J. D. Hunter. “Matplotlib: A 2D graphics environment.” In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[68] M. Galassi et al. *GNU Scientific Library Reference Manual (3rd Ed.)* ISBN: ISBN 0954612078.

[69] kubkon. URL: https://github.com/kubkon/zig-yaml.

[70] Lewis Van Winkle. *TinyExpr: tiny recursive descent expression parser, compiler, and evaluation engine for math expressions*. Version 1.1.1. https://github.com/codeplea/tinyexpr. Aug. 2026. DOI: 10.5281/zenodo.21817114. URL: https://codeplea.com/tinyexpr.

[71] The HDF Group. *Hierarchical Data Format, version 5*. https://github.com/HDFGroup/hdf5. DOI: 10.5281/zenodo.17808614. URL: https://www.hdfgroup.org/HDF5/.

[72] jbeder. URL: https://github.com/smnaugle/llfit_zig.

[73] M. Abreu et al. “Measurement of Reactor Antineutrino Oscillation at SNO+.” In: *Physical Review Letters* 135.12 (2025). ISSN: 1079-7114. DOI: 10.1103/gypt-lc9v. URL: http://dx.doi.org/10.1103/gypt-lc9v.

[74] C. Andreopoulos et al. “The GENIE neutrino Monte Carlo generator.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 614.1 (Feb. 2010), pp. 87–104. ISSN: 0168-9002. DOI: 10.1016/j.nima.2009.12.009. URL: http://dx.doi.org/10.1016/j.nima.2009.12.009.

[75] C. Lin. “Atmospheric Neutrinos in Partial Fill.”

[76] A. and Allega et al. “Initial measurement of reactor antineutrino oscillation at SNO+.” In: *The European Physical Journal C* 85.1 (Jan. 2025). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-024-13687-5. URL: http://dx.doi.org/10.1140/epjc/s10052-024-13687-5.

[77] Mario Schwarz et al. “Measurements of the lifetime of orthopositronium in the LAB-based liquid scintillator of JUNO.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 922 (Dec. 2018). DOI: 10.1016/j.nima.2018.12.068.

[78] G Peter Lepage. “A new algorithm for adaptive multidimensional integration.” In: *Journal of Computational Physics* 27.2 (1978), pp. 192–203. ISSN: 0021-9991. DOI: https://doi.org/10.1016/0021-9991(78)90004-9. URL: https://www.sciencedirect.com/science/article/pii/0021999178900049.

[79] John Skilling. “Nested sampling for general Bayesian computation.” In: *Bayesian Analysis* 1 (Dec. 2006), pp. 833–860. DOI: 10.1214/06-BA127.

[80] Gabriel Peyré and Marco Cuturi. *Computational Optimal Transport*. 2020. arXiv: 1803.00567 [stat.ML]. URL: https://arxiv.org/abs/1803.00567.

[81] cmglee. CC BY-SA 3.0, via Wikimedia Commons. 2019. URL: https://commons.wikimedia.org/wiki/File:Bilinear_interpolation_visualisation.svg (visited on 08/05/2026).

[82] M. C. Gonzalez-Garcia et al. “Status of direct determination of solar neutrino fluxes after Borexino.” In: *Journal of High Energy Physics* 2024.2 (Feb. 2024). ISSN: 1029-8479. DOI: 10.1007/jhep02(2024)064. URL: http://dx.doi.org/10.1007/JHEP02(2024)064.

[83] G. Milton. Private communication. Aug. 21, 2026.

[84] Steven D. Biller. "Combined constraints on Majorana masses from neutrinoless double beta decay experiments." In: *Physical Review D* 104.1 (2021). ISSN: 2470-0029. DOI: 10.1103/physrevd.104.012002. URL: http://dx.doi.org/10.1103/PhysRevD.104.012002.

[85] Greg Ashton et al. "Nested sampling for physical scientists." In: *Nature Reviews Methods Primers* 2.1 (May 2022), p. 39. ISSN: 2662-8449. DOI: 10.1038/s43586-022-00121-x. URL: https://doi.org/10.1038/s43586-022-00121-x.

[86] Radford M. Neal. *Slice Sampling*. 2000. arXiv: physics/0009028 [physics.data-an]. URL: https://arxiv.org/abs/physics/0009028.

[87] Optimsmoose. CC BY-SA 3.0, via Wikimedia Commons. 2011. URL: https://commons.wikimedia.org/wiki/File:Summary_of_slice_sampling.png (visited on 08/12/2026).

[88] George Marsaglia. "Choosing a Point from the Surface of a Sphere." In: *The Annals of Mathematical Statistics* 43.2 (1972), pp. 645–646. DOI: 10.1214/aoms/1177692644. URL: https://doi.org/10.1214/aoms/1177692644.

[89] P. Billingsley. *Probability and Measure*. 1995.

[90] Cédric Villani.

[91] Robert J. McCann. "A Convexity Principle for Interacting Gases." In: *Advances in Mathematics* 128.1 (1997), pp. 153–179. ISSN: 0001-8708. DOI: https://doi.org/10.1006/aima.1997.1634. URL: https://www.sciencedirect.com/science/article/pii/S0001870897916340.