UNIVERSITY OF CALIFORNIA,
IRVINE

New Methods to Reduce Cosmogenic Backgrounds of Super-Kamiokande in the Solar Neutrino Energy Regime

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Scott Locke

Dissertation Committee:
Professor Emeritus Henry Sobel, Chair
Professor Steve Barwick
Professor Mu-Chun Chen

2020

DEDICATION

This work is dedicated to my wife, Chengyan, whose love, patience, and support during this long and arduous process cannot be overstated. To my parents Sue, Dave, and Becky for their support and love over the years, starting well before I embarked on this adventure. To my grandpa Jerry Prosser and Anne Barnes who were some of the biggest supporters for me getting my PhD. To the rest of my family for their continued support. Bonnie and Chuck Ulman for being my Huntington Beach parents, as I was their Vancouver son. To Gil Ramirez for being the longest roommate I have ever had, and being a great one at that. For all of my friends in Irvine who helped me take a step away from work and enjoy the little things. Specifically to the Balboa crew for the lunchtime raids and random outings, Ben Avila for helping me with most iterations of Yuri's Fury, and Mike Milgie for the poker nights and Tuesday basketball. Lauren Anthony and Masacko Ishikura for organizing the Saturday dinners during tank open work, as well as helping keep my sanity that summer.

TABLE OF CONTENTS

Page

LIST OF FIGURES **vi**

LIST OF TABLES **xiv**

ACKNOWLEDGMENTS **xv**

VITA **xvi**

ABSTRACT OF THE DISSERTATION **xix**

1 Introduction **1**
- 1.1 Neutrino Detection . . . 3
- 1.2 Types of Neutrinos . . . 5
- 1.3 Neutrino Oscillations . . . 7
- 1.4 Neutrons . . . 14

2 Super Kamiokande Detector **15**
- 2.1 Overview of the Super-Kamiokande Experiment . . . 15
- 2.2 Detector Specifics . . . 18
- 2.3 PMTs . . . 19
- 2.4 SK Water System . . . 21
- 2.5 Air Purification Systems . . . 23
- 2.6 EGADS . . . 24
- 2.7 Super Kamiokande High Voltage, Electronics, and DAQ . . . 27
  - 2.7.1 SK-IV and QBEEs . . . 28
  - 2.7.2 High Voltage . . . 29
- 2.8 WIT . . . 31
- 2.9 2018 Detector Maintenance . . . 34
  - 2.9.1 SK Leak . . . 34
  - 2.9.2 Rust Removal . . . 36
  - 2.9.3 Water Flow Changes . . . 37
- 2.10 Physics within SK . . . 38

3 Calibration and Simulation in Super Kamiokande **40**
3.1 Event Simulation . . . 40
3.1.1 Cherenkov Photon Production . . . 41
3.1.2 Detector Response . . . 42
3.2 PMT Calibration . . . 43
3.2.1 Absolute Gain . . . 43
3.2.2 Relative Gain . . . 45
3.2.3 PMT Quantum Efficiency Measurement . . . 46
3.2.4 PMT Timing Calibration . . . 49
3.3 Water Transparency Measurement . . . 51
3.3.1 Measuring Water Transparency with Laser Light Injection . . . 52
3.3.2 Measurement of Water Transparency with Decay Electrons . . . 54
3.3.3 Position Dependence and Top-Bottom Asymmetry . . . 57
3.4 Absolute Energy Scale Calibration . . . 60
3.4.1 LINAC . . . 60
3.4.2 DT Generator . . . 62

4 SK Event Reconstruction **65**
4.1 Muon Track Fitter: Muboy . . . 65
4.2 Event Reconstruction . . . 69
4.2.1 Vertex Reconstruction . . . 69
4.2.2 Direction Reconstruction . . . 70
4.2.3 Energy Reconstruction . . . 70
4.3 Evaluating Fit Quality of BONSAI . . . 74
4.3.1 BONSAI Vertex Reconstruction Goodness . . . 74
4.3.2 KS test on Direction Reconstruction . . . 75

5 Backgrounds and Data Reduction **76**
5.1 Cosmogenic Backgrounds . . . 76
5.1.1 Cosmic Rays . . . 76
5.1.2 Spallation . . . 77
5.2 Detector Material Induced Backgrounds . . . 79
5.3 Trigger Rates . . . 81
5.4 Data Reduction Overview . . . 82
5.5 Pre-Reduction . . . 83
5.5.1 Vertex Position Cut . . . 84
5.5.2 Loose External Cut . . . 84
5.5.3 Energy cut . . . 84
5.5.4 Loose Event Quality Cut . . . 84
5.5.5 Pre-reduction Summary . . . 85
5.6 First Reduction . . . 85
5.6.1 Bad Run Removal . . . 85
5.6.2 Number of Hit PMTs . . . 88
5.6.3 Calibration Source Cut . . . 88
5.6.4 PMT Flasher Cut . . . 89

5.6.5 Muon Related Event Cut . . . 89
5.6.6 OD Event Cut . . . 90
5.7 Spallation Cut . . . 91
5.8 Ambient Background Cuts . . . 91
5.8.1 Event Quality Cut . . . 91
5.8.2 Hit Pattern Cut . . . 92
5.8.3 Tight Fiducial Volume Cut . . . 93
5.9 External Event cut . . . 94
5.10 Small Hit Cluster Removal . . . 95
5.11 ^{16}N from Muon Capture Cut . . . 96
5.12 Convection Period Cut . . . 97

6 Spallation Analysis **99**
6.1 Neutrons in Hadronic Showers . . . 99
6.2 Identifying Neutron Captures After Muons . . . 101
6.3 Search for Spallation Near Neutron Cloud Centers . . . 105
6.4 Neutron Cloud Cut Tuning . . . 112
6.5 Multiple Spallation . . . 117
6.6 Spallation Removal With Only Preceding Muon Information . . . 120
6.7 Muboy vs mufitpe . . . 121
6.8 Time Correlation . . . 122
6.9 Transverse Distance Correlation . . . 123
6.10 Residual Charge . . . 125
6.11 dE/dx . . . 127
6.12 Deadtime Estimation Method and Dead Volume Function . . . 135

7 Solar Neutrino Analysis in Super Kamiokande **137**
7.1 Signal Extraction . . . 137
7.2 Signal Events . . . 138
7.2.1 Extended Likelihood Method . . . 141
7.2.2 Simulation of Events . . . 142
7.3 Systematic Uncertainties . . . 142
7.3.1 Energy Scale . . . 143
7.3.2 Signal Extraction . . . 143
7.3.3 Spallation Systematic Uncertainty . . . 144
7.4 ^{8}B Flux Measurement . . . 146
7.5 Day/Night Asymmetry . . . 147
7.6 Energy Spectrum . . . 148
7.7 Solar Upturn . . . 149

8 Comparison of Results **154**

9 Conclusion and Outlook **159**

Bibliography **162**

LIST OF FIGURES

Page

1.1 Top: Outline of the *pp* chain reaction in the sun and the relative ratios for each process. Bottom: Outline of the CNO cycle with relative fractions for path. 8

2.1 Super Kamiokande's location within Japan 16
2.2 Diagram of the SK detector 17
2.3 Schematic view of Rn free air system. In March 2013, the cooling system of the last charcoal column was upgraded to go down to -60℃, achieving near 100% removal efficiency 24
2.4 Schematic and Picture of the QBEE board 29
2.5 Schematic for the SK-IV DAQ. 30

3.1 Left: Schematic view of the PMT response to an incident photon in the simulation. Right: Type of PMT response for 420 nm light as a function of incident angle. The black area represents photons absorbed into the PMT surface with no production of p.e. 43
3.2 Typical single p.e. charge distribution from Ni source calibration. Left (right) show the distributions in linear (logarithmic) scale. The colors in the left diagram represent normal PMT gain (black), double gain and half threshold (green), and the linear extrapolation (red). 45
3.3 The relative gain of ID PMTs obtained from laser calibration 46
3.4 Typical hit timing distribution of the Ni calibration. The on-time (off-time) window is set to 750-1050 ns (1200-1500 ns) 47
3.5 Incident angle dependence of the PMT acceptance. The red (black) curves represent the acceptance with (without) the acrylic cases used in SK II-IV. The greater discrepancy at high incident angles is partially a result of this. . 48
3.6 The position dependence of the normalized hit probability. The red (blue) points show the result from the calibration data (MC simulation). The upper panel shows barrel PMTs with teh horizontal axis as the z position of the PMTs. The lower left image is of the Top PMTs and the lower right is of the bottom PMTs, where the horizontal axis is now the r^2 from the detector central axis. 49

3.7 Visual representation of the time-walk effect. The red (blue) curve correspond to large (small) charge signal seen by a PMT, and the dashed line is the TDC discriminator threshold. The pulse height exceeds threshold faster for large charge signals compared to small charge signals. 50
3.8 Schematic view of the laser injection system for the WT parameter measurement (left). Typical hit rate distributions for the six regions (right). The horizontal axis of the hit rate distributions is the t-t.o.f. subtraction for PMT hit time. The left broad peak corresponds to light arriving from scattering, and the right peak corresponds to light from scattering. 54
3.9 The left are the typical observed coefficients of water parameters and fitted functions shown as dashed lines. The horizontal axis is the wavelength and the vertical axis are the coefficient values. The right is the time variation of the coefficients in the early dates of SK-IV as measured by the laser injection system. The horizontal axis shows is the date, with the symmetric scattering in purple, asymmetric scattering in black, and absorption shown in blue. . . 55
3.10 Left: An illustration of the various parameters used in the decay-e WT calculation. Right: A typical Δ^i_{eff} distribution plotted in d from the decay-e for a day. The red line is a linear fit to the distribution from 1200cm $\leq d <$ 3500 cm for $\ln(\Delta^i_{eff})$, which is used to obtain the WT value. 57
3.11 Time variation of the number of effective hit PMTs, N_{eff}, as calculated with constant (varying) WT values shown in black (red). 58
3.12 The z-position dependence for the LINAC (left) and DT (right) calibrations that determine the absolute energy scale of SK-IV. Both plots compare the data and MC information without the TBA correction applied. The LINAC shows z-dependence for two calibration holes and energy, and the DT data shows the z-dependence at all primary DT positions. 59
3.13 The time variation of TBA for Xe lamp (blue) and Ni source (red) data during the first half of SK-IV data taking. 59
3.14 Schematic of the LINAC geometry with respect to the detector. Data taking position are the numbered shapes, the dashed line shows the fiducial volume boundary, and labels for major beam tuning components. 61
3.15 Left: The difference in mean N_{eff} between the DT data and MC simulation for SK-IV, binned in zenith angle. The combined uncertainty due to position dependencies correlated to the zenith direction is $\sim \pm 0.1\%$ after subtracting the absolute offset. Right: The stability of the SK-IV energy scale using DT calibration data after gain correction. The data is from the center most calibration port corresponding to the z-axis in detector coordinates. The N_{eff} is stable within $\pm 0.5\%$ 64

5.1 Scatterplot of spallation products expected to occur in SK, showing the half-life and energy of each isotope. Note that the energy values are the maximum values, and the vertical axis is in a logarithmic scale[1] 79

5.2 The energy spectrum following each reduction step during the 34-hit SLE trigger data period of SK-IV. The final solar neutrino sample for SK-IV (Red) is contrasted with that of SK-III (Black). The pop-out of the low energy region shown right shows the effect of differing pre-reduction criteria for the two trigger thresholds (34 and 31-hit) of SK-IV 82
5.3 Defintition for d_{eff}, p_{wall}, f_{wall}, θ_{PMT}. d_{eff} is the distance from the reconstructed vertex to the ID wall along the inverted reconstructed direction vector, and d_{wall} is the point on the wall where this tracks back to. f_{wall} is the distance from the reconstructed vertex to the ID wall along the reconstructed direction vector. θ_{PMT} is the angle between the reconstructed direction vector and each PMT with respect to the reconstructed vertex. 86
5.4 Correlation of g_V^2 and g_V^2 for data (left) and solar neutrino MC (right). The black solid (dashed) lines corresponding to $g_R^2 = 0.10$ (0.25) 86
5.5 Comparison of a run with (top) and without (bottom) a significantly flashing PMT. The red box outlines the cut criteria based on maximum charge of a hit PMT (x-axis) and the number surrounding 24 PMTs with hits (y-axis). The uper right corner are events that are removed 89
5.6 Typical distribution of the time difference to previous low energy triggered event. Red dashed line represents the 50 μs cutoff, with events to the left cut. The 1 μs peak is due to PMT ringing, and the 15 μs peak is caused by PMT after pulsing. 90
5.7 The distribution of $\mathscr{L}_{pattern}$ for data (black) and solar neutrino MC (red), with the corresponding energy regions above each of the plots. 93
5.8 r^2 vs z vertex distributions for the three energy bins from the solar neutrino analysis affected by the TFV cut. The TFV cut removes events to the right of the black line in the left and middle vertex distribution, and below the black line in the right vertex distribution. The bin color and scale corresponds to events/day/bin 94
5.9 Relationship between N_{20rawT}/N_{eff} and $r02$ for background (left) and solar neutrino MC (right) for events in the edge region. The dashed line corresponds to the cut of $r02 \times (N_{20rawT}/N_{eff}) < 75.0$ 96
5.10 Event rate increase in lowest energy bins due to convection and other changing detector conditions. The events below 6 MeV have increased event rate during this time, and therefore are removed from the final sample. 98

6.1 Simple Cartoon illustrating common terms in this section and chapter. A muon track is shown in black, with candidate events (spallation or neutrons) are shown relative to the track. The transverse distance is shown in blue, while the distance along the track, x, is shown in green. Commonly ln will be used, which is a taken as $\ln = x_i - x_{ref}$, where x_{ref} is some reference point on the track 102

6.2 Comparison of events caused by after pulsing as seen in WIT. The sub-5μs peaks are believed to be from decay electrons and PMT ringing. The left figure is the Δt correlation without any other cuts performed besides the intial 5m lt requirement. The right image is after finalized event quality and energy cut are applied. Since these events are immediately after muons, most PMTs are hit and afterpulsing is somewhat random. Reconstruction requirements, but since they are not a real events the reconstruction quality is poor 103
6.3 The lt^2 correlation of neutron candidate events to the muon track, for different values of g_{vertex} (Sec 4.3.1). Blue is $g_{vertex} > 0.4$, red is $g_{vertex} > 0.5$, and green represents $0.4 \geq g_{vertex} < 0.5$. The purple lines from left to right correspond to lt$^2 = (1m)^2, (2m)^2, (3m)^2, (4m)^2$. Comparing the two goodness cuts showed that there was a distinct improvement to the signal to noise level of the sample. The green distribution illustrates this showing over half of the background events were removed while only losing $\sim 25\%$ of the signal events. 104
6.4 Vertex correlation distributions for neutrons after muons, applying the final vertex quality cuts. The left plot shows the ln only distribution, for multiplicities of 2 (blue), 3 (red), 4 (purple), 5 (cyan), and 6 (orange) neutrons. As multiplicity increases, width of distributions increases, as well as the distribution becoming more skewed to the right. The right plot shows the correlation of lt^2 and ln for the neutron clouds for all multiplicities. The distribution is strongly peaked in lt^2 and ln near 0, but the showers can be several meters long and away from the track. 105
6.5 Δt distribution for a $g_{vertex} < 0.5$ and tight lt (lt $<$ 1.5m) cut. The distribution was fit from 50μs to 500 μs to with an exponential plus constant. The resulting fit gives a decay constant of $\tau = 210.3 \pm 1.64$, slightly above AmBe source calibration data. 106
6.6 Energy distribution for all neutrons seen in WIT after final event reconstruction quality cuts applied with short Δt removed. The red line corresponds to energy cut value decided by MC. The apparent peak of the energy spectrum is from the product of a rapidly changing efficiency and the tail of the 2.2 MeV gamma spectrum. The true peak is expected to be $\sim$1.9 MeV. The extending tail is possibly from the early decay of spallation within 500 μs of the muon. 107
6.7 Neutron multiplicity distributions for varying cuts. Black represents raw multiplicity with no cuts applied and all others have the short Δt removed (Δt$>$20μs). Blue dashed corresponds to the $g_{vertex} > 0.5$ requirement in the initial study. Solid blue (red) correspond to loose (tight) cut based on g_{vertex} and g_{dir} 108

6.8 Vertex correlation distributions for spallation candidates to neutron cloud center. The left plot shows the total spallation signal region without background subtraction. Each red vertical line from left to right corresponds to a 2m, 3m, and 4m distance to center respectively. The tight correlation between spallation events to neutron cloud center can be seen. The right figure shows the dependence of the R^3 on neutron multiplicity. This is background subtracted and normalized. With a greater number of neutrons, the neutron cloud location is better known, and the likelihood that correlated spallation will be located nearby 109
6.9 The background subtracted spallation totals per multiplicity bin. 110
6.10 The background subtracted Δt distribution for spallation candidates found within 5m of a neutron cloud containing at least two events, and one event meeting quality event criteria. The red line is an exponential fit from 8s to 45s, corresponding to the ^{16}N dominated time region. The fit comes to roughly 1σ higher than the true decay constant for ^{16}N 111
6.11 Initial spallation sample energy distribution, with background subtraction. Peak of distribution is around 7MeV, which is consistent with expectations for the energy of most common spallation isotopes. 112
6.12 Solar peak distribution for the WIT period. The black line corresponds to the total solar neutrino events in the time period. The yellow (green) line corresponds to the former spallation cuts removed (remaining) events. The blue (red) line corresponds to removed (remaining) events from the neutron cloud identification. 113
6.13 Comparison of "bad" (left) event reconstruction and "good" (middle) reconstruction g_{vertex} vs g_{dir} distributions. 2.2 MeV gammas were simulated in WIT, with events with a reconstructed vertex more than 5m from truth considered bad, and those within 5m good. The yellow peak of the good distribution is significantly shifted to the lower right, and weighted regions take this into consideration. The right distribution shows the distribution for the entire WIT data period. 114
6.14 Figure illustrating the two main potential coordinate systems. Shown is a mock muon with neutrons (blue) and their center of mass position (red). Both use the muon direction as the +z direction, and the left one has the z-axis pass through the neutron cloud center of mass and the right one maintains the original muon track 115
6.15 Figures showing the six different neutron multiplicity categories and the corresponding lt^2 vs ln distributions for neutron cloud centered (top) and muon centered (bottom). Cuts based on the muon centered coordinate system have better spallation tracking, allowing for smaller volume cuts to be made, especially for low tagged-neutron multiplicity showers 116
6.16 Cumulative Δt distributions for signal (left) and background (right) assuming a 5m bubble cut for any cloud. Much of the spallation is removed in the first couple seconds, and extending the bubble cut for low multiplicity and small Δt adds minimal deadtime to the overall amount. 118

6.17 Comparison between events cut (left) from the final sample to those that remain after the cut (right). No evident peak in the left distribution shows the relatively small deadtime created by this cut. 119
6.18 Comparison of ability for Muboy (left) and mufitpe (right) to fit very bright muons, as measured by spallation transverse distance to the fitted track. Muboy successfully reconstructs more high energy muon, and the spallation tracking is significantly better than mufitpe. The PDF for lt in the former likelihood method was disregarded (set to 1) due to the poor tracking of mufitpe in this energy range 122
6.19 Signal (left) and background (right) distributions for Δt, with only a 3m lt cut applied. Both distributions are binned so the width of each bin is flat in log(Δt). Signal region is fit with the seven principal isotopes, with their decays constants fixed, and each isotopes contribution to the total fit allowed to vary. Background is expected to be flat, and fit with a linear equation to test this. The fit is 1σ consistent with flat, and is most likely a result of the binning and poor statistics in the low Δt regime. 123
6.20 Signal (a) and background (b) distributions for the six highest residual charge bins, and middle Δt (0.1 s$\leq \Delta$t $<$ 3 s. PDFs are overlaid on the distribution to show the fits used in the final likelihood. 126
6.21 Signal (Blue) and Background (Red) PDFs for Q_{res}. 127
6.22 This is the signal (blue) and background (red) PDF distributions for the ln parameter to the dE/dx peak. The bias in the BG peak reveals a preference of the dE/dx method to find a "peak" near the track end, if the muon did not shower. 131
6.23 Signal (left) and background (right) distributions for ln vs excess charge for a dE/dx bin. The plots are normalized in each ln bin to account for differing track lengths. The background is relatively flat, with edge effects at very large ln, while the signal distribution is strongly peaked near 0 ln. Fits were made for excess charge in slices of ln, and then fits on the fit parameters to have a 2D PDF for the signal and background distributions. 132
6.24 ROC curve comparing the former spallation log likelihood cut (blue) to the new spallation log likelihood (red) with updates to lt, Δt, Q_{res}, and ln. Choosing the desired efficiency (x-axis), the difference in livetime (y-axis) is seen. Although the curves are close, this results in a couple % difference in overall deadtime at the desired cut value. The desired value is near 20%, and depends on whether neutron cloud data is available. 133
6.25 Left shows the fraction of removed events for the non-WIT period spallation cut based on r^2 and z. The right is a two-dimensional fit to the left distribution, and used to correct the counting of MC events. 136

7.1 Feyman diagrams of W/Z boson exchange in ν-e elastic scattering 138
7.2 Differential cross section for an incident neutrino of 11 MeV shown left, and total cross section as a function of neutrino kinetic energy shown right. . . . 140

7.3 Normalized neutrino energy spectrum deduced from Winter (top). Dashed lines represent the ratio between the neutrino spectrum (bottom), and the $\pm 1\sigma$ experimental uncertainties, to the spectrum deduced by Winter[2] . . . 140
7.4 Angular distribution of recoil electrons from incident solar neutrinos. Horizontal axis shows teh angle of recoil electrons from incident solar neutrinos, and the vertical axis shows event rate in the fiducial volume. The red, green, blue, and purple lines correspond to neutrinos above 3.5, 4.5, 6.5, and 9.5 MeV respectively. 141
7.5 Expected ^{8}B energy spectrum using MC simulation after cuts are applied. The ^{8}B flux from the SNO result [3] and Winter spectrum [2] are used. Tight Fiducial Volume Cut lowers event rate for energies below 5.5 MeV (5.0 Mev Kinetic above) 143
7.6 Dot product of all of the muon track wiggles to the original muon fit. Y-axis is logarithm of the counts/bin. Most muons have relatively the same fit, with some differences depending on the case. The peak at 0 corresponds to a successful wiggle dotted with a failed old fit. The peak is amplified by a factor of 25, since it is filled for each wiggle. There is an excess peak at -1.0 and -0.9 corresponding to failed new fit and without (with) a successful old fit respectively. These are filled only once. Spikes located at approximately -0.7, 0.45, and 0.7 are from a "preferred" set of fits Muboy has a tendency to fall to when fitting for both the wiggled and former track fits. These spikes are a result of dotting these fits and corresponding wiggles together. 145
7.7 Zenith angle distribution for the ^{8}B solar neutrino flux measurement for the 4.0 to 20.0 MeV energy range. Horizontal axis has ALL, Day, and Night, then the five zenith day bins and six zenith night bins. solar neutrinos passing through the Earths core is in the rightmost bin. Error shown is statistical. . 148
7.8 $\cos\theta_{sun}$ distributions for each energy bin, with the total energy range (4-20) shown in the uppermost left plot, with the rest showing 0.5 MeV bins from 4 MeV to 9.5 MeV (left to right, top to bottom). Flat blue lines show background contribution and black lines are signal and background. 150
7.9 Highest energy $\cos\theta_{sun}$ distributions (>9.5 MeV). Flat lines show background shape (flatter) and signal (following data) with $\pm 1\sigma$ fits. 151
7.10 Solar upturn using only SK-IV data, with quadratic (light blue) and exponential (red) fits with 1σ band. Purple denotes the overlap of the two regions. Using the KAMLAND (dark blue), Solar best fit with 1670 days of SK-IV Day/Night data (green), and SK+SNO (robin egg blue). 152
7.11 Contour plots showing 1-5σ regions for the solar oscillation parameter fits for SK-IV only data (Green) compared to the 1-3σ KamLAND best fit regions (blue). The fits on the left (right) are performed without (with) taking into consideration the effect of the Day/Night Asymmetry. The shaded regions for both are the 3σ allowed region. 153

8.1 Lowest energy bin (left) compared to spallation region energy bin (right, 10.5-11.0 MeV). Near match in background is seen for the higher energy bin, but large separation in the lowest energy bin from increasing background from non-spallation sources as a result of deadtime reduction 156
8.2 Top: $\cos(\theta_{sun})$ final solar sample the new spallation cuts (red) overlaid with old (blue). Bottom: Difference between the old and new spallation cuts, taken as the sample with the new spallation cuts minus the sample with the old spallation cut. Lower energies have non-spallation dominant backgrounds and reducing the deadtime increases the background in these energies. Difference in background shape fits are shown as the blue line and difference in signal fits shown in black 157
8.3 Comparison of the $\cos(\theta_z)$ distributions for the new (red) and old spallation cut (black). The change in number of events increase the significance of the asymmetry. 158
8.4 Comparison of the energy spectrum for the final solar sample using the new (red) spallation cut against the old spallation cut (black). 158

LIST OF TABLES

Page

3.1 Cherenkov Radiation threshold for typical charged particles 42
3.2 Absolute gain for each SK phase 45

5.1 List of possible muon-induced spallation products that can occur in SK. The rates for ^{8}Li/^{8}B cannot be decoupled since their lifetimes and energy are too similar. This table is transcribed from [4]. Some results were consistent with 0, so a 90% confidence limit was placed on the yield (e.g. ^{8}He) 80
5.2 The different trigger thresholds and corresponding trigger rates and event timing windows 82
5.3 Example reduction efficiencies for each step of the pre-reduction process. The overall reduction is a factor of ∼100 86

6.1 Neutron Temperature ranges 100
6.2 Table showing the different cut conditions for the cloud cut. + is used to represent showers of at least that multiplicity. 117
6.3 Example energy dependent spallation log likelihood cut values for non-WIT period and choosing signal region to be $\cos\theta_{sun} > 0.8$. Looser cuts at very low and high energies result of spallation not being the main background (low) and less overall background (high). These do not include the ln likelihood. . 134

7.1 Cross section coefficients for ν-e elastic scattering. Broken into the type of interaction occuring 139
7.2 Percentage fraction of cases for comparison of muon fits for old fit vs new . . 145

8.1 Table showing increase in signal events and reduction in relative error for each energy bin of the solar neutrino analysis. 155

ACKNOWLEDGMENTS

I would like to thank all mentioned within the dedication. Without their support this journey would've been much harder and less rewarding. I would like to thank Dr. Henry Sobel and Dr. Michael Smy for taking me as a student, as well as for their mentoring, support, and understanding. I do not think I could've had a better pair of people to work for. I would like to thank Pierce Weatherly for helping me out within the group at UCI and as someone to bounce ideas off of. I would like to thank Dr. John Beacom and Dr. Shirley Li for their help on the theoretical side of my reasearch. I would like to thank Dr. Yukki Nakano, Dr. Yasuo Takeuchi, Dr. Lluis Marti Magro, Dr. Motoyasu Ikeda, Dr. Yasuhiro Nakajima, Dr. Kai Martens, Dr. Shunichi Mine, and all of the other Super-Kamiokande collaborators I have worked with over the years. I appreciate the leadership of Professor Masayuki Nakahata. I would like to thank Jeff Griskevich for all of his help with the UCI water system and for his support during the many trips to Japan. Thank you to all of my fellow graduate students, especially the many study groups in the early years of classes. Thank you to Professor Steven Barwick and Professor Mu-Chun Chen for taking time to be on my PhD Committee. Thank you to the US Department of Energy for their financial assistance through the group grant.

VITA

Scott Locke

EDUCATION

Doctor of Philosophy in Physics — **2020**
University of California Irvine — *Irvine, CA*

Bachelor of Science in Physics — **2012**
University of Washington — *Seattle, WA*

Minor in Mathematics — **2012**
University of Washington — *Seattle, WA*

RESEARCH EXPERIENCE

Graduate Research Assistant — **2015–2020**
University of California, Irvine — *Irvine, California*

Voluntary Laboratory Assistant — **2013–2014**
University of Washington, Rabinovitch Lab — *Seattle, WA*

TEACHING EXPERIENCE

Lab Teaching Assistant — **2014–2015**
University of California Irvine — *Irvine, CA*

Teaching Assistant — **2014–2015**
University of California Irvine — *Irvine, CA*

Volunteer Teaching Assistant — **2011–2012**
Garfield High School — *Seattle, WA*

REFEREED JOURNAL PUBLICATIONS

Indirect Search for Dark Matter from the Galactic Center and Halo with the Super-Kamiokande Detector **2020**
PhysRevD 102 072002

Search for proton decay into three charged leptons in 0.37 megaton-years exposure of the Super-Kamiokande **2020**
Phys.Rev.D 101 5, 052011

Search for Astronomical Neutrinos from Blazar TXS 0506+056 in Super-Kamiokande **2019**
Astrophys.J.Lett. 887 1, L6

Evaluation of gadolinium's action on water Cherenkov detector systems with EGADS **2019**
Nucl.Instrum.Meth.A 959 163549

Sensitivity of Super-Kamiokande with Gadolinium to Low Energy Anti-neutrinos from Pre-supernova Emission **2019**
Astrophys.J. 885 133

First Measurement of the Total Neutron Cross Section on Argon Between 100 and 800 MeV **2019**
PTEP 2019 (2019) 5, 053F01, Prog Theor Exp Phys

Atmospheric Neutrino Oscillation Analysis with Improved Event Reconstruction in Super-Kamiokande IV **2019**
Phys.Rev.D 99 3, 032005

Measurement of the tau neutrino cross section in atmospheric neutrino oscillations with Super-Kamiokande **2018**
Phys.Rev.D 98 5, 052006

Search for Boosted Dark Matter Interacting With Electrons in Super-Kamiokande **2018**
Phys.Rev.Lett. 120 22, 221301

Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV **2018**
Phys.Rev.D 97 7, 072001

Search for an excess of events in the Super-Kamiokande detector in the directions of the astrophysical neutrinos reported by the IceCube Collaboration **2017**
Astrophys.J. 850 2, 166

Search for nucleon decay into charged antilepton plus meson in 0.316 megaton·years exposure of the Super-Kamiokande water Cherenkov detector **2017**
Phys.Rev.D 96 1, 012003

REFEREED CONFERENCE PUBLICATIONS

Spallation Studies in Super-Kamiokande **2020**
Neutrino 2020

Spallation Studies in Super-Kamiokande **2019**
Topics in Astroparticle and Underground Physics 2019

First Measurement of the Neutron-Argon Cross Section Between 100 and 800 MeV **2019**
Topics in Astroparticle and Underground Physics 2019

Showering Muons in Super-Kamiokande **2017**
TeV Particle Astrophysics 2017

Solar and Supernova Neutrinos in Super-Kamiokande **2017**
American Physics Society Division of Particles & Fields

ABSTRACT OF THE DISSERTATION

New Methods to Reduce Cosmogenic Backgrounds of Super-Kamiokande in the Solar Neutrino Energy Regime

By

Scott Locke

Doctor of Philosophy in Physics

University of California, Irvine, 2020

Professor Emeritus Henry Sobel, Chair

Solar neutrinos pass through the Earth's surface with a flux on the order of $10^{11}/\text{cm}^2/s$. The Super-Kamiokande (SK) searches for neutrinos from the sun with an energy above 4.0 MeV. Neutrinos rarely interact with matter, and removing backgrounds that would otherwise overwhelm the signal is imperative. SK observes about 2 muons a second at its depth of 2700m water equivalent, and a fraction of these muons shower within the detector with a possibility to create radioactive isotopes (spallation) which live from milliseconds to tens of seconds. Radioactive isotopes from any source are by far the largest source of background in the solar neutrino energy region observed in SK, and spallation is the dominant background to neutrino interactions between 6 and 20 MeV. Spallation in SK is mostly caused by neutrons and pions interacting with the Oxygen nucleus. Recently, new techniques involving identifying neutron in the showers after muons and effective tagging of multiple spallation, along with improvements and additions to former spallation tagging methods have reduced deadtime by ∼55% (∼45%) where neutron data is available (unavailable). This increases measured solar neutrino signal by 12% and reduces relative statistical error by 6.6% for the entire SK-IV solar neutrino sample.

Chapter 1

Introduction

Early Particle Discoveries

The late 1800s and early 1900s led to an explosion of new discoveries particles and ideas in the world of quantum physics, from discoveries of the electron, β radiation, α radiation, X-rays, γ-rays, and even the concept of the atomic nucleus. Each new finding led to new questions, and trying to answer these questions and characterize different forms of radiation became one of the big efforts in the early 20th century. This is what led to the birth of particle physics. Particle physics is the branch of physics that deals with the properties, relationships, and interactions of subatomic particles. During these discoveries, one curious observation was made by James Chadwick in 1914 when he asserted the β-decay spectrum was continuous, contrary to α and γ decays which have a narrow peaks [5]. Later, Charles Ellis and William Wooster looked at the continuous β-decay spectrum of ^{210}Bi where each reaction is 0.35 MeV total. Lise Meitner and Wilhelm Orthmann verified the continuous spectrum using an improved method, and as the process was understood at the time, this meant the law of conservation of energy was violated.

This was remedied on paper by Wolfgang Pauli in 1930 in a letter to Meitner et al, where he postulated a "neutron" with little to no mass was also emitted in the process, with the sum of the β and "neutron" energies remaining constant [6]. Although not mentioned in the letter, Pauli also knew this may also resolve the issue of certain spin statistics anomalies. The discovery of what is known today as the neutron in 1932 was not the solution of the above postulate, due to its mass being close to the mass of the proton. Enrico Fermi later included the idea of Pauli's "neutron" in his 1934 paper outlining a successful model for β-decay, and changed the name to neutrino, or "small and neutral"[7]. This new particle would have spin ½, be electrically neutral, and have a mass similar to that of an electron (or no more than 1% the mass of a proton) allowing for the conservation of energy, momentum, and angular momentum from β-decay. The neutron, discovered by James Chadwick in 1932, would decay into a proton, β-ray, and a neutrino (later changed to anti-neutrino for this specific case).

Pauli, joking to a friend had said, "I have done a terrible thing. I have postulated a particle that cannot be detected"[8]. This statement would be partially confirmed by Hans Bethe and Rudolf Peierls in 1934, where they published a paper on the neutrino with a calculated cross section of roughly 10^{-44}cm^2[9]. In simpler terms, a neutrino could travel a vast distance without ever interacting with matter, and most would pass through the earth unimpeded. The pair concluded the neutrino would not be able to be seen by anyone. This statement was made without the knowledge of nuclear fission though, which would prove to be key to the initial discovery of the neutrino. The large flux of (anti-)neutrinos from a nuclear reactor would provide a flux large enough to overcome this small cross section.

1.1 Neutrino Detection

An early attempt at neutrino detection was made by Ray Davis in 1954. Trying to utilize the process suggested by Bruno Pontecorvo of[10]:

$$^{37}\mathrm{Cl} + \nu \rightarrow \ ^{37}\mathrm{Ar} + e^{-} \tag{1.1}$$

^{37}Ar is radioactive and decays back to ^{37}Cl with a half life of 35 days and the ^{37}Ar could be chemically separated out from the ^{37}Cl and counted with a gas-filled proportional counter. Davis would place initially a 1000-gallon tank of carbon tetrachloride at the Brookhaven Graphite Research Reactor[11]. However, this experiment failed to detect anything above background for two reasons. First, the reactor was too small for the size of target Davis was using, as well as the nuclear reaction creating anti-neutrinos instead of neutrinos, a difference not understood at the time of the experiment.

The next major attempt at seeing the elusive neutrino would be performed by Frederick Reines and Clyde Cowen. The two first tried to detect neutrinos in the Cowan-Reines neutrino experiment at the Hanford nuclear reactor in Washington in 1953[12]. They used 200 L of water doped with cadmium chloride ($CdCl_2$) sandwiched between two tanks of liquid scintillator and photomultiplier tubes (PMTs). The idea behind the experiment relied on the following reaction from Fermi's theory:

$$\overline{\nu}_e + p \rightarrow n + e^{+} \tag{1.2}$$

The two products from this interaction could be seen together in coincidence, creating a "double bang" signature for the process. The e^+ would annihilate on an electron, creating a pair of 511 keV γ-rays, and the neutron would capture on Cadmium (Cd), with a 9 MeV γ cascade. These two in coincidence would hopefully be a tell-tale signature for the process.

Unfortunately, at Hanford the cosmic rays passing through created too much of a background. In 1955, the experiment was relocated to the newly constructed Savannah River reactor in South Carolina. Here they were able to place their detector 11m away from the reactor with 12m of overburden [12].

The reactor had a regular on/off cycle for the production of enriched plutonium, and this allowed for the comparison of the two phases for whether or not any signal seen was from the reactor. This proved to be enough to show a difference between event rate with the reactor on and off, and $\sim$3 neutrinos/hour were seen. This was published in the July 20, 1956 issue of Science [13]. They had also found their measured cross section to be in close agreement with the predictions at the time.

Around the same time in 1953, E. J. Konopinski and H. M. Mahmoud published a paper applying the β-decay process to the as μ well. They concluded "that two like neutrinos are ejected in μ decay" is the correct process for the decay[14]. Specifically the two neutrinos output would be distinct from each other as follows:

$$\mu^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\mu \text{ and } \mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu \tag{1.3}$$

It was with this that the eventual framework for the lepton number and its conservation was established. Although this symmetry can be violated by chiral anomalies.

In 1962 Lederman, Schwartz, and Steinberger were able to verify the distinction between the two types of neutrinos by creating a pure neutrino beam[15]. The technique was developed by Schwartz, and the experiment was carried out at Brookhaven National Lab. Using the Alternate Gradient Synchrotron, energetic protons were impinged upon a beryllium target, creating a hadronic shower of particles. In the shower, π^+ would decay as:

$$\pi+ \rightarrow \mu^+ + \nu_\mu \tag{1.4}$$

on the way to a 5,000-ton steel wall made from harvested battleships. Only the ν_μ would pass completely through the steel and would be detected on the other side in a ten-ton spark chamber. The spark chamber was filled with neon gas and aluminum plates. The neutrinos would capture on the protons within the aluminum plates, and the out going charged leptons would be seen in the gas. Differentiation between ν_μ and ν_e could be made by comparing the track lengths of events if they were long enough. Muons typically would wholly pass through the detector as they are minimum ionizing particles, while the electrons would be contained due to stronger interactions within the chamber. This verified the neutrinos produced from the π-decay were different from those produced in β-decay. This work led to Ledermen, Schwartz, and Sternberger earning the Nobel Prize in 1988. The method of using accelerated protons impinging on a target is still used to this day to produce a pure neutrino beam.

1.2 Types of Neutrinos

Neutrinos are produced from many processes and sources can be found throughout the universe in space and time. These range from the origin of the universe, nuclear processes, showers, supernovae, and more [16]. With current detector technology, most observed neutrino interactions are from the sun, reactors, atmospheric decays, and beams, with those from the sun having the greatest flux for the whole Earth's surface. Most reactor experiments are in relatively close proximity to a nuclear reactor, allowing for much greater flux. Neutrinos from the sun (solar neutrinos) are also a gateway to understanding the different nuclear processes occurring within the sun.

The sun can be simplified to a gigantic nuclear fusion reactor. Within the core of the sun, most of the fusion process is carried out through the *pp* chain and the ultimate

fusion of four protons into a helium atom.

$$\begin{aligned} p + p &\rightarrow \mathrm{D} + e^+ + \nu_e \\ \mathrm{D} + p &\rightarrow {}^3\mathrm{He} + \gamma \\ {}^3\mathrm{He} + {}^3\mathrm{He} &\rightarrow {}^4\mathrm{He} + 2p \end{aligned} \tag{1.5}$$

This is the main engine process within the sun, with 98.4% of the solar luminosity produced in this manner. The latter parts of the chain introduce the start of heavier elements. With both ^{3}He and ^{4}He in the mix, some heavier elements are able to form, as follows:

$${}^3\mathrm{He} + {}^4\mathrm{He} \rightarrow {}^7\mathrm{Be} + \gamma \tag{1.6}$$

With the ^{7}Be possibly following two channels, the first being

$$\begin{aligned} {}^7\mathrm{Be} + e^+ &\rightarrow {}^7\mathrm{Li} + \nu_e \\ {}^7\mathrm{Li} + p &\rightarrow 2\ {}^4\mathrm{He} \end{aligned} \tag{1.7}$$

or

$$\begin{aligned} {}^7\mathrm{Be} + p &\rightarrow {}^8\mathrm{B} + \gamma \\ {}^8\mathrm{B} &\rightarrow {}^8\mathrm{Be}^* + e^+ + \nu_e \\ {}^8\mathrm{Be}^* &\rightarrow 2\ {}^4\mathrm{He} \end{aligned} \tag{1.8}$$

There is a smaller fusion chain of Carbon-Nitrogen-Oxygen (CNO) through which fusion and decay processes cycle between different isotopes of the three elements. This accounts for the remaining luminosity of the sun. To get the carbon necessary for this reaction, a rare fusion in the case of the sun of three helium atoms into an excited resonant state of ^{12}C, and an uncommon (0.04%) de-excitation to a stable ^{12}C. Recently, in June 2020, the BOREXINO collaboration published results for the first direct detection of neutrinos produced in this reaction[17]. It should be noted the ratio of *pp* to CNO for a star's luminosity

varies from star to star, and the CNO cycle can be a much bigger, or even dominant fraction of "hotter" stars. The total output for the CNO cycle can be summarized in the same way as the *pp* chain as:

$$4p \rightarrow \ ^4\mathrm{He} + 2e^+ + 2\nu_e \tag{1.9}$$

The CNO cycle and *pp* chain reactions within the sun were first theorized primarily by Hans Bethe in 1938, shortly after attending the Washington Conference of Theoretical physics[18]. Later, solar models in the mid-1960s from John Bahcall[19] and others were formulated in order to use neutrinos as a means to understand processes in the core of the sun. In 1965, an experiment was built by Ray Davis and his team in the Homestake mine in order to try and make measurements on the flux of solar neutrinos. The experiment used the same ^{37}Cl to ^{37}Ar process from his previous attempt to see neutrinos. With an overburden of 4,850 feet (1480 m), enough background from cosmic rays was eliminated to make a measurement. However, Davis only observed around a third of the expected number of neutrinos from the sun. This gave birth to what would be referred to as the "solar neutrino problem." At the time, there was debate between physicists on whether there was fault in the solar models or with Davis's experiment. This would go unresolved for more than 30 years until the Super-Kamiokande (SK)[20] and later the Sudbury Neutrino Observatory (SNO) experiment[21], found evidence of neutrino oscillations.

1.3 Neutrino Oscillations

Most experiments of the time were only measuring one type of neutrino flavor, generally the ν_e flavor, so only part of the whole picture could be seen. SK and SNO were able to prove through atmospheric (SK) and directly from solar neutrinos (SNO) that neutrinos on top of having three lepton flavors, also had mass eigenstates. The SK data showed neutrino disappearance consistent with oscillations for the ν_μ channel but was not able to do so for

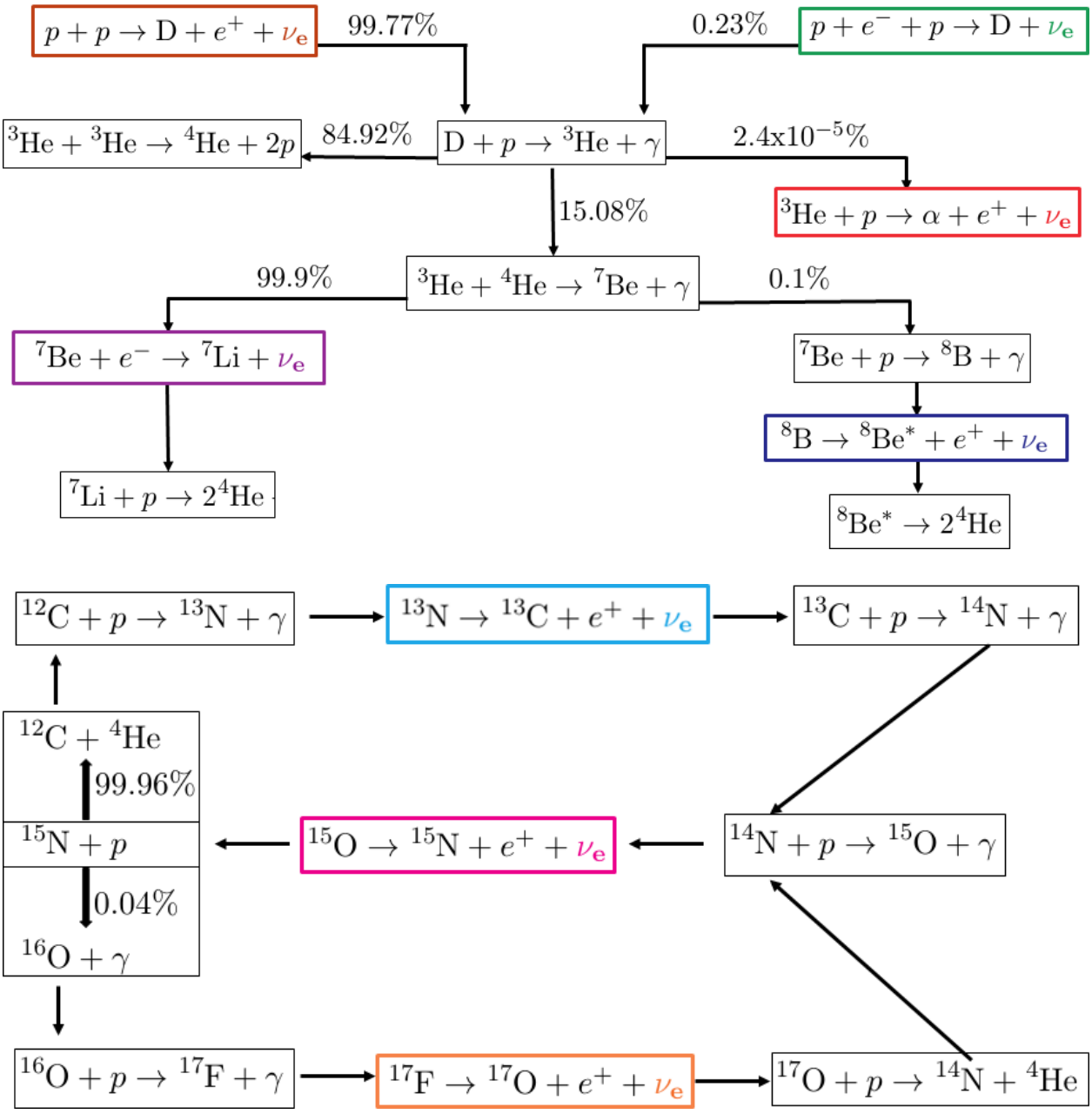


Figure 1.1: Top: Outline of the *pp* chain reaction in the sun and the relative ratios for each process. Bottom: Outline of the CNO cycle with relative fractions for path.

the ν_e channel. SNO's first results showed the ν_e charged current (CC) rate was inconsistent with the SK elastic scattering measurement. A year later a full measurement for the three CC channels was released, showing the full picture for solar neutrino oscillations. The SK atmospheric results alone was able to prove neutrinos having mass, as there was an observed change in the neutrinos. If all neutrinos were massless, then in their frame of reference time wouldn't pass and there for no changes would be observed. Since there were observations showing the properties of the neutrino changing, mass is necessary for time to pass for the neutrino.

The idea for neutrinos having mass eigenstates was first brought forth in 1957 and 1958 by Pontecorvo [22][23], when looking into the possible Majorana nature of the neutrino and exotic mixing, such as muonium to anti-muonium. In 1962 Maki, Nakagawa, and Sakata introduced a mixing matrix in order to explain neutrino oscillations as presented by Pontecorvo[24]. The resulting unitary matrix is now known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, or U_{PMNS}. With this, neutrinos are formed in a definite flavor state (ν_e, ν_μ, ν_τ) which is a superposition of mass eigenstates (ν_1, ν_2, ν_3) based on how the neutrino interacts. The flavor eigenstates are considered weak, and without other interactions the neutrino has a definite mass eigenstate. As the neutrinos propagate through vacuum, the superposition of the flavor eigenstates alters, thus resulting in the neutrino oscillation.

This phenomenon of neutrino oscillation is the reason for the solar neutrino problem observed by Davis at the Homestake experiment. The neutrinos were produced in the ν_e flavor state at the core of the Sun, but as the neutrinos traversed through the sun and on their way to Earth, there was a non-zero probability for the neutrino to change its flavor en route. The deficit in number of neutrinos seen here at Earth from the sun was due to a lack of sensitivity to the other flavor eigenstates that the neutrinos had oscillated into.

In the current understanding, there are three mass and flavor eigenstates as men-

tioned before. U_{PMNS} can be used to transform between the eigenstates in the following manner:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U_{\text{PMNS}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.10}$$

Where the U_{PMNS} can be further factorized into three more matrices which are tied to the types of neutrinos that they affect the most:

$$U_{\text{PMNS}} = \begin{bmatrix} c_{12}\Upsilon_1 & s_{12}\Upsilon_1 & 0 \\ -s_{12}\Upsilon_2 & c_{12}\Upsilon_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{12}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{12}e^{-i\delta_{CP}} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{13} & c_{23} \end{bmatrix} \tag{1.11}$$

With $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$. $\Upsilon_i = e^{i\alpha_i/2}$ corresponds to the the Majorana phase factor, if it exists. In order of the above, the first matrix corresponds to measurements from solar neutrinos, the second from reactor and atmospheric neutrinos, and the third from atmospheric and neutrino beams. While neutrinos propagate, their wave function can be written as:

$$|\nu_i(t)\rangle = e^{-i(E_i t - p_i x)} |\nu_i(0)\rangle \tag{1.12}$$

Which can be simplified in the ultra-relativistic limit (observed neutrinos always are at this limit) to:

$$E_i = \sqrt{p_i^2 + m_i^2} \approx p_i + \frac{m_i^2}{2p_i} \approx E + \frac{m_i^2}{2E} \tag{1.13}$$

Where E_i is the energy of the mass eigenstate, p_i is the three-momentum, m_i is the mass of the neutrino, and E is the total energy of the neutrino. Simplifying a little further with $t \approx L$ and dropping phase factors, our equation of state becomes:

$$|\nu_i(L)\rangle = e^{-i\frac{m_i^2 L}{2E}} |\nu_i(0)\rangle \tag{1.14}$$

It can be seen with this formulation how eigenstates with different masses will oscillate at different frequencies due to the m_i^2 dependence of the wave function, with the heavier ones (larger mass) oscillating at a higher frequency. Using this formulation, it can be calculated the transition probability for a neutrino to oscillate from one flavor to another in the following way:

$$\begin{aligned} P_{\alpha\to\beta} = |\langle\nu_\beta(L)|\nu_\alpha\rangle|^2 &= \left|\sum_i U^*_{\alpha i}U_{\beta i}e^{-i\frac{m_i^2 L}{2E}}\right|^2 \\ &= \delta_{\alpha\beta} - 4\sum_{i>j}\mathrm{Re}(U_\alpha i^* U_{\beta i}U_{\alpha i}U^*_{\beta i})\sin^2(\frac{\Delta m^2_{ij}L}{4E}) \\ &+ 2\sum_{i>j}\mathrm{Im}(U^*_{\alpha i}U_{\beta i}U_{\alpha i}U^*_{\beta i})\sin(\frac{\Delta m^2_{ij}L}{2E}) \end{aligned} \tag{1.15}$$

With $\delta_{\alpha\beta}$ being the standard Kronecker delta, and $\Delta m^2_{ij} = m_i^2 - m_j^2$, otherwise known as the squared mass splitting. If $\delta_{CP} = 0$ and there is no CP-violation, the second sum in the equation is dropped.

Currently, there are successful measurements for two of the three combinations for Δm^2_{ij}. From above, the first matrix, or those values corresponding to Δm^2_{21} and θ_{12} are also referred to as Δm^2_{solar} and θ_{solar}, because of its role in solar neutrino measurements. Δm^2_{32} and θ_{23} is often referred to as Δm^2_{atm} and θ_{atm}. From measurements of solar neutrinos, it has been found $m_2 > m_1$, however only the magnitude of Δm^2_{atm} is known and not the sign of the splitting. This has led to an issue in establishing the complete ordering of the masses, and two possibilities come out:

$$m_1 < m_2 < m_3 \quad or \quad m_3 < m_1 < m_2 \tag{1.16}$$

with the former called normal ordering, and the latter inverted ordering. This hierarchy is important due to its implications on different Grand Unifying Theories and the formation of the universe, and can be determined from neutrino beam experiments, as the placement of m_3 affects the characteristic lengths at which oscillation probabilities are the greatest for

those experiments.

The SNO experiment identified neutrino oscillations by making two simultaneous measurements of the total flux of all ^{8}B and *hep* solar neutrinos. ^{8}B neutrinos are produced by:

$$\begin{aligned} {}^7\mathrm{Be} + p &\rightarrow {}^8\mathrm{B} + \gamma \\ {}^8\mathrm{B} &\rightarrow {}^8\mathrm{Be}^* + e^+ + \nu_e \end{aligned} \tag{1.17}$$

and HEP neutrinos created by:

$${}^3\mathrm{He} + p \rightarrow {}^4\mathrm{He} + e^+ + \nu_e \tag{1.18}$$

SNO used a detector of 1 kton of highly pure heavy water (D_2O, D = ^{2}H) in an acrylic vessel[21]. The vessel was surrounded by an array of PMTs and light collectors to measure the Cherenkov radiation from interactions within the volume. The use of heavy water allowed for measurements of the Neutral Current (NC) interactions of all of the flavors, as well as the Charged Current (CC) interactions of neutrinos with the Deuterium in the detector, and neutrino-electron elastic scattering between neutrinos and atomic electrons. Through the NC channels the total flux of solar neutrinos could be measured independent of the flavor of the neutrinos. This was detected via:

$$\begin{aligned} \mathrm{D} + \nu &\rightarrow p + n + \nu \\ \mathrm{D} + n &\rightarrow {}^3\mathrm{H} + \gamma \ (\sim 6 \text{ MeV}) \end{aligned} \tag{1.19}$$

With the NC channel having no preference to neutrino flavor. The gamma from the neutron capture on Deuterium is then detected in the following electromagnetic shower by the surrounding PMTs. This allows for a measurement of the total flux of solar neutrinos from the sun. Secondly, SNO measured the ν_e flux from the CC interactions within the Deuterium

nucleus as follows:

$$\mathrm{D} + \nu_e \rightarrow p + p + e^- \tag{1.20}$$

And unlike the NC interaction, this is only sensitive to the ν_e interactions. With enough understanding, a direct comparison between the NC and CC rates allows for a direct measurement of the abundance of ν_e compared to the other flavor states.

Another curious observation made by different experiments was an energy dependence on the difference between expected solar neutrino rate and observed. Different experiments, including Homestake, GALLEX, GNO, SAGE, and Kamiokande had found differing values for the discrepancy between seen neutrinos and what was expected from the different solar models. The different experiments looked at different source processes for the neutrinos from the sun, which in turn had them looking at different neutrino energies, as well as a flavor dependence to the expected neutrino rates. Kamiokande specifically had seen $\sim$45% of the expected solar neutrino flux. In 1979, Lincoln Wolfenstein presented work on understanding neutrino oscillations in matter, and in 1985 Stanislav Mikheyev and Alexei Smirnov predicted neutrino mixing can undergo adiabatic conversion with a slow decrease in the density of matter. This is what would be come to known as the Mikheyv-Smirnov-Wolfenstein (MSW) effect[25]. Electrons in matter change the energy levels of the mass eigenstates of the neutrinos due to CC coherent forward scattering of the electron neutrinos through weak interactions. This acts in a similar manner to an index of refraction for the neutrinos passing through matter, but instead of a different effective speed of light, there is a difference in the effective mass of the neutrinos during propagation.

1.4 Neutrons

As mentioned earlier, within a couple years of the introduction of the neutrino as a concept the neutron was discovered by James Chadwick in 1932 [26]. The Rutherford model for the nucleus in 1911[27] had a massive positively charged nucleus surrounded by a cloud of negatively charged particles. This was expanded in 1920 by Rutherford when he suggested the nucleus consisted of positive protons and neutrally charged particles, with the neutral particles suggested to be a proton and electron in some closely bound state[28]. The sentiment for the neutral particles to be a bound state of a proton and electron would also help explain β-decay, as it was known the electrons escaping the atoms during β-decay were from the nucleus and not the orbiting electrons.

In 1930 Walther Bothe and Herbert Becker found using α emission of polonium to impinge on certain light elements resulted in an unusually penetrating radiation[29]. The fact that this radiation was electrically neutral, resulted in the idea of this being a different form of γ radiation. In 1932, it was shown if this new form of radiation fell on any hydrogen containing paraffin it would eject very high energy protons (5 MeV), inconsistent with the idea this was some form of γ-radiation. James Chadwick and Norman Feather used a polonium source on ^{9}Be to eject these unknown particles towards a paraffin wax[26]. The interaction with the wax would eject protons from hydrogen atoms in the target paraffin wax, which would be detected by a small ionization chamber. Chadwick would also test how the new radiation impacted various gases, and ultimately discovered this unknown radiation contained neutral particles of roughly the same mass as the proton. He won the 1935 Nobel Prize in Physics, for the discovery of the neutron.

Soon after, many physics theorists began to form theories including the newly discovered neutron as a fundamental particle. Making the neutron a spin $\frac{1}{2}$ particle helped to solve the nuclear spin puzzles of preceding theories.

Chapter 2

Super Kamiokande Detector

2.1 Overview of the Super-Kamiokande Experiment

The Irvine–Michigan–Brookhaven (IMB) [30][31] and Kamioka Nucleon Decay Experiment (Kamiokande)[32][33] experiments preceded Super-Kamiokande (SK) and made several discoveries but failed to achieve their primary goal and observe nucleon decay. The two collaborations joined together to build a larger, more sensitive detector Super-Kamiokande nearby Kamiokande in the Kamioka mine. The main focus of SK from Kamiokande and IMB shifted from nucleon decay to the study of neutrinos (the primary source of background in nucleon decay searches), in particular investigating the possibility of neutrino oscillations to explain the solar and atmospheric neutrino anomaly.

SK is a 50-kton ultrapure water Cherenkov detector built within the Japanese mountain, Ikenoyama. It is a right cylindrical detector (41.6m $\times$ 39.3m), with 1000m (2700m water-equivalent) of rock overhead[34]. It is located within the Kamioka mine, previously used to mine zinc by the Kamioka Mining and Smelting company. The detector first turned on April 1st, 1996 and has been running since with some downtime for detector maintenance

Figure 2.1: Super Kamiokande's location within Japan

and upgrades. These downtime periods divide SK experiment into six different experimental phases. The detector is optically separated into an Inner Detector (ID) and Outer Detector (OD), with the OD used as a cosmic ray veto for events seen in the ID. When a charged particle is moving faster than the speed of light in a medium it emits light in a cone around the direction of travel, this is Cherenkov light. The PMTs in both the ID and OD collect this Cherenkov light; their signals are recorded by the Data Acquisition system (DAQ).

The Kamioka mine is located near the city of Kamioka, Gifu Prefecture, Japan, with the detector located at 36°25'33" N, 137°18'37" E, 371.8m above sea level. To access the main experimental areas, users drive through a 1.8km mine drift. SK shares the Kamioka Mine with several other experiments: CANDLES[35], XMASS[36], KamLAND[37], EGADS[38], and nearby KAGRA[39], covering topics from the search for neutrinoless double β-decay, gravitational waves, dark matter searches, reactor neutrino physics, and more.

The six phases of SK are noted numerically from I-VI. The initial phase (SK-I) was ended in order to perform maintenance. An accident occurred during the pure water filling when the detector was $\sim$75% full: a PMT imploded near the bottom, causing a chain

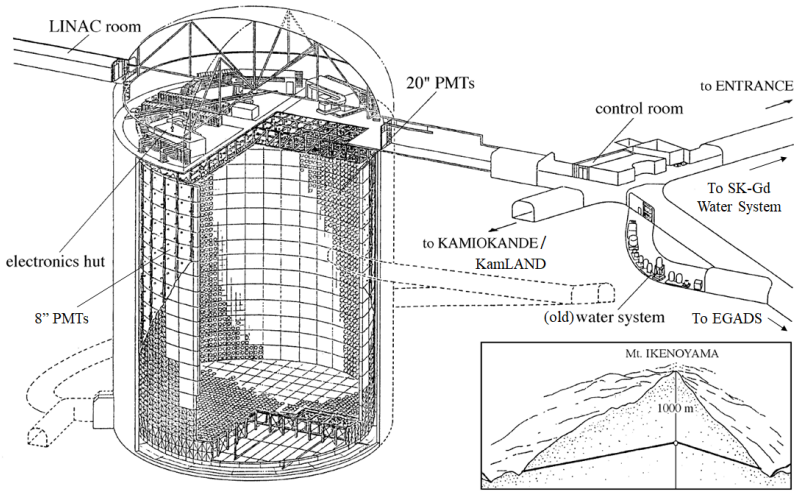


Figure 2.2: Diagram of the SK detector

reaction of implosions resulting in the loss of ∼60% of the PMTs. The surviving PMTs, along with spares and available stock, were redistributed with the resulting detector (SK-II) having a reduced photo cathode coverage (19%). SK-II ran while more PMTs were manufactured to replace the ones that were destroyed. Eventually, the full photo cathode coverage (40%) was restored in SK-III and has been maintained in the subsequent phases of SK. For SK-IV, the electronics and DAQ were upgraded to the QTC (charge to time converter)-Based Electronics with Ethernet (QBEE) boards, but there were no major detector hardware changes[40][41]. In preparation for Gadolinium loading, the small leaks of the SK tank were sealed and the detector was cleaned (Sec 2.9). With the end of this maintenance, SK-V was started. SK-VI began once $Gd_2(SO_4)_3$ was dissolved into the detector's water to enhance the neutron detection signal.

SK-I ran from April 1996 to July 2001, the implosion accident happened in November 2001. SK-II ran from October 2002 to October 2005, and the "full reconstruction" was completed in July 2006. SK-III ran form July 2006 until August 2008, SK-IV lasted from August 2008 until May 2018. SK-V began at the end of January 2019 and ended in July 2020 with the start of the Gd loading. There was a small modification to the detector water

injection during February 2020 to make modifications to the water system (Sec 2.9.3), but this did not warrant a new SK phase, and SK-VI started in July 2020 and continues to run at the time of this writing. The present Gadolinium concentration is, 0.02% $Gd_2(SO_4)_3$ resulting in the capture of 50% of thermalized neutrons on Gd rather than H[42]. Once predetermined bench marks for water transparency, detector stability, etc. are achieved (expected to happen in 2021-2022) the concentration will be increased further. To prevent another implosion chain reaction, the PMTs of SK-II onward are fitted with acrylic coverings in the front, and fiberglass reinforced plastic (FRP) coverings for the back to dissipate the energy of individual PMT implosions which may occur.

2.2 Detector Specifics

As mentioned previously, the detector is a 50 kton right cylinder stainless steel tank. The main structure within the detector is an assembly of stainless steel"super modules". The super modules are 2.1m $\times$ 2.8 m $\times$ .55 m, and are assembled to create a smaller cylinder within the tank to mount the PMTs for the ID and OD. The region from the tank wall to the super module assembly is the OD, and the region fully within the super modules is the ID. The space within the super modules themselves is deadspace.

Black sheet is used in the ID and black-white Tyvek in the OD to keep the two regions optically separated. White Tyvek is used to line the outermost walls of the OD in tandem with the white side of specially manufactured black-white Tyvek sheets on the super module to maximize the collection of light by the OD thanks to the exceptional reflectively of the white Tyvek. The white Tyvek helps to compensate for the limited photo-coverage in the OD. Since the OD is a cosmic ray veto, exact location of light production is not necessary. Also, the OD is not used for event reconstruction so there is no need for precise timing of OD hits as well, and differentiating reflected and directly collected light is not as

important.

2.3 PMTs

The PMTs are all hand blown borosilicate glass, and are produced by Hamamatsu. More information can be found in [43]. The PMTs of the ID and OD are mounted on the super modules, with full sized super modules holding 12 (4x3) 50cm diameter PMTs for the ID and 2 20.3 cm diameter PMTs mounted in wavelength shifting plates for the OD. There were 11,156 PMT 11,129 20-inch R3600 PMTs in the ID and 1885 8-inch PMTs in the OD for SK-I, The reduced coverage of SK-II had 5,182 ID PMTs, and 11,129 ID PMTS for SK-III and on. The R1408 and R5902 models comprise the PMTs in the outer detector. During SK-I, only the R1408 model was used, and were originally deployed in the IMB experiment. Many R1408 tubes were lost to the implosion, and a significant amount of the remaining surviving PMTs had continued to die through SK-IV from leaks in the PMT housing or damage to the coaxial cable from the implosion. Due to the failure rate of unpotted IMB housings, replacement PMTs had their bases potted to the PMT with epoxy by Hamamatsu, resulting in a much lower failure rate for the PMTs.

PMTs work by releasing a photo-electron (p.e.) when a photon strikes the surface, specifically the photocathode. A high voltage gradient within the PMT accelerates the electron towards a dynode, where it strikes and releases more electrons. This process is repeated from dynode to dynode multiplying the signal each time and eventually reaching the anode with enough current to output the signal to the to the DAQ, hence the name Photo Multiplier Tube. The PMTs are operated between 1.7 and 2 kV, and have a gain around 10^7.

Since the whole process for a PMT relies on a single electron to start a chain, they

are susceptible to false signals from thermally excited electrons from the PMT itself starting the chain reaction. These accidental signals are referred to as "dark noise" and happen at a rate of a few kHz per PMT. SK's cooler temperature helps to reduce the dark noise rate from being even higher.

In the OD, there is only ∼0.5% photocoverage provided by the PMTs. To help increase this, the PMTs are mounted to a 60 cm square doped acrylic wavelength shifting plate, with a thickness of 1.3 cm [44]. OD surface coverage is increased to ∼ 10% of the OD surface by the waveplates. The doping material, bis-MSB, absorbs UV light and re-emits 55% of the light in the blue-green region. The OD PMTs are also more sensitive to light in this region compared to the UV Cherenkov photons, and increases the light collection efficiency by 150%. The OD is a veto and calorimeter system for SK high energy analyses, but for SK low energy analyses the OD is used only as a veto for events and in muon fitters. The OD also acts as a buffer against γs and backgrounds coming from the surrounding rock from entering the ID.

The dark rate average for the ID PMTs depends partly on the manufacturing date of the PMTs and varies with time. For reference, in SK-IV the running averages for each batch's dark noise rate is between 3.8kHz and 7.4 kHz. For any given 200 ns time window, there will be ∼12 dark noise hits (this 200 ns window corresponds to whether an event will trigger or not, Sec 2.7.1). For detector simulations, the implemented dark rate is based on the measured average dark rate of ∼5.94 kHz, or the measured dark rate for a given period. The shifting dark rate is handled offline in affected analyses.

The inside of the tank is lined with 26 horizontal and vertical Helmholtz coils to compensate for the Earth's magnetic field, reducing it from 450 mG to 32 mG. Since the photoelectrons produced on the photo-cathode are collected to the first anode using an applied electric field, the geomagnetic field affects the collection efficiency of the PMT. Since the ID uses 20-inch PMTs, the electron has to travel significantly further than the smaller

PMTs commonly used in physics. The reduction of the geomagnetic field limits its affects on the collection efficiency of PMTs to ∼1-2%.

2.4 SK Water System

Maintaining water quality is paramount to the success of SK. Undesired impurities need to be removed from the SK water to increase water transparency as well as reduce radioactive isotopes being pumped into the detector. Clean water is needed for the Cherenkov light to be able to travel from the point of production to the PMTs on the wall of the ID or OD, and most impurities in the water within SK either contribute to absorption or scattering of the Cherenkov light. Bacteria can also be introduced to the detector through direct human contact or indirectly from touching objects to be inserted into the detector (calibration equipment for example). Bacteria can grow in the water and its waste can impact water transparency and create biofilms on the surfaces of the structures or PMTs. If a film forms on the surface of a PMT, the PMTs optical response to incoming light can be altered. Keeping the water cool helps to reduce PMT dark noise rate, as well as curbing bacterial growth within the detector.

The SK water system is specifically designed in order to try and handle theses factors and ensure desired water purity. The SK water supply is from a small river running through the mine that is fed by rainfall and snowpack. This water is initially treated to meet SK's standards before being fed into the detector. The water from the tank itself is circulated through the purification system as well, at a rate of ∼60 tons/h in SK-IV, and ∼60-120 tons/h in SK-V and SK-VI depending on the detector conditions and configuration.

The water system of SK-IV consists of:

- 1μm Filters: Polypropylene mesh filters with $\sim$ 1μm pore sizes to sift small particles

and dust.

- Heat Exchangers (HE): Cools the water passing through the system at various points as heat is input into the system in various parts throughout the process through friction, pumps, and surroundings.
- Exchange Resin: Ion exchange resin bed. Swapped in SK-I for a Cartridge Polisher. Removes light and heavy ions from the water. Polisher was able to also remove heavy ions that include radioactive species.
- UV sterilizer: 254 nm UV lamp used to kill bacteria
- Rn-Free-air Dissolving Tank: Dissolves very low Rn concentration air into the water to increase vacuum degasifier's radon removal efficiency
- Vacuum Degasifier: Removes dissolved gasses from the water that impact water transparency (CO_2), background event rate (Rn), and to discourage bacterial growth in water (O_2)
- Ultra Filter (UF): Membrane filtration system with membrane pore size $\sim$10nm the pressure filters ultra-fine particles from the water. What is missed in this is filtered by RO
- Reverse Osmosis (RO): Staged membrane filtration system with pore size of $\sim$ 1nm. Pressure filters out contaminants with 1000 times the molecular weight of water
- UF Reject Tank: Reject stream from UF and RO are sent to this tank. This water is cycled through another set of ROs, with the product being sent to the main system, and reject returned to this tank. Contaminants from the system are perpetually trapped here.

With all of these components, SK is able to maintain optical transparency for the desired wavelengths ($\sim 80 - 85\%$ remaining light intensity after 15 meters travel) through removal

of contaminates and dissolved gases (∼96% effective at gas removal). The water is held to be at ∼13℃, which is also useful for the reduction of convection near the water inlets of the detector. The result within SK is for water transparency on the order of 100m attenuation length in the 350-450 nm wavelengths.

2.5 Air Purification Systems

The Kamioka mine has high levels of radon gas, which is a health hazard for those passing through the mine and a significant background in the low energy analysis and many strategies have been limit radon in the experimental areas. To begin with, the walls of the mine are sealed with Mineguard[45] and the entrances between the mine and SK experimental area is simple air lock of two sets of doors.

At the entrance of the mine, the "Radon Hut" houses a large blower sending fresh air taken a distance from the mine entrance through a 10-micron filter and into the experimental areas. The air intake is far enough from the entrance to not be contaminated with air coming from the mine during the summer months. The mine is a constant temperature year round, and depending on the season the air flows into the mine from the entrance, or out of it with the changing outside air temperature.

A small amount of the air from the Radon Hut undergoes an additional step of being filtered through activated charcoal as shown in Fig 2.3[46]. The air is compressed to a pressure of around 7.5-8 atm, and passed through a 0.3 μm mesh filter to remove dust particles before being stored in a buffer tank. After the buffer tank, moisture is removed and passed through two activated carbon columns which absorb the Rn. The air is passed through two more mesh filters (0.1 and 0.01 μm) and finally through a chilled charcoal column. The final charcoal column's cooling system was upgraded in March 2013, achieving near 100%

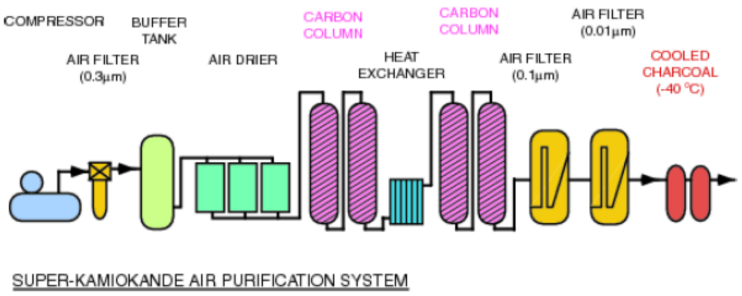


Figure 2.3: Schematic view of Rn free air system. In March 2013, the cooling system of the last charcoal column was upgraded to go down to -60℃, achieving near 100% removal efficiency

Rn removal. This additional step allows for the Radon concentration to be reduced to a $0.06\text{mBq} \pm 0.05\text{mbq/m}^3$, and pumps it to the top of the SK tank above the water. The supplied air for the tank top is added over-pressure in order to maintain a buffer between the water and the slightly more radon rich air located in the SK dome area. This air also is supplied to the water systems Rn-free-air Dissolving Tank, mentioned before.

2.6 EGADS

There is a 200-ton detector also in the experimental area called EGADS (Evaluating Gadolinium Action on Detector System) which was used as a test bed for the addition of Gadolinium to the main SK detector [38]. EGADS is instrumented with 240 20-inch PMTs along with its own DAQ system. The 240 PMTs are a combination of SK PMTs, with and without acrylic vessels, and Hyper-K prototype PMTs. EGADS main purpose was to prepare the SK detector for the addition of Gadolinium Sulfate ($Gd_2(SO_4)_3$). The addition of Gadolinium to SK was proposed to have a better means to see the capture of neutrons for inverse beta decay (IBD) interactions within the detector [42]. Without Gd, mostly thermal neutrons

capture on H and produce a 2.2 MeV:

$$\begin{aligned} \overline{\nu}_e + p &\rightarrow e^+ + n \\ \mathrm{H} + n &\rightarrow D + \gamma \ (2.2 \text{ MeV}) \end{aligned} \tag{2.1}$$

Gd is a neutron hungry nucleus, and provides an easier to see signal for neutron captures on the nucleus, with Gd captures of neutrons giving off an ∼8 MeV gamma cascade:

$$^{A}\mathrm{Gd} + n \rightarrow \ ^{A+1}\mathrm{Gd}^* \rightarrow \ ^{A+1}\mathrm{Gd} + \sim 4\gamma \ (\sim 8 \text{ MeV tot}) \tag{2.2}$$

This cascade has a much higher efficiency to be seen in SK, with roughly triple the light yield compared to the 2.2 MeV signal for capture on H and increases the signal to see neutrons after anti-neutrino captures. This is pivotal for searching for anti-neutrinos from different sources, especially in the search for the Diffuse Supernova Neutrino Background (DSNB). The average cross section for Gd, once the natural abundance of its different isotopes is taken into consideration, is 49,000 barns as compared to 0.3 barns for H [47].

For the use of Gd within SK, one of the main concerns was whether or not water transparency could be maintained with Gd-loaded water. To measure water transparency in EGADS, the Underground Device Evaluating Attenuation Lengths (UDEAL) was developed, and was designed by Michael Smy with the help of Jeff Griskevich and Andrew Renshaw. UDEAL fills an 8.2m pipe with water from EGADS and takes measurements of light intensity for different wavelength lasers. Cycling through the lasers, the beam is shown through a reflective neutral density filter, with the reflected light passed through the pipe and aligned to the opening of a 12" integrating sphere at the bottom of the device. The light passing through the neutral density filter enters a 4" integrating sphere at the top of the contraption. Using these spheres helps to stabilize the light intensity measurements performed by UV-enhanced Si photodiodes operated in reverse bias mode and amplified by op-amps. Intensity measurements from the top and bottom integrating spheres are made, and the ratio of

the intensities are compared for each wavelength at differing depths of water. Using this information, the attenuation length of light for the varying wavelengths can be made and a calculation on the % of light remaining after 15m (%LL15) can be made. 15m was chosen as it is the average distance Cherenkov light travels within SK. A weighted average of light loss for the seven different lasers is made, with the weights based on a convolution of the Cherenkov spectrum with the PMT response:

$$\%\text{LL15} = 100 \sum_{i}^{7} \omega_i \times e^{-L \cdot (\lambda_i)} \tag{2.3}$$

where $L = 15$m, and the sum runs over the 7 lasers of wavelength i, ω is the lasers weight, and α is the measured attenuation length for a given laser. The weights for all the wavelengths were determined from Monte Carlo, where the effective spectrum of Cherenkov light that leads to observed photo-electrons in low energy neutrino interactions was calculated. The different weights, ω for the different lasers are:

λ [nm]	337	375	405	445	473	532	595
ω	0.252	0.253	0.206	0.141	0.106	0.039	0.003

This system was modeled after IDEAL (Irvine Detector Evaluating Attenuation Lengths) located at UCI. EGADS successfully showed long term exposure of SK components to water doped with up to 0.2% $Gd_2(SO_4)_3$ had little to no effect on water quality, and that the system did not lose Gd concentration after long term running.

2.7 Super Kamiokande High Voltage, Electronics, and DAQ

SK has front end electronics that include the High Voltage (HV) system for the PMTs and the Data Acquisition system (DAQ) that takes the analog signal from the PMTs, integrating and digitizing it. The front end electronics for SK I-III used Analogue Timing Modules (ATMs) to acquire the signal (integrated charge of a pulse plus its timing) from the PMTs. The ATM would digitize this signal using a 12 bit Analog to Digital Converter (ADC), and handled 12 PMT channels each. These ATMs were contained within TKO (Tristan KEK Online) crates, with 12 TKO crates in each of the 4 perimeter electronics huts located on top of SK.

Each hut is responsible for a quadrant of the detector, and passes the collected information of PMT hits to the central hut. The central hut is home to the trigger system and control electronics. The trigger system then passes the information to online event builder computers located within the control room. The information sent are the ID and OD charge and time information along with trigger flags and time stamps. A completed full-detector event is formed by the online event builders from the merged data streams, and sent to the Kenkyuto via fiber optic lines for storage and analysis.

The basic information taken from individual PMTs hits and used for SK data analysis is the charge (Q), time (T), and PMT cable number. The time and cable number give the 4-vector of each PMT hit, and are the only information necessary for the reconstruction of low energy events ($<\sim$ 40 MeV) due to uncertainty in differentiating between minimal (1-3) p.e. Charge becomes more important for higher energy events and cosmic ray muon reconstruction.

2.7.1 SK-IV and QBEEs

With the start of SK-IV, new electronics and online system replaced the whole DAQ system. The ATMs were replaced with QTC (charge to time) based electronics with Ethernet (QBEEs)[40][41]. The two main benefits of upgrading the front end electronics from the ATMs to the QBEEs were the ability to have a continuous readout of the detector allowing for use of a software trigger instead of hardware trigger, as well as a dynamic range for recording charges. The QBEEs have three different amplifiers for three different gain ranges (small, medium, and large), covering a dynamic charge range of 0.2 to 2500 pC. This is an increase of ~ 5 compared to the old electronics, and will prove useful for this analysis.

Each QBEE has 8 charge-to-time (QTC) ASIC chips with three input channels per QTC chip. The QTC digitizes the PMT signal, with the leading edge of the waveform corresponding to arrival time of the event and the width of the waveform corresponding to the total charge seen by the PMT. The outputs of the QBEEs are sent to the front end PCs where they are sorted in time and sent to the merger PCs. The merger PCs merge all of the data and is checked against the software trigger program. In SK-I to III, QTCs were used to measure OD hit time and charge information, which was passed to time-to-digital converters (TDCs). With the introduction of QBEEs to the ID DAQ, QBEEs were also introduced to the OD DAQ, which have the OD QTCs integrated into them. The former ATMs used two channels per PMT, where one would be used for sending information while the second could take data while sending was occurring. If too much information was taken end, and both channels occupied, it would be lost. The increased bandwidth from the Ethernet and upgraded electronics allows for this to no longer be the cause of any bottlenecks. The bandwidth is increased by a factor of $\sim$100. There is also an inherent deadtime of $\sim$500 ns after each acquired PMT pulse is integrated over 400 ns. The QBEEs do have roughly the same T and Q resolution compared to the ATMs however, although both are much finer than the intrinsic resolution of the PMTs themselves and an upgrade wasn't necessary.

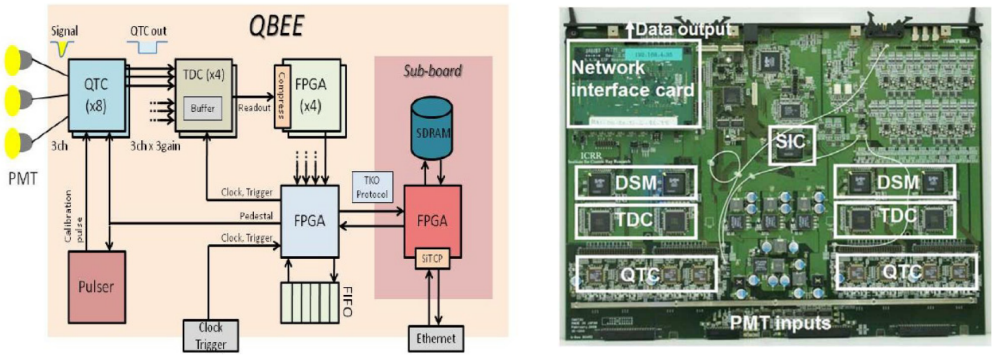


Figure 2.4: Schematic and Picture of the QBEE board

A new software trigger was implemented with the upgraded electronics, moving on from the former hardware trigger. With the former trigger, PMT hit pulses would be 200 ns wide and of standard height. All of the pulses on a board were added together, with all of the pulses from a board then added for a given hut, and finally the hit pulse sums for the huts would be added and run through a differentiator. If the total pulse sum (HITSUM) passed a discriminator threshold, then a global trigger was issued and the event data is saved. The threshold for the new system takes a 200 ns sliding window, looking for the number of coincident PMT hits in the window. This is now a counting exercise for individual PMTs hits, and when above a threshold (34 and 31 for different parts of SK-IV) the event data is saved with proper flags and information. The number of coincident hits were initially set to match the former HITSUM variable used with the old electronics. In the low energy regime, the number of hits is highly correlated with the energy since most hits for an individual event will be on the single p.e. scale. This will be discussed further in section (

2.7.2 High Voltage

The High Voltage (HV) supply system for the ID PMTs were originally from CAEN Components and were installed in 1995, but due to failure rate of the HV modules were upgraded to “iSeg” modules made by the German manufacturer iseg Spezialelektronic GmbH in 2013.

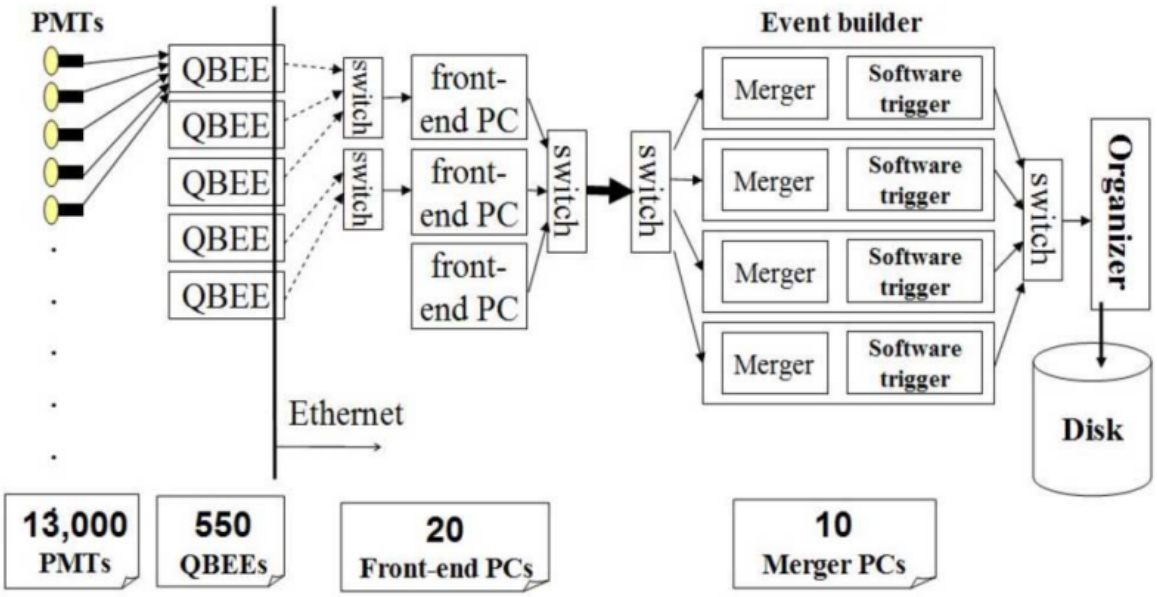


Figure 2.5: Schematic for the SK-IV DAQ.

Both systems allowed for voltages of all of the PMTs to be set individually, which is utilized to ensure all the PMTs have the same gain throughout the detector. The HV systems is able to monitor the status of the PMTs and periodically update the slow contorl monitor of possible issues that arise.

The OD HV system used a LeCroy 1454 HV Mainframe in each of the four electronics huts, through most of SK-IV. Near the end of SK-IV the OD HV system was replaced by iSeg EHS series HV modules in one of the huts, and the rest were replaced during the detector downtime in 2019 to finish the maintenance work. The new system allowed for better control of the OD HV from the control room, without having to enter the huts to perform certain control tasks as well as being able to check individual PMT HV status remotely.

Each huts HV unit has 40 channels connected to "paddle" cards, with each paddle card delivering HV to 12 PMTs. Paddle cards were used to cut cost by reducing the number of HV channels needed. The paddle cards have dip witches that allow for individual PMT channels to be locked out if they have failed. The nominal voltage applied to the PMTs is

1.8kV, however this varies PMT to PMT. In order to try and optimize PMT performance and have consistency, PMTs were matched based on what voltage gave the same single p.e. peak and grouped on the same paddle card. If there was too much of a discrepancy between desired PMT voltages, Zener diodes were placed in order for the voltages to be achieved across a card, with a single input voltage. If the signal from an OD PMT passes a threshold, the information is passed along to the OD DAQ.

The OD information is passed to the central hut, where the data is combined with relevant trigger and timestamp information, and passed along again to the online event builder computers to be combined with the ID data. The OD trigger threshold is 19 (22) hits within 200 ns for SK-I to III (SK-IV). After a global trigger is issued the OD data is saved for a window of 10μs before to 6μs after the issued trigger.

2.8 WIT

The Wideband Intelligent Trigger (WIT) system is a secondary trigger system within SK for extending the standard software trigger down to 3.0 MeV[48]. As discussed before, the normal SK software trigger, uses a coincidence trigger threshold of 34/31 hits within 200 ns to trigger on an event. The WIT system performs online reconstruction of calibrated time and charge from the QBEEs, allowing for the further reduction in threshold energy. The initial development goals for the WIT system was to use the lowered energy threshold ($4 \rightarrow 3$ MeV) to explore the lower energy edge of the Mikheyev-Smirnov-Wolfenstein (MSW) (Sec 1.3 effect in the sun. The analysis presented in this thesis is the first to use WIT for physics results.

WIT is placed into the normal data flow scheme after the SN Burst recorder, and initially consisted of a cluster of 8 PCs (currently 10, with PCs added or replaced as

needed). The information is sent in blocks of 1344 hardware trigger (HW) counts of data. HW counts are 17 μs sections of data corresponding to a time period of 32,768 (2^{15}) TDC counts. The 1344 hardware counts corresponds to $\sim$23 ms of data, with an overlap of 64 HW counts between adjacent blocks. The overlap ensures any 1ms of data to be wholly contained within a block. All but one of the PCs take merged data, from before the normal DAQs software trigger is applied, and apply the online reconstruction. If an event meets the WIT trigger conditions its information is passed to the final PC in the farm whose job is to organize the incoming streams of data in time and output into a single file.

WIT selects events by performing a series of hit cleanings and fits, eventually making a decision whether to save the event based on the final fit. WIT initially performs a pre-trigger selecting hits in a sliding 230 ns window, requiring 11 hits above dark noise rate, as measured using the time periods between WIT pre-triggers. 230 ns corresponds to the approximate amount of time it would take light to travel corner to corner of the detector, with a little wiggle room. These hits are processed with Software Triggered Online Reconstruction of Events (STORE). STORE treats the interaction as a point source since low energy electrons (<20 MeV) only travel a few cm. All pairs in the set of selected PMTs must have Δt $< \Delta x/c_w$, where Δt is the time difference between the two PMTs and $\Delta x/c_w$ is the distance between the two PMTs divided by the speed of light in water. STORE takes 4 hit combinations of PMTs from this set of selected PMTs and solves for a potential vertex ($\overrightarrow{x}$ and time (t_0)). The variable SC is the largest coincidence found by STORE from all these points and times, and is defined as follows:

$$SC(\overrightarrow{x}, t_0) = \text{Max} \sum_{|t-t.o.f-t_0| \leq 2\sigma} e^{-\frac{(t-t.o.f.-t_0)^2}{2\sigma^2}} \tag{2.4}$$

where $t - t.o.f. - t_0$ is the time residual for the PMT hit with respect to the test vertex and $\sigma \sim 5$ ns.

The accepted vertices are passed to the CLUSFIT (CLUSter FIT) algorithm, which acts as a pre-filter for super low energy events. Using the vertexes found by STORE, CLUSFIT will look for hit neighbors within space and time and combine them to test the fit, and selects the fit with the greatest goodness:

$$g(\overrightarrow{x}) = \mathrm{Max}\frac{1}{N}\sum_{|t-t.o.f-t_0|\leq 2\sigma}\left(1-\frac{(t-t.o.f.-t_0)^2}{2\sigma^2}\right) \tag{2.5}$$

where $t - t.o.f. - t_0$ is again the time residual for the PMT hit with respect to the fitted vertex, $\sigma \sim 5$ ns, and N is the number of selected hits for the tested vertex. The greatest goodness is chosen for the vertex fit, and checked for if it is in the fiducial volume. If the event passes CLUSFIT, it moves to an online version of BONSAI (Sec 4.2.1). A 1.5 μs window is placed around the trigger time, and BONSAI fits all time residuals across this whole window. After passing through BONSAI, the event is triggered depending on SC, CLUSFIT and BONSAI vertex locations, and whether or not $N_18 > 9$. N_18 is the number of time residuals whose values are between -6 and 12 ns. The minimum SC value to trigger an event is 6.6, necessitating events to have a minimum of 7 hits involved in the SC calculation.

When tested against Ni-Cf calibration data near the center of the detector, the WIT trigger efficiency was still near 100% in the lowest energy bin (3-3.5 MeV), while the SK 34 (31) hit trigger was ∼5% (∼40%) effective at triggering on the same events. When looking for n capture on H, the efficiency is close to 0% for the normal trigger, and ∼17% for WIT within the FV. The ability to still see n captures on H, albeit with poor efficiency, was pivotal for the analysis to be presented later.

2.9 2018 Detector Maintenance

Before adding Gd to the SK water, small leaks of the SK tank had to be sealed to ensure containment of all chemicals in the water. Therefore a detector maintenance period was scheduled, and was performed from June 1st 2018 until January 2019. In addition to sealing the SK tank, XX ID PMTS and ∼200 OD PMTs which had died over the past 12 years were replaced, and rust spots that had developed over the years (e.g. at nicks of the stainless steel caused by the implosion or weld seams) were removed. In addition, the PMT support structure was cleaned to remove dirt.

2.9.1 SK Leak

Previously the region of the Jinzu River basin experienced heavy metal poisoning (in particular Cadmium) affecting local wildlife, crops, and eventually humans. It lead to a disease dubbed "itai-itai" (literal translation is "it hurts, it hurts"). At the time of the detector maintenance, SK was leaking one ton of water a day and the plan to add the heavy metal Gadolinium in conjunction with the tank leak raised concerns. SK guaranteed to fix the leak in the detector to an acceptable level: the concentration in the Jinzu River basin would be below the legal limit for mercury, the most stringent criterion for any metal contaminant. This translates into a leak rate of at most 30 l/day. Studies performed by draining a pit located outside the detector determined the leak was most likely near the bottom of the tank, and the size of the hole would only be a couple mm wide (in the case of a single leak).

In order to accomplish this reduction in leak rate, all welding spots, which were thought to be the most likely location for a leak, were double coated in a compound called Mineguard-C6. Mineguard is typically used for underground rock support and applied as a spray in underground environments. Special formulations of Mineguard were developed

and samples were sent to the Kamioka Observatory and UCI to test the compounds effect in water. The compound had to: not dissolve in ultrapure water or water doped with up to 0.2% Gadolinium Sulfate $Gd_2(SO_4)_3$, maintain adequate optical transparencies to the level of SK-IV, any emanations from the sealing material be removable by the water system, low radioactivity, and maintain integrity and adhesion to stainless steel as it may stretch under distortions to the tank wall from change in pressure.

The plates were placed within the UCI Soak system, the prototype water system for EGADS and SK-Gd. The water system was a proof-of-concept for the possibility of using a band pass system to selectively maintain $Gd_2(SO_4)_3$ concentration in water, while being able to perform routine water cleanup to maintain transparency. After successfully showing water transparency similar to those of SK-III and SK-IV could be maintained, the system was used to test different materials that would potentially be placed in SK and their effects on water transparency.

To test the water transparency, the Irvine Device Evaluating Attenuation Lengths (IDEAL) was used. Designed by Michael Smy as well, IDEAL is the predecessor to UDEAL discussed earlier, and uses a 6.4m pipe instead of the 8.2m pipe in UDEAL. Different formulations of Mineguard coated steel plates were tested and a final formula decided upon. The changes to the initial formula include using fumed silica and calcium carbonate to reduce the epoxy's radioactivity and changing the polyurethane to polyurea to avoid hydrolysis issues as polyurethane breaks down in water.

During the detector maintenance of 2018, Mineguard was successfully applied to the welding lines along the tank, as well as new leaks that appeared during the tank draining process. There is water from within the mine that surrounds the back fills the cavity the tank is located in. As the detector drained, and before the pressure could be relieved, a number of additional minor leaks arose and sealed. Upon refilling of the detector, the leak fix effort was determined to be a success, as no water leakage was observed from the SK tank

within the accuracy of our measurement. An upper limit of 17 l/day was placed on the leak rate of the detector.

2.9.2 Rust Removal

Rust within the detector was a concern due to the effect Gd doped water had on lifting rust from surfaces and putting it into solution, drastically lowering water transparency. This was an issue in the past when the 1 kton detector at KEK was loaded with $GdCl_3$ to test Gd loading. KEK had an iron tank and part of the inside was coated with anti-rust paint, however the top section of the tank was not coated with the paint. With the addition of Gd, the top section which had rusted and the rust was lifted making the water opaque. The issue of Gd loaded water and rust was experienced again in the EGADS detector when a cable was not Gd safe (steel instead of stainless steel), and had completely rusted and dissolved into solution. EGADS had to be shut down and go through complete cleaning afterwards. Care was taken to inspect every surface of the Stainless steel tank and all of the super modules to look for rust spots. Rust would be removed and the steel re-passivized to prevent return of rust in the future. On top of the nicks from implosion shrapnel, any rusting pieces of equipment were searched for. As lower and lower levels were reached, the levels of rust increased as more of the implosion force was experienced there.

Since this was such a great concern for the success of SK-Gd, most of the human effort was placed on this aspect of the maintenance process. All Tyvek was removed, and the entire surface of the detector was searched for rust. During the initial rust removal process, it was found the surfaces were coated with residual diesel exhaust from when the modules were first installed more than 20 years prior. Added effort was made to clean all the surfaces on top of searching for rust, as it was an unknown how the residual diesel exhaust would react to Gd. This ended up being an added benefit for the SK low energy analysis, as this

exhaust happened to be part of the radioactive background coming from the PMTs and super modules.

The rust removal was also a success, as with the addition of $Gd_2(SO_4)_3$ in July 2020, changes to the water transparency were consistent with the expected system shock of the initial Gd loading. This shock had been seen in EGADS and the UCI Soak system whenever Gd was added.

2.9.3 Water Flow Changes

The final set of detector changes associated with the 2018 maintenance did not occur until a year after the detector turned back on. Many changes to the water system had to be made in order to load the detector with Gd, and one of the major functionality changes was refitting the water system to be able to increase the flow of water within the detector from 60 tons/hour to 120 tons per hour. For reference, results from EGADS had been produced with a flow rate equivalent to 250 tons/hour.

The increase in flow rate was desired to compensate for potential complications introduced by Gd loading. With Gd loaded water, expectations for more emanation from detector material, faster bacterial growth, and faster rusting if non-stainless steel materials remained in the detector.

However, early in SK-V, 120 tons/h flow rate was attempted, but a massive convection region had formed from the turbulence formed by the greater flow rate and removal of diffusers at the water inlets during the 2018 tank open work. Convection is not new to SK, but the convection region in SK-IV was limited to the volume below Z = -11m, while the convection observed in the early 120 tons/hour flow rate studies had the convection had extended to Z = +7m. Outside of this convection region, laminar flow is maintained. This

convection region is detrimental to the low energy of the solar analysis, as radioactive backgrounds, particularly Rn, emanating from different parts of the detector and surroundings are spread to a bigger region of the detector. ^{222}Rn has a half life of 3.8 days, which gives time for the Rn found within the convection region to fully spread throughout it. The diffusers worked to redirect the water flow from the inlets to the horizontal direction, instead of vertically, and did play a bigger role than previously thought in maintaining a smaller convection region.

Therefore, another small tank open effort was organized, with the use of a submersible to reintroduce the diffusers to water inlets at the bottom of the detector. A submersible was used instead of a diver due to safety concerns for diving to the bottom with the detector full, even though a submersible would have more difficulties placing the diffusers. When completed, reduction in convection region was achieved, and increase flow rate was able to be maintained.

2.10 Physics within SK

As briefly mentioned earlier, Cherekov light is produced when a charged particle travels faster than the speed of light for a given medium. Like a sonic boom when a plane travels faster than the speed of sound, a shockwave of photons are produced when they exceed the speed of light for the medium. The equation to meet this condition becomes

$$\beta > \frac{1}{n} \quad \text{where} \quad \beta = \frac{v}{c} \tag{2.6}$$

n is the index of refraction for the medium, v is the particles velocity, and c is the speed of light in a vacuum. This corresponds to a Cherenkov energy threshold for the most common charged particles in SK at: $e^{\pm} = 767\text{keV}; \mu(^{\pm} = 157.4\text{MeV}; \pi^{\pm} = 207.9\text{MeV}$. The light itself

will be emitted at an angle (θ_C) with respect to the particles trajectory, and this angle can be found by: $\cos\theta_C = \frac{1}{n\beta}$. For water, if the particle is ultra-relativistic, $\theta_C \sim 42$ degrees. This is usually the case for electrons, while heavier particles are more likely to be travelling slower than this and the angle becomes smaller.

If an incoming neutrino interacts within SK, it can create particles such as muon or electrons, or even transfer energy to orbiting electrons, and with enough energy that they are above Cherenkov threshold. Electrons will travel a short distance ($\sim$10 cm) while muons will travel a greater distance. The light emitted by these events within the detector are collected by some of the many PMTs on the wall of the ID, with the "ring" of light from the Cherenkov cone giving information on the location and type of particle observed. Neutrino interactions are not the only source of these kind of events within the detector and they aren't even the most common, which adds some difficulties into the mix. Muons can also be present from cosmic rays, while there are other sources for electrons such as radioactivity emanating from the PMTs and support structure, within the water supply itself, or from spallation resulting from cosmic ray muons.

It is crucial to be able to sift through all of the data, and try to separate as much of the desired signals from backgrounds as possible. The bulk of the effort of this thesis was in reducing the cosmogenic induced background within the solar energy region of SK.

Chapter 3

Calibration and Simulation in Super Kamiokande

All calibration in SK involves simulating different aspects of the detector and/or the detector in full. The behavior of Cherenkov light in water, behavior and response of the PMTs, and the relationship between Cherenkov light produced for an event and its energy are all pivotal to understanding what is observed in data. Details for calibration for SK-IV can be found in [49].

3.1 Event Simulation

The SK detector simulator is a heavily customized Geant3 Monte Carlo (MC) simulation[50], called SKDETSIM. With the addition of Gd, recent effort has been made to update SKDETSIM to run with Geant4 since this package has better simulation for n capture on Gd[51]. SKDETSIM fully simulates the detector's response, based on the calibration performed in the detector, to events that occur within the detector, such as recoil electrons traveling in

water and the emission and propagation of Cherenkov light. How different calibrations affect SKDETSIM will be discussed in the associated calibration sections.

3.1.1 Cherenkov Photon Production

The number of Cherenkov photons, N, produced per unit length, dx and unit wavelength, $d\lambda$, is represented as the following:

$$d^2N = \frac{2\pi\alpha}{n(\lambda)\lambda^2}(1 - \frac{1}{n^2\beta^2})dxd\lambda \tag{3.1}$$

where n is the refractive index for water, α is the fine structure constant, β is the velocity of the photon in the medium in units of c. Since the refractive index n depends on not only the wavelength, but also the water temperature, the water pressure, and so on. All of these are taken into consideration in the simulation.

The Cherenkov opening angle θ is represented as:

$$\cos\theta = \frac{1}{n\beta} \tag{3.2}$$

which can then be used to find the Cherenkov light emission threshold:

$$E_{thr} = \frac{nm}{\sqrt{n^2 - 1}} \tag{3.3}$$

where m is the mass of the charged particle. Using these formulas, the Cherenkov light emission is simulated in the SKDETSIM. Some example Cherenkov thresholds are listed in Table 3.1.

Particle Type	Energy Threshold (MeV)
$e^{\pm}$	0.767
$\mu^{\pm}$	157.4
$\tau^{\pm}$	207.9

Table 3.1: Cherenkov Radiation threshold for typical charged particles

3.1.2 Detector Response

To simulate the detector response, the PMT and the electronics should be taken into account. In the SK simulation, the three kinds of PMT responses considered are the absorption by the photo-cathode, reflection on the PMT surface, and transmission through the PMT. The left of Fig. 3.1 is a diagram of these three possibilities, and these responses depend on the incident angle of the photon, which is shown in the right of the figure.

The probability of the i^{th} PMT's to produce one p.e. when the generated photon reaches its surface is defined by:

$$P(\lambda, i) = \mathrm{QE}(\lambda) \times P_{obs}(\lambda, \theta_i) \times \mathrm{COREPMT} \times qetable(i) \tag{3.4}$$

where QE(λ) is the wavelength dependent quantum efficiency of the PMT, P_{obs} is the probability of selecting either of the PMT's responses for absorption, COREPMT is a common parameter used for all PMTs to correct for the average quantum efficiency (tuned with LINAC, Sec 3.4.1), $qetable(i)$ is the relative quantum efficiency for the i^{th} PMT (Sec 3.2.3). If/when p.e. are generated, the output charge is simulated. When it exceeds the proper thresholds, the photon is detected by the PMT in the simulation.

The Cherenkov photons can also reflect off of surfaces within the detector, such as the surface of the PMT or the black sheet. The reflectively of these surfaces are measured during certain calibrations and are taken into account in the SK simulation.

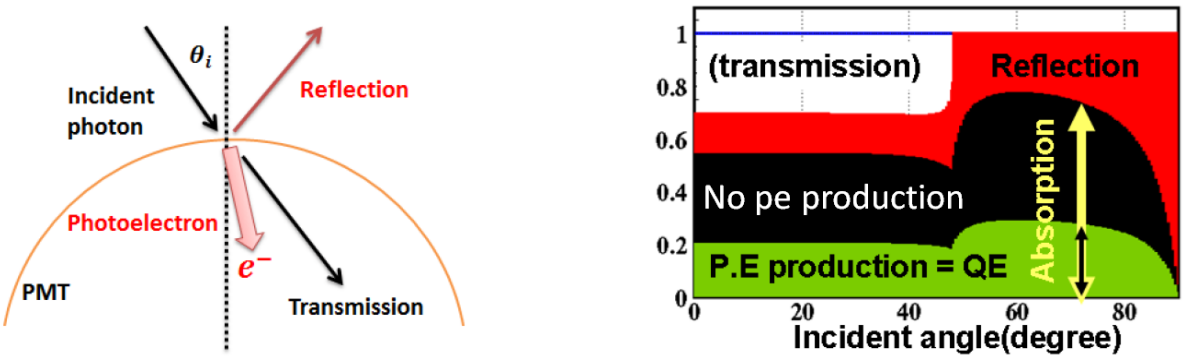


Figure 3.1: Left: Schematic view of the PMT response to an incident photon in the simulation. Right: Type of PMT response for 420 nm light as a function of incident angle. The black area represents photons absorbed into the PMT surface with no production of p.e.

3.2 PMT Calibration

Since the timing behavior of PMTs depends on the charge of the measured pulse. Since PMTs are the active detection equipment, it is important to understand and characterize the behavior of their timing (T) and charge (Q). For low energy events, T and Q are treated mostly independent of each other since they use single photoelectron (p.e.) statistics and occupancy is more important than Q. It is still important to characterize and understand what is seen in the data. To begin with, I will define a couple factors related to gain calibration. The first is "gain", which is the conversion factor from the number of p.e. to charge in units of pC. The second is "QE", which is the product of the quantum efficiency and the collection efficiency of the p.e. onto the first dynode of the PMT. These values are different PMT to PMT, so calibration is performed in order to account for this.

3.2.1 Absolute Gain

The absolute gain is applied to all PMTs, and this factor converts the number of p.e. to the output charge in pC. This value depends on the HV values applied to the PMTs, and the applied voltage for each PMT was determined at the beginning of SK-III to standardize

output charge using a Xe light source fed to a scintillation ball. The applied voltages were chosen so the observed charge response Q^{obs} was the same for all PMTs for a given intensity of the Xe light source.

For a the i^{th} PMT in SK, the Q_i^{obs} is proportional to the product of the PMTs QE_i, absolute gain A_i, and the total number of photons N_i^{γ}.

$$Q_i^{obs} \propto N_i^{\gamma} \times \mathrm{QE}_i \times A_i \tag{3.5}$$

After evaluating the Q_i^{obs} for all of the PMTs, supplied HV values were determined and applied.

With the supplied HV set for all the PMTs, the absolute gain was determined using a Ni-Cf calibration source, commonly referred to as "nickel (Ni) source" calibration. The Ni source is a polyethylene ball containing a ^{352}Cf source surrounded by NiO wool. It is compact enough to be lowered into any accessible calibration hole by hand crane. ^{352}Cf has a half-life of 2.65 years and decays with an α particle 96.9% of the time or spontaneous fission 3.1% of the time with an average of 3.8 neutrons. The neutrons thermalize in the water within the ball and capture on the Ni. . The The absolute gain was determined to measure the output charge distribution of single p.e. hits. Fig 3.2 shows the typical charge distribution of single p.e. hits using the Ni source calibration data in SK-III. With this calibration result, the absolute gain for SK-III was determined to be 2.243 pC/p.e. The process was repeated for SK-IV and found the value to be 2.645pC/p.e. Table 3.2 shows the absolute gain determined for each phase. On a long term scale, the gain will steadily increase very roughly about 2-3% per year. The dip from SK-II to SK-III is the result of introducing newly manufactured PMTs. With the length of SK-IV and increasing gain in the latter half, a time dependent gain correction was applied by the low energy group, and is explained in Sec. 4.2.3.2.

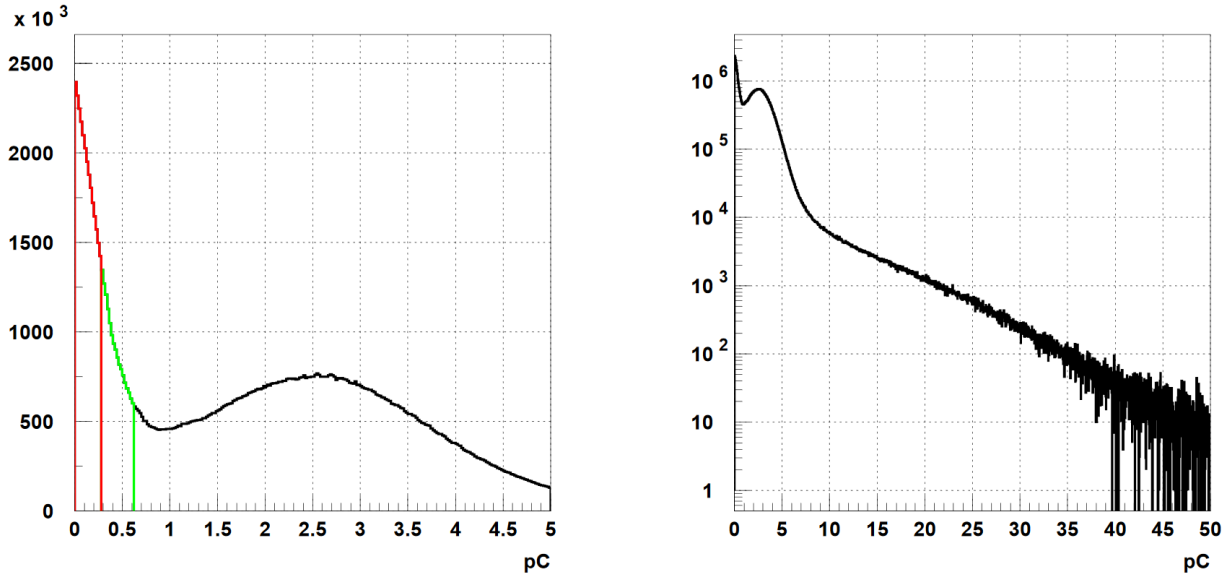


Figure 3.2: Typical single p.e. charge distribution from Ni source calibration. Left (right) show the distributions in linear (logarithmic) scale. The colors in the left diagram represent normal PMT gain (black), double gain and half threshold (green), and the linear extrapolation (red).

SK Phase	Absolute gain [pC/p.e.]
I	2.055
II	2.297
III	2.243
IV	2.645

Table 3.2: Absolute gain for each SK phase

3.2.2 Relative Gain

The relative gain adjusts for the relative difference among PMTs to ensure that all PMTs produce a uniform response for the output charge. To evaluate the relative gain for each PMT, a laser calibration system was implemented. The system uses a nitrogen-laser-driven dye laser with a neutral density filter wheel to allow for low and high intensity settings. For the low (high) output, each PMT detects 1 (∼50) p.e., with the Q_i^{obs} of the i^{th} PMT in the high intensity data defined as:

$$Q_i^{obs} = I_i^{high} \times \mathrm{QE}_i \times G_i^{rel} \tag{3.6}$$

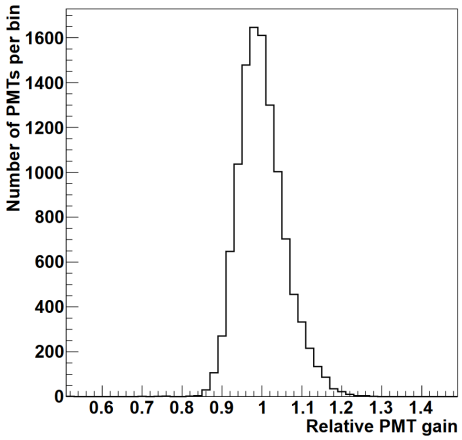


Figure 3.3: The relative gain of ID PMTs obtained from laser calibration

where I_i^{high} is the light intensity seen and G_i^{rel} is the relative gain factor for the i^{th} PMT. With the low intensity data, the number of hits (N_i^{hit}) is counted for the number of times the i^{th} PMT records a charge above threshold. Since Q_i^{obs} is proportional to both QE_i and G_i^{rel}, while N_i^{hit} is set up to only depend on QE_i, the ratio can be taken to get G_i^{rel}:

$$\frac{Q_i^{obs}}{N_i^{hit}} \propto G_i^{rel} \tag{3.7}$$

Fig. 3.3 shows the outcome of the relative gain measurement. With the RMS of this distribution being within 6%. This method is particularly useful since taking the ratio negates the effects of water transparency and geomagnetic field variations since they are canceled out.

3.2.3 PMT Quantum Efficiency Measurement

As mentioned before, QE refers to the product of the PMT's quantum efficiency and the PMT's collection efficiency. The PMT's quantum efficiency is the fraction of incident photons that produce photoelectrons in the photocathode via photoelectric effect. To determine the QE for each PMT, the effect from the gain for the PMT needs to be mitigated, as well

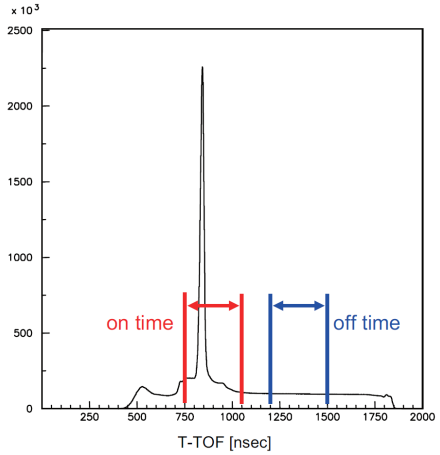


Figure 3.4: Typical hit timing distribution of the Ni calibration. The on-time (off-time) window is set to 750-1050 ns (1200-1500 ns)

reasonably as knowing the number of photons reaching the PMT. The nickel source satisfies these requirements. The hit probability should be proportional to the value of QE each PMT when low intensity light is injected.

Since dark noise hits and/or radioactive background hits can enter the event timing window during calibration, two regions are defined for the hit timing window, corresponding to "on" and "off" time hits determined with t - t.o.f. (time minus time of flight) subtraction to correct for travel time to each PMT from the source location. The on-time window covers 300 ns of the main peak of the distribution, while the off-time window corresponds to a 300 ns window starting 150 ns after the on-time window. These windows and distributions are shown in Fig. 3.4.

To obtain the number of true hits, the number of off-time hits are subtracted from the on-time hits. The hit rate is also dependent on the location of the source relative to the PMT, so the full definition of hit rate of the i^{th} PMT is as follows:

$$h^i_{corr} = \frac{h_i \times r_i^2}{a(\theta_i)} \tag{3.8}$$

where h_i is the observed hit rate for the i^{th} PMT, r_i is the distance from the nickel source to

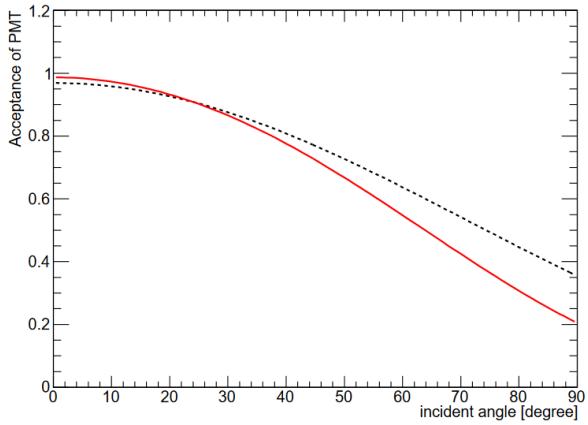


Figure 3.5: Incident angle dependence of the PMT acceptance. The red (black) curves represent the acceptance with (without) the acrylic cases used in SK II-IV. The greater discrepancy at high incident angles is partially a result of this.

the PMT, and $a(\theta_i)$ is the PMT acceptance which is extracted by MC simulation (Fig 3.5). The acceptance depends on the geometry of the PMT itself, and from SK-II on the acrylic cases covering the PMTs. The PMT angular dependence of the PMT acceptance is represented as:

$$a(\theta_i) = 0.205 + 0.524\cos\theta_i + 0.390\cos^2\theta_i - 0.132\cos^3\theta_i \tag{3.9}$$

where $theta_i$ is the incident angle of the photon to the PMT. Finally, the hit rate is normalized to obtain the hit probability:

$$P_i^h = \frac{h_i^{corr}}{\sum_i h_i^{corr}/N_{all}} \tag{3.10}$$

where N_{all} is the number of PMTs in the detector. The hit rate probability for the top, bottom, and barrel regions is shown in Fig. 3.6. The position dependence that remains is caused by the photon scattering and absorption in water, as well as from reflections coming from neighboring PMTs or the black sheets. To evaluate this position dependence, a MC simulation was prepared which considered absorption and scattering in water. For the MC simulation generation, the QE values for each PMT was set to the same value, and the results are shown in blue in Fig 3.6. There remains a difference between the MC and data on the few % level.

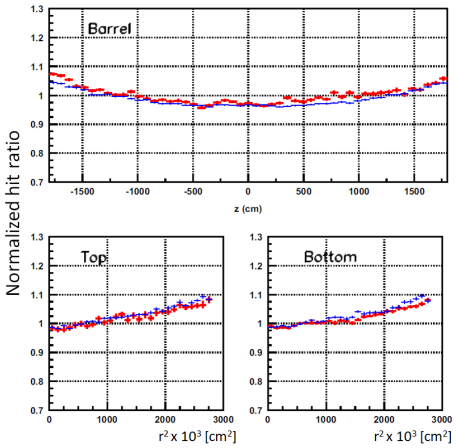


Figure 3.6: The position dependence of the normalized hit probability. The red (blue) points show the result from the calibration data (MC simulation). The upper panel shows barrel PMTs with teh horizontal axis as the z position of the PMTs. The lower left image is of the Top PMTs and the lower right is of the bottom PMTs, where the horizontal axis is now the r^2 from the detector central axis.

Taking the ratio between the data and the MC, the individual QE values for each PMT is calculated, and added to the "qetable." The qetable is a lookup table for the QE of each PMT.

3.2.4 PMT Timing Calibration

The PMT timing information is the most important item for vertex reconstruction. With possible differences in cable length and and variation in the response of electronics, timing response is not identical amongst all PMTs. Also, an inherent property of the PMT signal is "time-walk", which is the slight difference in PMT timing depending on signal charge, and is shown in Fig. 3.7. In short, larger charge will reach the TDC discriminator threshold faster than a smaller charge will, relative to when the signal started. In order to calibrate the PMT timing response for various charges, the laser calibration system is used again. The relation between timing and pulse height is called the "TQ-map". This map is fit with various polynomials, depending on the charge range, and the parameters of the fits are used as a

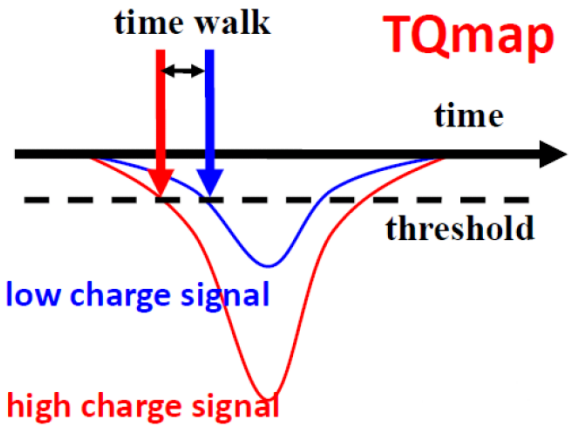


Figure 3.7: Visual representation of the time-walk effect. The red (blue) curve correspond to large (small) charge signal seen by a PMT, and the dashed line is the TDC discriminator threshold. The pulse height exceeds threshold faster for large charge signals compared to small charge signals.

correction to the time response for each PMT as a function of observed charge. To calibrate the TQ-map for each channel, the t-t.o.f. timing distribution is made for each PMT. The selected laser hits of each readout channel are divided into 180 bins of charge (Q_{bin}). Q_{bin} is the amount of charge in pC and defined linearly from 0 to 10 pC and logarithmically from 10 pC to 3981 pC. After separating into 180 bins, timing distributions are made for each Q_{bin}. The timing distributions are fitted with an asymmetric Gaussian function, as the timing distribution is not symmetric due to let hits from light scattering or reflecting. The timing peaks and standard deviations for all of the bins are then fit with various polynomials. These polynomials are split up by charge region, and are defined as follows:

$$\begin{aligned} F_1(x) &\equiv P^3(x) && Q_{bin} \leq 10 \\ F_2(x) &\equiv F_1(10) + (x-10) \times [F_1'(10) + (x-10) \times P^3(x-10) && 10 < Q_{bin} \leq 50 \\ F_3(x) &\equiv F_2(50) + (x-50) \times P^6(x) && Q_{bin} > 50 \end{aligned} \tag{3.11}$$

where x is a continuous parameter for Q_{bin}, F_1' is the first derivative of F_1 to enforce continuity between F_1 and F_2 at Qbin = 10. These are all constructed to be continuous at the boundaries, and result in a total of 15 fit parameters, 4 each from F_1 and F_2, and 7 from F_3

for each PMT, and saved as the TQ-map. The total timing resolution is evaluated for each Q_{bin} using the average fitted timing resolutions of all PMTs.

To monitor the TQ-map in real-time, the laser component of the timing calibration system was replaced with a N_2 laser better suited for continuous running. The N_2 laser emits 337.1 nm light, and the dye shifts it to 398 nm. For continuous running, the intensity is held at a constant level that achieves ∼99% occupancy in the ID PMTs with a maximum average Q of ∼20 p.e. per PMT. This occurs at a rate of approximately 0.03Hz and the long-term stability of the timing response in SK-IV is within ±0.1 ns.

3.3 Water Transparency Measurement

Knowing the purity and state of the water within SK is paramount to the success of all physics analyses in the collaboration. Small changes in the water transparency (WT) impacts the attenuation length of light in the water and affects physics results. Accurate simulation in MC requires understanding the behavior of photons in water altered by things like scattering, absorption, and reflection.

The WT is tracked in two ways, the first is using the N_2 laser discussed in the previous section as a means to understand photon propagation within the water. The other is using the decay electrons from stopping cosmic ray muons within the detector as an energy correction for water attenuation length in the solar neutrino analysis.

3.3.1 Measuring Water Transparency with Laser Light Injection

The water attenuation length is characterized by three coefficients, α_{abs}, α_{sym}, and α_{asym} in SKDETSIM, with the path length for photons modeled as an exponential:

$$\begin{gathered} I(\lambda) = I_0(\lambda) e^{L \times \alpha_{total}(\lambda)} \\ \text{where } \alpha_{total}(\lambda) = \alpha_{abs}(\lambda) + \alpha_{sym}(\lambda) + \alpha_{asym}(\lambda) \end{gathered} \tag{3.12}$$

where λ is the wavelength of the light. The wavelength dependent water transparency is defined as $L_{WT}(\lambda) = 1/\alpha_{total}(\lambda)$. The components of α_{tot} correspond to the total (absorption) absorption (α_{abs}), the symmetric component of scattering from Mie scattering and Rayleigh scattering (α_{sym}), and the asymmetric component of the Mie (forward) scattering (α_{asym}).

Functions are formed to describe each WT coefficient outside of calibration wavelengths. The functions are fit to the calibration data using sets of WT parameters A_i, S_i, and M_i that describe the α_{abs}, α_{sym}, and α_{asym} components respectively. The function to describe the WT α_{abs} coefficient is:

$$\alpha_{abs}(\lambda) = \begin{cases} A_0 \times [\frac{A_1(t)}{\lambda^4} + A_2(\frac{\lambda}{500})^{A_3}] & \text{if } \lambda < 463.9 \text{ nm} \\ \frac{A_0 \times A_1(t)}{\lambda^4} + \mathrm{PF}(\lambda) & \text{if } \lambda \geq 463.9 \text{ nm} \end{cases} \tag{3.13}$$

where the WT parameter $A_1(t)$ is varied with time based on decay electron data, described in the next section. The Function PF(λ) has been experimentally measured by Pope and Fry using their integrating cavity absorption meter[52]. The symmetric and asymmetric scattering WT coefficients are:

$$\alpha_{sym}(\lambda) = \frac{S_0}{\lambda^4}(1 + \frac{S_1}{\lambda^2}) \tag{3.14}$$

$$\alpha_{asym}(\lambda) = M_0[1 + \frac{M_1}{\lambda^4} \times (\lambda - M_2)^2] \tag{3.15}$$

These coefficients are measured using light injectors from the N_2 laser (337 nm) or an assortment of laser diodes (LD) (375 nm, 405 nm, 445 nm, and 473 nm). Before the laser diodes were installed in 2009, a dye-laser system was used to inject 365 nm, 400 nm, and 420 nm light. The light is injected every minute and the hit rate in six different regions of the SK ID are monitored.

The six regions are broken into one top region, and 5 barrel regions as shown in Fig 3.8. The scattered hit rate and observed total charge, Q_{tot}, are measured and the ratio of hits to Q_{tot} (hits/Q_{tot}) are obtained for each of the regions. The observed rate at the bottom PMTs are used as a reference to monitor the intensity of the laser. Various sets of MC simulations are generated varying the α coefficients and the same distributions for hits/Q_{tot} are obtained for the various regions. The distributions of the calibration data and the MC simulation are compared using the χ^2 defined as:

$$\chi^2 = \sum_r egion \frac{(Data - MC)^2}{\sigma^2_{data} + \sigma^2_{MC}} \tag{3.16}$$

where Data (MC) refers to the peak position of the hits/Q_{tot} distribution for data (MC) and the standard deviation σ_{data} (σ_{MC}). The WT coefficients are scanned over to find the minimum value of χ^2.

The typical coefficients are shown in the left of Fig. 3.9 with the time variation of the measured coefficients from early SK-IV data shown on the right. α_{sym} is stable within $\sim$3% during the period shown (purple). However, the α_{abs} and α_{asym} varied by 20-40% and 20-60% respectively over the same period. Since α_{abs} is much larger than α_{asym}, the time variation of the WT was mainly caused by the absorption process of water.

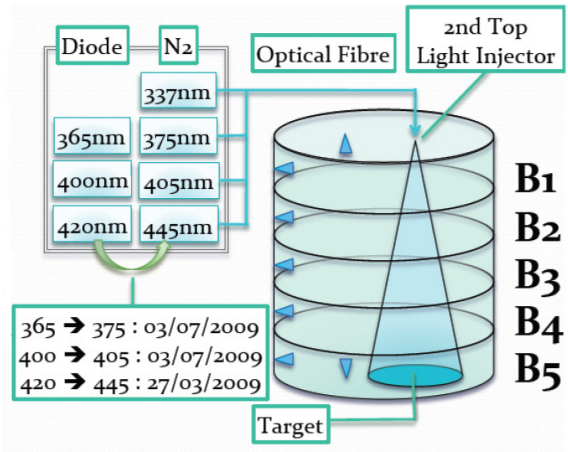


Figure 3.8: Schematic view of the laser injection system for the WT parameter measurement (left). Typical hit rate distributions for the six regions (right). The horizontal axis of the hit rate distributions is the t-t.o.f. subtraction for PMT hit time. The left broad peak corresponds to light arriving from scattering, and the right peak corresponds to light from scattering.

3.3.2 Measurement of Water Transparency with Decay Electrons

Although the SK water system has made the WT relatively stable through SK-IV, there is still ∼20% variation of the transparency in time. In order to monitor this time variation, WT is also continuously monitored using decay electrons (positrons) from cosmic ray muons stopping within the detector. Stopped muons decay into a neutrino, an antineutrino, and electron (positron):

$$\mu^- \rightarrow \nu_\mu + \overline{\nu}_e + e^- \text{ or } \mu^+ \rightarrow \overline{\nu}_\mu + \nu_e + e^+ \tag{3.17}$$

These decay electrons, also known as Michel electrons, have a well known energy spectrum, and the time dependence of the mean energy of the decay-e is measured to track the time variation of the WT. The selection criteria for these events are:

1. The time difference, Δt, between the parent muon event and the decay electron candidate event is between 3.0 μs $\leq \delta t \leq$ 8.0 μ

2. The reconstructed vertex of the decay electron candidates event is within the fiducial

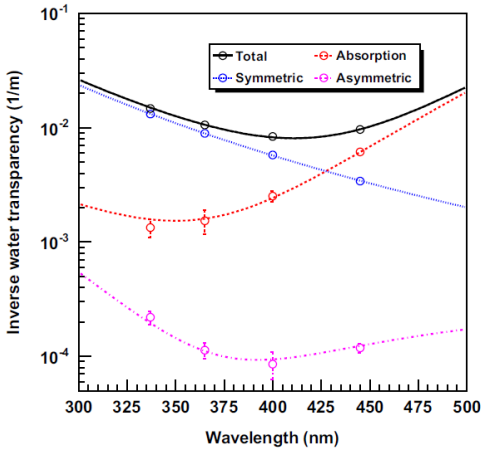


Figure 3.9: The left are the typical observed coefficients of water parameters and fitted functions shown as dashed lines. The horizontal axis is the wavelength and the vertical axis are the coefficient values. The right is the time variation of the coefficients in the early dates of SK-IV as measured by the laser injection system. The horizontal axis shows is the date, with the symmetric scattering in purple, asymmetric scattering in black, and absorption shown in blue.

volume (>2m from any wall) and within 250 cm of the stopping point of the muon.

3. The number of hit PMTs must be greater than 50

Around 1500 decay-e events per day pass these selection criteria. For these selected events, the distance between each hit PMT and the event vertex is calculated.

After selecting the candidate events, the PMT hits undergo a cleaning process to remove errant hits[53]. Hit PMTs must be between 32°and 52 °of the event's reconstructed direction as shown in Fig 3.10, corresponding to ±10°of the Cherenkov cone. Also, only PMTs with a t-t.o.f. calculation within 50 ns are considered. The ring is then split into 36 10°bins along the azimuth (also shown in Fig 3.10) with PMTs from each bin grouped together to get:

$$\Delta N^i_{eff} = \sum_{j=1}^{N_i} e^{d_j/\lambda_e} \tag{3.18}$$

where λ_e is the relative water transparency, N_i is the number of hit PMTs in the i^{th} 10°, and N^j_{eff} is the effective hit value of the PMT, and d_j is the distance of the j^{th} PMT in

the i^{th} slice of the cone. Assuming the size of each 10°bin is small, the average distance and total N_{Eff} are:

$$\overline{d}_i = \sum_{j=1}^{N_i} d_j/N \tag{3.19}$$

$$N_{eff}(\overline{d}_i) = \sum_{j=1}^{N_i} N_{eff}^j \tag{3.20}$$

and using the average distance, the exponential from Eq 3.3.2 can be factored out of the sum to get:

$$\Delta N_{eff}^i = N_{eff}(\overline{d}_i) \times e^{\overline{d}_i/\lambda_e} \tag{3.21}$$

With each of the 10°bins being corrected for the attenuation due to the WT, each of the ΔN_{eff}^i should be equal. With all of the ΔN_{eff}^i from a day's decay-e collected, they can be averaged over the Michel spectrum. The aggregate $\ln(\Delta N_{eff}^i)$ distribution can then be binned in d, with the resulting distribution then fit linearly to find the WT attenuation length, λ_e, from the slope. The fit is made between 1200 and 3500 cm. This λ_e is averaged over a week of time, and used as an input to track the WT coefficient α_{abs} through the WT parameter $A_1(t)$ discussed in the previous section.

For low energy events, the WT also plays an important role in the offline reconstruction of events. It affects all aspects of the reconstruction, and requires a week before and a week after the day one wants to simulate WT conditions. The WT measurement is also affected by the increasing PMT gain, as this changes the N_{eff}^i hit distributions for the decay-e. Fig 3.11 shows the effect of not correcting N_{eff} based on the WT.

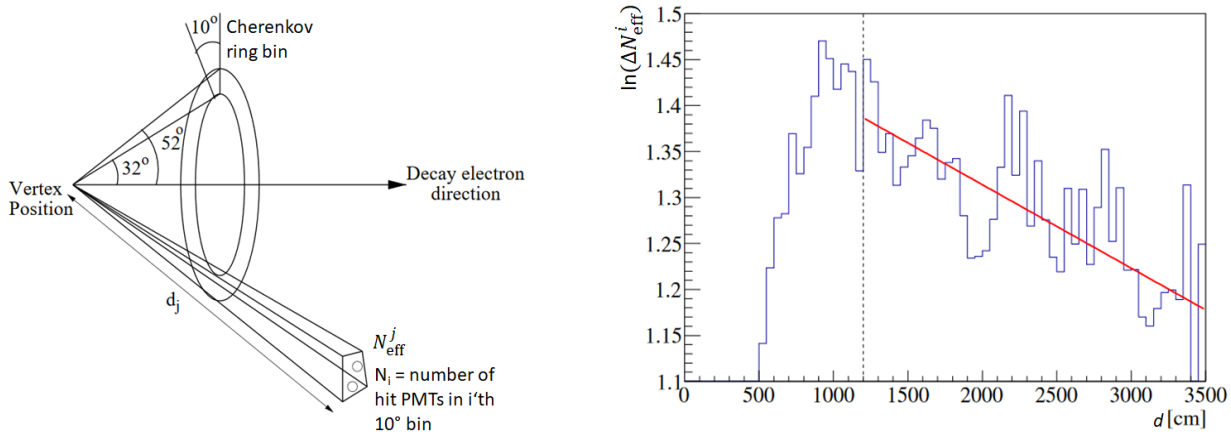


Figure 3.10: Left: An illustration of the various parameters used in the decay-e WT calculation. Right: A typical Δ^i_{eff} distribution plotted in d from the decay-e for a day. The red line is a linear fit to the distribution from 1200cm $\leq d <$ 3500 cm for $\ln(\Delta^i_{eff})$, which is used to obtain the WT value.

3.3.3 Position Dependence and Top-Bottom Asymmetry

After applying the corrections for WT to the different calibration sets, there still remains a difference between the number of hits seen at the top and bottom of the detector. Both the Ni source and Xe lamp calibration data show the average hit rate for the top region to be ∼5% less than the bottom region. This difference is called the Top-Bottom Asymmetry (TBA) and can vary in time. It is believed the reason for this is the bottom having the "freshest" water, as the water is injected into the detector at the bottom, and that the top of the detector has the oldest water, as the water flows upward and is removed from the top. The TBA is applied separate from the WT correction event though it is believed to be related. Fig 3.12 shows the effect the TBA has on the LINAC and DT calibrations (these calibrations are discussed in Sec 3.4.

The TBA is determined using the monthly Ni source calibration along with the real-time Xe lamp system which occurs once a second. The TBA is formed by taking the difference between the average hit rate (H) at top minus that at the bottom, divided by the

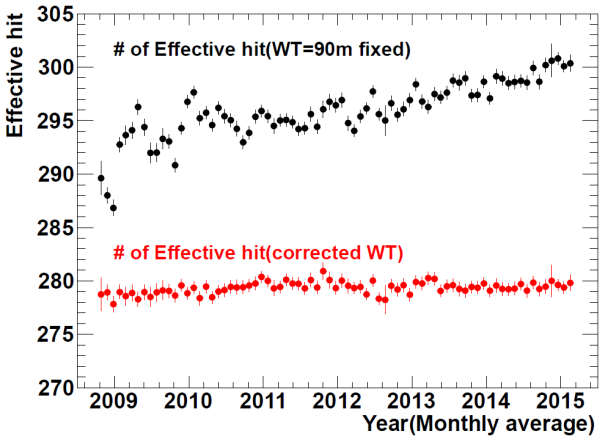


Figure 3.11: Time variation of the number of effective hit PMTs, N_{eff}, as calculated with constant (varying) WT values shown in black (red).

average in the barrel:

$$\text{TBA} = \frac{\langle H_{top} \rangle - \langle H_{bottom} \rangle}{\langle H_{barrel}} \tag{3.22}$$

For the Xe lamp data, the hit rate in the above equation is exchanged for the average charge Q. The analysis method is similar to that of the QE measurement described in Sec 3.2.3, the difference being the QE value is added to Eq. 3.2.3:

$$h^i_{corr,tba} = \frac{h_i \times r_i^2 \times \text{QE}_i}{a(\theta_i)} \tag{3.23}$$

Using $h^i_{corr,tba}$, the average hit rate at the top, barrel, and bottom are obtained and applied to Eq. 3.3.3. Fig. 3.13 illustrates the result of TBA using the laser and Ni source calibrations, and the two calibrations are consistent with each other.

In order to take into consideration the position dependence of the TBA, the absorption coefficient (α_{abs} is modified to:

$$\alpha_{abs}(\lambda, z) = \begin{cases} \alpha_{abs}(\lambda) \times (1.0 + \beta z) & z \geq -1100\text{cm} \\ \alpha_{abs}(\lambda) \times (1.0 - 1100\beta) & z < -1100\text{cm} \end{cases} \tag{3.24}$$

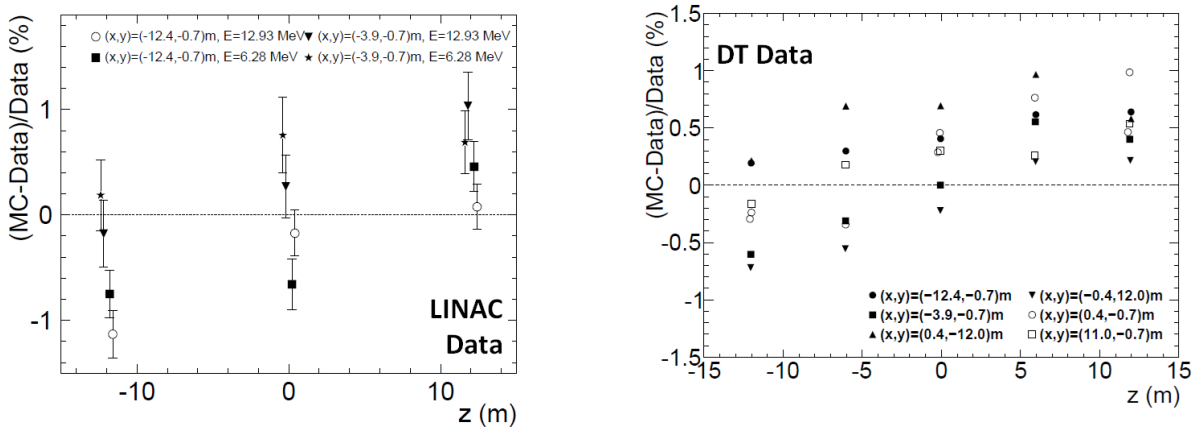


Figure 3.12: The z-position dependence for the LINAC (left) and DT (right) calibrations that determine the absolute energy scale of SK-IV. Both plots compare the data and MC information without the TBA correction applied. The LINAC shows z-dependence for two calibration holes and energy, and the DT data shows the z-dependence at all primary DT positions.

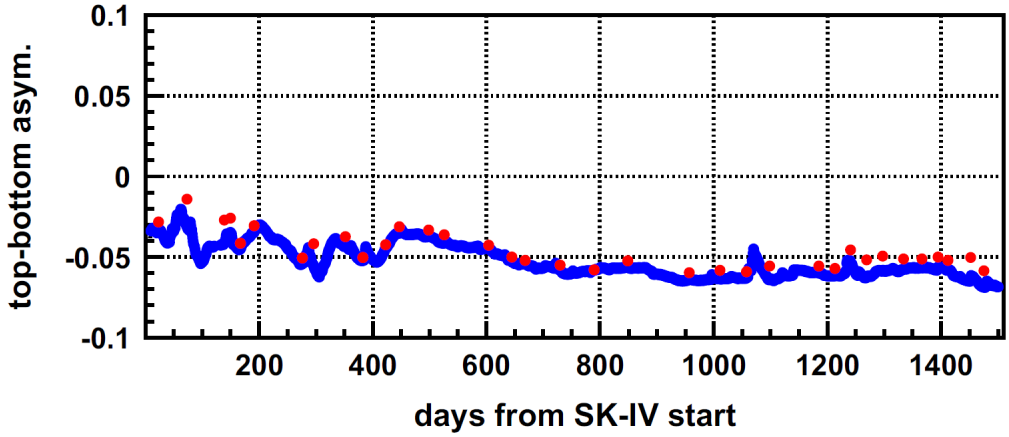


Figure 3.13: The time variation of TBA for Xe lamp (blue) and Ni source (red) data during the first half of SK-IV data taking.

The cutoff at -1100 cm in α_{abs} corresponds to the boundary of the convection that occurs at the bottom of the detector. Ni MC simulations with various values for the parameter β are prepared, and the results of the various MC simulations are compared to the Ni calibration data. Finally the paramter β is determined as:

$$\beta = (-0.164 \times \mathrm{TBA}^2 - 3.676 \times \mathrm{TBA}) \times 10^{-3} \tag{3.25}$$

3.4 Absolute Energy Scale Calibration

The calibration for the absolute energy scale is the most important item to measure the energy spectrum of solar neutrinos. In order to determine the absolute energy scale, the low energy group of SK uses LINAC and DT calibration. The information of the absolute energy scale is necessary to properly simulate the detector response, and the detector simulation should be tuned with it.

3.4.1 LINAC

A medical LINear ACcelerator (LINAC) was acquired from a a hospital and converted into a calibration tool for SK in 1996[54]. The LINAC has a permanent home in an area attached to the SK dome. The LINAC delivers electrons through a series of pipes and magnets into the detector at different locations. The setup is designed to deliver mono-energetic electrons into the detector at very low intensity.

The electron gun uses a filament at the beginning stage to produce the electrons for the beam. A series of magnets steer the beams direction with respect to the beam pipe, and eventually down and into the detector. The D magnets are more powerful magnets used to navigate the turns in the beam pipes, while most of the other magnets are used for fine

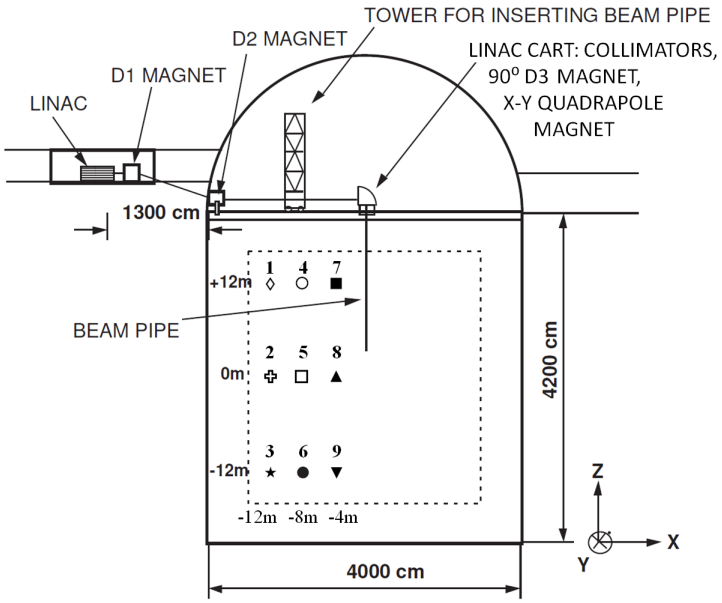


Figure 3.14: Schematic of the LINAC geometry with respect to the detector. Data taking position are the numbered shapes, the dashed line shows the fiducial volume boundary, and labels for major beam tuning components.

tuning of the beam direction as well as unfocusing and refocusing the beam for its 90°turn into the detector. The electron momentum is selected by the first D magnet (D1), and the D1 magnet is calibrated at the beginning of each LINAC data taking day and is fixed to set values based on desired momentum to ensure consistency. The true beam momentum is measured by germanium detector as well. The 2016 LINAC calibration required partial use of a second Ge detector while one was repaired, and also the electron gun filament required replacement.

Beam monitoring is performed with the use of scintillation counters before D2, D3, and at the end of beam pipe. Workers use these to find the beam center and adjust the magnets, collimators, and gun voltage to achieve a focused beam of the desired intensity. There are truly five counters at the end of the beam pipe, one trigger and four veto counters, used to trigger and veto events during data taking. Full LINAC calibration takes data at 9 positions along the -x-axis, corresponding to different combinations of X = -388.9cm, -813.1cm, -1237cm and Z = +1197cm, -6cm, -1209cm. LINAC calibrations are time and labor

intensive, usually occurring over weeks and some calibrations used only a subset of these positions, omitting the x=-813.1cm position, in order to . These combination of positions allow for many different areas of the detector to calibrated and check differences for example of events near the center of the detector and those near the sides.

The germanium detector is calibrated using several sources:

- Natural radioactivity (includes ^{40}K at 1.46 MeV and ^{208}Tl at 2.61 MeV)
- Sealed ^{40}K (1.17 MeV and 1.33 MeV)
- Sealed ^{40}K (0.662 MeV)
- Ni-Cf source used in Ni calibration (γs up to 9 MeV

The results from the Ge measurement of the beam at each position are used in SKDETSIM. The MC and LINAC data are compared to determine the uncertainty in the electron beam energy. For larger electron momenta, the electrons can fully penetrate the germanium crystal and reach inactive materials. This effect is also simulated in MC. LINAC calibration has measured the energy scale to 1%, and since the injection point is well defined, the data can be used to check the accuracy of vertex and direction reconstruction software. A drawback however is although energy and vertex can be varied, only a singular direction can be injected into the detector. More information can be found in [54].

3.4.2 DT Generator

A deuterium-tritium (DT) generator is used to check the absolute energy scale for low energy events on a more frequent basis than LINAC calibration[55]. A full LINAC calibration takes around 20 days, with 2 shifts/day, and 3-4 workers/shift, while DT calibration is performed in around 3 days, with 1 shift/day and 4 workers per shift. The associated deadtime with the

LINAC is much greater due to the work involved in changing beam pipe positions within the detector. The comparably compact size of the DT generator also allows for more calibration ports to be tested, and routine calibrations probing 25 positions within the detector. The DT generator generates fast neutrons at various locations within the detector, which in turn produces ^{16}N in the water.

The DT generator is lowered into the detector at various locations by cable and crane, and is controlled by computer. The main positions for DT calibration are at (X,Y) = (0,0), (±12,0), and (0, ±12), with five Z positions taken for Z = ±12, ±6, and 0. Near LINAC calibration, data will also be taken at the additional two calibration ports used by the LINAC at X = -813.1cm and -388.9cm. When the generator is "fired", it isotropically produces 14.2 MeV neutrons, and the neutrons interacting with the oxygen to produce ^{16}N:

$$\begin{aligned} {}^{2}\mathrm{H} + {}^{3}\mathrm{H} &\rightarrow {}^{4}\mathrm{H} + n \\ n + {}^{16}\mathrm{O} &\rightarrow {}^{16}\mathrm{N} \end{aligned} \tag{3.26}$$

The ^{16}N has a half-life of 7.13s. ^{16}N decays 66% of the time with a β (4.3 MeV) and γ (6.1) MeV and 28% of the time a single β with an endpoint of 10.41 MeV. A veto signal is activated just prior to firing of the generator, and remains active after firing while the torpedo is raised as much as possible to reduce shadowing. Due to space constraints, the torpedo cannot be raised as high at the Y = ±12m positions, and therefore shadowing is increased. Firing of the DT generator produces $\sim 10^6$ neutrons, which in turn produce $\sim 10^4$ ^{16}N events. Maintaining the veto during the initial firing of the generator also helps to prevent DAQ crashes from the high initial event rate. Once data taking is finished for a firing, the torpedo is lowered again and the process is repeated until 20 firings occur at a single vertex position.

Once the entire process is finished, the peak and shape of the reconstructed energy distribution is compared to MC to estimate how the energy scale depends on the position,

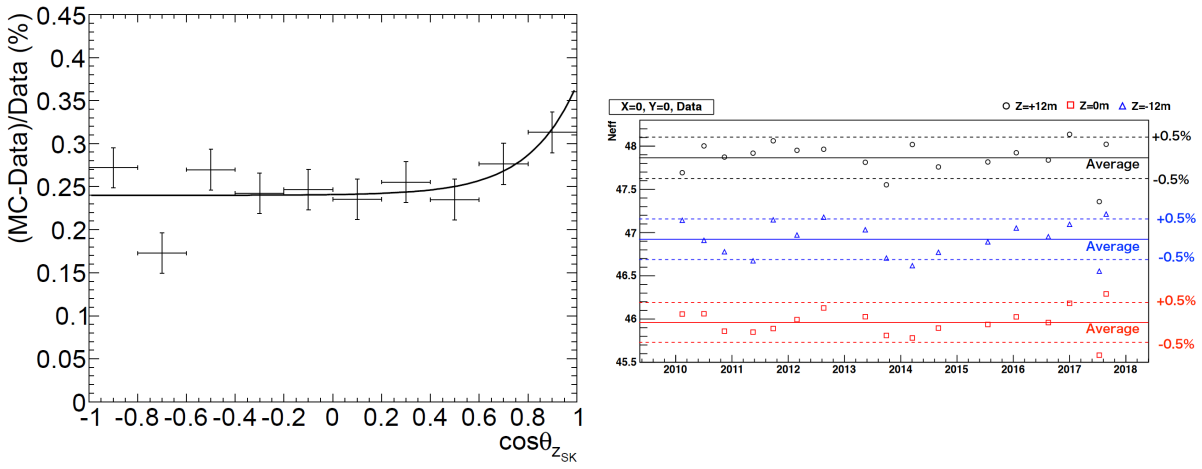


Figure 3.15: Left: The difference in mean N_{eff} between the DT data and MC simulation for SK-IV, binned in zenith angle. The combined uncertainty due to position dependencies correlated to the zenith direction is $\sim \pm 0.1\%$ after subtracting the absolute offset. Right: The stability of the SK-IV energy scale using DT calibration data after gain correction. The data is from the center most calibration port corresponding to the z-axis in detector coordinates. The N_{eff} is stable within $\pm 0.5\%$

direction, and time-varying detector qualities. More information can be found in [55].

Chapter 4

SK Event Reconstruction

4.1 Muon Track Fitter: Muboy

SK has a plethora of muon fitters, but only two of them are fast enough to keep up with the rate of downward going cosmic rays, mufitpe and Muboy. Mufitpe was used for almost all SK analysis, but for this analysis Muboy is used. Muboy gives more accurate results for the fitting of muons, and is able to differentiate between different categories. For non through-going muons mufitpe struggles to successfully fit a track, and as a result there is a higher rate of failed or unreliable fits. The muon parameters from Muboy were used in order to tag neutrons after a muon, as well as the update to the standard spallation cut. Muboy is a general-purpose muon fitter that will give both entrance and direction parameters for a muon, but also fits four different cases of muons. The different categories of muons are single through-going (single muon passing wholly through the detector, ∼82% of muons), stopping muon (single muon that stops within the detector, ∼7% of muons), multiple muons (muon bundle passing through the detector, can have multiple track fits, ∼7% of muons), and corner clippers (very short track length, clipping the corner of the detector, ∼4%). It

also is not dependent on OD hits to make a fit but will use OD information when determining whether a muon stops within a detector.

Muboy works by first applying a PMT cleaning algorithm. An initial p.e. cut is made based on the number of hits within an event, with a minimum requirement of 2.0 p.e. for <8000 hits, 2.5 p.e. for 8000-10000 hits, and 3.0 for >10000 hits. This is followed by a nearest neighbor cut where each PMT is required to have at least N nearest neighbors hit within 300cm of it as for the following number of p.e. (minimum number of hit PMTs): >7500 (5), 5000-7500 (4), 2500-5000 (3), 500-2500 (2), and <500 (1).

If there are less than 10 PMTs left after cleaning, the event is classified as a non-fit. For the OD, only the charge needs to be positive.

After cleaning, Muboy takes the entry point of the muon as the earliest PMT with at least three nearest neighbors hit. If there is none with three nearest neighbors hit, the required number of nearest neighbors is reduced by one until a match is found. The first guess for entry time, t0, is the timing of the PMT hit selected as the entry point. The first guess for direction is taken to be the patch of 9 PMTs (3x3 grid) with the highest number of p.e. Generally, when a muon is leaving the detector, the highest concentration of Cherenkov light is seen. This is because of the rings of light being emitted along the track get progressively smaller as they are directed into the exit point. If the patch of 9 PMTs is centered on the entry PMT, the q-weighted center of mass of tube hits is instead used to get the initial direction. A causality cut is put in place as a further cleaning process. Taking the entry point and time, all PMTs hit must have hit times no faster than the speed of light (29.94 cm/ns), and no slower than light speed in water (29.94/1.395 cm/ns) from the entry point. Scattered light will be late, and tubes hit by multiple muons will have early light. Some leeway is given to these values, and light must arrive at a PMT within 18-34 ns/cm with respect to the entry PMT and time. If the number of hits >33 cm/ns from the entry point and time is more than 45, they are classified as the first category of multiple

muons, and the causality window is tightened to 19.5-32 cm/ns if the muon is categorized as a multiple muon in this way to try to better isolate the primary track. As with the previous stage, if <10 tubes are hit it is a Muboy non-fit.

The initial direction fit is wiggled and checked against a likelihood dependent on the fraction of PMT hits consistent with the Cherenkov cone and the timing residual for the PMT hit with respect to the expected timing of the track. Events previously not classified as multiple muons are allowed to have t0 changed, and it is allowed to do so by calculating it from an 800 cm time window of the test track. With this new track, Muboy checks again for too much early light, this time if more than 35 PMTs are early by the 33cm/ns criteria, the muon is classified as a type two multiple muon. For non-multiple muons, the time and direction fitting are complete. Multiple muons have hits associated with the primary track removed by removing all events within a 400 cm time residual of the primary track. A fit similar to what is described above is repeated for the remaining hits varying the entry parameters and maintaining direction. If another reasonable fit is found, the procedure is repeated until up to 10 tracks are found. There is a chance for the muon to be classified as a multiple muon and only one track is found. This could be due to the tightness of the bundle, or if the muon showers when passing through the detector. Showers would produce out of time light inconsistent with the principal direction.

For single muons, they are checked for if they are corner clippers or stopping muons. For corner clippers, events with less than 2000 ID hits near the edge of the top or bottom and leave the adjacent edge are checked. If they are at the top of the detector ($z_{entry} > 1750$cm) a formula based on entry and direction x and y components is used with track length to determine if it is a corner clipper, and at the bottom ($z_{entry} < -1600$cm), if track length is <4m it is classified as a corner clipper. For stopping muons, the amount of light seen at the ID exit PMT and light seen in the OD is used to classify the case. If <300 p.e. are seen within 400 cm of the exit ID PMT or <400 p.e. and <30 OD p.e. near the exit

point. If <200 p.e. is seen in the ID near the exit PMT, it is more or less always a stopping muon, but 300 p.e. is used as a good guess. If the muon is classified as a stopping muon, the stopping point of the muon is determined by forming a dE/dx (light loss along the track) for the muon. Going along the track, if the average dE/dx for a 150 cm segment falls below 40% of the average dE/dx for the first 150 cm of the track, then that point is considered the stopping point for the muon.

Studies were performed to judge the accuracy of Muboy for fitting muons. A sample of ∼2000 muons were both fit by Muboy and also scanned by eye. 1728 of the events were determined by Muboy to be single through-going muons, and only 9 of those were thought to be something else by eye scan. 74 were found by Muboy to be corner clippers, only one was found to be a stopping muon by hand scan. For the 129 events classified by Muboy as multiple muons, 17 were thought to be through going by eye-scan. And finally, the 113 found by Muboy to be stopping muons, roughly a third were thought to be through-going muons, this being the biggest discrepancy.

Resolution studies found the entry point resolution to be around 100cm for all types of muons, with the exception of multiple muons with more than 3 tracks. The direction resolution on the order of 6 degrees for all types. Muboy overall functions well as a robust fitter with reasonably accurate fits, with faster fits than the other precision fitters within SK. More information can be found in [56] and [57].

4.2 Event Reconstruction

4.2.1 Vertex Reconstruction

Low energy electrons (<20 MeV) from interactions will only travel a short distance in the water (∼10 cm) and are therefore treated as a point source in the case of vertex fitting, especially when compared to the size of the detector and spacing between PMTs. The main fitting algorithm used within the low energy analysis of SK is BONSAI (Branch Optimization Navigating Successive Annealing Iterations)[58] with a vertex resolution of ∼50cm for events >16 MeV. BONSAI works by cleaning up PMT, hits and then fitting many vertices to four hit combinations from PMTs to make an initial guess. A t - t.o.f. (time minus time of flight) subtraction is made for all PMT hits for a given test vertex, and a likelihood is formed and tested against is made for the expected timing residuals. t - t.o.f. works off the assumption the light from the event travels directly to the PMT from event vertex. Due to scattering, the peak of the t-t.o.f. distribution is smeared towards later times, requiring for a larger acceptance window compared to what is expected from solely hardware constraints. The light The likelihood for the timing residuals was formed from using LINAC (LINear ACcelerator) calibration data for the signal PDF, while the background PDF is assumed to be flat. Each guess on initial vertex is compared against 12 other nearby points on a grid, and the result with the largest likelihood is made the new center of the search grid for that guess. A range of acceptable likelihoods is used as a criteria for whether a path is pursued further at each iteration. This is repeated for shrinking grid size and likelihood ranges until differences around 1cm are reached. This fitter is used both for the solar analysis as well as the neutrons found after muons in WIT.

4.2.2 Direction Reconstruction

As mentioned before, incoming neutrinos interacting within the detector through elastic scattering create recoil electrons. These recoil electrons mostly retain directionality with respect to the incident neutrino (within about 15 degrees). For direction fitting, only PMTs with a t-t.o.f. within a 30 ns window are considered. From the reconstructed vertex, we assume azimuthal symmetry of the Cherenkov light and many direction vectors are tested against a likelihood formed from MC. The likelihood function is broadly peaked around 42 degrees, corresponding to the Cherenkov angle for an ultra-relativistic particle in water. The reconstructed direction is the one that returns the greatest likelihood value when summing over hits. Each hits likelihood contribution is the angular likelihood, as well as a QE correction for the individual PMT due to the angle of incidence. The directional resolution improves with event energy, as more hits allows for a more precise estimation of the event direction. However this is limited by multiple Coulomb scattering for all events, setting a limit to the overall angular resolution to ∼20 degrees. For the lowest energy events analyzed in SK, the directional resolution is ∼45 degrees.

4.2.3 Energy Reconstruction

4.2.3.1 $Neff$ Calculation

For the energy reconstruction, simple counting of PMT hits is used. Low energy events, like those in the solar analyses, it is typical for an event to have PMTs only hit by a single photon. Noise is reduced by requiring the time residuals of PMT hits (t - t.o.f.) to be within a 50 ns window. The window is chosen to select the most amount of event hits, starting slightly before the event timing and extending past it. Most of the window situated after the event time, due to light more likely being late due to scattering. This number of hits is referred to

as N_{50}. The actual energy estimator, N_{eff}, is a correction to the N_{50} calculation, which takes into consideration geometry and other factors for the individual hit PMTs to lessen position dependence and other systematic effects. N_{eff} represents the number of photoelectrons (pe) seen for an event, with ∼6pe = 1 MeV reconstructed. To calculate the individual N^i_{eff} for a single PMT, the following equation is used and the total N_{eff} equal to the sum over all hits comprising N_{50}:

$$N^i_{eff} = (X_i + \epsilon_{tail} - \epsilon_{dark}) \times \frac{N_{all}}{N_{alive}} \times \frac{R_{cover}}{S(\theta_i, \phi_i)} \times e^{r_i/\omega(t)} \times \frac{1}{QE_i} \tag{4.1}$$

$$N_{eff} = \sum_{i}^{N_{50}} N^i_{eff} \tag{4.2}$$

The occupancy correction term (X_i) is used to correct for events in which the single photon assumption for hit PMTs would be violated. This could happen possibly from an event occurring near a PMT and the Cherenkov cone direction being pointed towards the PMT. In a case like this, it would be expected neighboring PMTs would be hit as well, and the more neighboring PMTs hit, the brighter the spot located at the i^{th} PMT, or missed photoelectrons. If no neighboring PMTs are hit, then it would be a dim region and therefore the correction is roughly equal to 1. If all 8 neighboring PMTs are hit, then the correction is set to a value of 3, which was determined by MC studies. The full correction is as follows, where the log term was determined from Poisson statistics:

$$X_i = f(x) = \begin{cases} -\log(1 - x_i) & x_i < 1 \\ 3.0 & x_i = 1 \end{cases} \tag{4.3}$$

$$x_i = \frac{\#\ PMT\ hits\ in\ 3x3}{9} \tag{4.4}$$

The two ϵ terms correspond to the corrections from late hits and dark noise. The dark noise correction takes the average dark noise rate for the PMTs in a run to calculate the number

of dark hits within the 50 ns. This correction is then spread out over the 50 hits:

$$\epsilon_{dark} = \frac{N_{alive} \times R_{dark} \times 50ns}{N_{50}} \tag{4.5}$$

while the tail correction adds up the additional hits in an extended time window of 100 ns (N_{100}), minus the expected dark noise rate again, and spread equally to the N_{50} contributing hits:

$$\epsilon_{tail} = \frac{N_{100} - N_{50} - N_{alive} \times R_{dark} \times 50ns}{N_{50}} \tag{4.6}$$

N_{all} and N_{alive} are the correction for the time dependent PMT coverage. Over time, especially over the length of an experiment like SK, PMTs will die or become faulty. This corrects the ratio of the actual coverage and the expected coverage.

S is the correction for the effective photocathode coverage for the PMT. Depending on the direction of the incident photon on a PMT, there can be shadowing effects from neighboring PMTs. The function is derived from MC simulation of each phase of SK and is used the determine the directional cross sectional area of each PMT as viewed from angles θ and ϕ, where θ and ϕ are traditional spherical coordinates centered at the PMT. R_{cover} corresponds to the overall photocathode coverage of the detector depending on the phase of SK. For SK-IV, $R_{cover} = .404$.

The exponential term is the water transparency correction, where r_i is the distance of the reconstructed vertex to the PMT, and $\omega(t)$ is the time dependent water transparency value for that run. Finally, the QE_i term is the quantum efficiency factor for the i^{th}.

4.2.3.2 SK-IV Gain Correction

Over time, the absolute gain for the PMTs in SK-IV has steadily increased by $\sim 2-3\%$ per year relative to the initial SK-IV value. With the length of SK-IV and the extent of the drift

gain in the latter half of the phase, the low energy group of SK applied a time-dependent gain correction[59]. The fractional difference in the hit rate (HR) and the gain (G) are treated as linearly correlated since the gain increase does not exceed 20% throughout the SK-IV period, and is expressed as:

$$C_{\mathrm{G}} \cdot F_{\mathrm{G}} = F_{\mathrm{HR}} \tag{4.7}$$

$$F_{\mathrm{HR}} \equiv \frac{\mathrm{HR}(t) - \mathrm{HR}(0)}{\mathrm{HR}(0)} \text{ and } F_G \equiv \frac{\mathrm{G}(t) - \mathrm{G}(0)}{\mathrm{G}(0)} \tag{4.8}$$

This gain correction parameter C_{G} is assumed to be uniform across time and PMT batches, even though the gain increasing amplitudes vary. It is then used to to make a correction to the two components of the N_{eff} calculation that would be affected by this increasing gain, the occupancy (X_i) and PMT dark noise parameter (ϵ_{dark}). To obtain the C_{G} factor, change in the decay-e spectrum from muons is examined and applied to N_{eff}, assuming an unchanged $\omega(t)$ function. The occupancy term $X(x_i)$ is modified to account for the different types of PMTs and includes their individual QE values:

$$X(x_i) \rightarrow \frac{X(x_i/\mathrm{QE}_i)}{1 + F_{\mathrm{G}} \times C_{\mathrm{G}}} \tag{4.9}$$

And the PMT dark rate term is also modified on a PMT by PMT basis instead of an averaged dark rate for the whole detector, with the additional application of the gain correction.

$$\epsilon_{dark} \rightarrow \frac{\epsilon^i_{dark}}{1 + F_{\mathrm{G}} \times C_{\mathrm{G}}} \tag{4.10}$$

The water transparency function, $\omega(t)$ if then recalculated with a new N_{eff} using the corrected occupancy and dark noise rate to make $\omega^{'}(t)$, and N_{eff} is calculated a final time with all three corrections for a relatively time-invariant N_{eff} distribution. This whole process stabilizes the effect of the changing gain on the water transparency, energy scale, and other affected parameters. Fig 3.11 shows the change in the N_{eff} for the decay-e spectrum of

stopping muons, previously discussed in Sec 3.3.2.

4.2.3.3 Calculating Energy from N_{eff}

With the calculation of N_{eff} and its corrections, the reconstructed energy of strictly electrons and positrons is acquired. There are two energy regions to the calculation, separated at ~ 25 MeV, or an N_{eff} of 189.8. The equations for the two regions are as follows:

$$\mathrm{E}_{rec} = \begin{cases} [0.82 + 0.13N_{eff} - 1.11 \cdot 10^{-4}N_{eff}^2 & \\ \quad + 1.25 \cdot 10^{-6}N_{eff}^3 - 3.42 \cdot 10^{-9}N_{eff}^4] & N_{eff} < 189.8 \\ 25 + 0.138(N_{eff} - 189.8) & N_{eff} \geq 189.8 \end{cases} \tag{4.11}$$

The values for the coefficients above are for the SK-IV data set, and are obtained through LINAC calibration data and associated MC.

4.3 Evaluating Fit Quality of BONSAI

There are two main parameters used in different aspects of the solar analysis to test how well the fit for BONSAI was performed. The other is a Kolmogorov-Smirnov (KS) test variable on the direction vector of the fit, and tests azimuthal symmetry of PMT hits.

4.3.1 BONSAI Vertex Reconstruction Goodness

The goodness parameter for the BONSAI vertex reconstruction ($g_{VT}$is based on the timing information of hit PMTs. It uses two timing residual Gaussians to form a weighted average

Gaussian:

$$g_{vertex} = \sum_{i}^{N_{hits}} \frac{G_i(\omega)}{\sum_i^{N_{hits}} G_i(\omega)} \times G_i(\sigma)$$
$$G_i(u) \equiv \exp\left[-\frac{1}{2}\left(\frac{\tau_i - t_0}{u}\right)^2\right] \tag{4.12}$$

where τ_i is the t - t.o.f. calculation for the individual PMT hit to the reconstructed vertex, t_0 is the peak of the t - t.o.f. timing distribution, $\omega = 60$ns and is used to encompass all of the PMT hits for the event, and $\sigma = 5$ns and corresponds to the characteristic PMT timing resolution for single-pe statistics. The goodness calculation ranges from 0 to 1, with higher values corresponding to a better vertex reconstructions.

4.3.2 KS test on Direction Reconstruction

The goodness parameter for the directional fit (g_{dir}) is a KS test on directional symmetry checks to see if all PMT hits are evenly spaced from 0 to 2π in the azimuthal angle (ϕ_i) around the direction vector. The hits are ordered from 0 to 2π and checked against a uniform distribution for angles, $\phi_i^{uniform} = \frac{2\pi i}{N_{hits}}$. The minimum difference between the uniform set and the data set is subtracted from the maximum difference and normalized to retrieve the answer:

$$g_{dir} \equiv \frac{\max[\phi_i^{uniform} - \phi_i^{data}] - \min[\phi_i^{uniform} - \phi_i^{data}]}{2\pi} \tag{4.13}$$

Larger differences in g_{dir} corresponds to divergence from a Cherenkov-like events hit pattern, and indicates background events or noisy events such as hit cluster events. The value for g_{dir} also ranges from 0 to 1, however 0 indicates a better result of the test statistic as it corresponds to a more even Cherenkov ring.

Chapter 5

Backgrounds and Data Reduction

Ignoring detector efficiencies or energy thresholds, SK expects ~ 386 ^{8}B ν_e elastic scattering interactions a day within the detector. However, once detector efficiencies, energy thresholds, and everything else are taken into consideration, this number decreases by more than an order of magnitude. For the SK final sample, roughly 20 solar neutrino events are seen a day. One of the major limiting factors are the remaining backgrounds that are unable to be removed physically, despite experimenters best efforts. Since these backgrounds cannot be entirely be removed physically, effort has to be made with software analysis to try and remove these and improve the purity of the final sample.

5.1 Cosmogenic Backgrounds

5.1.1 Cosmic Rays

Cosmic rays are high-energy protons and nuclei traveling through space near the speed of light. Some originate from the sun, within our galaxy, and even distant galaxies[60]. The

full source of cosmic rays is not well known, and many experiments today search for their origins. For the sake of SK, the cosmic rays originating from within our galaxy are what are responsible for most of the cosmogenic backgrounds. The origin of these cosmic rays are believed to be from supernova remnants shock acceleration. Once the particles reach the Earth's atmosphere, the particles produce showers of secondary particles, primarily pions, with a smaller chance for kaons and baryons. If the pion is not charged, it will decay into two photons, with the photons creating an electromagnetic cascade. If the pion is charged, it will overwhelmingly decay leptonically into a muon and muon neutrino:

$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu} \text{ or } \pi^{0} \rightarrow \gamma + \gamma \tag{5.1}$$

For π^{+} the result will include a neutrino, while for $\pi-$ it will be an anti-neutrino. The branching ratio for π to μ is .999877. The muons produced in these showers are able to penetrate deep underground to the detector, despite its 1000m overburden. However, the rate is greatly reduced, as the detector only observes muons at a rate of 2.2Hz [$6\times10^{-8}/(\text{cm}^2\times s\times sr)$]. Muons need an energy greater than 1.3 TeV in order to penetrate to SK. This allows for SK to have the ability to see solar neutrinos. However, even with the much-reduced rates of muons within the detector, there still is a sizable background

5.1.2 Spallation

Spallation is the process in which a particle impacts with matter, and the result of the collision is the expulsion of large numbers of nucleons. At SK, even though the muon rate is low the muons discussed above have a chance to interact and shower within the detector. When these showers occur, like the air showers, there are many secondary particles produced as the shower cascades. The hadronic components of the shower can break apart and capture on nuclei, creating unstable isotopes. This process is an example of spallation. These

radioactive isotopes are a source of background in the solar neutrino energy region for SK. Spallation for the rest of this paper will refer specifically to the radioactive isotopes located within water produced through the spallation process.

Spallation from these muons is the dominant background between 5.49 and 19.49 MeV kinetic, and the spallation concerned with in SK usually β-decay

$$ {}^{\mathrm{M}}_{\mathrm{N}}\mathrm{X} \rightarrow {}^{\mathrm{M}}_{\mathrm{N}\mp 1}\mathrm{Y} + \beta^{\pm} \tag{5.2} $$

which will feign a signal similar to a low energy neutrino event. The spread of energies from spallation also covers most of the solar neutrino spectrum within SK, so proper tagging and removal is imperative.

Muons passing through will always shed energy in the form of delta rays. The more energy the muon deposits within the volume the muon can produce an Electromagnetic (EM) shower through pair production, and eventually with high enough energy deposition can make hadrons. Whether or not the shower is primarily a hadronic shower or an EM shower, there will be a large component of light produced within the shower. $\sim$⅓ of pions produced in a shower are neutrally charged, and decay into two photons:

$$ \pi^0 \rightarrow \gamma + \gamma \tag{5.3} $$

These gammas create an EM cascade in the water, producing a significant amount of light. This light is what is traditionally used in order to tag potential spallation causing showers. Discussed later in this analysis, details of incoming muons and candidate events are used to form a likelihood to decide whether or not to cut an event as likely spallation.

Beacom and Li had made calculations to the expected spallation product yield within SK in [61] and a study was performed to measure the radioactive isotope production

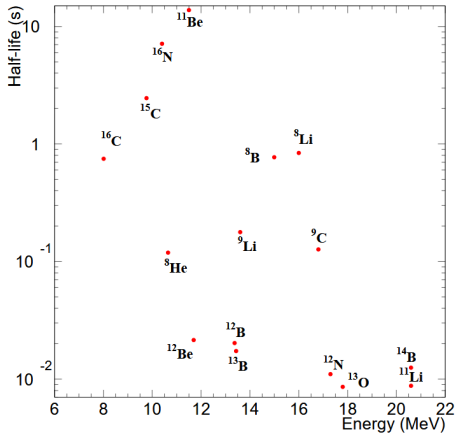


Figure 5.1: Scatterplot of spallation products expected to occur in SK, showing the half-life and energy of each isotope. Note that the energy values are the maximum values, and the vertical axis is in a logarithmic scale[1]

from spallation in SK-IV by Zhang [4]. Table 5.1 shows basic information for spallation that can be found in SK, with measured rates from Zhang if they were obtained. Fig 5.1 shows the energy and decay half-life for many isotope. Of note from the table and figure is ^{16}N, which is the most abundant spallation isotope within SK, while also having one of the longest lifetimes. This makes its removal troublesome for many analyses.

5.2 Detector Material Induced Backgrounds

Although effort was made to minimize radioactivity of the equipment and materials used in the construction of SK, the detector is not completely radiopure. There are still radioactive backgrounds from the PMTs and emanating from the support structure in the detector. Radioactive isotopes that decay with β's or γ's mimic solar neutrinos interactions, and cutting these events comes at the cost of part of the solar neutrino signal. The first major source of radioactive background is from Radon (Rn) from the Uranium decay chain being introduced to the detector, from either the water system, PMTs, and/or the support structure. Uranium is contained in detector components containing heavy elements, and eventually emit

Isotope	$\tau_{1/2}$ [sec]	Decay Mode	Kinetic Energy [MeV]	Primary Process	Measured Rate [kton^{-1}day^{-1}]
^{8}He	0.119	β^-	9.67 + 0.98(γ)	^{16}O(π^-, ^{3}H + 4p + n)^{8}He	< 1.4
		β^- n	16%		
^{8}Li	0.838	β^-	$\sim$13	^{16}O(π^-, α + ^{2}H + p + n)^{8}Li	$8.3 \pm 0.3 \pm 0.3$
^{8}B	0.77	β^+	13.9	^{16}O(π^+, α + 2p + 2n)^{8}B	
^{9}Li	0.178	β^-	13.6 (50.5%)	^{16}O(π^-, α + 2p + n)^{9}Li	$0.9 \pm 0.3 \pm 0.3$
		β^- n	49.5%		
^{9}C	0.127	β^+ n	3$\sim$15	^{16}O(n, α + 4n)^{9}C	< 1.4
^{11}Li	0.0085	β^-	16$\sim$20 ($\sim$50%)	^{16}O(π^+, 5p + π^0 + π^+)^{11}Li	-
		β^- n	$\sim$16 ($\sim$50%)		
^{11}Be	13.8	β^-	11.51 (54.7%)	^{16}O(n, α + 2p)^{11}Be	< 16.9
			9.41 + 2.1(γ) (31.4%)		
^{12}Be	0.0236	β^-	11.71	^{18}O(π^-, α + p + n)^{12}Be	-
^{12}B	0.0202	β^-	13.37	^{16}O(n, α + p)^{12}B	$19.8 \pm 0.1 \pm 1.0$
^{12}N	0.0110	β^+	16.32	^{16}O(π^+, 2p + 2n)^{12}N	$2.8 \pm 0.1 \pm 0.1$
^{13}B	0.0174	β^-	13.44	^{16}O(π^-, 2p + n)^{13}B	-
^{13}O	0.0086	β^+	13.2 or 16.7	^{16}O(π^-, μ^- + p + 2n + π^-)^{13}O	-
^{14}B	0.0138	β^-	14.55 + 6.09(γ)	^{16}O(n, 3p)^{14}B	-
^{15}C	2.449	β^-	9.77 (36.8%)	^{16}O(n, 2p)^{15}C	< 6.7
			4.47 + 5.30(γ)		
^{16}C	0.747	β^- n	$\sim$4	^{18}O(π^-, n + p)^{16}C	-
^{16}N	7.13	β^-	10.42 (28.0%)	^{16}O(n, 2p)^{16}N	$39.7 \pm 3.3 \pm 2.8$
			4.29 + 6.13(γ) (66.2%)		

Table 5.1: List of possible muon-induced spallation products that can occur in SK. The rates for ^{8}Li/^{8}B cannot be decoupled since their lifetimes and energy are too similar. This table is transcribed from [4]. Some results were consistent with 0, so a 90% confidence limit was placed on the yield (e.g. ^{8}He)

the Radon from its decay chain into the water. The second major radioactive background is from Thorium (Th) decay chain introduced from the PMT glass. As mentioned before, one of the improvements to the low energy analysis from the 2018 detector maintenance was the cleaning of the PMT support structure of contaminants (e.g. diesel fumes).

For the Radon, the main issue for the low energy analysis is the eventual β-decay of ^{214}Bi, which has an endpoint energy of around 3.27 MeV. From the Thorium chain, the troublesome isotope is ^{208}Tl with a $\beta\gamma$ decay, with the total energy of the decay being around 5.00 MeV. These are a bulk of the irreducible radioactive backgrounds that come from components of the detector itself. One way to help reduce these effects is to maximize the laminar flow within the detector and reduce convection. The elements do not travel far through the water through diffusion, so reduction of turbulence prevents the spread further inward in the detector.

5.3 Trigger Rates

For the low energy group, the radioactive backgrounds from detector and cosmogenic sources are the most troublesome backgrounds in the <6 MeV and 6-20 MeV range respectively.

SK triggers on events based on the number of hits within a 200 ns window (N_{200}, Sec 2.7). The different triggers are based on the value of N_{200}, and their levels, rates, and values are shown in Table 5.2. Without any energy thresholds or accounting SK expects ~326 ^{8}Bν_e elastic scattering interactions a day. Once thresholds and efficiencies are accounted for, the rate reduces by more than an order of magnitude and is on the sub-mHz scale.

Because these event rates are so high, steps are taken in the low energy online processing remove obvious backgrounds, and more stringent cuts are applied offline depending on the analysis being performed.

Trigger Type	N_{200} Threshold	Event Rate (Hz)	Event Timing (μs)
SLE	34/31	4400/11000	1.5 (-.5 $\sim$ 1.0)
LE	47	45	40 (-5 $\sim$ 35)
HE	50	11	40 (-5 $\sim$ 35)
SHE	70/58	4	40 (-5 $\sim$ 35)

Table 5.2: The different trigger thresholds and corresponding trigger rates and event timing windows

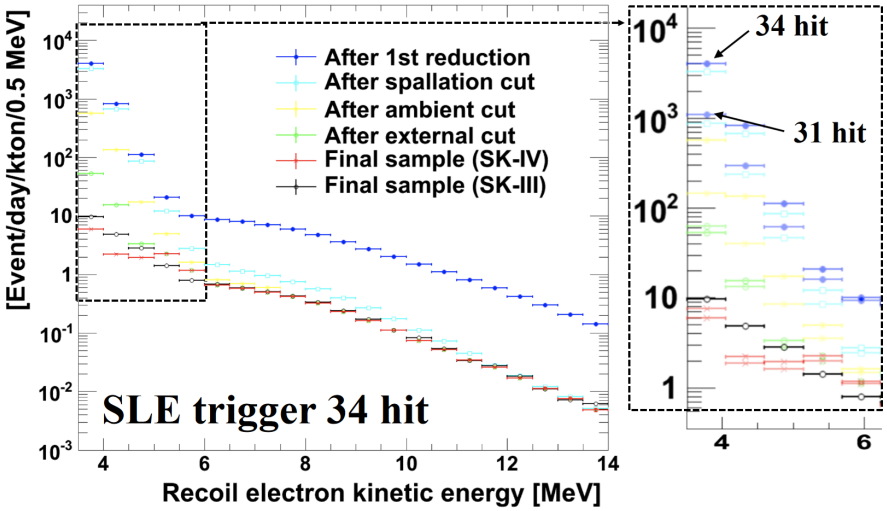


Figure 5.2: The energy spectrum following each reduction step during the 34-hit SLE trigger data period of SK-IV. The final solar neutrino sample for SK-IV (Red) is contrasted with that of SK-III (Black). The pop-out of the low energy region shown right shows the effect of differing pre-reduction criteria for the two trigger thresholds (34 and 31-hit) of SK-IV

5.4 Data Reduction Overview

The following is an outline of the data reduction steps as they pertain to the solar analysis. The overall effect each step has on the recoil electron spectrum is shown in Fig. 5.2

After the pre-reduction, the solar analysis systematically removes back ground events with:

- First reduction removes (Sec 5.6)
  - Bad runs and subruns

  - Very high energy
  - Calibration triggered/source events
  - Flashing PMTs
  - Muon-related Events
  - OD triggered events
- Spallation event removal (Sec 6)
- Ambient backgrounds cuts:
  - Poor event reconstruction
  - Patlik (pattern likelihood) cuts
  - Tight Fiducial Volume (TFV)
- External event cut
- Small hit cluster removal
- ^{16}N from muon capture

Each of these cuts will be explained in the following sections.

5.5 Pre-Reduction

The DAQ and online computers are tasked with properly managing the high trigger rate of low energy events. If all events triggered by the DAQ were kept in their entirety, the disk space required to store 1 day of data would exceed 100 GB. The goal of the pre-reduction is to filter out as much of the extremely-likely background events though a series of cuts.

5.5.1 Vertex Position Cut

As discussed in Sec 5.2, a large source of radioactive backgrounds is from the PMTs and support structure. To avoid the backgrounds associated with this, a Fiducial Volume (FV) cut is applied to the data requiring that events are more than 200 cm from the wall of the ID. This reduces the total volume of the ID from 33.5 kton to 22.5 kton.

5.5.2 Loose External Cut

To further reduce background from radiation sources near the PMT structure, a larger, direction dependent wall cut is applied. For pre-reduction purposes, if the distance from the reconstructed vertex following the inverted direction vector, d_{eff} (Fig 5.3), is <400cm the event is cut. Later, an energy dependent cut is applied on d_{eff}.

5.5.3 Energy cut

On the low end, if the reconstructed energy of the event is <3 MeV, the event is removed. If the event has more than 2000 pe, it is considered a muon and removed from the low energy data. It is stored elsewhere for reconstruction for spallation tagging and other studies.

5.5.4 Loose Event Quality Cut

Using the vertex reconstruction goodness, g_V, from Sec 4.3.1 and the directional goodness, g_V, from Sec 4.3.2 are used to assess the quality of the reconstructed events. The difference of the square of these parameters ($g_R^2 = g_V^2 - g_A^2$) is used to apply an event quality cut. Data and MC were compared for these parameters, and the peak from solar neutrino signal (MC) was significantly shifted from the background (data). This is shown in Fig. 5.4 for the two

samples. The pre-cut removes events with $g_R^2 < 0.10$.

5.5.5 Pre-reduction Summary

The summary of the effect of each step can be found in Table 5.3. In total, the pre-reduction reduces the number of events by a factor of ∼100, with the solar neutrino signal loss, determined from MC, being <5% for energies above 5MeV.

5.6 First Reduction

Some of the following cuts are optimized in order to maximize significance, as measured by:

$$Significance = \frac{Signal}{\sqrt{Signal + Background}} \tag{5.4}$$

For the lowest energy regions, the background dominates the signal, and the significance can be approximated to $Signal/\sqrt{Background}$. These cuts are tuned and applied offline after accounting for the various detector environment details from calibrations.

5.6.1 Bad Run Removal

The SK data stream is broken up into "runs" to help with the organization of the data based on real-world time and detector conditions. Detector conditions refers to the varying conditions that SK can be running under, whether calibration is being performed, T2K beam is on, different trigger logic, studies are being performed, etc. Runs are limited to a maximum of 24 hours, and each run is subdivided into subruns which range from ∼30-90s in length.

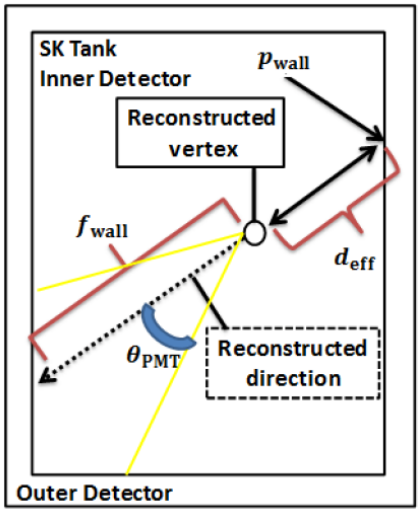


Figure 5.3: Defintition for d_{eff}, p_{wall}, f_{wall}, θ_{PMT}. d_{eff} is the distance from the reconstructed vertex to the ID wall along the inverted reconstructed direction vector, and d_{wall} is the point on the wall where this tracks back to. f_{wall} is the distance from the reconstructed vertex to the ID wall along the reconstructed direction vector. θ_{PMT} is the angle between the reconstructed direction vector and each PMT with respect to the reconstructed vertex.

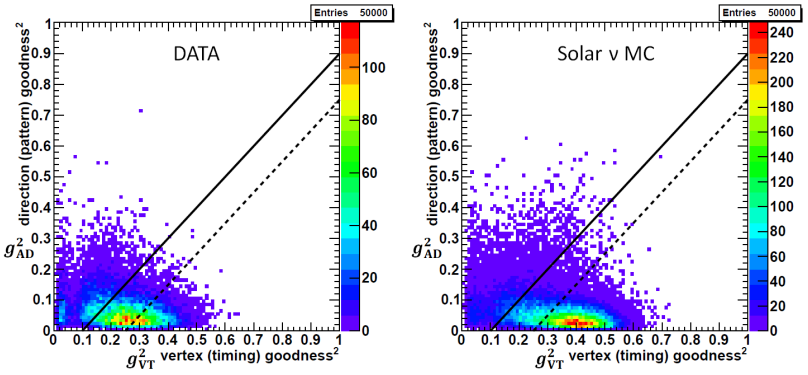


Figure 5.4: Correlation of g_V^2 and g_V^2 for data (left) and solar neutrino MC (right). The black solid (dashed) lines corresponding to $g_R^2 = 0.10$ (0.25)

Pre-reduction Step	Number of Events Remaining	Efficiency
Fiducial Volume	28,161,	14.2
Loose external cut	14,864	7.5
Energy Cut	6,872	3.5
Loose Event Quality Cut	2,057	1.0

Table 5.3: Example reduction efficiencies for each step of the pre-reduction process. The overall reduction is a factor of ∼100

To ensure data quality used in many analyses, several criteria are considered after the data has been taken and saved when deciding if a specific set of data should be included in the final sample. "Bad" runs or subruns are flagged if the following criteria is met:

- Too short of run

  If the run is less than five minutes, the run is rejected due to insufficient pedestal data to judge run quality.

- Close to ID HV being turned on

  Data within 15 minutes of the ID HV being turned on is not used due to increased dark rate of the PMTs during this time.

- Calibration Work

  Periodic calibration work occurs within SK during normal running, such as LED supernova burst testing, and the corresponding subruns are removed.

- Hardware or software issues

  Issues with hardware or software, for example HV issues or DAQ crashes, require nearby subruns in time to be discarded.

Usually, when something happens with the detector requiring data to be removed, only part of the run is discarded and the rest can be used for analysis. In addition to these, there are solar neutrino analysis specific cuts to further filter out bad runs or subruns.

The first set are requirements on the number of bad ID channels in a subrun as follows:

- If the number of bad channels is less than 10 the subrun is removesd. This can happen for short runs, and there is not enough information gathered to determine the (sub)run qualilty.

- When a QBEE has problems, 24 PMTs connected to the board are affected and considered bad channels. If the number of bad channels is too high above the monthly average, the (sub)runs are removed: $N_bad \geq N_{bad,avg} + 1.5 \times 24$.

The third and final requirement is applied to the OD PMTs. The OD trigger rate can fluctuate similarly to that of the ID, and if this fluctuation is too much, then the (sub)run is removed. The criteria for this is a $+5\sigma$ cut on the expected OD trigger rate of 1.954 ± 0.180Hz, and corresponds to 2.847 Hz.

5.6.2 Number of Hit PMTs

A further cut is placed on high energy events, since solar neutrinos do not produce recoil electrons above ∼20 MeV. If the total number of hit PMTs for an event is more than 400, the event is cut. This corresponds to ∼60 MeV, and helps to remove high energy events such as decay electrons, remaining cosmic-ray muons, and atmospheric neutrinos.

5.6.3 Calibration Source Cut

There are multiple types of automated calibrations performed routinely within SK, events from these sources are tagged and removed. The frequency of these calibrations can occur from every second (laser calibration) to a couple times an hour (pedestal events) to once every several weeks (LED supernova burst detection system testing). A vertex cut is also placed on the region surrounding these devices, as there is a possibility of light being emmitted from the equipment itself, even outside of operation. A 2m vertex cut is placed from the calibration device and a 1m cut is placed on device cables.

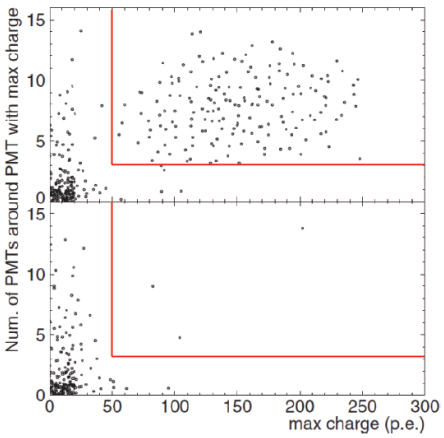


Figure 5.5: Comparison of a run with (top) and without (bottom) a significantly flashing PMT. The red box outlines the cut criteria based on maximum charge of a hit PMT (x-axis) and the number surrounding 24 PMTs with hits (y-axis). The uper right corner are events that are removed

5.6.4 PMT Flasher Cut

During normal operation, occasionally an arc discharge occurs on the dynode of a PMT. Nearby PMTs see a flash caused by this, and flashing PMT itself sees a large number of pe. If the charge seen by a PMT exceeds 50 pe while the number of hits seen by neighboring PMTs in a 5×5 grid centered on the suspect PMT exceeds 3, the event is considered a flasher and removed from teh sample. Fig. 5.5 shows example runs with (top) and without (bottom) a significantly flashing PMT. Sometimes flasher events will occur only a few times within a run, and the PMT will cease the flashing behavior. The flashing can also become more troublesome and the PMT is turned off.

5.6.5 Muon Related Event Cut

Cosmic ray muons not only cause radioactive backgrounds through spallation, but also through decay electrons and PMT ringing and after-pulsing. A tight time cut on Δt ($\Delta t = t_{SLE} - t_{muon}$) can be placed on events in relation to preceding muons to account for these

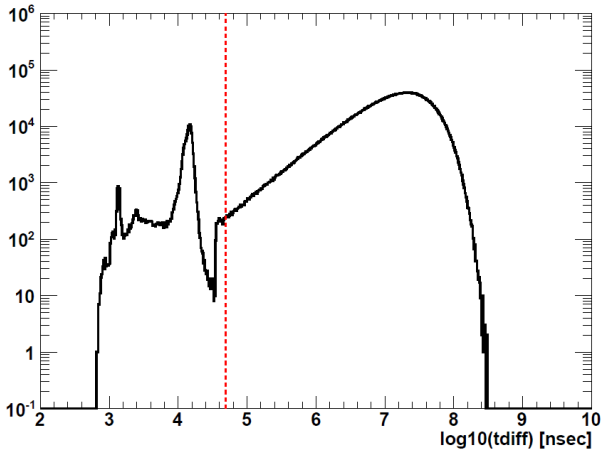


Figure 5.6: Typical distribution of the time difference to previous low energy triggered event. Red dashed line represents the 50 μs cutoff, with events to the left cut. The 1 μs peak is due to PMT ringing, and the 15 μs peak is caused by PMT after pulsing.

effects. Fig 5.6 shows a distribution of the time difference to the previous triggered low energy event. PMT ringing occurs around 1 μs after a muon and PMT after pulsing occurs around 15 μs, both peaks seen in the figure. Events with $\Delta t < 50\mu$s are removed from the sample.

5.6.6 OD Event Cut

Recoil electrons are only expected to travel a few cm in distance, and events originating in the ID FV should be wholly contained within the ID. Since the ID and OD are optically separated, the light from these events should also be only seen within the ID. If the OD is triggered, which corresponds to a total charge measured in the OD to be $\gtrsim$ 20 pe, the event is flagged. Additionally, if the OD sees at least 20 hits, the event is flagged regardless of OD trigger.

5.7 Spallation Cut

The main topic of this thesis is on the spallation cut, and will be throughly explored later in Chapter 6.

5.8 Ambient Background Cuts

The following cuts are performed in the offline analaysis, and are done to further remove the persistent backgrounds found in the detector.

5.8.1 Event Quality Cut

Using the same event quality criteria as Sec 5.5.4, the event quality cut is tuned for different energy regions to maximize the efficiency of extracting the solar neutrino signal using the significance criterion introduced at the beginning of this section. The energy dependent cut on g_R^2 that maximizes the significance is as follows:

$$g_R^2 > \begin{cases} 0.30 & 4 \leq E < 5.5MeV \\ 0.26 & 5.5 \leq E < 7.5MeV \\ 0.22 & E \geq 7.5MeV \end{cases} \tag{5.5}$$

with E being the total energy of the recoil electron. The cut becomes more stringent as the energy is lowered, as the fraction of background in the sample increases with decreasing energy.

5.8.2 Hit Pattern Cut

Sometimes, the decays from the radioactive backgrounds (more so for spallation) decay with multiple βs and/or γs, resulting in distortion or smearing of the Cherenkov light. The Cherenkov ring pattern in the hit PMTs is better defined when a single electron is the source of Cherenkov light. Multiple β and/or γ events can somewhat be distinguished from single electron events using the difference in hit patterns.

The hit pattern likelihood function, "patlik" is defined as

$$\mathscr{L}_{pattern}(E, \overrightarrow{v}) \equiv \frac{1}{N_{50}} \sum_{j}^{N_{50}} \log[P_j(E, \cos\theta_j, f_{wall})] \tag{5.6}$$

with E being the total energy of the event, N_{50} is the same as in Sec 4.2.3.1, P_j is the probability density function (PDF) of the hit pattern at the j'th PMT, and is a function of E and the event variables θ_j and f_{wall} defined in Fig. 5.3.

The cut values are evaluated by the same significance calculation mentioned earlier this section, and the energy dependent results are as folows:

$$\begin{aligned} &\mathscr{L}_pattern > -1.88, 6.5 - 8.0MeV \\ &\mathscr{L}_pattern > -1.86, 8.0 - 12MeV \\ &\mathscr{L}_pattern > -1.95, 12 + MeV \end{aligned} \tag{5.7}$$

Fig. 5.7 shows the distributions for these three energy regions, with the blue dashed line marking the corresponding cut and grey shaded region showing the removed events. The cut is not applied to events below 6.5 MeV total energy. At this energy level, the small amount of PMT hits makes it difficult to make a meaningful cut.

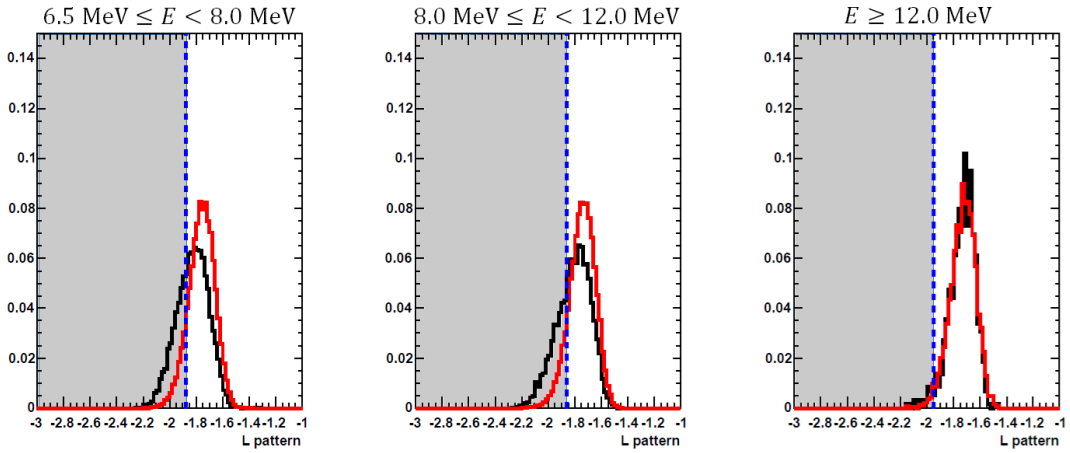


Figure 5.7: The distribution of $\mathscr{L}_{pattern}$ for data (black) and solar neutrino MC (red), with the corresponding energy regions above each of the plots.

5.8.3 Tight Fiducial Volume Cut

Solar neutrinos are expected to interact uniformly within the water of SK, regions with high event rates are likely due to radioactive backgrounds. The sources of these backgrounds are part of what was discussed in Sec. 5.2. As mentioned, the energy from these events are 3.27 MeV from the ^{214}Bi β-decay and 5.00 MeV for the ^{208}Tl decay with β and γ(s). Because of the poor energy resolution of $\sim$20% for SK in the $E \simeq 4.5$ MeV, the reconstructed energy of the events can be well above the 4.0 MeV threshold for the solar analysis.

With the non-uniform backgrounds within the ID, further energy-dependent fiducial volume cuts are applied to better maintain the uniformity of vertex distributions. Uniform vertex distributions help to keep the shapes of the background distributions more uniform, allowing for better modeling. This in turn helps to minimize the associated systematic uncertainties.

The standard fiducial volume cut covered in Sec 5.5.1 corresponds to 22.46 ktons of volume by cutting all events, regardless of energy, within 200cm of the PMT structure. For events with reconstructed energies between $4.0 \leq E \leq 5.0$ MeV and $5.0 \leq E \leq 5.5$ the accepted volume is even smaller, with a tight fiducial volume (TFV) cut applied to the data.

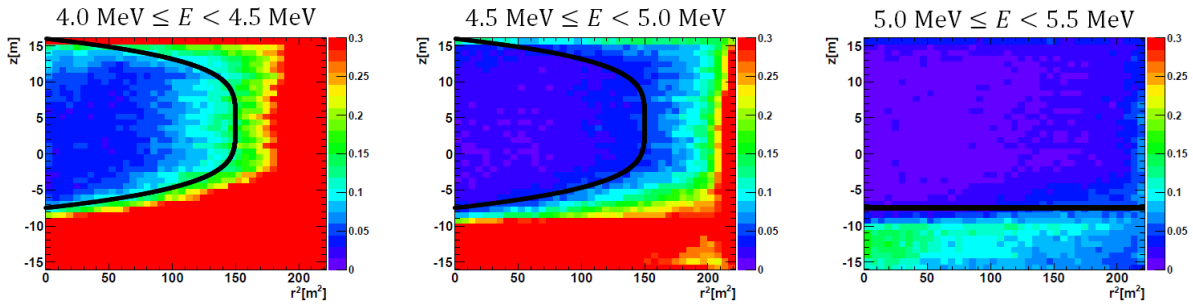


Figure 5.8: r^2 vs z vertex distributions for the three energy bins from the solar neutrino analysis affected by the TFV cut. The TFV cut removes events to the right of the black line in the left and middle vertex distribution, and below the black line in the right vertex distribution. The bin color and scale corresponds to events/day/bin

The boundaries for the TFV cuts were determined in order to maximize the allowed volume and the significance at low energies. The TFV cuts events in the following regions:

$$\begin{aligned} z &< -7.5m, & 5.0 \leq E \leq 5.5 \\ r^2 + (\frac{150.0}{11.75^4} \times \mid z - 4.25 \mid^4) &< 150.0, & 4.0 \leq E \leq 5.0 \end{aligned} \tag{5.8}$$

which results for a total volume of 8.85 kton for Events with $E < 5.0$ MeV and a total volume of 16.5 kton for events with $5.0 \leq E < 5.5$ MeV. Fig. 5.8 shows the vertex distributions for the three lowest energy bins in the solar analysis, with the TFV boundary for the regions shown as a black line.

5.9 External Event cut

As mentioned in Sec. 5.5.2, radioactivity and γ-rays comes from the PMT glass and housing, PMT structure, and other detector components. The reconstructed direction of these events generally point inward from the wall and can be removed using a cut on d_{eff}. There is a slight difference in the d_{eff} calculation from before, using a more accurate WT value for the event reconstruction. For the pre-reduction, the WT value was assumed to be 90m. With the updated information, events with the following d_{eff} values are removed for corresponding

energy ranges and p_{wall} locations (Fig 5.3):

$$\begin{aligned}
d_{eff} &< 400cm, && E \geq 8.0\text{MeV} \\
d_{eff} &< 650cm, && 5.5 \leq E < 8.0\text{MeV} \\
d_{eff} &< 1000cm, && 4.0 \leq E < 5.5\text{MeV and } p_{wall} \text{ at top} \\
d_{eff} &< 1200cm, && 4.0 \leq E < 5.5\text{MeV and } p_{wall} \text{ at barrel} \\
d_{eff} &< 1300cm, && 4.0 \leq E < 5.5\text{MeV and } p_{wall} \text{ at bottom}
\end{aligned} \tag{5.9}$$

Since the TFV cut applied to the low energy events below 5.5 MeV, the d_{eff} was optimized with this taken into consideration and using p_{wall} as an added parameter. The larger high-background volume at the bottom of the detector results in a more stringent d_{eff} cut at the bottom compared to the top for p_{wall}, as was seen in the TFV cut as well.

5.10 Small Hit Cluster Removal

When a radioactive decay occurs at a PMT glass or an FRP cover while coinciding with a dark noise fluctuation, an event can be triggered. When this happens the reconstructed vertex is close to the wall and the hit pattern becomes a small cluter. To characterize these clusters, two parameters are introduced:

- $r02$: Minimum radius containing more than 20% of hit PMTs within a 20 ns time window
- N_{20rawT}: Maximum number of hits within a 20 ns raw timing window (no t.o.f. subtraction)

Background events expect to have the ratio of N_{20rawT}/N_{eff} to be smaller than that of solar neutrino signals.

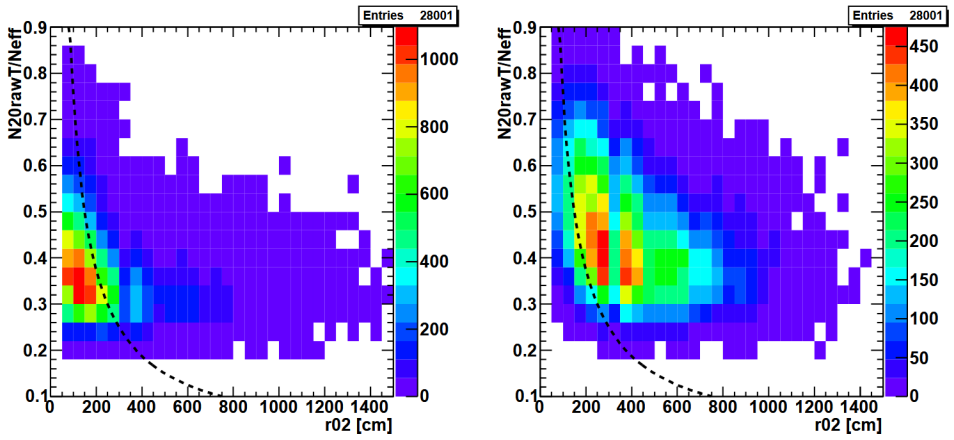


Figure 5.9: Relationship between N_{20rawT}/N_{eff} and $r02$ for background (left) and solar neutrino MC (right) for events in the edge region. The dashed line corresponds to the cut of $r02 \times (N_{20rawT}/N_{eff}) < 75.0$

Fig 5.9 shows the relationship pattern $r02$ and N_{20rawT}/N_{eff} for background (left) and solar neutrino MC (right). After evaluating the significance, the cut criteria is as follows:

$$r02 \times (N_{20rawT}/N_{eff}) < 75.0 \begin{cases} r^2 > 155\text{m}^2 & z < -7.5\text{m}(5.0 - 5.5\text{MeV}) \\ r^2 > 120\text{m}^2 & (z < -3.0\text{m} or z > 13.0\text{m}(4.0 - 5.0\text{MeV}) \end{cases} \tag{5.10}$$

5.11 ^{16}N from Muon Capture Cut

^{16}N is produced in the SK water generally from fast neutrons, whether it is produced through spallation or the DT generator. However, another method of production is from the capture of a muon on an oxygen nucleus:

$$\mu^- + {}^{16}\text{O} \rightarrow {}^{16}\text{N} + \nu_\mu \tag{5.11}$$

Muon-induced ^{16}N occurs uniformly throughout the detector, and has a half-life of 7.13s. ^{16}N has two primary decay modes, with 66% of the time decaying with a β (4.3 MeV) and γ (6.1) MeV and 28% of the time a single β with an endpoint of 10.41 MeV, with all others being a combination of a γ(s) with a single β emission. All of these modes mimic the signal of a solar neutrino and need to be removed. The normal spallation cut removes some of these events, but the long half life of the events paired with the typically low energy of stopping muons, makes these events harder to tag. To ^{16}N events produced by muon capture, the correlation between muons tagged as stopping and the event reconstructed vertex is considered. If the following criteria is met, the event is removed:

1. Muons classified as stopping do not have an accompanying event within 100 μs of the muon.

2. The muon stopping point and the reconstructed vertex are less than 250 cm apart and between 100μs and 30s in time.

This results in a deadtime of $\sim$0.53% for this cut, estimated from taking samples with normal time correlation between events as signal, and inverted time correlation as background. This is the final reduction step for the solar neutrino analysis.

5.12 Convection Period Cut

During February of 2018 SK-IV began a convection period within the detector, where the convection zone was extended to the entire detector in order to test certain detector responses in preparation of the 2018 detector maintenance. However, the increased convection led to an increase in the background for the lowest energy bins. The low energy group removes events below 6 MeV for this period from the final solar sample. An additional small test was performed in May 2018, and the affected energy regions were removed for those runs as well.

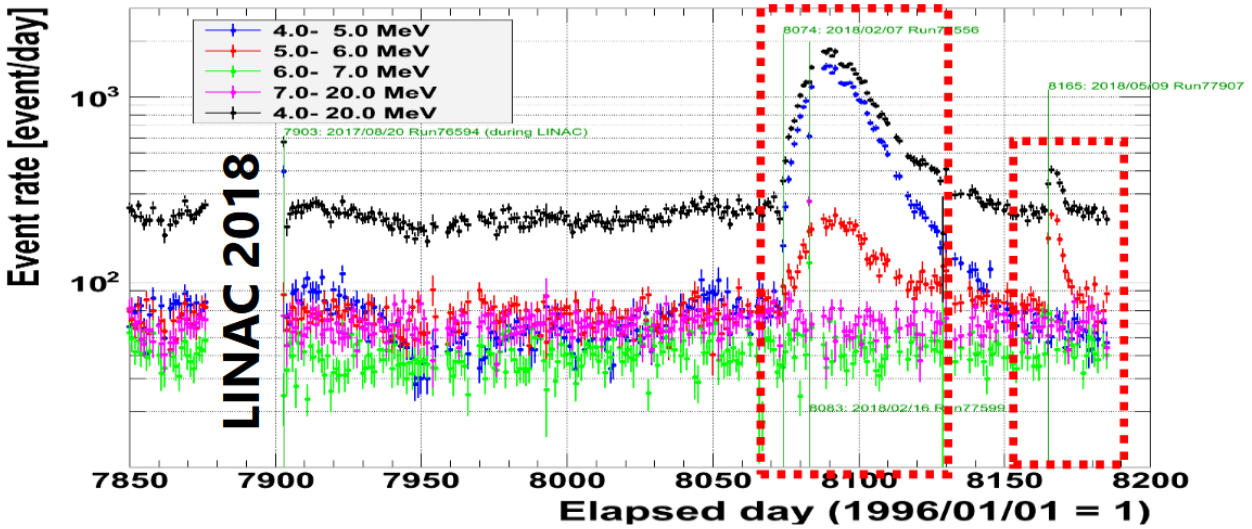


Figure 5.10: Event rate increase in lowest energy bins due to convection and other changing detector conditions. The events below 6 MeV have increased event rate during this time, and therefore are removed from the final sample.

Chapter 6

Spallation Analysis

6.1 Neutrons in Hadronic Showers

Cosmic-ray muons and the secondary particles from their showers break apart nuclei and create fast neutrons and beta-decay isotopes [62]. These isotopes (spallation) are the dominant background in Super-Kamiokande in the 6-20 MeV energy range. A large portion of spallation products have energies below 6 MeV, but the abundance of other radioactive backgrounds begins to dominate. It was noted by Beacom and Li [63] that fast neutrons in showers producing ^{16}N would be of high enough multiplicity that they would be able to be detected in Super-Kamiokande. Neutron captures on Hydrogen produce a deuterium nucleus and a 2.2 MeV gamma:

$$n + \mathrm{H} \rightarrow \mathrm{D} + \gamma \ (2.2 \text{ MeV}) \tag{6.1}$$

The 2.2 MeV γ produces on average 7 detectable photons in SK during the process, resulting in weak sensitivity for SK to adequately detect the interaction. The normal SK trigger system allows for all data to be saved after a high energy event (SHE+AFT) in order to search for Diffuse Supernova Neutrino Background events. When an SHE trigger is done, 500μs of

Energy Range	Neutron Energy
0.0–0.025 eV	Cold neutrons
0.025 eV	Thermal neutrons
0.025–0.4 eV	Epithermal neutrons
0.4–0.5 eV	Cadmium neutrons
0.5–1 eV	EpiCadmium neutrons
1–10 eV	Slow neutrons
10–300 eV	Resonance neutrons
300 eV–1 MeV	Intermediate neutrons
1–20 MeV	Fast neutrons
>20 MeV	Ultrafast neutrons

Table 6.1: Neutron Temperature ranges

data immediately after is saved to search for secondary events, however this is disabled after muons in order to not overwhelm the software trigger.

Neutrons can be classified by their energy, otherwise known as their temperature. There are ten different neutron energy categories which are shown in Table 6.1. When produced in showers, neutrons are generally in the fast or ultra-fast energy range, and the largest capture cross-section occurs once the neutron has undergone enough nuclear collisions to make the neutrons "thermal". The energy for thermal neutrons is the mode of the Maxwell-Boltzman distribution at the temperature of the system. When neutrons are thermal, they have their largest capture cross section on a nucleus, although capture can still happen at higher energies.

The separate WIT system does not have this AFT trigger as it processes the data differently from the normal DAQ stream. As discussed in Sec 2.8, WIT will typically only save events with 11 hits above dark noise in a 230 ns window, as well as other online reconstruction requirements outlined in Sec 2.8. This decreases the ability to tag the 2.2 MeV γs from neuton captures on H to be 17.2% (13.5%) (measured with MC generate 2.2 MeV γs) within the fiducial (full) volume. This efficiency decreases to 10.0% (7.5%) when a cut on vertex distance to truth and event quality is applied. However, the number of fast neutrons

in the spallation causing showers is large enough to compensate for the small efficiency. In order to look for these neutrons, a simple search of events was performed within 500 μs of a muon, with the event being reconstructed again offline. After offline reconstruction was performed, a cut of 5m was placed on the neutron distance to the track, with the primary muon track used if it was a multiple muon. All events passing this 5m and 500 μs cut, would be included into the sample.

6.2 Identifying Neutron Captures After Muons

The investigation of the neutron clouds was broken into two pieces. An initial study into the nature of the neutron clouds after muons was performed, and provided insight into what was being seen, and the likelihood of success. The second stage was to refine the variables for the most effective cut to be made on the clouds to balance detector livetime and spallation tagging efficiency. The main characteristics investigated in the outset were the clustering of events after a muon, multiplicity, and timing correlation of the candidate events to the muon track, specifically for proof of the ability to tag neutrons after muons in Super-Kamiokande. For a consistent comparison of events, the coordinate system is shifted to cylindrical coordinates, with the muon track as the z-axis, entry position as the origin, and direction of the muon the direction for positive z. In this coordinate system, the radius correlates to the transverse distance (lt) of the candidate neutrons to the muon track, and longitudinal distance is the distance to the average z position for an event with greater than 2 candidates. It is assumed there is no preference in azimuthal direction overall when a shower is produced. With no angle dependence, any studies with the use of lt generally use its square (lt^2) to flatten out the phase space and make it easier to separate the background of the distributions. A few key features arose when performing this initial investigation. First was within the Δt distribution without cuts. There were peaks of events in the 10-20 μs range .

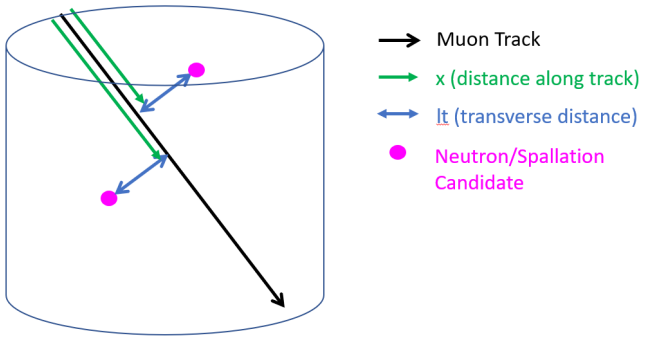


Figure 6.1: Simple Cartoon illustrating common terms in this section and chapter. A muon track is shown in black, with candidate events (spallation or neutrons) are shown relative to the track. The transverse distance is shown in blue, while the distance along the track, x, is shown in green. Commonly ln will be used, which is a taken as $\ln = x_i - x_{ref}$, where x_{ref} is some reference point on the track

With the high energy of the muons needed in order to produce secondary events, PMTs have the possibility of after pulsing on the time scale of 10-20 μs after muons. With the relatively low threshold needed for WIT to trigger above dark noise, a pileup occurs in that region due to fake and mis-reconstructed events. Using just the goodness of the fit at the reconstructed vertex and applying a cut on it removed most of these events, as well as providing for a higher purity sample. Values were checked for g_{vertex} greater than 0.4 and 0.5. This is successful in removing the event pile up in the region of 10-20 μs, but also makes a deficit in the number of events seen due to the mis-reconstructed events having poor a mix of good and bad hits. A 20 μs cut was placed for this analysis to avoid this region and also avoid muon decay electrons. With full Gd-loading in SK-GD, this cut will need to be revisited as the capture time drops to $\sim 30\mu$s. The lt^2 distributions showed distinct improvement when moving the cut on g_{vertex} greater than 0.4 to 0.5. A sophisticated algorithm may have the potential to clean up the noisy hits to recover some of the events, but for captures on H, the capture time is so long comparatively, it was deemed not worthwhile to pursue these changes. However, as the concentration of Gd increases in the detector the fraction of neutron captures occurring in this time range will grow and be harder to ignore.

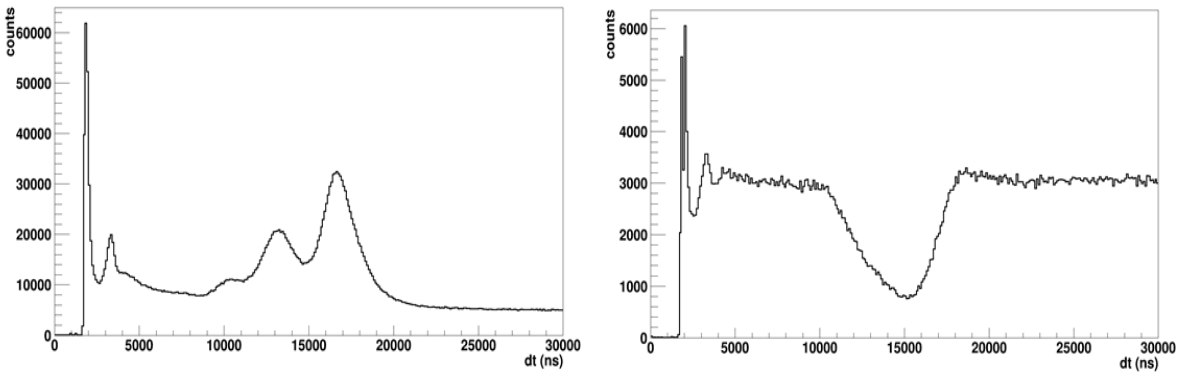


Figure 6.2: Comparison of events caused by after pulsing as seen in WIT. The sub-5μs peaks are believed to be from decay electrons and PMT ringing. The left figure is the Δt correlation without any other cuts performed besides the intial 5m lt requirement. The right image is after finalized event quality and energy cut are applied. Since these events are immediately after muons, most PMTs are hit and afterpulsing is somewhat random. Reconstruction requirements, but since they are not a real events the reconstruction quality is poor

Neutron cloud shapes were investigated with these initial purity cuts in place on neutron capture candidates. ln distributions show there is more smearing along the muons direction compared to distance traveled away from the track in lt^2. This is seen looking at the two distributions, where most of the candidate events were contained within 2m of lt, the ln had a spread of up to 8m and was strongly dependent on multiplicity. As showers grew in size, so did the flattening of the ln distribution to cover a larger ln range.

To increase purity further, a cut of 1.5 m in lt was made to make an accurate measurement of capture time. This capture time was compared to previous neutron capture studies performed in SK. SK published AmBe data in 2015 quoting a neutron lifetime of 203.7 $\pm$ 2.8 μs. This was used as a benchmark to see if the events observed after a muon were consistent with neutron captures on Hydrogen. A fit was performed from 50-500 μs, of the form

$$\mathrm{N} = \mathrm{A}e^{\frac{-\Delta t}{\tau}} + \mathrm{C} \tag{6.2}$$

where N is the event rate after a muon, A is the amplitude of the fit, and τ is the time

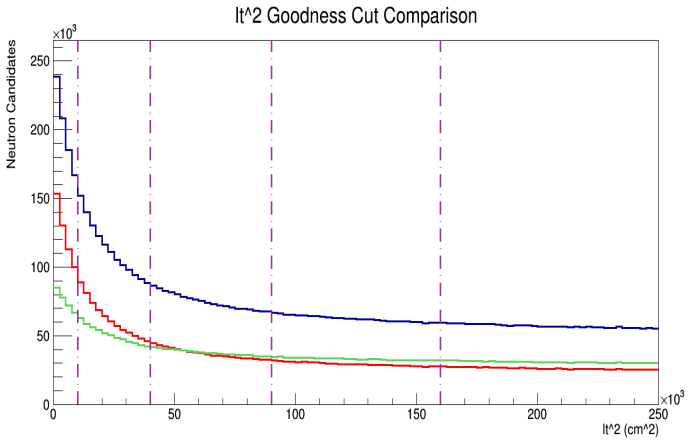


Figure 6.3: The lt^2 correlation of neutron candidate events to the muon track, for different values of g_{vertex} (Sec 4.3.1). Blue is $g_{vertex} > 0.4$, red is $g_{vertex} > 0.5$, and green represents $0.4 \geq g_{vertex} < 0.5$. The purple lines from left to right correspond to $\mathrm{lt}^2 = (1m)^2, (2m)^2, (3m)^2, (4m)^2$. Comparing the two goodness cuts showed that there was a distinct improvement to the signal to noise level of the sample. The green distribution illustrates this showing over half of the background events were removed while only losing $\sim 25\%$ of the signal events.

constant for the capture. The constant C is there to compensate for an assumed constant background, as a background sample was not made for the neutron cloud data. τ is the time constant for the capture process and the fit value is 210.3 $\pm$ 1.6 μs, 2σ away from the AmBe measurement of 203.7 $\pm$ 2.8. While statistically there is some tension with the AmBe source, deviations may come from systematics not taken into consideration, or slight spallation contamination from early decays of very short lived spallation.

The 2.2 MeV γ energy is not well reconstructed since the energy scale assumes it is an electron and an energy of 2.2 MeV below trigger threshold, so an up fluctuation in detected light must occur. When looking at the energy, the events are sharply peaked, while also falling off completely by 6 MeV reconstructed energy. A cut at 6 MeV was placed moving forward, in order to further reduce background and incidentals.

Finally, the neutron multiplicity was investigated to see relative fractions of showers for different shower sizes. When deciding on cuts to be made with the before mentioned variables, balancing deadtime and cut efficiency is important. Deadtime is is the time the

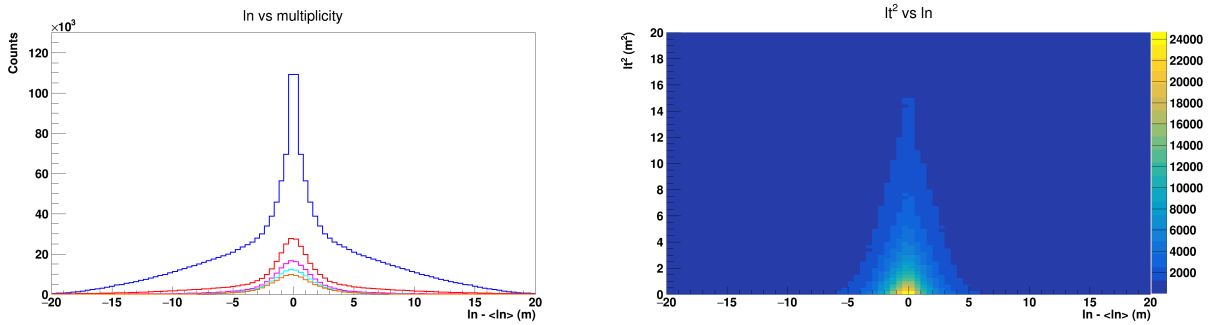


Figure 6.4: Vertex correlation distributions for neutrons after muons, applying the final vertex quality cuts. The left plot shows the ln only distribution, for multiplicities of 2 (blue), 3 (red), 4 (purple), 5 (cyan), and 6 (orange) neutrons. As multiplicity increases, width of distributions increases, as well as the distribution becoming more skewed to the right. The right plot shows the correlation of lt^2 and ln for the neutron clouds for all multiplicities. The distribution is strongly peaked in lt^2 and ln near 0, but the showers can be several meters long and away from the track.

detector (or part of the detector) is not active for any reason, and is the complement to relative livetime. Spallation cut efficiency is how well the spallation cut removes spallation background. For reference, the former solar spallation tagging method had a deadtime of 20% with a spallation cut efficiency of 90%. This means 20% of solar neutrino signal was lost in order to remove 90% of the spallation background, leaving 80% of the solar neutrino signal and 10% of the spallation background. Multiplicity informs of the rate of different sized showers plays and provides information on how to better balance these two factors.

6.3 Search for Spallation Near Neutron Cloud Centers

With the existence of neutron clouds after muons proven, spallation from the neutron clouds was explored. Based on the observed cloud size, a search for spallation within a 5m sphere from the center of the neutron cloud was performed. In addition to the 5m cut, a 60s is chosen as a means to try and mostly contain the ^{16}N events, which have a half-life of 7s and decay constant of 10.3s. 60s accounts for just under 6 lifetimes of ^{16}N, or less than 0.3%

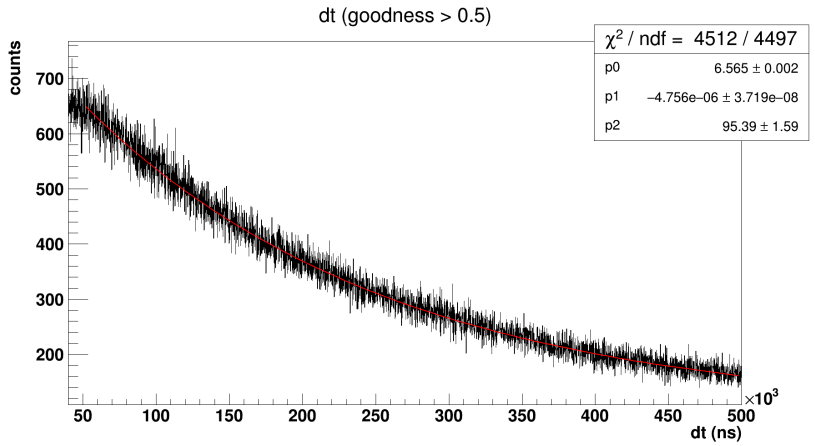


Figure 6.5: Δt distribution for a $g_{vertex} < 0.5$ and tight lt (lt < 1.5m) cut. The distribution was fit from 50μs to 500 μs to with an exponential plus constant. The resulting fit gives a decay constant of $\tau = 210.3 \pm 1.64$, slightly above AmBe source calibration data.

remaining. In order to make a signal and background comparison, the sample was run over in two directions of time. The first direction was normal time ordering, where spallation events were looked for after a muon with a neutron cloud. The second direction reversed the time ordering, looking for spallation preceding the muon in the area of the cloud.

For a cloud to be considered, it would have to follow a "2 + 1" criterion. Where the 2 represents the minimum number of events seen in WIT within 5m and 500 μs for the cloud to be considered, and the 1 is the minimum number of "good" events with additional requirements of $\Delta t >$20 μs, $g_{vertex} >$0.5, reconstructed within the ID, and energy <6 MeV. The cloud center was only determined using events that met the good event quality cut. Neutron multiplicity is defined as the number of events that are seen by WIT after the muon within 5 m and 500 μs. Some figures include multiplicity of 1, even though it violates the 2+1 criteria. This is included to emphasize the motivation to maintain the 2+1 criteria as a means to increase sample purity.

An energy cut was also placed on the spallation events to be between 6-20 MeV kinetic to look for the events in the spallation dominated range. 18.8 MeV is the endpoint for hep solar neutrinos, however energy resolution requires a slightly larger energy region

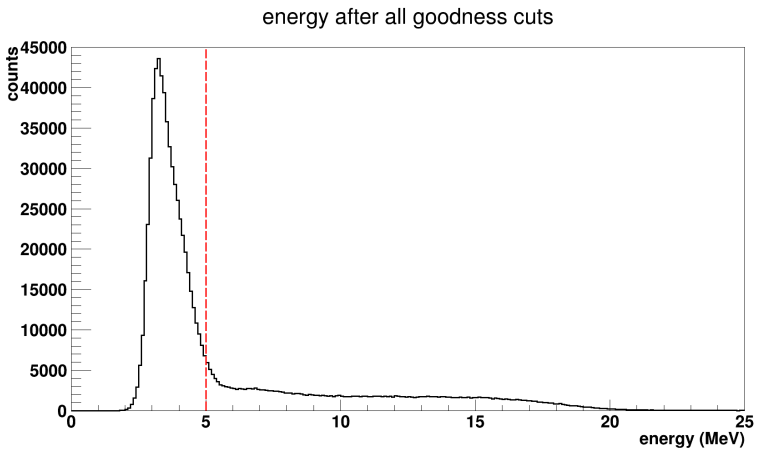


Figure 6.6: Energy distribution for all neutrons seen in WIT after final event reconstruction quality cuts applied with short Δt removed. The red line corresponds to energy cut value decided by MC. The apparent peak of the energy spectrum is from the product of a rapidly changing efficiency and the tail of the 2.2 MeV gamma spectrum. The true peak is expected to be $\sim$1.9 MeV. The extending tail is possibly from the early decay of spallation within 500 μs of the muon.

to explored. Spallation beyond 20 MeV is minimal, and with most having their endpoint energy between roughly 10-16 MeV. ^{1}2N with an energy of 16.9 MeV is the highest energy spallation with significant yield. This 16.4 has a chance to be reconstructed at a higher energy. There are two very short lived spallation isotopes with an endpoint around 20 MeV Kinetic (^{14}B at 15.06+6.09γ MeV and ^{11}Li at 21.13 MeV), but their yield is insignificant and lifetime make them easy to tag with any form of Δt cut.

In order to draw conclusions from the sample, all investigated variables would be compared by the normal time ordering (SIG) - reverse time ordering (BG) vs reverse time ordering (BG). For vertex correlation, the coordinate system was taken to be spherical with no zenith or azimuthal dependence. The distance of spallation from the neutron cloud center, R, was the only spatial correlation variable used, and its cube (R^3) was examined to flatten the phase space.

A sharp peak in the SIGNAL distribution for the vertex correlation showed spallation was successfully being tagged. The normalized SIG -BG vertex correlation, shows how

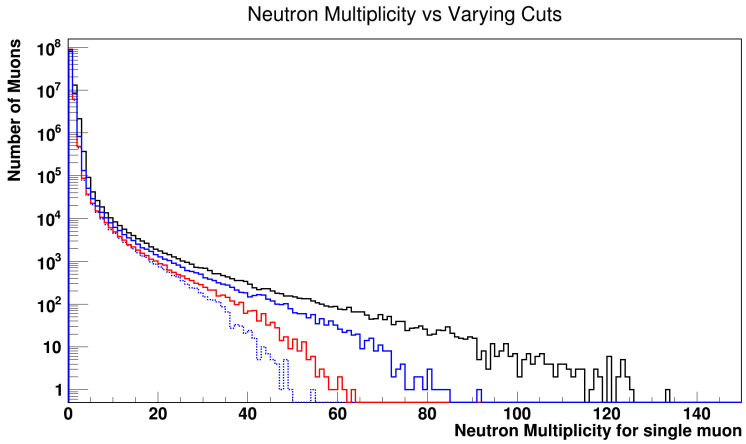


Figure 6.7: Neutron multiplicity distributions for varying cuts. Black represents raw multiplicity with no cuts applied and all others have the short Δt removed (Δt>20μs). Blue dashed corresponds to the $g_{vertex} > 0.5$ requirement in the initial study. Solid blue (red) correspond to loose (tight) cut based on g_{vertex} and g_{dir}

accuracy to tag spallation increases with neutron multiplicity. The reason is two-fold. First, due to the likelihood of accidentals in the neutron cloud with low tagging efficiency, the actual cloud center is better defined when the proportion of desired events increases. Second, with neutron multiplicity increasing, the shower size and energy deposit is also larger, increasing the likelihood for spallation to be produced.

The second point is backed when plotting neutron multiplicity against spallation multiplicity and counting occurrences. There is a positive correlation with increasing spallation yield as an increase in neutron multiplicity occurs. The deadtime accrued by each bubble around the neutron cloud center is essentially identical in this initial study, with only detector boundaries having the potential to shrink volumes. The integral of the neutron multiplicity plot is roughly proportional to the deadtime accrued, with the largest single contributing bin being the first. The total amount of potential spallation removed by the single candidate bin is less than that of the second bin, while accruing significantly more deadtime. This reinforced the motivation to use the 2+1 criteria for events.

When looking at the timing delay, it is necessary to note that there are three major

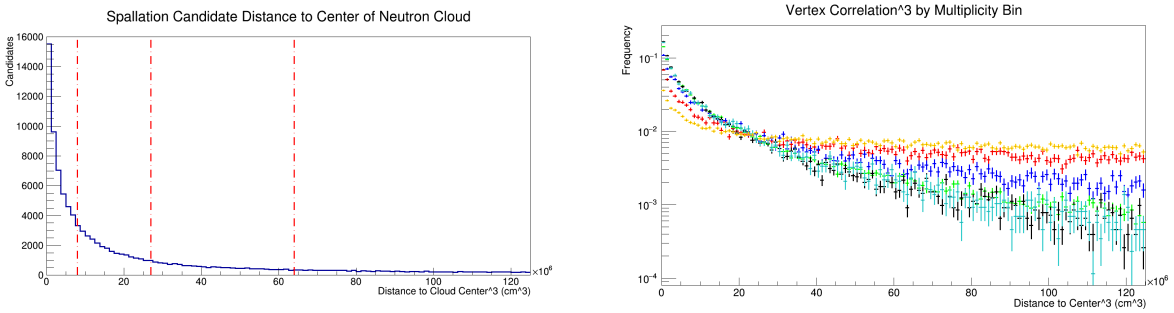


Figure 6.8: Vertex correlation distributions for spallation candidates to neutron cloud center. The left plot shows the total spallation signal region without background subtraction. Each red vertical line from left to right corresponds to a 2m, 3m, and 4m distance to center respectively. The tight correlation between spallation events to neutron cloud center can be seen. The right figure shows the dependence of the R^3 on neutron multiplicity. This is background subtracted and normalized. With a greater number of neutrons, the neutron cloud location is better known, and the likelihood that correlated spallation will be located nearby

spallation regions:

1. Short-lived ($\tau \sim 10$ms)
2. Medium lived ($\tau \sim 1$s)
3. Long lived ($\tau \sim 10$s)

^{16}N being the main concern as it is the highest yielding spallation in out detector, with decay constant $\tau = 10.3$s and being produced from the n-p interaction on Oxygen:

$$^{16}\mathrm{O} + n \rightarrow \ ^{16}\mathrm{N} + p \tag{6.3}$$

A multi-exponential fit can be made to the data as follows

$$\mathrm{N}(\Delta t) = \sum A_i e^{\frac{-\Delta t}{\tau_i}} \tag{6.4}$$

where A_i is the decays/second of isotope i, with its time constant τ_i. If τ_i is left free, we can

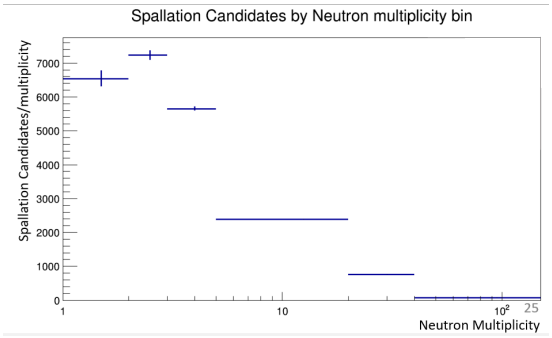


Figure 6.9: The background subtracted spallation totals per multiplicity bin.

restrict the time range to see if the expected isotopes are present. In the long lived region, we expect most contribution from ^{16}N with some contribution from ^{11}Be. The fitted region for 10s-45s, with a single fit comes out with a decay constant of:

$$\tau(^{16}\mathrm{N}) = 9.94 \pm 0.24 \text{ s} \tag{6.5}$$

just over 1σ from the ^{16}N expected decay constant. Other time regions are also in agreement with expectation.

As alluded to before, spallation products as a background in the solar analysis are generally an issue in the 6-20 MeV range, with the most prevalent form of spallation product,

With spallation successfully found after a neutron cloud, the overall effectiveness of the cut was checked. In order to check this, $\cos(\theta_{sun})$ was used in tandem with the two time direction samples. The signal direction would be able to give a spallation removal effectiveness calculation, while the reverse time cut would be able to give a deadtime calculation. To minimize influence from the solar peak for deadtime calculation, the negative $\cos(\theta_{sun})$ region would be fit with a constant and compared to the total sample. The former spalla-

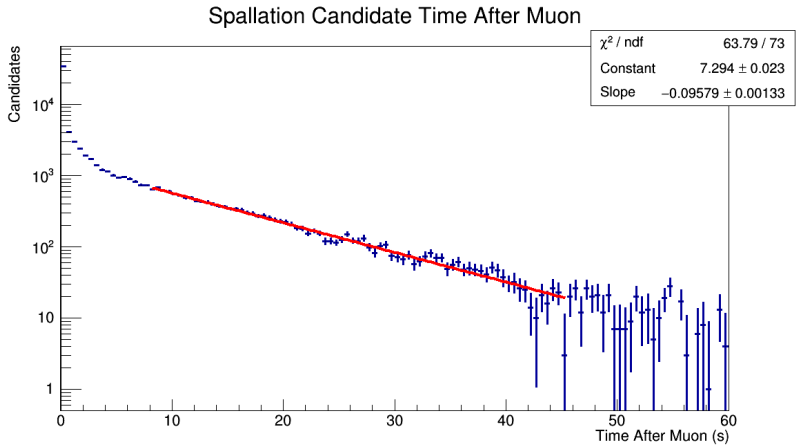


Figure 6.10: The background subtracted Δt distribution for spallation candidates found within 5m of a neutron cloud containing at least two events, and one event meeting quality event criteria. The red line is an exponential fit from 8s to 45s, corresponding to the ^{16}N dominated time region. The fit comes to roughly 1σ higher than the true decay constant for ^{16}N

tion log likelihood was used as a baseline for the effectiveness. FIg6.12 shows a comparison between new and old events. When using the 60s 5m cut, 56% of spallation is removed with 4.1% deadtime, compared to 90% effectiveness and 20% deadtime for the former cut. Looking in at the figure, the small deadtime for WIT is seen by a lack of peak (corresponds to removed solar signal events) in the blue plot. The large deadtime for the solar spallation removal shows a small peak in the removed events in the forward $\cos\theta_{sun}$ region.

With minimal tuning, the results show promise as to the ability to tag a significant amount of spallation without much deadtime. However, the neutron cloud is not able to compare in overall spallation removal effectiveness. Due to the low efficiency of the WIT system to trigger on neutrons and provide a reasonable fit, there aren't enough tagged events to be able to completely replace the old method. Therefore, a combined method with a tuned neutron cloud cut and while updating other spallation tagging techniques was pursued.

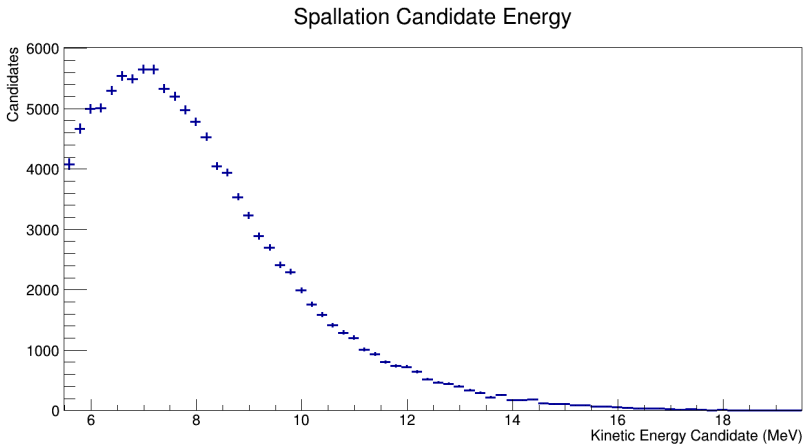


Figure 6.11: Initial spallation sample energy distribution, with background subtraction. Peak of distribution is around 7MeV, which is consistent with expectations for the energy of most common spallation isotopes.

6.4 Neutron Cloud Cut Tuning

With proof of neutron clouds and subsequent spallation, neutron selection criteria were refined to implement a cut. The goodness criterion was tuned by comparing the location of truth vertex positions for Monte Carlo generated 2.2 MeV γs and the reconstructed vertex. A cutoff was made at 5m vertex difference for whether a neutron reconstruction was "good" or "bad", to be on scale with the 5m cylinder cut for searching for neutrons after a muon. Using the goodness calculation and g_{dir} as explained in Sec. 4.3.2, the two samples were compared.

Circles were placed over the peak of the "good" neutron distribution, with two radii corresponding to different proportions of good to bad neutrons. Since the "bad" neutron distribution does not extend to the higher goodness, lower g_{dir} corner of the distribution, the circles were made into quarter circles and their maximum g_{dir} and minimum g_{vertex} values extended in a straight line. This resulted in three sections, with each event's contribution to the cloud position reconstruction being weighted on which region it was located in. The center of the circles were placed at ($g_{vertex} = .53$, $g_{dir} = .23$), with the circle radii being 0.08

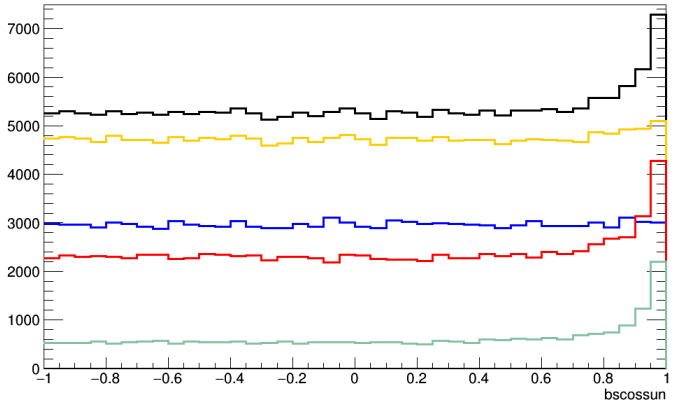


Figure 6.12: Solar peak distribution for the WIT period. The black line corresponds to the total solar neutrino events in the time period. The yellow (green) line corresponds to the former spallation cuts removed (remaining) events. The blue (red) line corresponds to removed (remaining) events from the neutron cloud identification.

and 0.12:

I Poorly reconstructed events (weight = 0)

II Moderately reconstructed events (weight = 1)

III Likely well reconstructed vertices (weight = 2)

The reconstruction energy was examined in the MC as well to better tune the energy cut for a good neutron. No event was found to have an energy greater than 5 MeV, the good neutron cut point was reduced.

To this point spheres had been used to parameterize the neutron clouds seen after muons. Showers should however have some form of direction due to momentum conservation. For the shape of the shower, three approaches were investigated in order to gauge which would be most effective at tagging spallation. Each of them would have the goal of transforming into cylindrical coordinates, with one major thrust axis, and the minor axes treated radially symmetric. The methods were formed from taking combinations of muon centered vs neutron centered clouds, and neutron shower direction vs muon direction:

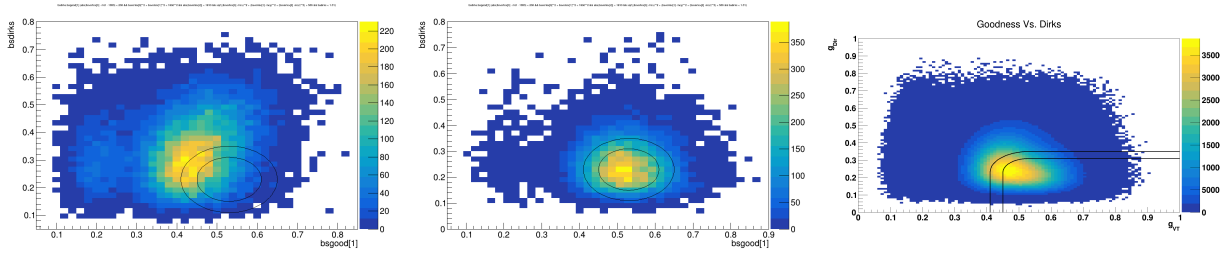


Figure 6.13: Comparison of "bad" (left) event reconstruction and "good" (middle) reconstruction g_{vertex} vs g_{dir} distributions. 2.2 MeV gammas were simulated in WIT, with events with a reconstructed vertex more than 5m from truth considered bad, and those within 5m good. The yellow peak of the good distribution is significantly shifted to the lower right, and weighted regions take this into consideration. The right distribution shows the distribution for the entire WIT data period.

I Neutron centered, major shower thrust axis

II Muon track centered, major shower thrust axis

III Neutron cloud centered, muon axis

IV Muon track centered, muon axis

Method I followed the idea of the using the thrust axis of the neutron cloud, independent of the muon, to parameterize the shower shape. The number of good neutrons needed to adequately define the shower direction and center was too great. This method would only be able to be used on 5% of muons. With this in mind, method II was bypassed as the shower thrust direction suffered more from poor neutron statistics. Method III followed the notion most showers would have a relatively similar thrust direction as the muon. If the principal axis was taken to be the same as the muon's, the multiplicity requirement of method I would be unnecessary. Method IV relied on the facts that most neutrons are not tagged and reconstruction is generally poor. It used the muon track as the axis for the center of the cloud, while using the relative neutron information to set the origin of the cloud on the axis.

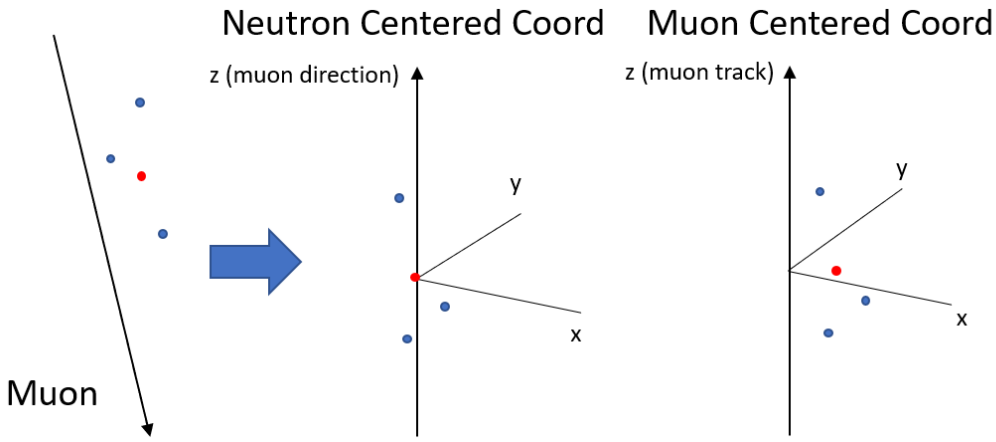


Figure 6.14: Figure illustrating the two main potential coordinate systems. Shown is a mock muon with neutrons (blue) and their center of mass position (red). Both use the muon direction as the +z direction, and the left one has the z-axis pass through the neutron cloud center of mass and the right one maintains the original muon track

Separating the clouds into 5 groups, based on multiplicity, background subtracted spallation profiles were made for method III and IV. The five multiplicity groups were (2,3,4-5,6-9, 10+). For the time being the 60s was maintained for searching for this spallation. For all of the groups, looking at the ln v. lt^2 plot for the distributions showed tighter correlation in the method IV, with difference being bigger the less neutrons were in the cloud. The uncertainty in the neutron position caused too much uncertainty to the center. When the multiplicity became large enough, cloud position was more well defined, and it could compete with the muon centered method.

With the muon centered coordinate system, tuning was performed in the individual multiplicity bins. The fraction of each multiplicity bin is:

Depending on the shower size, an ellipsoid is cut around where the neutron cloud center translates back to the muon track. Most gain to the reduction of deadtime to be made was in the 2 neutron case. The category has the most uncertainty in cloud location and accounts for almost ¾ of the showers. The 60s cut from the previous section is maintained for all categories of clouds, except for the 2-neutron case where it is reduced to a 30s cut.

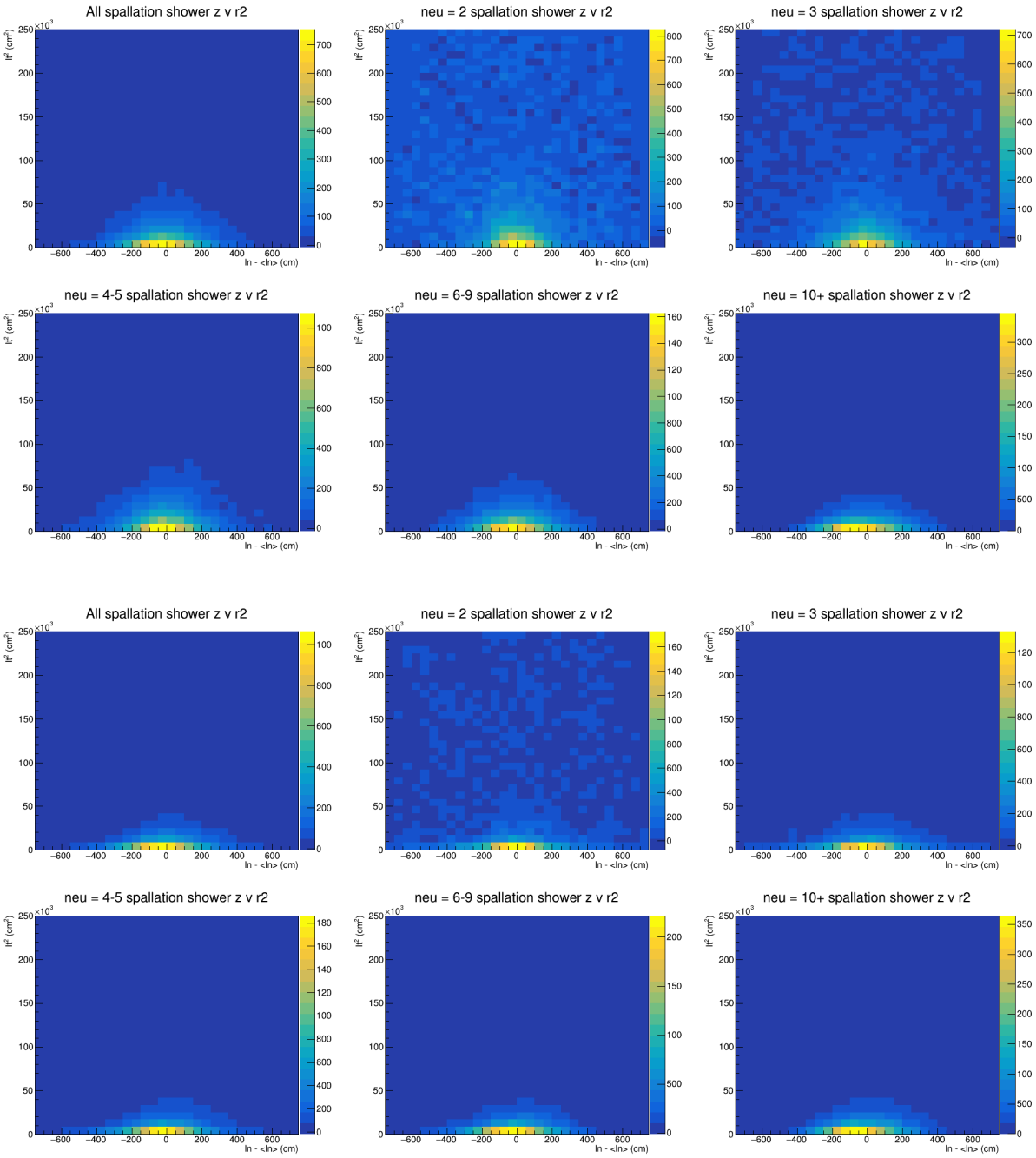


Figure 6.15: Figures showing the six different neutron multiplicity categories and the corresponding lt^2 vs ln distributions for neutron cloud centered (top) and muon centered (bottom). Cuts based on the muon centered coordinate system have better spallation tracking, allowing for smaller volume cuts to be made, especially for low tagged-neutron multiplicity showers

Multiplicity	2+	2+	2	3	4-5	6-9	10+
% of Showers	100	100	73	15	6	3	3
Δt	0.2	2	30	60	60	60	60
ln [cm]	750	500	350	500	550	650	700
lt [cm]	750	500	200	245	346	447	500

Table 6.2: Table showing the different cut conditions for the cloud cut. + is used to represent showers of at least that multiplicity.

An ellipse was made in ln and lt^2 for the five categories, increasing in both lt^2 and ln for increasing multiplicity. Table 6.2 outlines the various cut conditions. The shower multiplicity is proportional to the energy deposited and overall shower size. With larger showers, there is greater confidence subsequent spallation is related. A sphere of 7.5m was used to cut events within 0.2s of the muon and neutron cloud, and a 5m sphere was still used within the first 2 seconds. These values were chosen by comparing cumulative accidentals to the cumulative spallation removed. 0.2s covers the short-lived spallation lifetime and 2s covers a large fraction of the medium lived spallation. These spheres accrue minimal additional deadtime compared to the multiplicity specific cuts.

With all of these together, a similar efficiency to the 5m universal bubble cut is achieved at 54%, while reducing the deadtime to 1.38%. As stated before, this is not able to cut a satisfactory amount of spallation by itself, other improvements will need to be made in order to get similar efficiency to the old method. The period for which neutron data in WIT was available in SK-IV is 388 days.

6.5 Multiple Spallation

A novel idea is to use solar neutrino candidate events to tag themselves and other events as spallation if multiple events are found close to each other in space and time. With approximately 24 solar neutrinos a day, the likelihood for two solar neutrino events to be

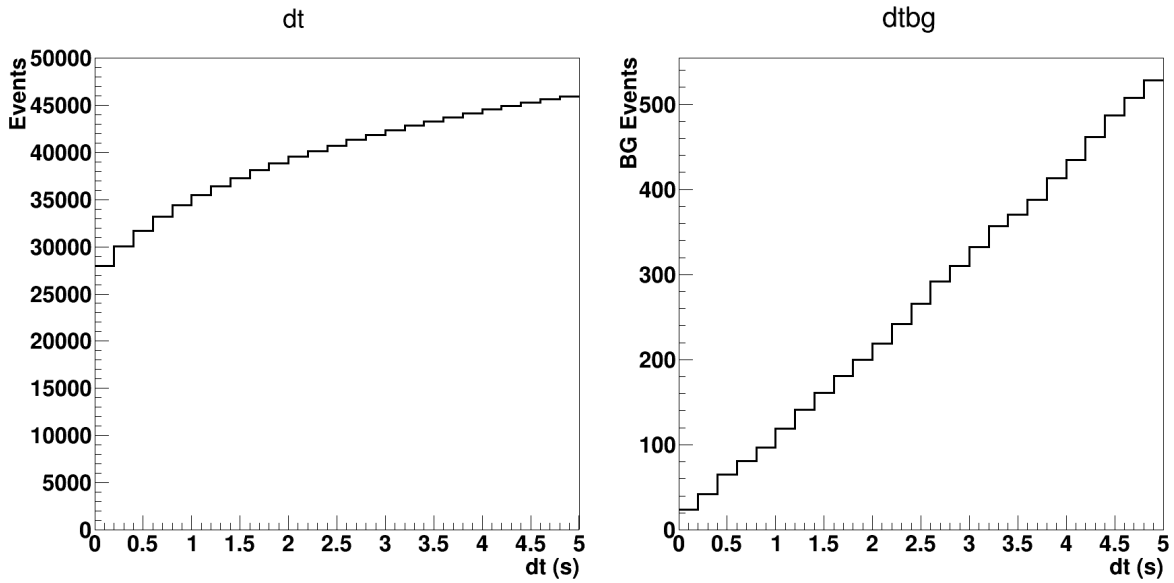


Figure 6.16: Cumulative Δt distributions for signal (left) and background (right) assuming a 5m bubble cut for any cloud. Much of the spallation is removed in the first couple seconds, and extending the bubble cut for low multiplicity and small Δt adds minimal deadtime to the overall amount.

adjacent to each other in time and space is relatively low. If events passing solar neutrino sample cuts are used to tag other events, then most multiple spallation events can be cut without a significant increase in deadtime. This method is also useful for two main reasons, it can be applied to all of SK-IV (as well as SK-I, II, and III) and does not rely on the muon. Previously in all phases of SK, only information pertaining to muons had been used to cut spallation events. This is troublesome when a muon has a poor fit or no fit at all. Being able to independently tag spallation unrelated to the muon helps to compensate for some of the short comings of both the spallation log likelihood method, as well as the neutron cloud.

The overarching goal was to have events of good quality and not due to other backgrounds. Criteria were determined for what event would be used to tag multiple spallation. Events making the final solar sample, minus two cuts would be candidates for a multiple spallation check. The two removed cuts were the spallation log likelihood, as well as the hit pattern likelihood cut. Spallation log likelihood by definition cuts spallation. The hit

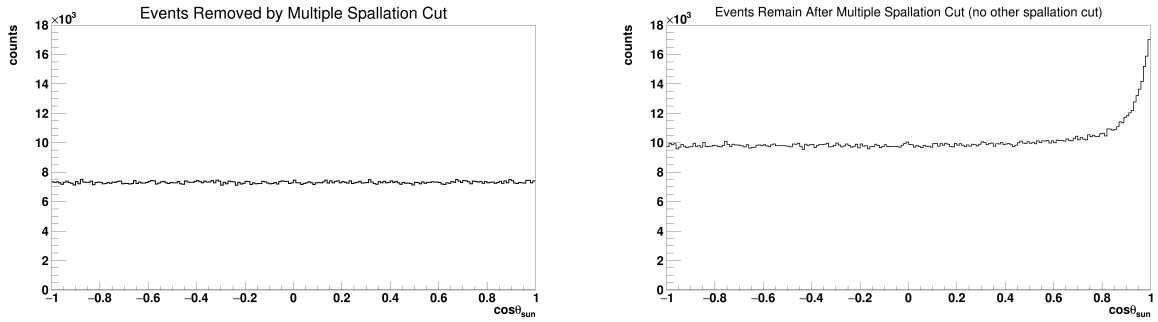


Figure 6.17: Comparison between events cut (left) from the final sample to those that remain after the cut (right). No evident peak in the left distribution shows the relatively small deadtime created by this cut.

pattern likelihood was removed due to its sensitivity to cut $\beta\gamma$ events. A significant amount of the spallation decays include a $\beta\gamma$ branch, so to remove them too early would reduce the cut's efficiency the cut. To avoid radioactive background that would be found in the lower energy regions, a 6 MeV cut was put on the events used to veto others. The cut was tuned to be a 4m spherical bubble for 60s and ultimately was able to remove 47% of spallation, with a deadtime of 1.3%.

Multiple spallation deadtime cannot be calculated with the reverse time cut and is unable to have a direct $\cos(\theta_{sun})$ comparison for the deadtime. A rough idea for the deadtime can be seen however trying to find an indication for a solar peak when looking at which events are cut. Since events tag each other, reversing the time direction for normal data doesn't remove the correlation between the events like it does for muons. The method for calculating the deadtime for multiple spallation, as well as the final deadtime estimations is covered in Sec 6.12.

6.6 Spallation Removal With Only Preceding Muon Information

With the cloud and multiple spallation cut together, roughly 65% of spallation is cut at only 2.4% deadtime. As this is significantly less than the desired 90% removal efficiency, an additional cut similar to the former spallation log likelihood based on the preceding muons' information is needed. Using the previous log likelihood as a template, different modes of improvement were investigated. The PDFs were tuned on the remaining spallation sample not cut by the multiple spallation and cloud cut. The former method cuts events from the sample using a log-likelihood ratio method, with the cut point for the log-likelihood-ratio decided by setting a desired spallation removal effectiveness and deadtime combination:

$$\mathscr{L}_{TOT} = \Pi_i \mathscr{L}_{sig,i}(x_i)/\mathscr{L}_{BG,i}(x_i) \tag{6.6}$$

where $\mathscr{L}_{sig,i}$ and $L_{BG,i}$ are the probability density functions (PDFs) for the spallation signal and background for a given variable x_i. When the logarithm of $\mathscr{L}_{TOT}$ is taken, the product becomes a difference of sums:

$$\log_{10}[\mathscr{L}_{TOT}] = \log_{10}[\prod_i \mathscr{L}_{sig,i}(x_i)/\mathscr{L}_{BG,i}(x_i)] \tag{6.7}$$

$$= \sum_i \log_{10}(\mathscr{L}_{sig,i}(x_i)) - \log_{10}(\mathscr{L}_{BG,i}(x_i)) \tag{6.8}$$

For some of the background likelihoods (e.g. Δt), the background is considered constant and the second term in that section of the sum can be dropped. If the individual PDFs are broken into multiple categories, the PDF needs to be normalized in order for the different categories to be treated equally. If there is only one category to be treated, then normalization can be

skipped because the cut value can absorb it since the addition is a constant:

$$\log_{10}(L_{norm}) = \log_{10}(NL) = \log_{10}(L) + \log_{10}(N) \tag{6.9}$$

where N is a normalization constant and L is the original PDF. Using this method, correlation between the variables is somewhat neglected by basing the individual PDFs on only one variable. This is mitigated by breaking up a PDF into sub categories where appropriate. The former likelihood was dependent on Δt, residual charge (Q_{res}), and lt. Residual charge is is the excess light seen within the detector compared to expectations for a minimum ionizing particle (MIP):

$$\mathrm{Q}_{res} = Q_{tot} - L_{muon} \cdot (\frac{\mathrm{d}Q}{\mathrm{d}L})_{MIP} \tag{6.10}$$

where Q_{tot} is the total charge for the muon, L_{muon} is the muon length, and $(\frac{\mathrm{d}Q}{\mathrm{d}L})_{MIP}$ is the expected charge deposition rate for a minimum ionizing particle. The MIP charge loss rate is typically near 26 p.e./cm, and varies in time due to detector conditions.

6.7 Muboy vs mufitpe

There are two main muon fitters used within the SK solar analysis, mufitpe and Muboy. The former spallation log likelihood was tuned using mufitpe. A brief study into the effectiveness of the two fitters was conducted to see which performed better. The first comparison was the lt^2 correlation of possible spallation events to the two fitters. To ensure a relatively high signal to noise, only spallation candidates within 100 ms of a muon was used. The resulting lt^2 distribution shows a sharper peak for Muboy at lower lt^2 and approaching 0 at a faster rate.

There was also a tendency for mufitpe to perform poorly at higher energy muons. This was even taken into account for the Q_{res} likelihood, as once the event muon had more

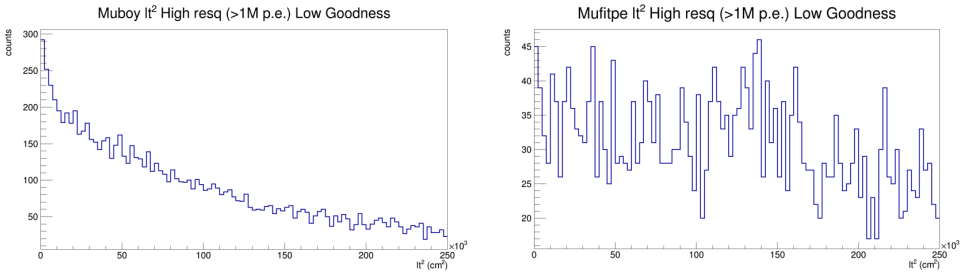


Figure 6.18: Comparison of ability for Muboy (left) and mufitpe (right) to fit very bright muons, as measured by spallation transverse distance to the fitted track. Muboy successfully reconstructs more high energy muon, and the spallation tracking is significantly better than mufitpe. The PDF for lt in the former likelihood method was disregarded (set to 1) due to the poor tracking of mufitpe in this energy range

than 1 Mpe of energy and a mufitpe goodness less than 0.4, the lt contribution to the likelihood was ignored. In order to translate the goodness cutoff between Muboy and mufitpe, the goodness of Muboy events over 1 Mpe was plotted, and there was a separation in the distribution, allowing for a poor goodness to be decided to be around 0.25. For mufitpe, there was poor correlation in this category, while for Muboy there was a strong correlation.

6.8 Time Correlation

The time difference PDF was constructed by taking events within 3m of the muon track. The background subtracted distribution was then fit with a sum of seven exponentials corresponding to the main decay constants:

$$N(\Delta t) = \sum_i A_i e^{-\Delta t/\tau_i} \tag{6.11}$$

Here the decay time constants, τ_i, were fixed to match the decay times of expected isotopes. It is of note ^{11}Be production is consistent within 0.25σ of 0. Of the three component PDFs to the former full likelihood, this is the only one of the same form as before. The background distribution was fit with a linear function, and found to be consistent with having no slope.

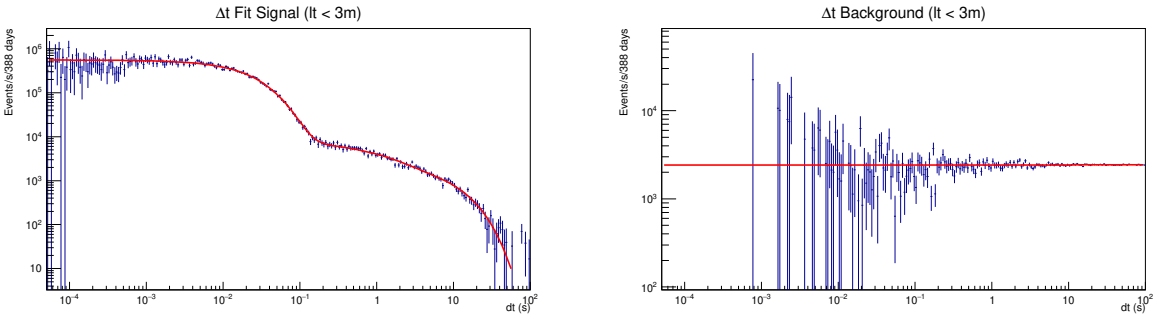


Figure 6.19: Signal (left) and background (right) distributions for Δt, with only a 3m lt cut applied. Both distributions are binned so the width of each bin is flat in log(Δt). Signal region is fit with the seven principal isotopes, with their decays constants fixed, and each isotopes contribution to the total fit allowed to vary. Background is expected to be flat, and fit with a linear equation to test this. The fit is 1σ consistent with flat, and is most likely a result of the binning and poor statistics in the low Δt regime.

Therefore, for the log likelihood calculation, only the signal PDF is used.

6.9 Transverse Distance Correlation

The former likelihood used a triple exponential in the lt phase space:

$$PDF_{old}(lt) = p_0 e^{p_1 \cdot lt} + p_2 e^{p_3 \cdot lt} + p_4 e^{p_5 \cdot lt} \tag{6.12}$$

Consistent with the previous sections, the lt distribution was translated to lt^2 and PDFs were developed from there. For the signal region the, the equation followed the same form, but with lt^2 instead:

$$PDF_{new,SIG}(lt^2) = e^{p_0 + p_1 \cdot lt^2} + e^{p_2 + p_3 \cdot lt^2} + e^{p_4 + p_5 \cdot lt^2} \tag{6.13}$$

The background PDF is approximately flat for short lt^2, but due to detector size constraints, falls to 0 at large lt^2. The volume at large lt^2 is not a cylinder, but the intersection of

the cylinder surrounding the muon track with the cylinder which describes the ID fiducial volume. To deal with this the background was fitted with a piecewise function for the different categories:

$$PDF_{new,BG} = \begin{cases} p_0 & lt^2 \leq x_0 \\ p_0 e^{p_1(lt^2-x_0)+p_2(lt^2-x_0)^2} & lt^2 > x_0 \end{cases} \tag{6.14}$$

x_0 was only allowed to take discrete values, and ten fits were made for $x_0 = \{n^2 \text{ m}^2 : n \text{ is an integer, } 0 \leq n \leq 9\}$ with the p_i constants allowed to vary. x_0 is the point in which the function switches from a constant to the Gaussian tail, and was chosen to be discreet to obtain better function convergence across all categories. The x_0 chosen for each case was the one whose fit had the lowest $\chi^2/d.o.f.$

The prior method subdivided the lt region into seven different categories based on Q_{res}. With larger charge deposit, the lt^2 distributions is expected to change. The more energy a shower deposits, the larger the shower and the higher the energy of individual particles will be. This increases likelihood for spallation being produced farther from the track. The categories were broken into the nominal time intervals for the three lifetimes of spallation:

I Short lived ($0s \leq \Delta t \leq 0.1s$)

II Medium lived ($0.1s \leq \Delta t \leq 3s$)

III Long lived ($3s \leq \Delta t \leq 30s$)

A significant difference between the distributions was found, with tighter track correlation for shorter time ranges. The number of lt^2 categories was tripled to accommodate for this. Category III was used for times greater than 30s. The bins for residual charge

remained consistent with the former method, with the bins being {$-\infty$-0, 0-25, 25-50, 50-100, 100-500, 500-1000, 1000-∞} kpe.

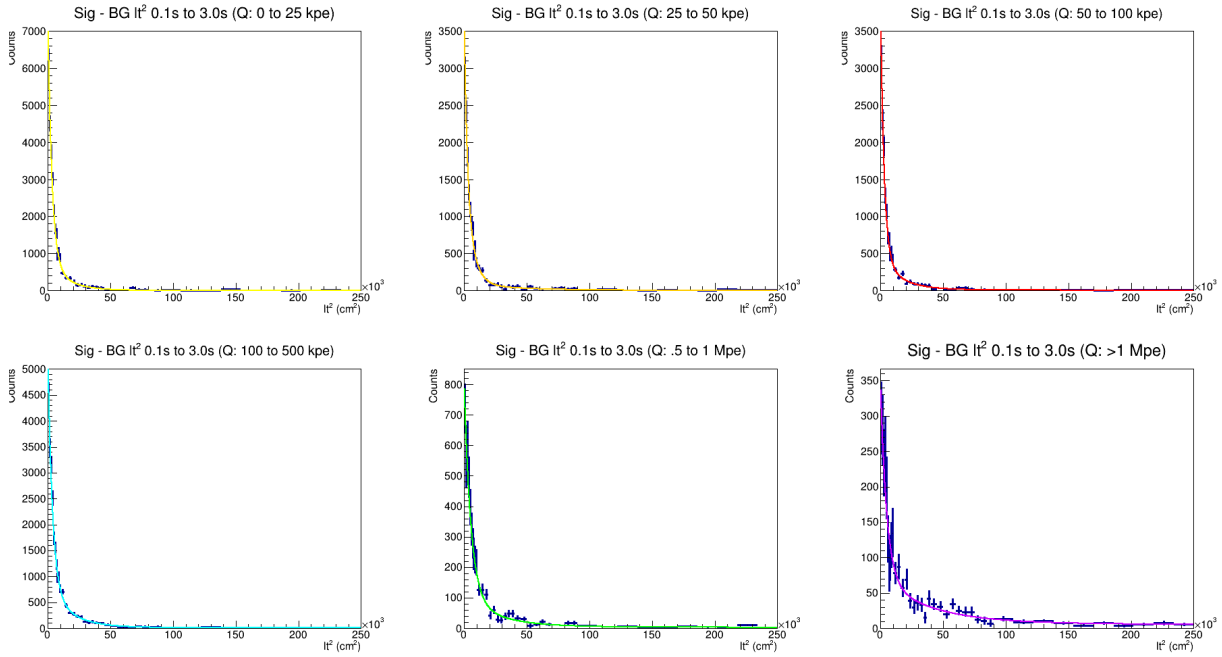
6.10 Residual Charge

The third and final refitting was to the residual charge distribution. There are two major changes coming with the recalculation. As discussed in Sec 2.7.1 before SK-IV, the electronics could only handle up to $\sim$250 p.e. for a single hit. With the update to the QBEEs, the charge saturation of a channel would raise to around 4000 p.e. The former likelihood was developed with the saturation effect, and for SK-IV the saturation was mimicked to keep calculations consistent. For this work, the artificial saturation was removed.

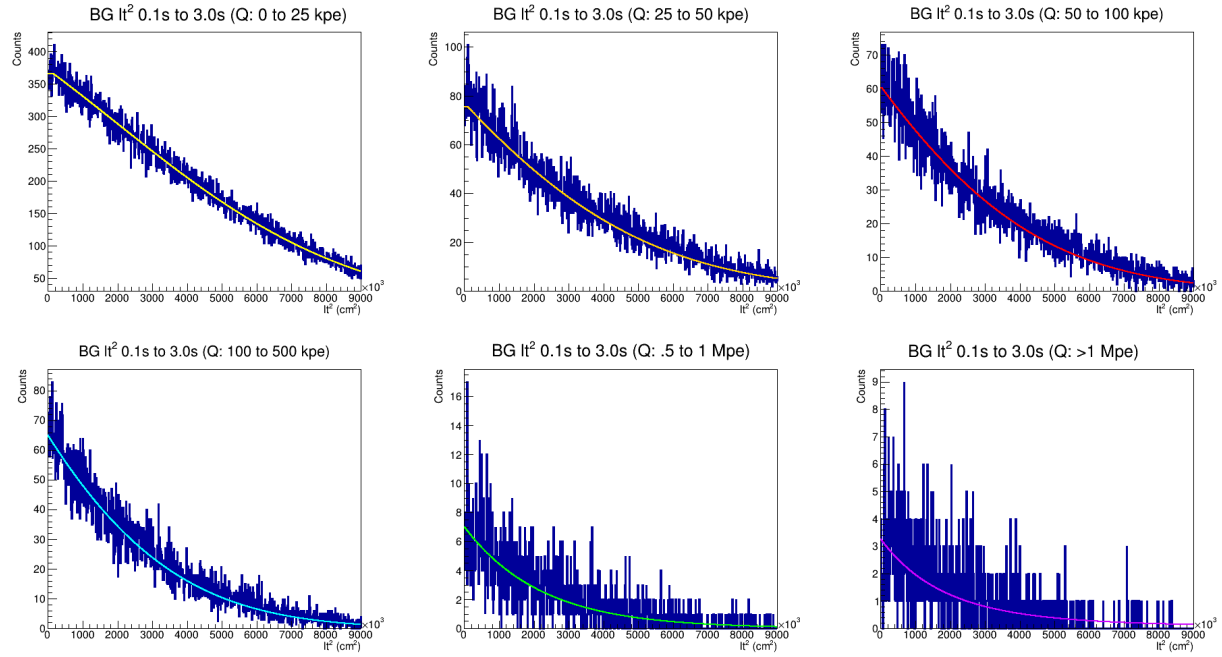
To create the sample to fit to, all muons within 10s of an event were taken. There was no cut on the lt distribution for the muon. Although Muboy was better at maintaining track fits, it still has some issues at very high residual charge.

The PDF was again defined piecewise, because of the poor behavior in the negative residual region. The signal and background region were reasonably similar in form, and therefor the PDF took the same form for the two regions. In the distribution, there were three negative residual charge bins. Linear interpolation was used from the center of these bins for the sake of the calculation. The positive region is a quintuple exponential as follows:

$$\begin{aligned} PDF_{new}(Q_{res}) = e^{p_0+p_1\cdot Q_{res}} + e^{p_2+p_3\cdot Q_{res}} \\ + e^{p_4+p_5\cdot Q_{res}} + e^{p_6+p_7\cdot Q_{res}} + e^{p_8+p_9\cdot Q_{res}} \quad (Q_{res} \geq 0) \end{aligned} \tag{6.15}$$



(a) Signal distributions and PDFs



(b) Background distributions and PDFs

Figure 6.20: Signal (a) and background (b) distributions for the six highest residual charge bins, and middle Δt (0.1 s$\leq \Delta$t $<$ 3 s. PDFs are overlaid on the distribution to show the fits used in the final likelihood.

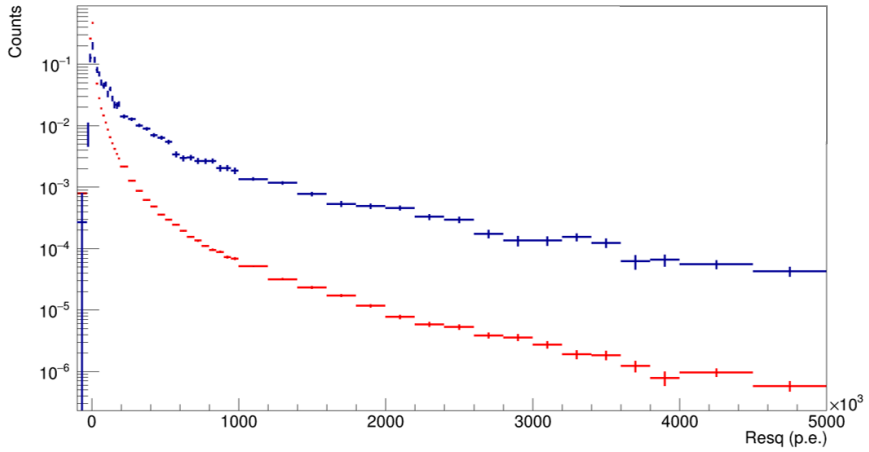


Figure 6.21: Signal (Blue) and Background (Red) PDFs for Q_{res}.

$$PDF_{new}(ln) = \sum_{i=1}^{3} A_i e^{\frac{(ln-ln_i)^2}{2\sigma^2}} \tag{6.16}$$

6.11 dE/dx

The fourth change to the likelihood was the addition of another parameter to characterize events. dE/dx profiles try to check for the energy loss of a muon as a function of position along the track. The method is used partially in Muboy for trying to measure track length for stopping muons. The dE/dx method was further developed in the Diffuse Supernova Neutrino Background (DSNB) search performed by Kirk Bays [1]. The Bays' method used PMT timing and hit geometry to map where along the track light would need to be emitted to solve the proper time equation:

$$T_{pmt} = l \cdot c_{vac} + d \cdot c_{water} + T_{muon} \tag{6.17}$$

where l is the length from entry point to emission point on the track, and d is the distance from the emission point to the PMT. d is dependent on l, and the solution in terms of l

reduces to a quadratic:

$$\frac{c^2(t-t_0)^2}{n^2} - r^2 + 2\left[\overrightarrow{\mathbf{r}} \cdot \widehat{\mu} - \frac{(t-t_0)}{n^2}\right] X + \left[\frac{1}{n^2} - 1\right] X^2 = 0 \tag{6.18}$$

where t_0 is the time the muon entered the detector, as fit by Muboy, t is the PMT time, $\overrightarrow{\mathbf{r}}$ is the vector from the muon entry point to the hit PMT, $\widehat{\mu}$ is the muon direction from Muboy, and X stands for $c(t_e - t_0)$ with t being the emission time of the Cherenkov photon from the muon track. t_e is the quantity that is being solved for in this equation.

In Bays' analysis 0, 1, or 2 solutions would be found. The track was broken into 50 cm bins, corresponding to roughly the low energy vertex resolution of SK, and the amount of charge placed into the dE/dx would be equal to the charge seen in the PMT. If there were two solutions, two positions would be filled with the PMT charge.

In 2015, Beacom and Li proposed improvements to Bays' dE/dx method, addressing what they described as the three improvable limitations. From their paper they state "First, when Δ is large, taking both solutions includes wrong information and over counts PMT hits. Second, when Δ is negative due to detector resolution, taking zero solutions ignores correct information and under counts PMT hits. Third, when Δ is small, the sensitivity of the solutions to detector resolution dilutes better information." [62] The Δ from the quote is in reference to the discriminant from Eq. 6.11. To address their concerns a new dE/dx method was developed.

The first issue addressed was the second point from Beacom and Li, which was allowing for solutions where the quadratic does not intersect with the track to be included. At the same time, the amount of charge contributed by a PMT was normalized to its charge amount, removing double counting and most non-counting instances.

The PMT hit times corresponding to each dE/dx bin boundary were determined

using Eq. 6.11, and the charge was spread across the dE/dx bins based on the bins' fraction of the PMT time residual distribution. That distribution was assumed Gaussian with a width depending on the PMT charge based on the PMT timing calibration data varying from a few ns down to <0.5ns. To integrate the Gaussian, the error function was evaluated at the bin boundaries or at extrema (see below) of Eq. 6.11. A typical PMT charge will distribute across bins with two peaks of roughly Gaussian shape corresponding to the two solutions of Eq. 6.16. This is not a true Gaussian along the muon track because of the non-linearity of the time as a function of l.

When doing the integral in this basis, the independent variable is non-monotonic in most instances and there are commonly three local extrema: two local maxima and one local minimum. The integral is broken up at the local extrema to maintain monotonicity on the interval of integration. If there are two true solutions to the time equation, this case is encountered. A bin contributes to the overall integral so long as one of the two bin boundaries are less than 5σ from the PMT time. This is also the cutoff for when the 0-solution case is forced or not. As long as one bin boundary on the length of the track is less than 5σ, some amount of charge will be filled into the distribution.

Since the sum of the bin contributions is not inherently normalized due to the non-linearity of t along the track, the bin contributions are scaled to ensure that their sum equals the measured PMT charge. This is what becomes the normalization factor for a PMT hit.

SK does not have full photo-cathode coverage and PMTs shadow each other, and therefore the angle of which the light hits the PMT is taken into consideration for a charge correction. However, for the dE/dx, there will be many angles of incidence to take into consideration. To do this, the distance for the bin center is used for the correction calculation. An incident zenith and azimuthal angle is found for light emitted from each bin center to the PMT, and the correction calculated. There is also a correction for the water transparency.

The calculated water transparency for the run is used to make a correction for light lost traveling from the point until the PMT due to attenuation.

Once the dE/dx has been constructed, the peak is used in order to identify spallation. The center of the bin with the greatest charge represents the shower peak. In Bays' method, a nine bin average was needed to identify the shower peak, but with the Gaussian charge distribution across bins applied an average was no longer needed [1]. This location is then used to calculate a longitudinal distance (ln) for the candidate spallation event to the peak. A likelihood was formed for the signal and background in the same way as previous sections. For the distributions a triple Gaussian was used to fit the signal, and a double Gaussian was used to fit the background.

$$PDF(ln) = \sum_{i}^{3} A_i e^{-\frac{1}{2}\frac{(ln-ln_i)^2}{\sigma_i^2}} \tag{6.19}$$

Where A_i, ln_i, and σ_i are the fit parameters for the triple Gaussian. [1] For the purity of the sample, a cut of 2m and 10s was made for all muon tracks to candidate events and the fit was performed.

There are two key features to note in the signal distribution. First, the sharp peak centered around 0m. The width of the peak can be seen to have most of its events within $\sim$4m, showing the ability of using the longitudinal distance to shower peak as a means to tag spallation. However, the second feature to note is the long tails in both directions. In the background subtracted distributions, the ideal situation would have this tail fall off much faster in larger ln. This shows that in a non-negligible amount of cases, the shower peak is not near the spallation candidate. This can be explained by several reasons, such as the muon can shower multiple times, artifacts of the dE/dx method, poor track fit, EM showers obscuring the spallation causing hadronic shower, and more.

[1] The fit center and width of the third Gaussian are nearly the same as one of the first two for the background fit.

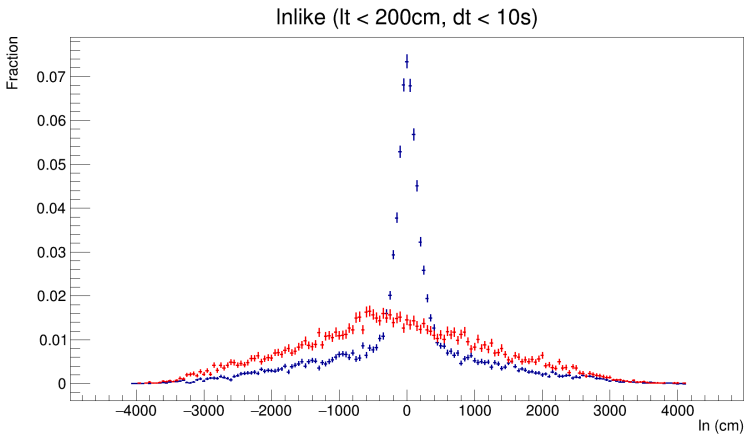


Figure 6.22: This is the signal (blue) and background (red) PDF distributions for the ln parameter to the dE/dx peak. The bias in the BG peak reveals a preference of the dE/dx method to find a "peak" near the track end, if the muon did not shower.

An alternative method to use the whole dE/dx and bin charge to make a two-dimensional likelihood to replace the former residual charge PDF and the ln PDF outlined above proved less effective than treating the two separate. A sample of likely non-showering muons was made, choosing from the peak of the p.e./cm distribution formed for all muons. A range was selected to contain the minimum ionizing peak for the muons, at 18 to 30 p.e./cm, chosen by looking at the energy loss for typical muons. For reference, for the normal SK Q_{res} calculation, normally around 26 p.e. is used as the p.e./cm for the calculation. 18-30 p.e./cm was chosen independent of the other value. This allowed for a library to be constructed of what a typical shower would look like for varying lengths of muon. Due to typical shower shape and other systematics, the library was non-uniform and typically had a rise in the tail end of the distribution motivating the grouping by track length. Using all the dE/dx for muons within 60s and 1m of a candidate event, a two-dimensional distribution was formed. The distribution was made by shifting the dE/dx plots from distance along track to ln, and filling the bins residual charge for the track compared to the MIP library, $Q_{bin} - Q_{lib}$.

Although individually this performed much better than the old residual charge calculation, once it was combined with the two other likelihoods, lt and Δt, the new fit

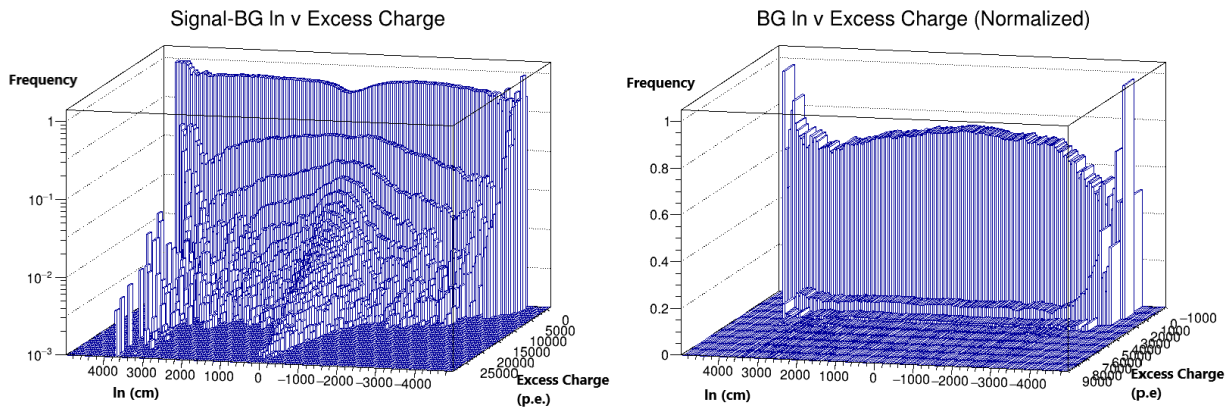


Figure 6.23: Signal (left) and background (right) distributions for ln vs excess charge for a dE/dx bin. The plots are normalized in each ln bin to account for differing track lengths. The background is relatively flat, with edge effects at very large ln, while the signal distribution is strongly peaked near 0 ln. Fits were made for excess charge in slices of ln, and then fits on the fit parameters to have a 2D PDF for the signal and background distributions.

under-performed compared to old. Using an updated Q_{res} and using the ln likelihood was the only way in which the dE/dx was able to successfully improve the performance.

When combining all of the new changes, the different potential likelihoods were compared using Receiver Operand Curves (ROC). All events passing the quality checks for for the solar analysis minus spallation cut discussed in the previous chapter had their spallation log likelihood calculated in 54 ($3 \times 3 \times 3 \times 3$) different ways. This corresponded to (none, new, old) PDF combinations for Q_{res}, lt, and Δt, and eventually (none, new with library, new without library) for ln. Comparing the fractions cut in the signal and background distributions as the axes for the ROC, it was visualized which method performed better when comparing to others.

The ROCs were also used to choose the cut value of the likelihood for the solar analysis. Depending on the analysis, there may be a preference of whether to increase livetime, decrease relative background, or somewhere in between. To decide, multiple samples were made with different conditions, relative to the solar sample. The first sample made used the ROC to match the former efficiency of the spallation cut, making improvements only in

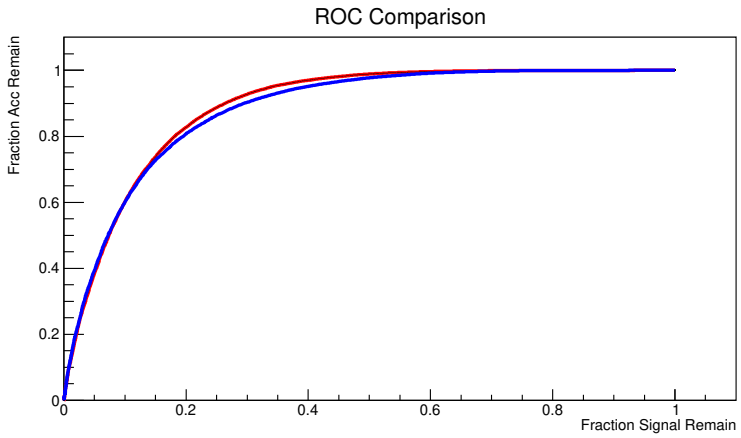


Figure 6.24: ROC curve comparing the former spallation log likelihood cut (blue) to the new spallation log likelihood (red) with updates to lt, Δt, Q_{res}, and ln. Choosing the desired efficiency (x-axis), the difference in livetime (y-axis) is seen. Although the curves are close, this results in a couple % difference in overall deadtime at the desired cut value. The desired value is near 20%, and depends on whether neutron cloud data is available.

livetime.

It was explored to use a Figure of Merit (FoM) to maximize in different energy regions. This took the form of:

$$FoM = \frac{\epsilon S}{\sqrt{\lambda B + \epsilon S}} \tag{6.20}$$

Where ϵ was the fraction of signal remaining, λ the fraction of background remaining, B is the fit to the background for negative $\cos\theta_{sun}$ values, and S is the Total events - Background for the very forward region. ϵ is taken as the compliment of the x values of the ROC, and λ is the compliment of the y-axis. The different FoM were defined using three criteria for the signal region:

i Varying $\cos\theta_{sun}$ to contain whole peak based on energy

ii $\cos\theta_{sun} > 0.8$

iii $\cos\theta_{sun} > 0.95$

The background fit (BG_{fit}) was again constant, and multiplied by the width of the signal

Energy	Cut Value
4-5	5.155
5-6	4.415
6-8	4.41
8-10	4.22
10-16	4.57
16-20	5.045

Table 6.3: Example energy dependent spallation log likelihood cut values for non-WIT period and choosing signal region to be $\cos\theta_{sun} > 0.8$. Looser cuts at very low and high energies result of spallation not being the main background (low) and less overall background (high). These do not include the ln likelihood.

region to obtain B, and signal was the total number of events in the region minus background:

$$B = (1 - \cos_{cutoff})BG_{fit}S = \sum_{sig} n_i - B \tag{6.21}$$

$\cos_{cutoff}$ is the range of $\cos(\theta_{sun})$ being checked for signal events from the list above and n_i was the total number of events in the bins within the $\cos(\theta_{sun})$ range. After performing these four fits, a full solar sample was made to test which of these performed better. For lower energies, the FoM desired a looser spallation cut value since spallation wasn't the dominant background, and for higher energies a stricter spallation cut value was desired (Table 6.3). However, when using energy dependent cut points there was only marginal improvement. Because of this, to be consistent with former data updates and keeping things simpler, the effectiveness was maintained at the level of the previous updates. For the energy regions with the most solar events, the change to FoM did not shift the cut value drastically. Since the value was close to the former one, it the overall result did not lead to great enough reduction in error to warrant change.

6.12 Deadtime Estimation Method and Dead Volume Function

Earlier, deadtime was taken as taking the reverse time cut for the same sample of events. For a more complete estimation all events between the energy of 3.49 MeV and 4.99 MeV kinetic are chosen from a mostly non-reduced sample. Their time is used, while taking three random variables and scrambling the vertex in SK:

$$x = 1690 \cdot \sqrt{R_1} \cdot \cos(2\pi R_2)\text{cm},\ y = 1690 \cdot \sqrt{R_1} \cdot \sin(2\pi R_2)\text{cm},\ z = 3610 \cdot R_3 - 1810\text{cm} \quad (6.22)$$

Where $R_i = Rand(0, 1)$. Since the spallation region extends into energies less than 4.99 MeV kinetic, backwards time correlation is still used to try and maintain event time independence. With a scrambled vertex, the time correlation between event and muon remains. Reversing the time direction removes this correlation to try and construct a proper deadtime calculation. All the muon information is unchanged, and normal likelihoods and cuts are made with original muon information and scrambled event vertex information. This allows for a function of deadtime to be made for throughout the detector.

A dead volume (deadvol) function is made to parameterize the likelihood that an uncorrelated event will be removed accidentally by the spallation cut by making a sample with the cuts outlined above. The deadtimes quoted are for the region of the fiducial volume (FV). The function is counted or integrated from the boundaries of the FV and are used for the deadtime estimates. Depending on the analyses though, these numbers may be different due to the non-uniformity of the sample. For example, in the lowest solar neutrino energy bins, the tightened fiducial volume count, making the volume compared to different and altering the relative deadtime. The external event cut also slightly alters the volume at the edges. Events are still allowed there, but some will be cut, altering the fraction of events from

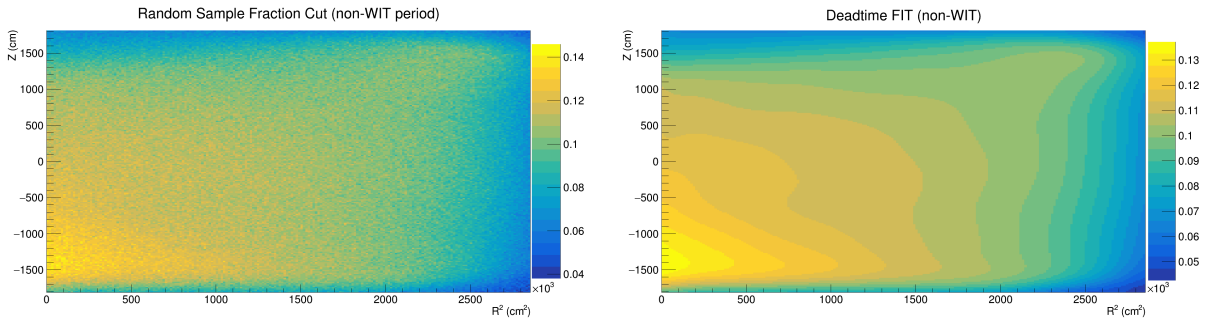


Figure 6.25: Left shows the fraction of removed events for the non-WIT period spallation cut based on r^2 and z. The right is a two-dimensional fit to the left distribution, and used to correct the counting of MC events.

a specific region. These effects are taken into consideration in the solar neutrino analysis.

The former deadvol function was two one-dimensional polynomials fit to the z deadtime (t^z_{dead}) and r^2 deadtime ($t^{r^2}_{dead}$)distributions for the cut events from the random sample, where $t^z_{dead}(z) = P^6(z)$ and $t^{r^2}_{dead}(r^2) = P^4(r^2)$, where P^i is an i^{th} order polynomial. This was updated to be a combined deadtime, where:

$$t_{dead}(z, r^2) = \sum_{i=0}^{n} \sum_{j=0}^{i} A_{ij} \times z^j \times (r^2)^{i-j} \tag{6.23}$$

where n is the order of the fit and A_{ij} are the fit parameters. Fits were made for increasing order of polynomials and the $\chi^2/d.o.f.$ were evaluated. A 16th order polynomial chosen as it was when the change of this test statistic tapered off. This allowed for a more precise representation of any features in the deadvol function. The non-WIT period random sample and the corresponding deadvol fit are shown in Fig. 6.25.

Chapter 7

Solar Neutrino Analysis in Super Kamiokande

7.1 Signal Extraction

Solar neutrinos are seen in SK through elastic scattering of neutrinos and electrons within the water. The neutrino and electron exchange a $W^{\pm}$ or Z boson and energy is transferred from the neutrino to the electron. For the ν_{μ} and ν_{τ} neutrinos, only the Z boson can be exchanged. The electron "recoils" and with enough energy transfer, the electron will give off Cherenkov light which is detected by the PMTs. The threshold for an electron to give off Cherenkov light is given in Table 3.1

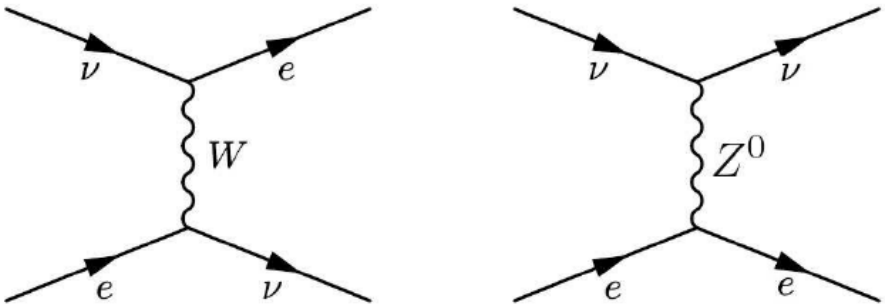


Figure 7.1: Feyman diagrams of W/Z boson exchange in ν-e elastic scattering

7.2 Signal Events

Based on theoretical expectation the elastic scattering is simulated using the differential cross section described by:

$$\frac{d\sigma}{dT_e} = \frac{G_F^2 m_e}{2\pi}\left[A_0 + B_0(1 - \frac{T_e}{E_\nu})^2 + C_0\frac{m_e T_e}{E_\nu^2}\right] \tag{7.1}$$

Where G_F is the Fermi coupling constant, m_e is the electron mass, E_v is the energy of the incident neutrino, T_e is the kinetic energy of the recoil electron, and $A_0/B_0/C_0$ are parameters dependent on the reaction type. The kinetic energy, T_e, is based on basic kinematics of the interaction with

$$T_e = \frac{2\cos^2\theta(E_\nu/m_e)}{(1-\cos^2\theta)(E_\nu/m_e)^2 + 2E_\nu/m_e + 1}E_\nu \tag{7.2}$$

θ is the angle of the direction of the recoil electron to the direction incident neutrino and the above equation is maximal when $\theta = 0$ (as $\theta = \pi$ isn't a true solution) and simplifies to

$$T_{e,max} = \frac{E_\nu}{1 + (m_e/2E_\nu)} \tag{7.3}$$

The $A_0/B_0/C_0$ parameters are the different representations of $g_V = I_3 - 2Q\sin^2\theta_W$ and $g_A = I_3$, where I_3 is the third component of the target particle's weak isospin, Q is the target particle's charge, and θ_W is the Weinberg angle. For the electron, $I_3 = -1/2$, $Q = -1$,

making $g_V \sim -0.04$ and $g_A = -1/2$. Table 7.1 summarizes the $A_0/B_0/C_0$ parameters for elastic scattering between $\nu_{e,\mu}$ and an electron. The ν_τ cross section is the same as the ν_μ.

Interaction	A_0	B_0	C_0
$\nu_e + e^- \to \nu_e + e^-$	$(g_V + g_A + 2)^2$	$(g_V - g_A)^2$	$-(g_V + 1)^2 + (g_A + 1)^2$
$\overline{\nu}_e + e^- \to \overline{\nu}_e + e^-$	$(g_V - g_A)^2$	$(g_V + g_A + 2)^2$	$-(g_V + 1)^2 + (g_A + 1)^2$
$\nu_\mu + e^- \to \nu_\mu + e^-$	$(g_V + g_A)^2$	$(g_V - g_A)^2$	$g_A^2 - g_V^2$
$\overline{\nu}_\mu + e^- \to \overline{\nu}_\mu + e^-$	$(g_V - g_A)^2$	$(g_V + g_A)^2$	$g_A^2 - g_V^2$

Table 7.1: Cross section coefficients for ν-e elastic scattering. Broken into the type of interaction occuring

The left of Fig 7.2 shows the differential cross section of the elastic scattering in the case of an 11 MeV neutrino incident on an electron. The total cross section is obtained from integrating Eq. 7.2. The recoil electron energy distribution is determined as follows:

$$F(T_e)dT_e = \left(\int_0^{E_{\nu,max}} \frac{d\sigma_{\nu-e}(E_\nu, T_e)}{dT_e} \Phi(E_\nu) dE_\nu \right) dT_e \tag{7.4}$$

where $\Phi(E_\nu)$ is the solar neutrino flux as a function of neutrino energy. $\Phi(E_\nu)$ depends on the neutrino energy spectrum shape of the ^{8}B decay, and the solar analysis of SK-IV uses the spectral shape of the ^{8}B neutrinos taken as a result of measuring the α spectrum:

$$^8\mathrm{B} \to \ ^8\mathrm{Be}^* + e^+ + \nu_e \ ^8\mathrm{Be}^* \to 2\alpha \tag{7.5}$$

performed by Winter *et al*[2], which is shown in Fig 7.3.

The angular distribution of the recoil electron relative to the neutrino direction is described by:

$$F(\theta)d\theta = \left(\int_0^{E_{\nu,max}} \Phi(E_\nu) \frac{d\sigma}{dT_e} \frac{dT_e}{d\theta} dE_\nu \right) d\theta \tag{7.6}$$

and the calculated angular distribution of solar neutrinos is shown in Fig 7.4. From the figure, if the recoil electron energy is above 4.0 MeV the scattered angle will be less than 25°. The low energy group of SK uses the cosine of the angle between the reconstructed vertex

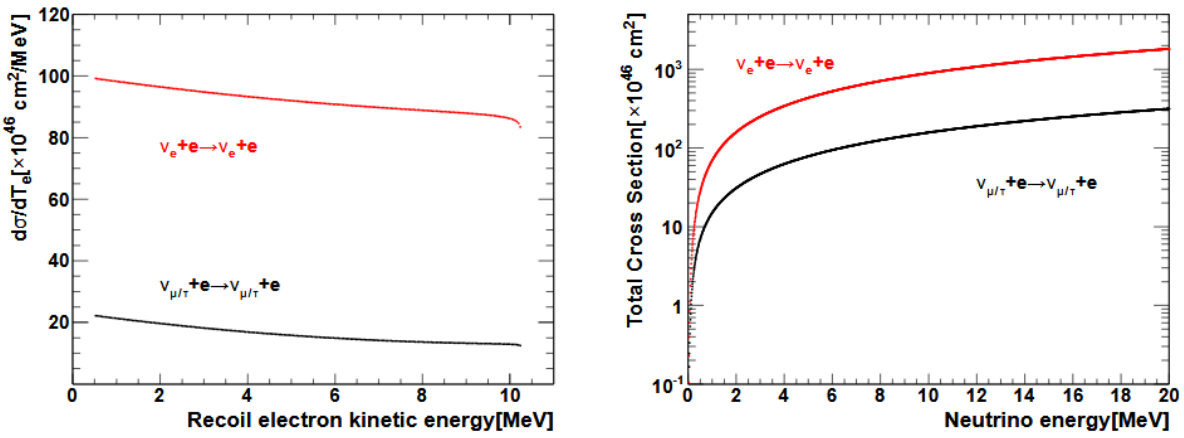


Figure 7.2: Differential cross section for an incident neutrino of 11 MeV shown left, and total cross section as a function of neutrino kinetic energy shown right.

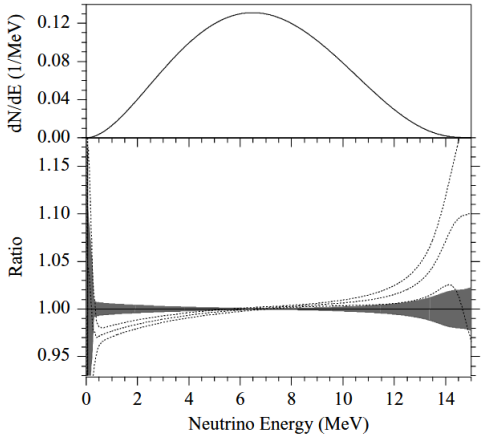


Figure 7.3: Normalized neutrino energy spectrum deduced from Winter (top). Dashed lines represent the ratio between the neutrino spectrum (bottom), and the $\pm 1\sigma$ experimental uncertainties, to the spectrum deduced by Winter[2]

and the solar neutrino direction to extract the solar signal. With this, $\cos(25) = 0.91$ would put the cutoff at the upper most bins. Due to directional reconstruction resolution not being so precise, this peak will smear into smaller $\cos(\theta_{sun})$ values, especially at low energies.

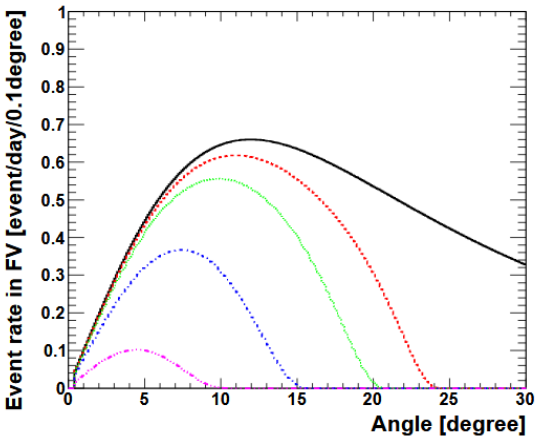


Figure 7.4: Angular distribution of recoil electrons from incident solar neutrinos. Horizontal axis shows teh angle of recoil electrons from incident solar neutrinos, and the vertical axis shows event rate in the fiducial volume. The red, green, blue, and purple lines correspond to neutrinos above 3.5, 4.5, 6.5, and 9.5 MeV respectively.

7.2.1 Extended Likelihood Method

In order to precisely measure the number of solar neutrino elastic scattering events, the $cos(\theta_{sun})$ distributions are fit with an extended maximum likelihood method. This method was first developed in SK-I and has been used in all phases of SK so far. The likelihood function is defined as:

$$\mathscr{L} = e^{-(\sum_i B_i + S)} \prod_{i=1}^{N_{bin}} \prod_{i=1}^{n_i} (B_i \cdot b_{ij} + S \cdot Y_i \cdot s_{ij}) \tag{7.7}$$

with N_{bin} being the number of energy bins, this is 23 bins for full SK-IV analysis corresponding to (20) 0.5 MeV bins for 4-14 MeV, (2) 1.0 MeV bins for 14-16 MeV, and (1) 4 MeV bin from 16-20 MeV. n_i is the number of observed of observed events in the i^{th} bin, Y_i is the fraction of signal events expected in teh i^{th} bin based on the solar neutrino MC simulation using the information outlined in the previous section, s_{ij} and b_{ij} are the signal and background probability density functions for the expected solar neutrino signal and expected background shape respectively, B_i is the background in the i^{th} energy bin, and finally S is the total number of solar neutrino events in the entire energy region. s_{ij} is

a two-dimensional function of recoil electron energyand $\cos\theta_{sun}$, $s_{ij} = s(E_{ij}, \cos\theta_{ij})$, while b_{ij} is function of $\cos\theta_{sun}$, $b_{ij} = b_i(\cos\theta_{ij})$, with both binned in energy and normalized with respect to $\cos\theta_{ij}$. The background shape is obtained from the detector zenith and azimuthal angular distributions of the real data after all solar neutrino reduction cuts.

7.2.2 Simulation of Events

One of the main goals of the solar neutrino MC simulation is to take into account the detector response. In the simulation, recoil electrons are uniformly generated in the entire ID. The energy and the scattering angle of the solar neutrino MC event are determined using Eq 7.2 and 7.2 outlined in Sec 7.2. Events are generated taking into account the direction of the Sun, in which the detector operation time is precisely reflected. The same reconstruction methods from Sec 4.2 are used to reconstruct the vertex, direction, and energy of the simulated events.

All of the data reduction steps are then applied to the simulated events and the final MC simulation events are obtained. With these simulation events, the expected fractions for Y_i are obtained. The expected energy distribution of the MC simulation recoil electron events is shown in Fig. 7.5.

7.3 Systematic Uncertainties

Most of the calculated uncertainties can be found in [64], and for the purpose of this thesis only the major systematic uncertainties will be explored at depth as well as the spallation cut's contribution to the systematic uncertainty.

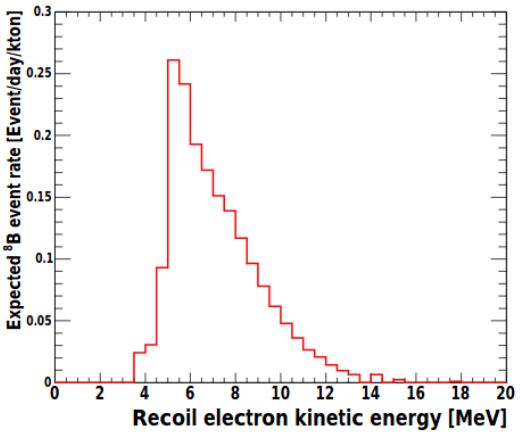


Figure 7.5: Expected ^{8}B energy spectrum using MC simulation after cuts are applied. The ^{8}B flux from the SNO result [3] and Winter spectrum [2] are used. Tight Fiducial Volume Cut lowers event rate for energies below 5.5 MeV (5.0 Mev Kinetic above)

7.3.1 Energy Scale

The systematic uncertainty accounted for to the total flux due to the energy scale is estimated using the systematic uncertainty of the absolute energy scale. Since the systematic uncertainty of the absolute energy scale is ±0.54%, this is used to shift the energy of the MC simulated events and evaluate the extracted total number of events. The new $Y_i^{\pm scale}$ values are obtained and the number of observed events changes by ±0.6% and taken as the energy scale contribution to the total uncertainty. This is the largest contribution to the total systematic uncertainty. The correlated shift going energy bin by energy bin is also determined and used throughout the solar analysis.

7.3.2 Signal Extraction

The systematic uncertainty in the signal extraction is estimated using dummy data whose signal and background events are known. After applying the extraction method to the dummy data, the extracted number of events and the true number of events are compared. The difference is taken as the systematic uncertainty and is found to be ±0.7%

7.3.3 Spallation Systematic Uncertainty

Prior spallation systematic uncertainty was performed by looking at the drift of the 45 day average for spallation deadtime. The spallation deadtime in these 45 day average bins was found to be stable within about $\pm 0.3\%$ for the entire SK-IV period, and therefore the systematic uncertainty was estimated to $\pm 0.2\%$.

A different method was employed to try and estimate the systematic uncertainty due to the updated spallation cut. The fits from Muboy were varied to try and track for uncertainty in the track fit. Muon fits with Muboy have their entrance fixed to a PMT location, and their exit is fixed to a PMT or a Q-weighted average is performed (Sec 4.1). If the exit was not a PMT, the track exit was moved to the closest PMT. From here the track was allowed to vary, within a 3x3 grid centered on the tracks entrance and exit PMTs (81 possible outcomes). The wiggling of the track was weighted with a 2D Gaussian with the width for both being the PMT spacing.

25 random outcomes were chosen, with recalculated dE/dx, and the deadvol functions were recalculated 25 times. The deadvol calculations were performed using 1 of the 25 new muon tracks each time, an additional 1 time for the former fit for 26 in total. For this calculation, the effect of the wiggling of tracks by up to one made minimal difference, but an additional source of systematic error was discovered.

Due to the fitting of all muons being computationally intensive, rerunning of fits is avoided when possible. To perform this study and apply new dE/dx fits, Muboy needed to be rerun to recover the TQ information which is not stored long term outside of SK reformatted files due to immense disk space required. Since the previous running of Muboy updates, fixes, and/or changes have happened to the tqmap and information, making for minor changes in the resulting fits for Muboy. Fig. 7.6 shows the dot product between the former muon track to the wiggled muon tracks on both linear and logarithmic scales, with

Category	Valid Old Fit	Valid New Fit	% fraction
1	Yes	Yes	99.76
2	No	Yes	0.0005
3	Yes	No	0.022
4	No	No	0.21

Table 7.2: Percentage fraction of cases for comparison of muon fits for old fit vs new

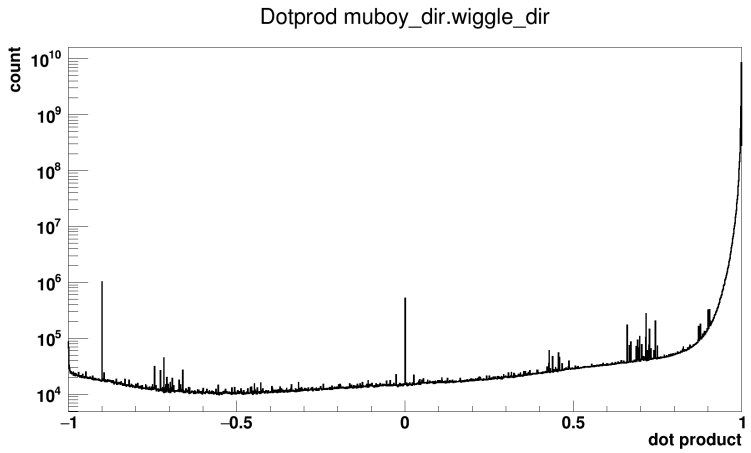


Figure 7.6: Dot product of all of the muon track wiggles to the original muon fit. Y-axis is logarithm of the counts/bin. Most muons have relatively the same fit, with some differences depending on the case. The peak at 0 corresponds to a successful wiggle dotted with a failed old fit. The peak is amplified by a factor of 25, since it is filled for each wiggle. There is an excess peak at -1.0 and -0.9 corresponding to failed new fit and without (with) a successful old fit respectively. These are filled only once. Spikes located at approximately -0.7, 0.45, and 0.7 are from a "preferred" set of fits Muboy has a tendency to fall to when fitting for both the wiggled and former track fits. These spikes are a result of dotting these fits and corresponding wiggles together.

the fraction of track categories listed in Table 7.2. The single peaks at 0, -0.9, and -1.0 correspond to no old fit and valid new fit (category 2), no fit for either (3), and valid old fit with no new fit (4) respectively in the logarithmic plot.

Running the fits through the random sample with this information, has negligible spread amongst the 25 iterations of shifting the muon fit, and an approximate 0.2% shift in the total livetime for both periods. This shift is attributed to the difference in muon fits, and the discrepancy of the results was decided to be added to the systematic error.

7.4 ^{8}B Flux Measurement

Using the full 2970.1-day data for SK-IV, the ^{8}B solar neutrino flux measurement is performed with the updated spallation cut. The energy ranges from 4.0 to 20.0 MeV in total electron energy and the extracted number of solar neutrino events is 65417$\pm$390(stat.)$\pm$916(sys).

In order to determine the ^{8}B solar neutrino flux, the number of the expected signal events is calculated as:

$$N_{exp} = \frac{N_{MC,survival}}{N_{MC,generated}} \times T_{Live} \times R_{exp} \tag{7.8}$$

where the N_{MC} terms are the generated events and surviving events after all of the reduction cut in the solar neutrino MC simulation, T_{Live} is the SK-IV livetime, and R_{exp} is calculated assuming the number of electrons in the ID is 1.086×10^{34} and using the spectrum shape measured by Winter *etal*. Using the SNO ^{8}B flux measurement using neutral current (NC) interactions, $\Phi_{^8\mathrm{B,SNO,NC}} = 5.25 \times 10^6/\mathrm{cm}^2/\mathrm{sec}$, the expected rate within the entire ID becomes 294.75 event/day/32.5 kton [3]. Finally, the expected signal events for the entire time period in the analyzed region of 4-20 MeV and after all cuts are applied is N_{exp} = 148,243 events.

Taking the observed signal and the expected signal in tandem, and correcting for the effect of the eccentricity of Earth's orbit on the neutrino flux, the corresponding ^{8}B solar neutrino flux at one AU is measured to be:

$$\Phi_{^8\mathrm{B,SKIV}} = 2.313 \pm 0.014(\mathrm{stat}) \pm 0.032(\mathrm{sys.}) \times 10^6 \mathrm{cm}^2/\mathrm{s} \tag{7.9}$$

7.5 Day/Night Asymmetry

Using the location of the sun throughout the day, the flux of solar neutrinos is measured in SK to study the effects of the Earth's matter on neutrino oscillations, and was earlier introduced in Sec 1.3. The variable used to measure the suns location with respect to the detector is $\cos\theta_z$ and is defined as the cosine of the angle between the detector's z-axis and the vector of the solar position to the reconstructed vertex's position.

The day/night asymmetry is extracted from two energy regions, 5.0 to 20 MeV and 4.0 to 20 MeV for the reconstructed energy. Using this definition for $\cos\theta_z$, daytime data corresponds to $\cos\theta_Z > 0$ and nighttime data corresponds to $\cos\theta_z \leq 0$. For each subrun, the direction of the Sun is calculated, and this is used to separate them into day and night bins. The calculated livetime for all of SK-IV is:

$$\cos(\theta_z) > 0, T_{Day} = 1433.9\text{days}\,\cos(\theta_z) \leq 0, \quad T_{Night} = 1536.2\text{days} \tag{7.10}$$

Using this and performing the same calculation as above the ^{8}B flux for both day and night is found to be:

$$\begin{aligned}
\Phi_{^8\text{B,SKIV,Day}} &= 2.284 \pm 0.020(\text{stat}) \pm 0.032(\text{sys.}) \quad 4.0 \leq E < 20.0 \text{ MeV} \\
\Phi_{^8\text{B,SKIV,Night}} &= 2.340 \pm 0.019(\text{stat}) \pm 0.033(\text{sys.}) \quad 4.0 \leq E < 20.0 \text{ MeV} \\
\Phi_{^8\text{B,SKIV,Day}} &= 2.284 \pm 0.020(\text{stat}) \pm 0.032(\text{sys.}) \quad 5.0 \leq E < 20.0 \text{ MeV} \\
\Phi_{^8\text{B,SKIV,Night}} &= 2.337 \pm 0.019(\text{stat}) \pm 0.033(\text{sys.}) \quad 5.0 \leq E < 20.0 \text{ MeV}
\end{aligned} \tag{7.11}$$

With the observed Day/Night Asymmetry calculated as:

$$\begin{aligned}
A_{DN,SKIV} &= \frac{\Phi_{\text{SKIV,Day}} - \Phi_{\text{SKIV,Night}}}{0.5 \times \left(\Phi_{\text{SKIV,Day}} + \Phi_{\text{SKIV,Night}}\right)} \\
&= -2.4 \pm 1.2(\text{stat})\%
\end{aligned} \tag{7.12}$$

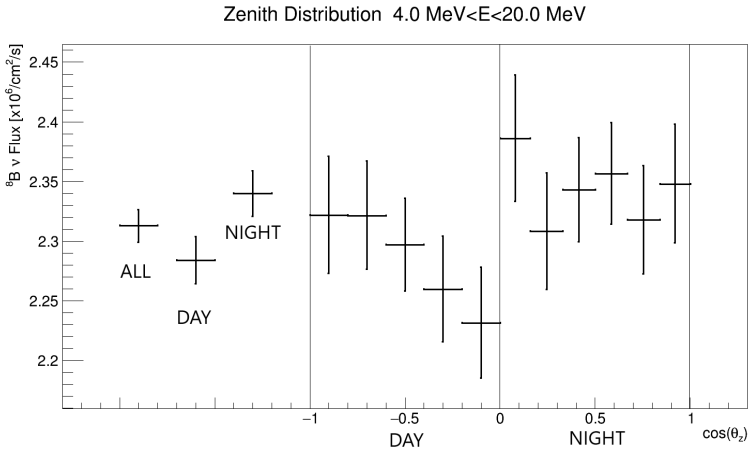


Figure 7.7: Zenith angle distribution for the ^{8}B solar neutrino flux measurement for the 4.0 to 20.0 MeV energy range. Horizontal axis has ALL, Day, and Night, then the five zenith day bins and six zenith night bins. solar neutrinos passing through the Earths core is in the rightmost bin. Error shown is statistical.

The day/night flux uncertainty for all of SK-IV is 2σ from zero.

7.6 Energy Spectrum

The method of the flux measurement is applied to events in all 23 energy bins, and the obtained flux is plotted on the spectrum. The energy bins below 8 MeV are also subdivided based off of their Multiple Scattering Goodness (MSG). MSG uses the PMT hit pattern and the resulting Cherenkov light cones to reconstruct multiple scatterings of electrons, and is another form of direction fit goodness. Poor direction reconstruction goodness can come from multiple Couloumb scattering of electrons while it emits Cherenkov light or multiple electrons in $\beta-\gamma$decays. Smaller MSG values have more background contamination of ^{208}Th decay with $\beta-\gamma$. Solar neutrino recoil electrons are expected to scatter less, and thus have a better goodness.

For each MSG bin (0 to 0.35, 0.35 to 0.45, and 0.45 to 1) below 8 MeV, $s_{i,j}$ and $b_{i,j}$ are obtained from MC and data respectively. The signal extraction process is repeated

for each scenario and for the new $s_{i,j}$ and $b_{i,j}$ values, and a simultaneous fit across all bins for MSG and energy is performed. This method helps to reduce the statistical error of the measurement by around 10%. Fig. 7.8 shows the solar peak for all energy bins.

7.7 Solar Upturn

The SNO experiment parameterized the energy dependence of the solar neutrino flavor composition using a quadratic fit on the survival probability centered at 10 MeV:

$$P_{ee}(E_\nu) = c_0 + c_1\left(\frac{E_\nu}{MeV} - 10\right) + c_2\left(\frac{E_\nu}{MeV} - 10\right)^2 \tag{7.13}$$

where c_0, c_1, and c_2 are polynomial parameters. This was done in order to discuss the energy dependence in a general way, and the SNO collaboration had found that this simple method also introduces no bias when determining the oscillation parameters. SK fits its energy dependence with the same formula, as well as fitting it with an exponential. The functional form of the exponential fit is as follows:

$$P_{ee}(E_\nu) = e_0 + \frac{e_1}{e_2}\left(e^{(\frac{E_\nu}{MeV}-10)} - 1\right) \tag{7.14}$$

Writing it in this way allows for a direct comparison of e_0 and e_1 to c_0 and c_1 from the quadratic fit while e_2 controls the steepness of the exponential rise. Using only the full SK-IV data set, both fits favor an upturn by more than 1σ. The solar oscillation parameter fits for only SK-IV, both with and without taking Day/Night Asymmetry into consideration, are shown in Fig 7.11. These are overlaid with the KAMLAND results for the same parameters. The method for calculating the χ^2 is outlined in the SK-IV solar neutrino paper on 1664 days of data [64].

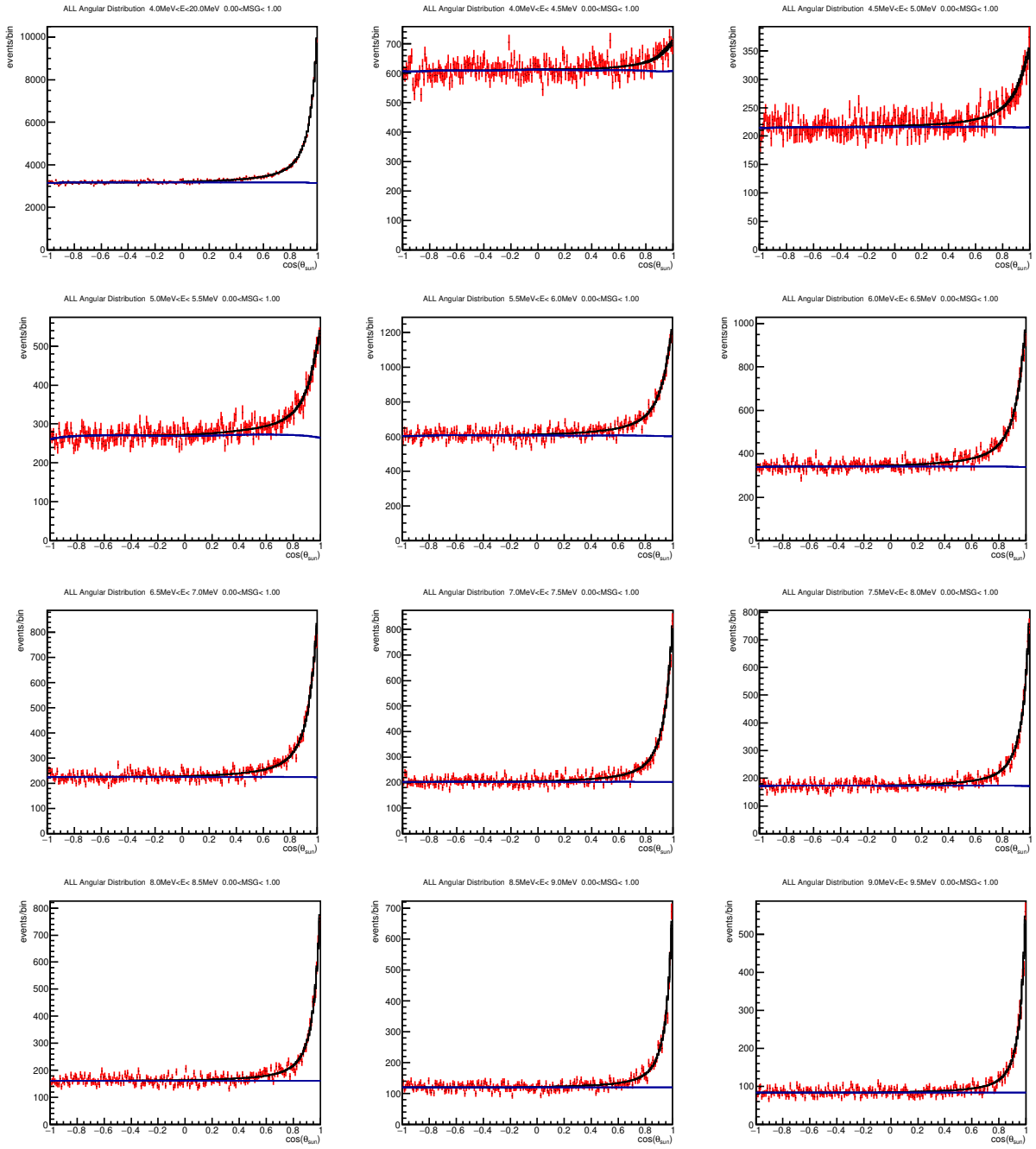


Figure 7.8: $\cos\theta_{sun}$ distributions for each energy bin, with the total energy range (4-20) shown in the uppermost left plot, with the rest showing 0.5 MeV bins from 4 MeV to 9.5 MeV (left to right, top to bottom). Flat blue lines show background contribution and black lines are signal and background.

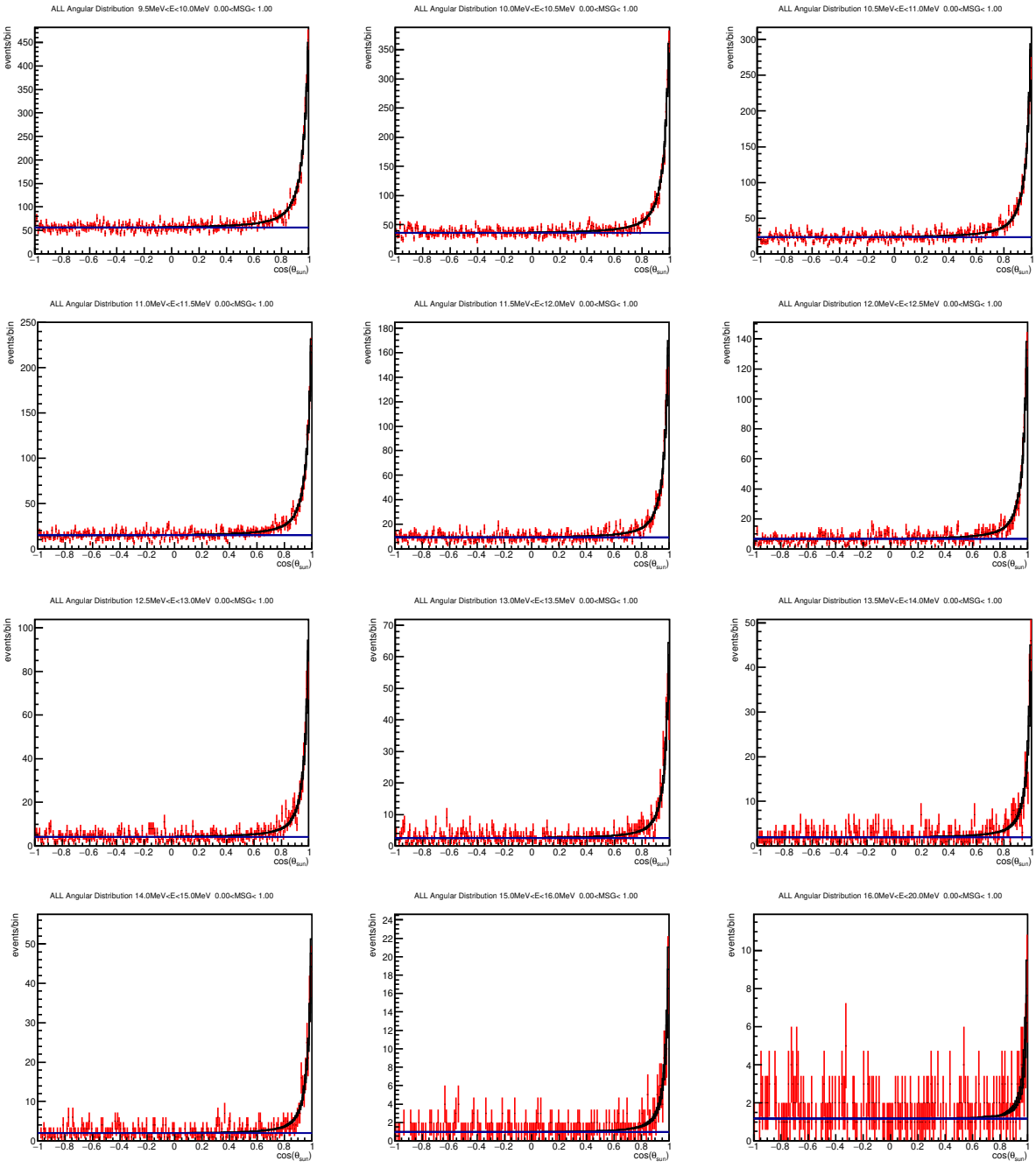


Figure 7.9: Highest energy $\cos\theta_{sun}$ distributions (>9.5 MeV). Flat lines show background shape (flatter) and signal (following data) with $\pm 1\sigma$ fits.

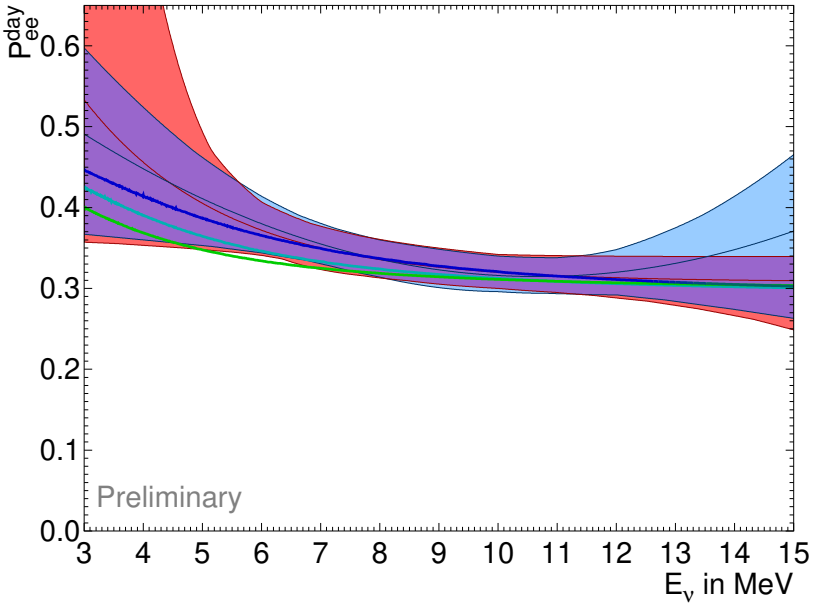


Figure 7.10: Solar upturn using only SK-IV data, with quadratic (light blue) and exponential (red) fits with 1σ band. Purple denotes the overlap of the two regions. Using the KAMLAND (dark blue), Solar best fit with 1670 days of SK-IV Day/Night data (green), and SK+SNO (robin egg blue).

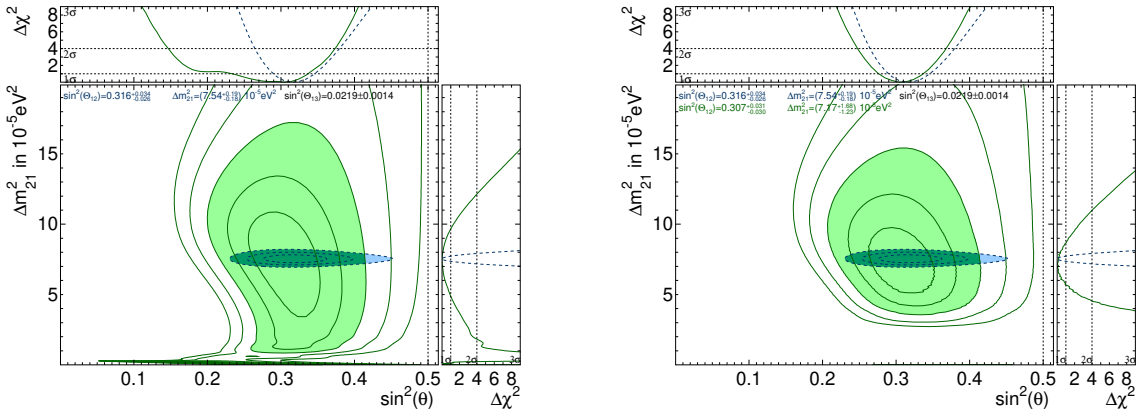


Figure 7.11: Contour plots showing 1-5σ regions for the solar oscillation parameter fits for SK-IV only data (Green) compared to the 1-3σ KamLAND best fit regions (blue). The fits on the left (right) are performed without (with) taking into consideration the effect of the Day/Night Asymmetry. The shaded regions for both are the 3σ allowed region.

Chapter 8

Comparison of Results

Since SK is such a long running experiment, reducing deadtime can add months or even years worth of data. With the improvements to the spallation cut, the amount of recovered solar neutrino signal events is 12% (7027 events) and a 6.6% reduction in relative error. The change in signal and relative error for each energy bin is shown in Table 8.1. With the 2970.1-day data, this is equivalent to adding 330 days of livetime, or roughly a year of data if downtime is included. Fig 8.2 shows the $\cos(\theta_{sun})$ distributions for the old and new spallation cuts overlaid on each other (left), with the difference between the two (bottom). As discussed in Sec 5.1.1 and 5.2, spallation is the dominant background in the energy ranges above 6 MeV, but when the energy is lowered backgrounds from the detector materials begin to take over. However, the spallation cut is applied regardless of the energy of the event seen for the solar sample. The spallation cut tuning was performed to have equal performance between 6.5 and 20.0 MeV while minimizing the deadtime.

This deadtime reduction is seen in the higher overall background in the new data set, caused by non-cosmogenic sources in the lowest energy bins. Fig. 8.1 shows the lowest energy bin compared to an energy bin in the spallation dominant region.

Energy Bin (MeV)	Event Increase	% Event Increase	% Reduction in Rel. Err.
4.0 - 4.5	211.6	12.8	5.7
4.5 - 5.0	306.1	15.3	7.9
5.0 - 5.5	447.4	12.1	6.6
5.5 - 6.0	1011.8	13.7	7.5
6.0 - 6.5	805.2	11.5	6.7
6.5 - 7.0	676.1	11.3	6.6
7.0 - 7.5	649.8	12.0	6.8
7.5 - 8.0	561.2	11.8	6.7
8.0 - 8.5	556.0	12.2	7.2
8.5 - 9.0	313.8	8.1	4.6
9.0 - 9.5	413.1	13.9	7.8
9.5 - 10.0	267.7	11.0	6.3
10.0 - 10.5	257.6	14.0	7.3
10.5 - 11.0	229.0	16.0	7.8
11.0 - 11.5	130.1	11.5	5.7
11.5 - 12.0	73.0	8.7	3.9
12.0 - 12.5	68.0	10.4	4.9
12.5 - 13.0	57.6	13.8	6.8
13.0 - 13.5	21.0	7.1	4.1
13.5 - 14.0	27.3	14.9	6.4
14.0 - 15.0	10.4	4.7	0.9
15.0 - 16.0	6.3	7.8	4.1
16.0 - 20.0	4.0	16.6	6.8

Table 8.1: Table showing increase in signal events and reduction in relative error for each energy bin of the solar neutrino analysis.

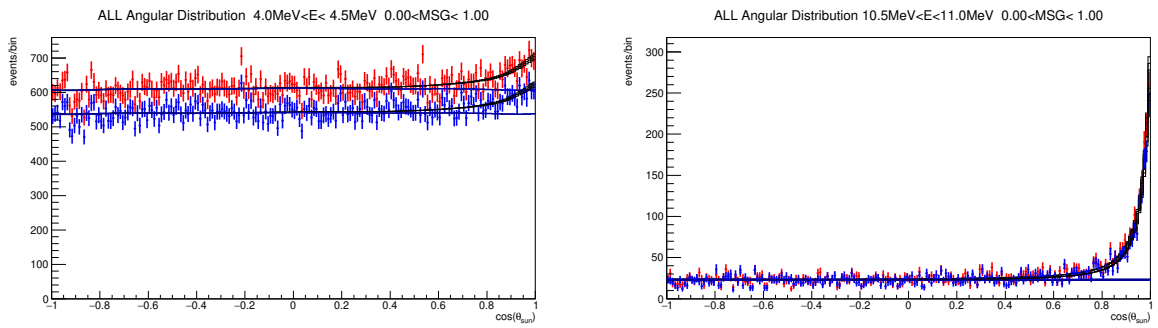


Figure 8.1: Lowest energy bin (left) compared to spallation region energy bin (right, 10.5-11.0 MeV). Near match in background is seen for the higher energy bin, but large separation in the lowest energy bin from increasing background from non-spallation sources as a result of deadtime reduction

The improvement in signal and relative error impacts the results for the Day/Night Asymmetry flux measurement and the overall spectrum analysis. For the Day/Night asymmetry, the recovery of events and reduced error has resulted in a change from 1.1σ favoring of the existence of a Day/Night Asymmetry with the old spallation cut, to a 2.0σ favoring of a Day/Night Asymmetry.

There is a relatively small change to the general shape of the energy spectrum, but the overall reduction in error across all energies can be seen in the shrinking of the error bars. The newest results are enough to have both the exponential fit and quadratic fit favor an upturn by more than 1σ, but the fit has not been calculated for the final sample using the old spallation cut. This was not the case for the results presented in [64].

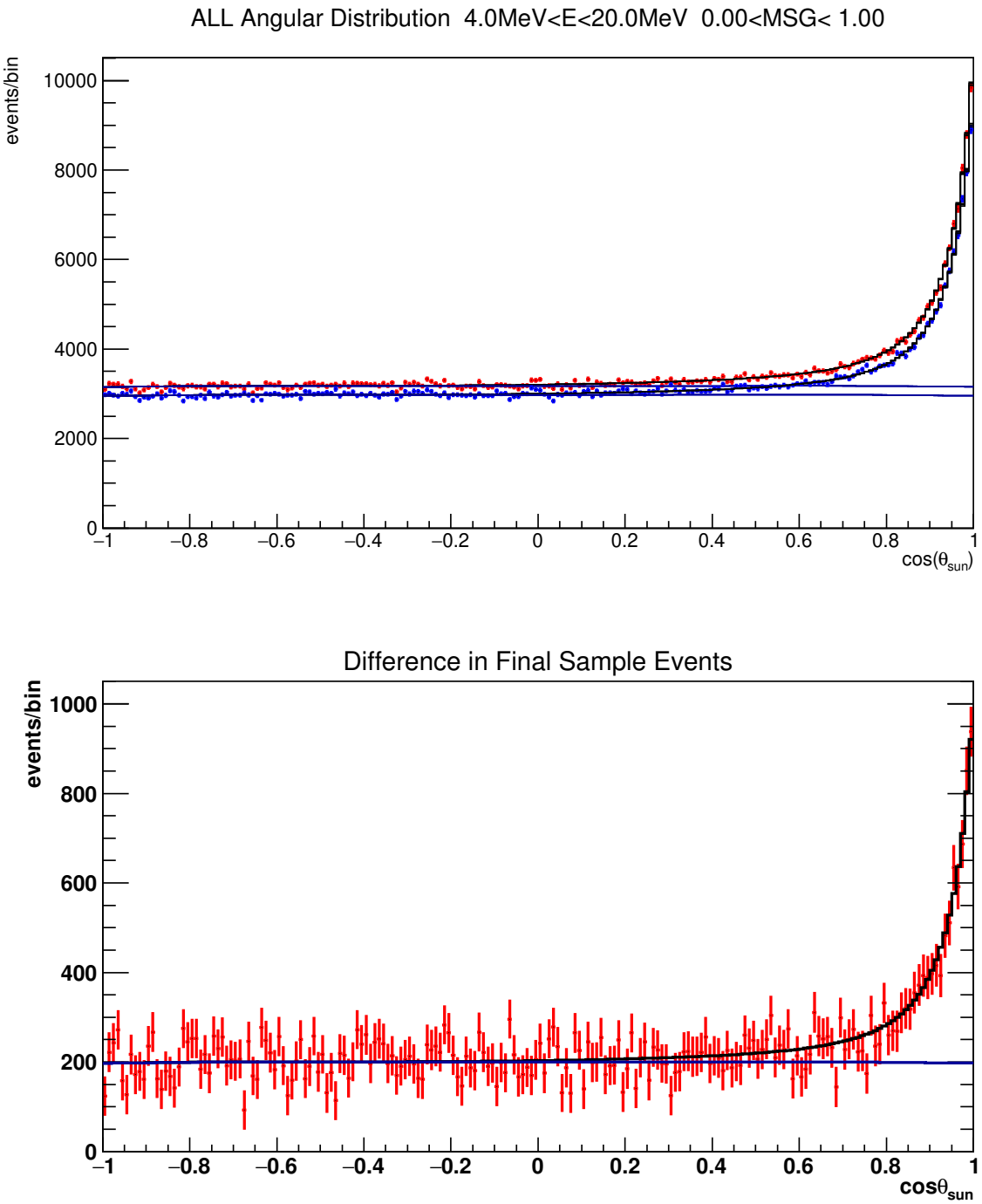


Figure 8.2: Top: $\cos(\theta_{sun})$ final solar sample the new spallation cuts (red) overlaid with old (blue). Bottom: Difference between the old and new spallation cuts, taken as the sample with the new spallation cuts minus the sample with the old spallation cut. Lower energies have non-spallation dominant backgrounds and reducing the deadtime increases the background in these energies. Difference in background shape fits are shown as the blue line and difference in signal fits shown in black

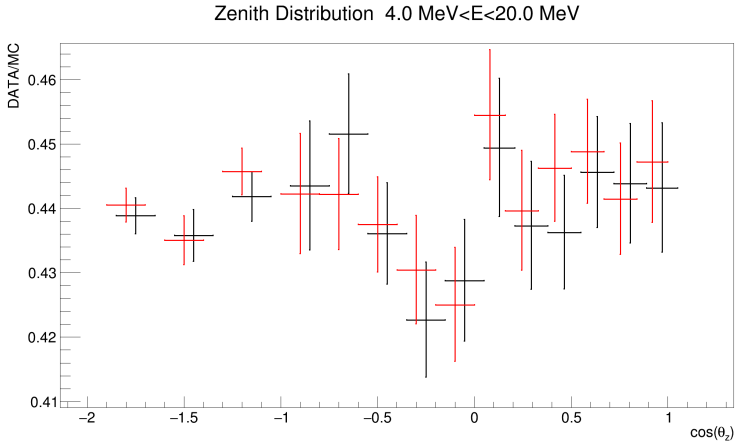


Figure 8.3: Comparison of the $\cos(\theta_z)$ distributions for the new (red) and old spallation cut (black). The change in number of events increase the significance of the asymmetry.

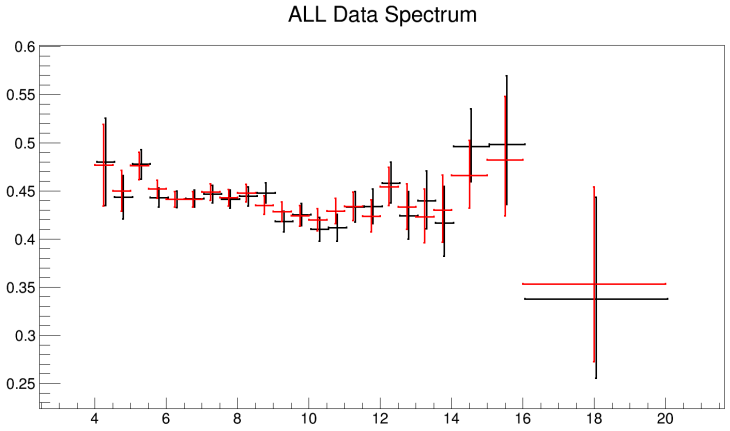


Figure 8.4: Comparison of the energy spectrum for the final solar sample using the new (red) spallation cut against the old spallation cut (black).

Chapter 9

Conclusion and Outlook

Updates were made to the solar neutrino analysis methods for background removal in the fourth phase of Super-Kamiokande, SK-IV. Spallation products, whose origin is from cosmic ray muons interacting and showering within the detector, is the largest background in the 6-20 MeV region of the solar analysis, and cannot be reduced with physical modifications to the detector short of physically moving the experiment deeper underground.

To reduce the cosmogenic induced backgrounds of the solar analysis, three avenues were taken:

I Direct identification of hadronic showers

II Tagging multiple spallation

III Using muon parameters to make a likelihood based cut

The first avenue was to directly tag hadronic showers after muons to isolate spallation. This was the first time in which the hadronic shower after a muon had successfully been tagged in a water Cherenkov detector. After tuning, this method alone tags 55% of spallation while

only accruing 1.3% deadtime. The low efficiency to see neutrons with just capture on H limits the spallation tagging efficiency, and requires additional methods to compensate.

The second new method of using spallation to tag other spallation is a novel approach that departs from using any muon information to make a cut on spallation, helping to compensate for possible issues with muon fits, especially for higher energy muons. However, not all showers causing spallation produce multiple radioactive isotopes. Still, 47% of spallation products in the solar sample with the final cut tuning are tagged this way while introducing only 1.3% deadtime.

Finally, re-tuning of the spallation log likelihood with the introduction of a new parameter, ln, further reduces the deadtime. Since the cut point for the log likelihood was tuned specifically for the final cut to have the same spallation tagging effectiveness as the former method, comparison of the total deadtime is used to evaluate improvements. Where all three methods are available (388 days), there is a 55% reduction in deadtime from the spallation cut compared to the old method with maintained spallation tagging effectiveness, and where neutron cloud data is not available (2582 days) deadtime is reduced by 45%. .

With the running length of SK-IV (2970 days of data), these improvements add close to a year of data, which results in an increase of 7000 solar neutrino signal events, with a reduction in the relative error of the analysis by 6.6%. These changes will be used in the final solar analysis results for SK-IV, to be published in the coming months.

The spallation tagging methods will also be useful for efforts in Hyper Kamiokande (HK) to perform any solar neutrino analysis. HK is the next generation water Cherenkov detector after SK to be built nearby in Japan [65]. However, the planned location of HK is expected to have less overburden, making cosmogenic induced backgrounds to be an even larger problem with an expected spallation rate to be a little less than an order of magnitude more. Effective methods to remove spallation will play a major role in the ability of HK to

perform solar neutrino analysis.

With the recent addition of Gd into the SK detector, the use of neutrons to tag spallation should be enhanced. More neutron captures should be seen as the efficiency to see the 8 MeV γ cascade from captures on Gd is much higher than the efficiency to see the 2.2 MeV γ from captures on H. The vertex resolution for captures on Gd is also better when compared to captures on H, allowing for better shower tracking. Initial studies looking for muons after a shower in the Gd era have shown an increase in the number or showers, the size of showers, and the overall distribution for showers to be more closely reconstructed to each other.

References

[1] Kirk Ryan Bays. *Search for the diffuse supernova neutrino background at Super-kamiokande*. PhD thesis, UC, Irvine, 2012.

[2] W.T. Winter, S.J. Freedman, K.E. Rehm, and J.P. Schiffer. **The B-8 neutrino spectrum**. *Phys. Rev. C*, **73**:025503, 2006.

[3] B. Aharmim et al. **Measurement of the ν_e and total ^{8}B solar neutrino fluxes with the Sudbury Neutrino Observatory phase-III data set**. *Phys. Rev. C*, **87**:015502, Jan 2013.

[4] Y. Zhang et al. **First measurement of radioactive isotope production through cosmic-ray muon spallation in Super-Kamiokande IV**. *Phys. Rev. D*, **93**(1):012004, 2016.

[5] J. Chadwick. **Intensitätsverteilung im magnetischen Spektrum von β-Strahlen von Radium B+C**. *Verhandlungen der Deutschen Phys. Gesellschaft*, **16**:383–391, 1914.

[6] Wolfgang Pauli. **Open letter to the group of radioactive people at the Gauverein meeting in Tübingen.** December 1930.

[7] E. Z. Fermi. **Versuch einer Theorie der β-Strahlen. I**. *Physik*, **88**, 1934.

[8] Christine Sutton. *Spaceship Neutrino*. Cambridge University Press, 1992.

[9] H. Bethe and R. Peierls. **The 'neutrino'**. *Nature*, **133**:532, 1934.

[10] B. Pontecorvo. **Inverse beta process**. *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, **1**:25–31, 1991.

[11] Raymond Davis. **A Half-Century with Solar Neutrinos**. *International Journal of Modern Physics A - IJMPA*, **18**:3089–3108, 07 2003.

[12] F. Reines. **The neutrino: From poltergeist to particle**. *Rev. Mod. Phys.*, **68**:317–327, 1996.

[13] C.L. Cowan, F. Reines, F.B. Harrison, H.W. Kruse, and A.D. McGuire. **Detection of the free neutrino: A Confirmation**. *Science*, **124**:103–104, 1956.

[14] E.J. Konopinski and H.M. Mahmoud. **The Universal Fermi interaction**. *Phys. Rev.*, **92**:1045–1049, 1953.

[15] G. Danby, J.M. Gaillard, Konstantin A. Goulianos, L.M. Lederman, Nari B. Mistry, M. Schwartz, and J. Steinberger. **Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos**. *Phys. Rev. Lett.*, **9**:36–44, 1962.

[16] Edoardo Vitagliano, Irene Tamborra, and Georg Raffelt. **Grand Unified Neutrino Spectrum at Earth: Sources and Spectral Components**. 10 2019.

[17] M. Agostini et al. **Experimental evidence of neutrinos produced in the CNO fusion cycle in the Sun**. *Nature*, **587**:577–582, 2020.

[18] H. A. Bethe. **Energy Production in Stars**. *Phys. Rev.*, **55**:103–103, Jan 1939.

[19] John N. Bahcall. **Solar neutrinos. I: Theoretical**. *Phys. Rev. Lett.*, **12**:300–302, 1964.

[20] Takaaki Kajita. **Nobel Lecture: Discovery of atmospheric neutrino oscillations**. *Rev. Mod. Phys.*, **88**(3):030501, 2016.

[21] Arthur B. McDonald. **Nobel Lecture: The Sudbury Neutrino Observatory: Observation of flavor change for solar neutrinos**. *Rev. Mod. Phys.*, **88**(3):030502, 2016.

[22] B. Pontecorvo. **Mesonium and anti-mesonium**. *Sov. Phys. JETP*, **6**:429, 1957.

[23] B. Pontecorvo. **Inverse beta processes and nonconservation of lepton charge**. *Sov. Phys. JETP*, **7**:172–173, 1958.

[24] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. **Remarks on the Unified Model of Elementary Particles**. *Progress of Theoretical Physics*, **28**(5):870–880, 1962.

[25] S. P. Mikheyev and A. Y. Smirnov. **Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos**. *Sov. J. Nucl. Phys.*, **42**:913, 1985.

[26] J. Chadwick. **The Existence of a Neutron**. *Proc. Roy. Soc. Lond. A*, **136**(830):692–708, 1932.

[27] E. Rutherford. **The scattering of alpha and beta particles by matter and the structure of the atom**. *Phil. Mag. Ser. 6*, **21**:669–688, 1911.

[28] Sir Ernest Rutherford, F.R.S. **Bakerian Lecture: Nuclear Constitution of Atoms**. **97**(686):374–400, July 1920.

[29] Valery Nesvizhevsky and Jacques Villain. **The discovery of the neutron and its consequences (1930–1940)**. *Comptes Rendus Physique*, **18**(9):592 – 600, 2017. Science in the making: The Comptes rendus de l'Académie des sciences throughout history.

[30] Todd Haines et al. **Neutrinos From {SN1987A} in the Imb Detector**. *Nucl. Instrum. Meth. A*, **264**:28–31, 1988.

[31] John M. LoSecco. **The History of "Anomalous" Atmospheric Neutrino Events: A First Person Account**. *Phys. Perspect.*, **18**(2):209–241, 2016.

[32] K.S. Hirata et al. **Observation in the Kamiokande-II Detector of the Neutrino Burst from Supernova SN 1987a**. *Phys. Rev. D*, **38**:448–458, 1988.

[33] K. Inoue. **Results from Kamiokande and status of super-Kamiokande**. *Nuclear Physics B - Proceedings Supplements*, **59**(1):267 – 276, 1997. Proceedings of the Fourth KEK Topical Confernce on Flavor Physics.

[34] S. Fukuda et al. **The Super-Kamiokande detector**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **501**(2):418 – 462, 2003.

[35] T. Iida et al. **The CANDLES experiment for the study of Ca-48 double beta decay**. *Nuclear and Particle Physics Proceedings*, **273-275**:2633 – 2635, 2016. 37th International Conference on High Energy Physics (ICHEP).

[36] K. Abe et al. **XMASS detector**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **716**:78 – 85, 2013.

[37] S. Abe et al. **Production of Radioactive Isotopes through Cosmic Muon Spallation in KamLAND**. *Phys. Rev. C*, **81**:025807, 2010.

[38] Magro, Lluís Martí, The EGADS Collaboration. **EGADS approaching GADZOOKS!** *EPJ Web of Conferences*, **95**:04041, 2015.

[39] Yoichi Aso, Yuta Michimura, Kentaro Somiya, Masaki Ando, Osamu Miyakawa, Takanori Sekiguchi, Daisuke Tatsumi, and Hiroaki Yamamoto. **Interferometer design of the KAGRA gravitational wave detector**. *Phys. Rev. D*, **88**(4):043007, 2013.

[40] H. Nishino, K. Awai, Y. Hayato, S. Nakayama, K. Okumura, M. Shiozawa, A. Takeda, K. Ishikawa, A. Minegishi, and Y. Arai. **High-speed charge-to-time converter ASIC for the Super-Kamiokande detector**. *Nucl. Instrum. Meth. A*, **610**:710–717, 2009.

[41] S. Yamada et al. *Commissioning of the New Electronics and Online System for the Super-Kamiokande Experiment*, **57**. April 2010.

[42] John F. Beacom and Mark R. Vagins. **GADZOOKS! Anti-neutrino spectroscopy with large water Cherenkov detectors**. *Phys. Rev. Lett.*, **93**:171101, 2004.

[43] A. Suzuki, M. Mori, K. Kaneyuki, T. Tanimori, J. Takeuchi, H. Kyushima, and Y. Ohashi. **Improvement of 20 in. diameter photomultiplier tubes**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **329**(1):299 – 313, 1993.

[44] L. P. Boivin and D. S. Hanna. **Studies of bis-MSB wavelength shifter for application in a supernova neutrino detector**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **273**(1):397 – 402, 1988.

[45] A.G. Thompson, E. Villaescusa, and C.R. Windsor. *Rock Supprot and Reinforcement Practice in Mining*. CRC Press, 1999.

[46] Y. Nakano, H. Sekiya, S. Tasaka, Y. Takeuchi, R.A. Wendell, M. Matsubara, and M. Nakahata. **Measurement of radon concentration in super-Kamiokande's buffer gas**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **867**:108 – 114, 2017.

[47] Kaito Hagiwara et al. **Gamma Ray Spectrum from Thermal Neutron Capture on Gadolinium-157**. *PTEP*, **2019**(2):023D01, 2019.

[48] Giada Carminati. **The new Wide-band Solar Neutrino Trigger for Super-Kamiokande**. *Physics Procedia*, **61**:666 – 672, 2015. 13th International Conference on Topics in Astroparticle and Underground Physics, TAUP 2013.

[49] K. Abe et al., The Super-Kamiokande Collaboration. **Calibration of the Super-Kamiokande detector**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **737**:253 – 272, 2014.

[50] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini. **GEANT3**. 1987.

[51] S. Agostinelli et al. **Geant4—a simulation toolkit**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **506**(3):250 – 303, 2003.

[52] Robin M. Pope and Edward S. Fry. **Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements**. *Appl. Opt.*, **36**(33):8710–8723, Nov 1997.

[53] Erik Karl Blaufuss. *Study of the Energy Spectrum of Solar Neutrinos above 5.5 MeV*. PhD thesis, Louisiana State University, 2000.

[54] M Nakahata et al., The Super-Kamiokande Collaboration. **Calibration of Super-Kamiokande using an electron LINAC: The Super-Kamiokande Collaboration**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **421**(1):113 – 129, 1999.

[55] E Blaufuss et al., The Super-Kamiokande Collaboration. **16N as a calibration source for Super-Kamiokande**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **458**(3):638 – 649, 2001.

[56] Shantanu Desai. *High energy neutrino astrophysics with Super-Kamiokande*. PhD thesis, Boston U., 2004.

[57] Zoa Conner. *A Study of solar neutrinos using the Super-Kamiokande Detector*. PhD thesis, Maryland U., 1997.

[58] Michael Smy, The Super-Kamiokande Collaboration. **Low Energy Event Reconstruction and Selection in Super-Kamiokande-III**. *Proc. of the 30th Int. Conf. on Cosmic Rays*, **5**:1279–1282, 2007.

[59] Linyan Wan. *Experimental Studies on Low Energy Electron Antineutrinos and Related Physics*. PhD thesis, Tsinghua University, 2018.

[60] G.T. Zatsepin and V.A. Kuzmin. **Upper limit of the spectrum of cosmic rays**. *JETP Lett.*, **4**:78–80, 1966.

[61] Shirley Weishi Li and John F. Beacom. **First calculation of cosmic-ray muon spallation backgrounds for MeV astrophysical neutrino signals in Super-Kamiokande**. *Phys. Rev. C*, **89**:045801, 2014.

[62] Shirley Weishi Li and John F. Beacom. **Tagging Spallation Backgrounds with Showers in Water-Cherenkov Detectors**. *Phys. Rev. D*, **92**(10):105033, 2015.

[63] J. Beacom and S. Li. personal communication.

[64] K. Abe et al. **Solar Neutrino Measurements in Super-Kamiokande-IV**. *Phys. Rev. D*, **94**(5):052010, 2016.

[65] K. Abe et al. **Hyper-Kamiokande Design Report**. 5 2018.