Measurement of $\sin^2 2\theta_{13}$ Using Delayed Neutron Capture on Hydrogen in Double CHOOZ

by

Kazuhiro Terao

Submitted to the Laboratory for Nuclear Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
SEP 0 4 2013
LIBRARIES

©Massachusetts Institute of Technology 2013. All rights reserved

Author .....
Laboratory for Nuclear Science
April 2, 2013

Certified by .....
Janet M. Conrad
Professor
Thesis Supervisor

Accepted by .....
John Belcher
Associate Department Head for Education

Measurement of $\sin^2 2\theta_{13}$ Using Delayed Neutron Capture on Hydrogen in Double CHOOZ

by

Kazuhiro Terao

Submitted to the Laboratory for Nuclear Science
on April 2, 2013, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Double CHOOZ is a reactor antineutrino experiment built to measure θ_{13}. The experiment is designed to use two detectors at different baselines (400 m and 1 km) to precisely measure the disappearance of $\bar{\nu}_e$ from the Chooz reactor cores in Ardenne, France. Our inverse beta decay (IBD) signal is a two-fold coincidence of a prompt positron followed by a delayed neutron capture on gadolinium. The delayed neutron capture releases 8 MeV of energy from multiple gamma rays and is easily distinguished from natural radioactive backgrounds. While delayed neutron capture on gadolinium is used in all the reactor-based θ_{13} measurements [1, 2, 3], it is also possible to detect IBD via delayed neutron capture on hydrogen. In Double CHOOZ the Hydrogen detection channel has twice the signal statistics as the gadolinium detection channel, and provides an independent data sample with which to cross-check the gadolinium analysis result. The first result [4] is found to be $\sin^2(2\theta_{13}) = 0.097 \pm 0.034$ (stat) $\pm$ 0.034 (syst). The author lead the collaboration to make this measurement possible, and the detail of the work is presented in this document.

Thesis Supervisor: Janet M. Conrad
Title: Professor

Acknowledgments

First of all, let me thank you for reading my thesis.

My deepest gratitude goes to the MIT Double CHOOZ group. Professor Janet Conrad, my advisor, a mentor, and a supporter in many aspects of my life since I joined MIT. Thank you for providing me with a lot of opportunities to meet many physicists and showing me how to be a professional researcher. Dr. Lindley Winslow, you are truly the great leader for me. Thank you for always being there when I needed a help. I greatly appreciate your outstanding leadership, providing me with protection and independence in my research work. Dr. Joshua Spitz, you gave me one of final and important lessons in my Ph.D: ask if I need it for my thesis, and don't do it if not. I always enjoyed our trip to Meadhall for pints of beer after hours. Dr. Chris Jones, how could I have survived the waiting time for batch queues to start without the manga resource you provided? Ben and Christina, although we have never really worked together, thank for being there and always cheerful.

I thank all the group members of the Astro-Particule et Cosmologie (APC) laboratory for accommodating me there for more than two years! Professor Herve de Kerret always made sure I had enough resources to work comfortably. Alessandra was always there and provided practical help to set myself up over there. Jaime always cared and helped me: you are like a big sister in France to me. Without you, I am afraid who can cover for me for scratching a car. Davide and Stefano contributed in the dominance of Italian physicists at the APC, which I enjoyed a lot because of you guys! Michel, you have been a great czar for the collaboration and also for my life in France. I do not think the French government extended my visa without your help and patience with me at the police station. Dr. Didier Kryn, thank you very much for providing a lot of extra help for me to work on the on-site machines. Dr. Anatael Cabrera, I cannot thank you enough for walking the path of commissioning and data analysis together. You showed me the most exciting part of the experiment, and also taught me an importance of enjoying my life. I will keep your lessons forever. Dr. Alberto Remoto, we are meant to be friends over pints of beer. We finally did not call your Italian friends to execute our plan of the "Italian job." I look forward to this in the future.

My research experience in France was great because of my European colleagues. Pau, I hope

you are still my friend after my never-ending talk over pints of lager. Anselm and Sebastian, we will see a pink elephant again in Brussels, or I shall visit you in Aachen to chill with Banana-Weizen in summer! Bernd, you are like my big brother. Thank you for being very, very patient with me and for listening to me when I needed to talk (a lot). Luis, I hope you don't mind me to include you in Europe. I enjoyed working with you in the labs!

It has been always my pleasure to work with Japanese collaborators in this experiment. Professor Masaki Ishitsuka, thank you for being always available on Skype for any discussion and for paying lots of attention to my work. I appreciate Professor Fumihiko Suekane with his keen support for the first publication of the hydrogen analysis. Professor Hara and Sumiyoshi, thank you both for taking a shift together at Chooz. I really enjoyed our dinner there. Speaking of meals, I cannot forget to name Dr. Tomoyuki Konno for our good time in Chooz and also in Paris! Dr. Kyohei Nakajima and Junpei Maeda, thank you for your great contribution to the hydrogen analysis and also your patience with me for asking you to hit "one more bar" during my stay in Japan. Dr. Tsunayuki Matsubara, thank you for your hard work to take over my CT task. Your ability of being able to sleep at any time is a true gift!

Next up are my colleagues from the States. Matt Toups, thanks for joining Janet's group at MIT. I can skip a long list of items to thank you here (I will speak to you later). Arthur, thanks for sharing your amazing beer. Yes, that comes before my gratitude in work! Camillo, I thank you very much for being there as the analysis coordinator and supporting the hydrogen analysis. Leslie and Mike, I am very thankful to both of your attention to the hydrogen analysis from the very beginning. Brandon, Erica, and Cara, you guys have been always awesome. I miss our Skype conversation (including the status). Keith and Rachel, thanks for your great work in the hydrogen analysis! Deepak and Pi-Jung, it has been great to work with nice and kind colleagues like you. Glenn, thank you for being always available to discuss everything about software in general! Igor, in reply to your thesis acknowledgement, yes, I do not know what to do without our endless chit chat over Skype. You seem to have found a great position in EXO, however, wearing a bright orange T-shirt and standing in the very center in the collaboration picture!

Presence of friends outside research has been invaluable. I would like to thank all of my MIT

friends. Steve and Dan, thanks for taking me out to nice bars around and sharing your time. Mike Betancourt and Matt Walker, I really enjoyed playing kickball with you guys. Shawn, your great effort to bring in a dinosaur is very much appreciated by all students. Matthew, Keigo, Seung, Jeremy, Kai and Juven, thanks for helping me to try out all those restaurants in the Cambridge and Boston areas. Great thanks to Michel Tomasik for introducing Jacob Wirth! Arghavan, thank you for hosting a potluck and serving awesome dishes. Your extra bed has saved me a lot of money, and your kindness saved my mind during my busy visits to Cambridge. Apratim, thank you for always organizing a dinner for my return! Noah, thank you very much for keeping me company in the office. You have been a great source of knowledge for me.

Coming toward the end, I would like to thank my colleagues from Berkeley. Professor Stuart Freedman, who past away several months ago, who guided me along the path into neutrino physics. Joining KamLAND experiment in your group was the beginning of my journey as a physicist. You are and will be greatly missed. Professor Herbert Steiner, your Quantum Mechanics class was the beginning of my research experience in KamLAND. Tommy and Lindley (again!), you are like my big brother and sister. I hope to see you both together next time with a pint of Guinness for Tommy, Cosmopolitan for Lindley, and any alcohol for me. I cannot forget to thank Adrian Kaho Chan for being an awesome friend and colleague. You are still by far the slowest eater I have ever met. Tom Banks, thank you very much for joining me for a beer at Hotel Nido! It is always fun to see you. Brian and Richard, you have been a great resource of knowledge and skills. I cannot forget, of course, Ana Ovcharova. I badly miss studying and working with you at the LBL. My cooking knife you gave me reminds me. It is, by the way, still very sharp and cuts tomatoes like tofu.

Finally, I am very, very grateful to my parents, Izumi and Norihiro Terao for always being there and caring about me. My three older sisters, Akiko, Miyoko, and Yasuko, thank you for taking care of me ever since I started my life in the States. David, thank you for welcoming me for a dinner when I had no food in Berkeley! To my wife's family, Chihiro, Maki, and Miho, huge thanks for always welcoming me. I found another home in Japan! Lastly but not the least, I appreciate my wife, Yumi, for just being with me. Our marriage is the best thing

that happened during my Ph.D. I am sorry for not being able to choose it as this thesis topic. Without you, I could never come this far.

Contents

1 Introduction **27**

2 Neutrino Physics **29**

2.1 Neutrinos in the Standard Model 29

2.2 Neutrino Oscillation 30

2.2.1 Survival Probability of $\bar{\nu}_e$ from Nuclear Rectors 33

2.2.2 Two-Flavor Oscillation Model 34

2.2.3 Measurement of Mixing Angles 35

2.3 Measurement of θ_{13} 38

2.3.1 Past Short-Baseline Reactor Experiments 39

2.3.2 Hints From Long-Baseline Accelerator Experiments 40

2.3.3 Recent Results From Short-Baseline Reactor Experiments 40

2.3.4 Future Prospects for Double CHOOZ 42

3 Double CHOOZ Experiment **45**

3.1 Goal of the Double CHOOZ Experiment 45

3.1.1 Two Detectors Concept 45

3.1.2 Experiment Signal 47

3.1.3 Detection of Delayed Neutron Capture 49

3.2 Double CHOOZ Detector 51

3.2.1 Inner Detector (ID) 51

3.2.2 Inner Muon Veto System (IV) 54

3.2.3 Outer Muon Veto System (OV) 55

3.2.4 Trigger and Readout Electronics 55
3.2.5 Measurement of the Number of Protons 58
3.3 Calibration 60
3.3.1 Calibration Sub-systems 60
3.3.2 ID Photo-Multiplier Tube Gain and Timing Calibration 62

4 Event Reconstruction **67**
4.1 Low-Level Waveform Reconstruction 67
4.2 Vertex and Track Reconstruction 69
4.2.1 Vertex Reconstruction 70
4.2.2 Track Reconstruction 70
4.3 Energy Reconstruction 71
4.3.1 Energy Scale: Photo-electron to MeV Calibration 72
4.3.2 Energy Non-Uniformity Correction 73
4.3.3 Time Instability Correction 74
4.3.4 Residual Energy Non-Linearity 75
4.3.5 Summary of Energy Reconstruction 76

5 Signal Selection **77**
5.1 Analysis Data Sample 77
5.2 Pre-Selection 77
5.2.1 PMT Charge Spreadness 78
5.2.2 PMT Pulse Time Uniformity 80
5.2.3 After-Muon Veto 81
5.3 Neutrino Event Candidate Selection 82
5.3.1 Prompt Candidate Selection 82
5.3.2 Delayed Coincidence Candidate Selection 83
5.3.3 Isolation Cut 84
5.4 Summary 84

6 Detection Efficiency and Systematics **87**

6.1 Introduction to Efficiency Estimation 87

6.1.1 Simulation Based θ_{13} Analysis 87

6.1.2 ^{252}Cf Calibration Source Data 89

6.1.3 $\bar{\nu}_e$ Sample for Calibration 89

6.2 Delayed Neutron Capture Detection Efficiency 91

6.2.1 Fraction of Neutron Capture on Hydrogen 92

6.2.2 ΔT, ΔR, and E_{delayed} Efficiency Study Using ^{252}Cf 95

6.2.3 Spill In/Out 101

6.2.4 Overall Delayed Selection Uncertainty 105

6.2.5 Inefficiency on Delayed Events From Muon Veto 105

6.2.6 Combining All 106

6.3 Pre-Selection Efficiency 106

6.3.1 Charge Uniformity (MQTQ) Cut 107

6.3.2 Photon Arrival Time Cut ($T_{\text{start}}^{\text{RMS}}$) 109

6.4 Summary 110

7 Background Prediction **113**

7.1 Accidentals 113

7.1.1 Off-Time Sample Selection 114

7.1.2 Inefficiency Correction Factor and Rate Estimation 115

7.1.3 Analytical Cross-Check 118

7.2 Lithium 9 119

7.2.1 Rate Estimation 121

7.2.2 Rate Cross-Check Without Using Muon Track 126

7.2.3 Simulation Based E_{prompt} Shape Estimation 128

7.3 Fast Neutrons 130

7.3.1 Low-Energy IV-Tag Analysis 131

7.3.2 Rate and Shape of Fast Neutrons 136

7.3.3 Fast Neutron Analysis Cross-Check 137

7.4 Stopping-Muon Veto . . . 139
7.4.1 Selection Cuts . . . 140
7.4.2 Verification of Zero Stopping-Muon Background . . . 140
7.5 Light Noise . . . 142
7.5.1 Separation of the Correlated Light Noise Background . . . 143
7.5.2 The Rate of Contamination . . . 145
7.5.3 The Shape of Contamination . . . 146
7.6 Summary . . . 147
7.7 Validation with Reactor OFF-OFF . . . 147
7.7.1 Data Sample . . . 147
7.7.2 Accidentals Rate and Time Stability . . . 148
7.7.3 OFF-OFF Event Rate and Shape After Accidentals Subtraction . . . 149

8 Signal Prediction **153**
8.1 Simulation of the Fission Rate in the Chooz Reactor Cores . . . 153
8.1.1 Core Simulation . . . 154
8.1.2 Neutrino Spectrum at Reactor Core . . . 155
8.1.3 $\bar{\nu}_e$ Flux in the Detector . . . 158
8.1.4 BUGEY4 Anchor Point . . . 159
8.1.5 Summary . . . 160
8.2 Detector Simulation . . . 160
8.2.1 Event Generation . . . 160
8.2.2 Geant4: Particle-Tracking Simulation . . . 161
8.2.3 Read-Out Electronics Simulation . . . 162
8.2.4 A la DATA Reconstruction . . . 162
8.3 Summary . . . 162

9 Oscillation Analysis **165**
9.1 Chi-Square Approach . . . 165
9.2 Oscillation Analysis Result . . . 167

9.3 Summary and Future Prospect 174

A Further Contributions **177**

A.1 Generic Production Framework 177

A.1.1 `DCProd` 177

A.1.2 Implementation Case: Common Trunk 179

A.2 Detector Commissioning 182

A.2.1 On-Site Data Quality Check 182

A.2.2 Data Stream Monitoring 183

A.3 Monitoring Data Processing and Quality Control 184

A.4 Analysis of Detector Response After Muon 184

B Time Projection Chamber for Detecting Fast Neutrons **187**

List of Figures

2-1 Atmospheric (black) and MINOS (blue) result. The combined result is shown in the red-shaded region for normal hierarchy at 90 % C.L. (dashed) and 3σ (solid) [5]. 36

2-2 Allowed region for neutrino oscillation parameters from KamLAND and solar neutrino experiments [6, 7]. The figure is taken from Ref. [8]. 37

2-3 L/E oscillation pattern of $\bar{\nu}_e$ flux from nuclear reactors in Japan, taken from Ref. [8]. 38

2-4 Allowed region of $\sin^2(2\theta_{13}) - \delta$ parameter space for T2K, MINOS, DC, Daya Bay (DB), and RENO combined fit from Ref. [9]. The left figure is normal and the right figure is for an inverted hierarchy. 42

3-1 The Chooz reactor complex in Ardennes, France 46

3-2 A simulated energy spectrum shape of $\bar{\nu}_e$ which is produced from a ^{235}U fission chain and undergoes an IBD interaction (blue). An IBD interaction cross section as a function of energy (red-dashed) shown with the original energy spectrum of $\bar{\nu}_e$ emitted from the fission process (black-dashed). The figure is taken from Ref. [10] 49

3-3 Measurements from various θ_{13} measurements [11, 1, 12, 3, 13, 14]. 50

3-4 Detector overview including the ID, IV, and the OV sub-systems [1]. The ID consists of the Target, γ-catcher, and Buffer shown in the figure. 52

3-5 Absorption and emission spectra of aromatic components in the Target and γ-catcher liquid scintillators [15]. 53

3-6 The readout diagram for the ID and IV electronics from the PMTs to the Data Acquisition (DAQ) [1]. 56

3-7 The ID and IV readout trigger rate stability during the data taking. A typical trigger rate was found to be around 130 Hz. 57

3-8 Left: a single photo-electron-like waveform from physics data. Right: a large amplitude waveform from a muon-like event. The digitization is done at 500 MHz. 58

3-9 Trigger efficiency curve as a function of visible energy (for the definition of visible energy, see Sec. 4.3). 59

3-10 A toy drawing of Z-Axis and Guide-Tube. The Guide-Tube runs from the Chimney along the Target wall down to the equator, then comes back along the γ-catcher wall. Z-Axis is a simple fish-line which can move a radioactive source along the vertical axis inside the Target 61

3-11 Example plot of PMT time-offset calibration using the ID Light Injection (IDLI) system. Each data point represents a PMT installed in the opposite side of the Buffer wall from the light source. The vertical axis shows a reconstructed pulse timing after subtraction of a reference time from the external trigger signal. The horizontal axis shows the distance between each PMT and the light source. . . . 63

3-12 Single photo-electron (SPE) distribution from PMT 180, one of the ID PMTs installed in the middle of the Buffer side-wall. The horizontal axis shows the integrated electronic charge from reconstructed pulses. The red line shows a fitted spectrum defined in Eq. 3.7. 64

3-13 Example ID PMT gain extracted from the IDLI data points. The gain variation at low charge due to SPE inefficiency can be seen [1]. 65

4-1 Left: a pedestal baseline mean value estimated from 8 hours' worth of physics data. The orange histogram shows the distribution from normal physics events while the red line shows that for externally triggered events. Right: a pedestal baseline RMS distribution shown in the same style as the left plot. 68

4-2 A result of applying `RecoPulse`to the waveform shown in the left of Fig. 3-8. The dotted line shows the range of reconstructed waveform in time (horizontal) and the maximum amplitude in amplitude (vertical). The shaded region on the left includes the time samples used for pedestal analysis. 69

4-3 Left to right: x, y, and z reconstructed position vertex distribution by RecoBAMA using calibration data with a ^{60}Co source deployed at the Target center. 71

4-4 A photo-electron distribution of a gamma ray from a neutron capture on hydrogen. Left is DATA taken with a ^{252}Cf neutron source while the right represents the MC, fitted by a function defined in Eq. 4.4. 72

4-5 Energy uniformity map from spallation neutrons [1]. 73

4-6 A neutron capture peak for hydrogen (left) and Gd (right) nuclei from spallation neutron data sample during the first 5 days of data taking. Extracted peak positions make up data points in Fig. 4-7. 74

4-7 A time stability of a relative shift in neutron capture peak energy position. Red data points are from a neutron capture on hydrogen and blue for ^{157}Gd. Each data point is made from a consecutive 5 days of data taking. Error bars represent an uncertainty from the fitting of the energy distribution. 75

5-1 Detector run-time and livetime for analysis on each calendar day since the beginning of data taking. The cumulative livetime is 240.06 days over 335 calendar days. 78

5-2 MQTQ distribution of delayed neutron capture from $\bar{\nu}_e$ MC sample (red) and ^{252}Cf data from the source deployment at γ-catcher edge (black). Blue data points show the MQTQ distribution of all events from a typical physics run. The gray dotted line indicates the default light noise rejection cut value for an IBD candidate pre-selection cut. 79

5-3 T_{start}^{RMS} distribution of delayed neutron capture from $\bar{\nu}_e$ MC sample (red) and ^{252}Cf data from the source deployment at γ-catcher edge (black). Blue data points show the T_{start}^{RMS} distribution of all events from a typical physics run. The gray dotted line indicates the default light noise rejection cut value for an IBD candidate pre-selection cut. 80

5-4 Number of IBD candidates with Gd selection cuts against the duration of after-muon veto time window. The gray line is drawn at 1 ms, indicating the default cut value used in pre-selection of candidates. A statistical uncertainty is shown in blue colored error band. 81

5-5 Daily muon rate stability during the data taking period. 82

5-6 DATA and MC E_{delayed} efficiency distribution as a function of the lower boundary of *ed* cut from the ^{252}Cf calibration source in the Target center. For details of how the efficiency is studied, see Sec. 6.1. Note that the MC correction factor introduced in the Sec. 6.1 is not applied to the MC efficiency shown in this plot. 83

5-7 Accidentals-subtracted IBD candidates' spectrum shape. From the left to right, E_{delayed}, ΔR, and ΔT distributions are shown. 85

6-1 Example plots from ^{252}Cf source calibration data sample. Yellow-filled histograms are from the MC and black points are from DATA after applying selection cuts described in the main text. Top row shows distributions from Guide-Tube and the bottom from Z-Axis (Target center) source deployment. From left to right, each column shows ΔT, ΔR, and E_{delayed} distributions. 90

6-2 Example plots of $\bar{\nu}_e$ candidates with the modified cut of $4 < E_{\text{prompt}} < 12.2$ MeV for DATA (blue) and MC (red). Statistics in DATA and MC are not normalized, and DATA points carry backgrounds in addition to $\bar{\nu}_e$ events. 91

6-3 E_{delayed} distribution from an $\bar{\nu}_e$ sample selected with a modified set of cuts (see the main text). The majority of neutron captures making this distribution are expected to be from the γ-catcher volume. 94

6-4 From top to bottom row: ΔT, ΔR, and E_{delayed} efficiency plotted against Z for the Target (left column) and R for the γ-catcher (right column). For each plot, the MC correction factor is applied to match the overall volume efficiency. . . . 96

6-5 The γ-catcher liquid scintillator volume distribution as a function of $R = \sqrt{x^2 + y^2 + z^2}$. The vertical axis shows the volume in arbitrary units. 97

6-6 DATA/MC ΔT efficiency ratio over Z-Axis. The sampling points taken to estimate the overall MC correction factor are shown in red. 97

6-7 From top to bottom: ΔT, ΔR, and E_{delayed} efficiency distributions in the Target (left column) and the γ-catcher (right column), respectively. 99

6-8 Resulting efficiency distributions from 10,000 fake experiments using a ΔT distribution shown in Fig. 6-4. Left is for DATA and the right is for MC. 100

6-9 Variation of DATA-MC discrepancy as a function of the lower boundary cut value to define the full sample for the efficiency study. A fit to the distribution yields a consistent time constant with a muon lifetime of 2.2 μs. 103

6-10 MQTQ DATA (blue) and MC (red) efficiency distribution from Z-Axis (left) and Guide-Tube (right). There is no MC correction factor applied for the plotted data points. 107

6-11 MQTQ DATA (blue) and MC (red) distributions for the Target center ^{252}Cf (left) and γ-catcher edge point (right). 108

6-12 $T_{\text{start}}^{\text{RMS}}$ cut efficiency along Z-Axis (left) and Guide-Tube (right). Blue data points denote DATA and red colored dots denote MC. 109

6-13 $T_{\text{start}}^{\text{RMS}}$ DATA (blue) and MC (red) distributions for the Target center ^{252}Cf (left) and γ-catcher edge point (right). 110

7-1 From left to right: accidentals E_{Prompt}, $E_{Delayed}$, and ΔT distributions. 115

7-2 The left plot shows ΔR distributions for on-time and off-time data samples in blue and red, respectively. The total off-time data sample is scaled down by 125, the number of sampling windows. The right plot shows the bin-by-bin ratios of two distributions for $1100 < \Delta R < 3500$ mm. 117

7-3 The ΔR ratio plot after increasing the muon veto period from 1 ms to 2 ms. The flat line fit indicates a $10.03 \pm 0.25\%$ excess of accidentals in the on-time sample. 118

7-4 Left: A daily comparison of two accidentals rate estimation methods. The blue data points are the rate measured from the off-time selection, and the red data points are the rate derived by analytical estimation. The methods agree well over the full data-taking period. Right: Ratio of rates show in the left plot. . . . 119

7-5 Day-to-day stability of parameters used in the analytical estimation. 120

7-6 ΔT_μ distribution for IBD candidates and all preceding high-energy muons with ID energy deposition > 600 MeV. 121

7-7 ΔT_μ distribution for ^{9}Li rate estimation. Black dots are all IBD candidates while the gray histogram shows the off-time accidentals sample for the ^{9}Li analysis. Blue data points are obtained by subtracting the off-time distribution from the IBD candidates. 123

7-8 FLUKA simulation result made for the KamLAND detector [4]. The distribution shows the distance of the closest approach for ^{9}Li generation vertex from the parent muon track. 123

7-9 Prompt energy spectrum with expected neutrinos and backgrounds for the ^{9}Li rate estimation. 127

7-10 Decay branches of ^{9}Li through ^{9}Be excited states from Ref. [16]. 129

7-11 Prompt energy spectrum of ^{9}Li MC sample. 130

7-12 A toy model of multiple fast neutrons produced by a muon interacting separately in ID and IV liquid scintillator (yellow). 131

7-13 E^{IV}_{Prompt} spectrum of IV-tagged fast neutron candidates. The different colors represent different IV PMT hit conditions: 0, 2, 4, 6 minimum IV PMT hit conditions are shown in black, red, magenta, and green, respectively. 132

7-14 E^{ID}_{Prompt} (left) and $E^{ID}_{Delayed}$ (right) spectra of IV-tagged FN candidates. The different colors represent different IV PMT hit conditions: 0, 2, 4, 6 minimum IV PMT hit conditions are shown in black, red, magenta, and green, respectively. 133

7-15 ΔT_{Start}^{ID-IV} (left) and E_{Prompt} vs. ΔT_{Start}^{ID-IV} (right) distribution. The ΔT_{Start}^{ID-IV} distribution is off-time subtracted. The right plot shows on-time sample in red and off-time sample in blue. 135

7-16 E_{Prompt} of fast neutron sample. The black data points contain the ID-IV correlated and uncorrelated component while the red data points show the ID-IV uncorrelated component. Blue data points are obtained by subtracting the red from black data points; they are the fast neutron E_{prompt} spectrum. 136

7-17 The fit result of the function in Eq. 7.16 to the fast neutron E_{prompt} spectrum. . 137

7-18 Prompt energy spectra of the low Q_{IV} (left) and high Q_{IV} cuts (right). Dotted lines show specific energy points, 12 MeV and 30 MeV from left. Thc high Q_{IV} spectrum is normalized to the low Q_{IV} spectrum by matching the integral of both spectra in 12 to 30 MeV. 138

7-19 A function defined in Eq. 7.16 is fitted to the E_{prompt} spectrum of the high-Q_{IV} sample 138

7-20 Comparison of the fast neutron sample spectrum (left) and fitted function shape (right) between the nominal fast neutron sample and high-Q_{IV} sample. 139

7-21 ΔT spectrum from the stopping-muon sample. 141

7-22 The ΔT spectrum fit with a triple exponential for ΔT spectrum of hydrogen IBD candidates after subtraction of accidental backgrounds' component. 142

7-23 ρ^2 vs. Z vertex distribution of coincidence pairs selected with a modified IBD selection cut, $0 < \Delta T < 15$ μs. Also only data after day 190 is used to enhance the existence of the background. 144

7-24 Subtracted ΔT distribution for the three volume cuts, V_1, V_2, and V_3 (see the main text for the definition). From the left to right, ΔT distribution is made by a volume-wise subtraction of $V_1 - V_2$, $V_1 - V_3$, and $V_2 - V_3$. 145

7-25 Subtracted ΔT distribution of an enlarged V_1 and V_2 definition for the cross-check. 146

7-26 Subtracted E_{prompt} spectrum shape for $V_1 - V_2$ (left) and $V_1 - V_3$ (right). Both spectra are scaled to the corresponding rate estimate. 146

7-27 A comparison of accidentals' daily statistics estimation between the off-time window (blue) and analytical method (red). A very good agreement is found over all OFF-OFF period, as was the case for the DC2ndPub data sample. . . . 149

7-28 The left figure shows ΔR distribution for IBD candidates and off-time selected data sample in blue and red color, respectively. The off-time data sample is scaled using a derived scaled factor for the OFF-OFF period. The right figure shows the bin-by-bin ratio of two distributions on the left for $1100 < \Delta R < 3500$ mm. 150

7-29 OFF-OFF data E_{prompt} is shown in the red data points. The stacked histograms represent ^{9}Li (cyan) and fast neutrons (orange). 150

8-1 The average thermal power data monitored and provided by the electrocity company in France as a function of days since the beginning of $\bar{\nu}_e$ data taking. . . . 155

8-2 Fission rates from the main fission isotopes, ^{239}Pu, ^{241}Pu, ^{235}U, and ^{238}U, simulated by MURE with input data from the EDF. 156

8-3 Error budgets for the main fission isotopes shown in Fig. 8-2 as a function of time based on the simulated burn-up by MURE. 157

8-4 The predicted energy spectrum of $\bar{\nu}_e$ at the reactor core from main fission isotopes shown in the legends. 158

8-5 The IBD candidate rate stability plots for the Gd analysis (top) and the hydrogen analysis (bottom). The horizontal axis shows the number of elapsed days since the beginning of the data taking period. The vertical axis shows the number of events per day. The predicted rate is shown in blue while the observed rate is shown in black. Backgrounds are subtracted. 163

A-1 CT data processing flow chart at `CCIN2P3`. File storage servers and batch servers are monitored and managed by multiple daemon processes constantly as described in the main text. Upon successful ending of the CT processing, output files become accessible to the analyzers and also are ready for inter-cluster data transfer. 180

A-2 Data stream monitoring scheme and example viewer. 183
A-3 The averaged pedestal mean value over all channels (Z-axis) for different muon energy (Y-axis) and time after muon (X-axis). One can see pedestal bias changes as a function of these two parameters. 185

List of Tables

3.1 Location of the Double CHOOZ near and far detectors. Values for the near detector will be measured once the installation is finished. 47

3.2 A summary of number of protons estimations in each detector sub-volumes shown with the MC correction factors 59

3.3 A summary of deployed calibration sources. ^{68}Ge goes through an electron capture to emit two 0.511 keV gamma rays. The energy of ^{252}Cf is a summed energy of gamma ray(s) from a delayed neutron capture only. The prompt events include multiple gamma rays and neutrons from the fission process. 61

4.1 Energy-scale systematic errors [1]. 76

6.1 Summary of the MC correction factors, efficiencies and uncertainties from the ^{252}Cf neutron source study. $\mathcal{F}$ and ϵ represent the MC correction factor and efficiency, respectively. The subscripts T and GC denote the relevant volume, which is either Target or γ-catcher, respectively. Numbers inside the parentheses indicate the systematic error component of the quoted total error. 98

6.2 Summary of the event fraction of various Spill In/Out sub-categories. The total to define a fraction is the total number of IBD candidates that are selected in the MC sample. 101

6.3 MC correction factors, $\mathcal{F}$, and DATA-MC systematic uncertainty, $\delta(\epsilon)$. The expected $\delta(\epsilon)$ is derived from ^{252}Cf source data. The MC event fraction is used to determine the weight of each parameter between Target and γ-catcher values. The observed values are derived from the $\bar{\nu}_e$ data sample. The quoted uncertainty on the observed value represents the statistical uncertainty in the data set. . . . 103

6.4 A summary table for the hydrogen capture fraction MC correction factors and uncertainties. The uncertainty is a quadrature sum of the DATA-MC discrepancy for the pure γ-catcher events and the difference in the MC hydrogen capture fraction quantified with the MC true vertex information. The Target and γ-catcher uncertainties are taken from Sec. 6.2.1. 104

6.5 DATA and uncorrected MC efficiencies for MQTQ cut. Errors are statistical only. 107

6.6 $T_{\text{start}}^{\text{RMS}}$ cut efficiencies for DATA and MC without a correction factor. Errors are statistical only. 109

7.1 Summary of parameter values for the ^{9}Li rate scaling from the Gd analysis. . . . 125

7.2 Summary of ϵ_{IVT} for different IV-tag cut conditions. 133

7.3 Time constants used for the fit to the ΔT spectrum from the $\bar{\nu}_e$ sample. 141

7.4 A summary table of backgrounds in the hydrogen analysis. As a reference, values from the Gd analysis are also shown. 147

7.5 Summary of parameter values to be used for OFF-OFF analytical rate estimation. 148

8.1 Uncertainties in the $\bar{\nu}_e$ flux prediction. 160

9.1 A comparison of uncertainties in the hydrogen and Gd analyses [1, 4]. Here, the variance is defined as: $\sqrt{\sum_{i,j} \mathcal{M}_{ij}}/N_{signal}$, where $\mathcal{M}_{ij}$ is the relevant covariance matrix, evaluated at $\sin^2 2\theta_{13} = 0$ and N_{signal} is the total number of signal events predicted at $\sin^2 2\theta_{13} = 0$. 167

Chapter 1

Introduction

The last few decades have been a truly exciting period in neutrino physics. Since an observation of a deficit in neutrino flux from the sun from the prediction [17], the so-called the solar neutrino problem, physicists have come a long way to better understand this elusive particle. From the evidence of neutrino flavor oscillation from many experiments [8, 6, 18, 19, 14, 2, 3, 11], now the notion of massive neutrinos is widely accepted, and the measurement of mixing parameters is almost complete. Among these parameters, the most recently measured is the last mixing angle θ_{13}. This mixing angle has been known to be small, and only an upper bound was given [20, 21] prior to the recent measurements [11, 2, 3, 14, 13]. A precise measurement of this parameter is the goal of the modern reactor-based neutrino experiments such as Daya Bay, RENO, and Double CHOOZ.

This document presents a pioneering work in modern reactor-based neutrino experiments, which is a measurement of neutrino mixing angle θ_{13} using delayed neutron capture on hydrogen. The result suggests this new approach is feasible for the Double CHOOZ experiment in which the author is involved. The Double CHOOZ collaboration has decided that a future measurement of θ_{13} should be made from a combined analysis of the standard method and this new approach. The combined analysis work is, however, to be done in the near future and is not a part of this thesis work.

Chapter 2 starts with a brief introduction to neutrino physics, including the present knowledge about neutrino mixing parameters, and ends with the current knowledge about the mea-

surement of the last mixing parameter, θ_{13}, from the modern reactor-based neutrino experiments, including Double CHOOZ in which the author is involved for this thesis work. The experimental concept and its goal are presented in chapter 3 with a detailed description of the detector. Chapter 4 briefly describes reconstruction of physics events. The signal extraction method is described in chapter 5, and the associated systematic uncertainty is described in chapter 6. The extracted signal candidate sample includes backgrounds whose prediction is given in chapter 7. Chapter 8 describes a prediction of the signal with a hypothesis of $\theta_{13} = 0$. Finally, chapter 9 describes a strategy of fitting the predicted background and signal to the extracted candidate sample to quantify the value of θ_{13}. The chapter ends with a manuscript published in a public journal.

Chapter 2

Neutrino Physics

This chapter briefly describes the modern picture of neutrino physics, focussing on the measurement of θ_{13}. Section 2.1 introduces neutrinos in the Standard Model (SM) and the mixing matrix that relates the mass and flavor eigenstates of neutrinos. Current knowledge of parameters in the mixing matrix is summarized in Sec. 2.2. Finally, Sec. 2.3 presents the status of results on θ_{13} measurements.

2.1 Neutrinos in the Standard Model

In 1930, Pauli first postulated the presence in nuclear beta decay of an electrically neutral fermion with mass much lighter than an electron. This third "invisible" particle in beta decay was invoked to explain the continuous electron energy spectrum. The name "neutrino" was given to the particle by Fermi, who provided a theory of the point-like weak interaction for the nuclear beta decay process in 1933. The direct detection of neutrinos was carried out by Reines and Cowan in 1956 by detecting $\bar{\nu}_e$ from a nuclear reactor. Since then, physicists have come a long way to reveal the deep nature of this elusive particle. In the SM, the most successful theoretical description of fundamental particles and their interactions today, neutrinos are considered massless neutral spin 1/2 leptons that come in three weak flavor eigenstates: ν_e, ν_μ, and ν_τ.

The SM assumption of a very low mass upper bound, of the order of $\mathcal{O}(1)$ eV or smaller,

is supported by the measurement of a tritium beta decay end-point energy [22]. It is possible, however, to introduce mass eigenstates for neutrinos with very tiny masses into the theoretical framework. In this case, at least three mass eigenstates are expected for neutrinos, to span the space of three weak flavor eigenstates. These are labeled as ν_1, ν_2, and ν_3. Historically, ν_3 is considered the heaviest and ν_1 as the lightest in the "normal" ordering of mass eigenstates with associated masses.

As with the case of quarks, these mass eigenstates do not need to coincide with the mass eigenstates in general. The Cabibbo-Kobayashi-Masakawa (CKM) matrix describes the mixing of strong and weak eigenstates of quarks. An analogous description in the lepton sector is called the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix,

$$\mathcal{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & e^{i\delta}\sin\theta_{13} \\ 0 & 1 & 0 \\ -e^{i\delta}\sin\theta_{13} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.1}$$

or, equivalently,

$$\mathcal{U} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \tag{2.2}$$

where c_{ij} and s_{ij} represent $\cos\theta_{ij}$ and $\sin\theta_{ij}$, respectively. As shown in Eq. 2.1, the PMNS matrix is simply a product of three unitary rotation matrices that introduce three rotation angles, known as "mixing angles," called θ_{12}, θ_{23}, θ_{13}. A CP-violating phase factor, δ, may also be included. Understanding mixing in the lepton sector requires quantifying the parameter values of the PMNS matrix.

2.2 Neutrino Oscillation

An introduction of a unitary rotation matrix is generic and simple enough, but the experimental challenge is to measure those parameters in Eq. 2.2 in order to demonstrate that this extension

to the SM theory is motivated. This can be done through investigating a physics phenomenon called neutrino oscillation, a manifestation of the existence of massive neutrinos.

Consider a very simple quantum mechanical Hamiltonian, with a simple plane wave solution to a flavor eigenstate α,

$$|\nu_\alpha\rangle = \sum_\ell \mathcal{U}^*_{\alpha\ell} e^{-iE_\ell t} |\nu_\ell\rangle. \tag{2.3}$$

Here, t represents time and E_ℓ represents the energy of mass eigenstate ℓ, which runs from one to three in the three-mass-eigenstates framework. The probability of observing ν_β after it exists for some time (or, equivalently, travels a certain distance) is, then,

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) &= |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 \\ &= \left| \sum_{k,\ell} \langle \nu_\ell | \mathcal{U}^\dagger_{\ell,\beta} \mathcal{U}_{\alpha,k} e^{-iE_k t} | \nu_k \rangle \right|^2 \\ &= \left| \sum_k \mathcal{U}^*_{\beta,k} \mathcal{U}_{\alpha,k} e^{-iE_k t} \right|^2 \\ &= \sum_{k,\ell} \mathcal{U}^*_{\beta,k} \mathcal{U}_{\alpha,k} \mathcal{U}_{\beta,\ell} \mathcal{U}^*_{\alpha,\ell} e^{-i(E_k - E_\ell)t} \end{aligned} \tag{2.4}$$

Our available sources of neutrinos–the sun, atmospheric interactions with cosmic rays, particle accelerators, reactors, and the like–all produce neutrinos through the weak interaction with energy scales from the keV to GeV level. At such an energy scale, given sub-electron-volt scale masses, neutrinos are ultra-relativistic to a very good approximation. This motivates Taylor expanding the energy term in Eq. 2.4, using the approximation:

$$E_i = \sqrt{|\vec{p}|^2 + m_i^2} \approx |\vec{p}| + \frac{m_i^2}{2|\vec{p}|}. \tag{2.5}$$

Note the above expression assumes a common momentum among different mass eigenstates. Further, an ultra-relativistic neutrino travels distance L in time approximately $\approx ct$. Under these assumptions, Eq. 2.4 becomes

$$P(\nu_\alpha \to \nu_\beta) = \sum_{k,\ell} \mathcal{U}^*_{\beta,k} \mathcal{U}_{\alpha,k} \mathcal{U}_{\beta,\ell} \mathcal{U}^*_{\alpha,\ell} \exp\left(-i \frac{\Delta m^2_{k\ell} L}{2E}\right) \tag{2.6}$$

where $\Delta m^2_{k\ell}$ is the squared mass splitting $m^2_k - m^2_\ell$. Let's further rewrite the summation into a more enlightening form. One sees that α and β are interchangeable if $k = \ell$. Also it can be seen that the expression for the case of $k < \ell$ is a complex conjugation of $\ell < k$. From these facts, one can write

$$P(\nu_\alpha \to \nu_\beta) = \sum_{k=\ell} \mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta\ell}\mathcal{U}^*_{\alpha\ell} + 2\sum_{k>\ell} Re\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta\ell}\mathcal{U}^*_{\alpha\ell}\exp\left(-i\frac{\Delta m^2_{k\ell}L}{2E}\right)\right] \tag{2.7}$$

Using Euler's formula and an identity for multiplication of complex numbers, one obtains

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) = \quad & \sum_{k=\ell} \mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\alpha\ell}\mathcal{U}^*_{\beta\ell} + 2\sum_{k>\ell} Re\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta\ell}\mathcal{U}^*_{\alpha\ell}\right] \\ & -4\sum_{k>\ell} Re\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta\ell}\mathcal{U}^*_{\alpha\ell}\right]\sin^2\frac{\Delta m^2_{k\ell}L}{4E} \\ & +2\sum_{k>\ell} Im\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta\ell}\mathcal{U}^*_{\alpha\ell}\right]\sin\frac{\Delta m^2_{k\ell}L}{2E}. \end{aligned} \tag{2.8}$$

The first two terms can be simplified to $\delta_{\alpha\beta}$ because it is the probability of observing ν_β at $L = 0$. This is true only if $\mathcal{U}$ is a unitary rotation matrix, meaning there is an assumption that the oscillation phenomenon can be described by only three eigenstates. Under this assumption, one gets

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) = \quad & \delta_{\alpha\beta} \\ & -4\sum_{k>\ell} Re\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta\ell}\mathcal{U}^*_{\alpha\ell}\right]\sin^2\frac{\Delta m^2_{k\ell}L}{4E} \\ & +2\sum_{k>\ell} Im\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta\ell}\mathcal{U}^*_{\alpha\ell}\right]\sin\frac{\Delta m^2_{k\ell}L}{2E} \end{aligned} \tag{2.9}$$

The final expression suggests a useful identity: $P(\nu_\alpha \to \nu_\beta; \mathcal{U}) = P(\nu_\beta \to \nu_\alpha; \mathcal{U}^*)$. CPT invariance ensures $P(\nu_\alpha \to \nu_\beta) = P(\bar{\nu}_\beta \to \bar{\nu}_\alpha)$. Accordingly, for the $\bar{\nu}_\alpha$ oscillation probability,

one finds

$$
\begin{aligned}
P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) = \quad & \delta_{\alpha\beta} \\
& -4 \sum_{k>\ell} Re\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta \ell}\mathcal{U}^*_{\alpha \ell}\right] \sin^2 \frac{\Delta m^2_{k\ell} L}{4E} \\
& -2 \sum_{k>\ell} Im\left[\mathcal{U}^*_{\beta k}\mathcal{U}_{\alpha k}\mathcal{U}_{\beta \ell}\mathcal{U}^*_{\alpha \ell}\right] \sin \frac{\Delta m^2_{k\ell} L}{2E}
\end{aligned} \tag{2.10}
$$

Both Eq. 2.9 and 2.10 show that the probability oscillates as a function of L/E. For a neutrino with energy E, it suggests that its flavor state oscillates as the neutrino travels in space. The phenomenon is called, therefore, "neutrino oscillation."

With these equations, the experimentalists' task is clear: pursue measurements to test these equations, and, if demonstrated, extract the mixing parameters and squared-mass splittings. For this, the experimentalists can tune the value of L/E and select the weak eigenstates α and β through appropriate beam and experimental design choices. A brief history of successful measurements in pursuit of neutrino oscillations is provided in section 2.2.3.

In the derivation of Eq. 2.9 and 2.10, two major assumptions are made: all mass eigenstates share the same momentum, and both $|\vec{p}|$ and L are treated as if they have single values. A more rigorous derivation can be done using a wave packet rather than a simple plane wave solution. This is shown in Ref. [23], and it essentially obtains the same result as what is shown above. A third aspect of our derivation is that it omits the presence of a possible complex Majorana phase factor which comes from an additional diagonal matrix element that should be multiplied in Eq. 2.1. This phase factor, even if it is present, does not affect the oscillation probability expression, and hence it is omitted. Thus, while our appraoch to the derivation has been greatly simplified, the overall conclusion is valid.

2.2.1 Survival Probability of $\bar{\nu}_e$ from Nuclear Rectors

As described later in this manuscript, the focus of this thesis is a measurement of the mixing angle θ_{13} through an observation of $\bar{\nu}_e$ disappearance in the flux originating from a nuclear reactor core. For this reason, the $\bar{\nu}_e$ survival probability expression is useful. We derive this

here for later reference. From Eq. 2.10 and 2.2, one obtains

$$\begin{aligned} P(\bar{\nu}_e \to \bar{\nu}_e) &= 1 - \frac{|\mathcal{U}_{e2}|^2}{|\mathcal{U}_{e2}|^2 + |\mathcal{U}_{e1}|^2} \sin^2(2\theta_{13}) \sin^2\left(\frac{\Delta m^2_{32} L}{4E}\right) \\ &\quad + \left(1 - \frac{|\mathcal{U}_{e2}|^2}{|\mathcal{U}_{e2}|^2 + |\mathcal{U}_{e1}|^2}\right) \sin^2(2\theta_{13}) \sin^2\left(\frac{\Delta m^2_{31} L}{4E}\right) \\ &\quad - \cos^2\theta_{13} \sin^2(2\theta_{12}) \sin^2\left(\frac{\Delta m^2_{21} L}{4E}\right) \end{aligned} \tag{2.11}$$

2.2.2 Two-Flavor Oscillation Model

Under certain conditions, the generic oscillation probability expression derived above can be further simplified into an expression that involves mixing of only two mass generations. This simplification is possible when either of two following cases is applicable: (1) one of the mixing angles involved in the relevant weak eigenstates is very small, or (2) one of the three squared-mass splittings is much smaller than the other two (i.e. $\Delta m^2 \langle\langle \Delta M^2$), and the experiment's L/E is optimized to the larger squared-mass splitting. In the first case, a small mixing angle implies $\sin\theta_{small} \approx 0$ and $\cos\theta_{small} \approx 1$. In the second case, $\sin^2 \delta m^2 \frac{L}{E} \approx 0$.

If one of the three mixing angle is very small, and hence only two flavors participate in the oscillation, one can write a very simple 2×2 unitary rotation matrix

$$\mathcal{U} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{2.12}$$

The corresponding oscillation probability expression becomes as simple as

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} + (-1)^{\delta_{\alpha\beta}} \sin^2(2\theta_{\alpha\beta}) \sin^2(\frac{\Delta m^2 L}{4E}) \tag{2.13}$$

In the second case, assuming $\delta m^2_{21} \langle\langle \Delta m^2_{32} \approx \Delta m^2_{31} \approx \Delta M^2$, an experiment with $\Delta M^2 L/4E \approx \mathcal{O}(1)$ sees the following oscillation probability

$$P(\nu_\alpha \to \nu_\alpha) = 1 - 4\,|\mathcal{U}_{\alpha 3}|^2 (1 - |\mathcal{U}_{\alpha 3}|^2) \sin^2(\frac{\Delta M^2 L}{4E}) \tag{2.14}$$

$$P(\nu_\alpha \to \nu_\beta; \alpha \neq \beta) \;=\; 4\left|\mathcal{U}_{\alpha 3}\mathcal{U}^*_{\beta 3}\right|^2 \sin^2\left(\frac{\Delta M^2 L}{4E}\right) \tag{2.15}$$

2.2.3 Measurement of Mixing Angles

The first observation which revealed neutrino flavor change was done by Ray Davis et al. in 1965. They performed a measurement of the neutrino flux originating from the Sun's internal fusion reactions and found a deficit from the predicted flux from the standard solar model by John Bahcall et al. [17]. This observation has been reproduced by multiple other experiments, including Sudbury Neutrino Observatory (SNO) [24, 25, 6], and is understood today as neutrino oscillations modified by "matter effects" [26] that arise from the high electron density of the sun. Detailed discussion about matter effects and neutrino oscillation is not included in this thesis as this has a negligible effect on the current θ_{13} measurement, but readers may see details in Ref. [27, 28]. Although the solar environment leads to a much more complex phenomenology for neutrino flavor change, the measured solar parameters should also be manifested in the simple oscillation phenomena we have derived above. This connection was demonstrated by the KamLAND reactor-based experiment [8]. This was an important sequence of events in that it demonstrates that our new beyond-SM theory is capable of making predictions–the solar parameters could be used to predict parameters measured in KamLAND. This motivates continued effort to quantify all of the mixing parameters in Eq. 2.1 so that further tests can be made.

The first component of the PMNS matrix in Eq. 2.1 has been measured by the Super Kamiokande (SK) and MINOS experiments through an observation of ν_μ disappearance probability. SK employs a 50-kiloton water Cherenkov detector in the underground mine with 2700 meter-water-equivalent (m.w.e.) overburden to detect ν_μ produced from a shower of cosmic rays. In particular, muons, pions, and kaons produced by cosmic-ray interactions decay to produce ν_μ. A simple survival probability formula

$$P(\nu_\mu \to \nu_\mu) \;=\; 1 - \sin^2(2\theta_{atm}) \sin^2\left(\frac{\Delta m^2_{atm} L}{4E}\right) \tag{2.16}$$

is used where $\theta_{atm} \approx \theta_{23}$ and $\Delta m^2_{atm} \approx \Delta m^2_{32} \approx \Delta m^2_{31}$. Obtained results are $\sin^2(2\theta_{atm}) > 0.82$

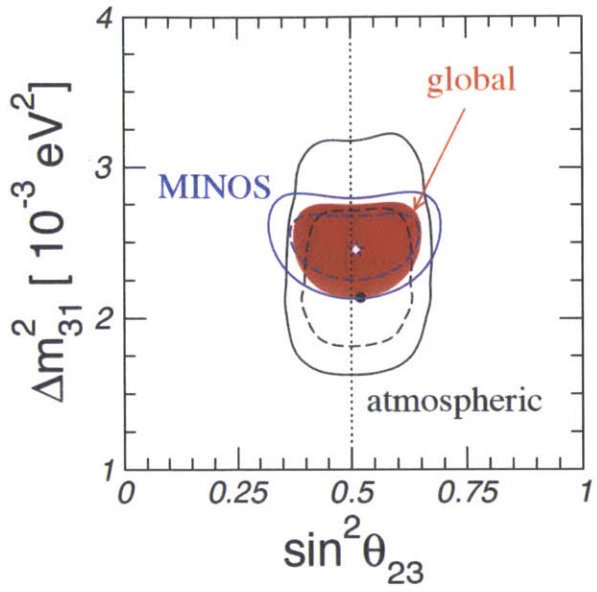

Figure 2-1: Atmospheric (black) and MINOS (blue) result. The combined result is shown in the red-shaded region for normal hierarchy at 90 % C.L. (dashed) and 3σ (solid) [5].

and $5 \times 10^{-4} < |\Delta m^2_{atm}| < 6 \times 10^{-3}\text{eV}^2$ at 90 % confidence level (C.L.) [18]. The MINOS experiment is a long-baseline accelerator experiment that uses a proton beam impinging on a fixed target to generate a high-purity ν_μ beam through charged pion and muon decay in flight. MINOS employs near and far detectors to compare an unoscillated flux from the near detector with the oscillated flux measured by the far detector. MINOS results are $|\Delta m^2_{atm}| = 2.32^{+0.12}_{-0.08} \times 10^{-3}\text{eV}^2$ and $\sin^2\theta_{atm} > 0.90$ at 90 % C.L., which is in a good agreement with the SK results [19]. The combined allowed region of atmospheric and MINOS experiments is shown in Fig. 2-1 taken from Ref. [5].

The third component of the PMNS matrix was precisely measured by the reactor phase of KamLAND and combined fits with SNO [6] and solar neutrinos measurements [7], as mentioned above. KamLAND is a kiloton spherical liquid scintillator detector sensitive to neutrinos from the Sun, the earth's internal crust, and numbers of nuclear reactors in Japan. In the measurement of $\bar{\nu}_e$ from nuclear reactors, the flux-weighted average distance to all nuclear reactors in Japan is about 180 km. The site is the same as the SK experiment, and hence comes with 2700 m.w.e. overburden. It uses an inverse beta decay (IBD) interaction, $\bar{\nu}_e + p \rightarrow e^+ + n$, to identify a signal by a two-fold coincidence of e^+ and neutron events.

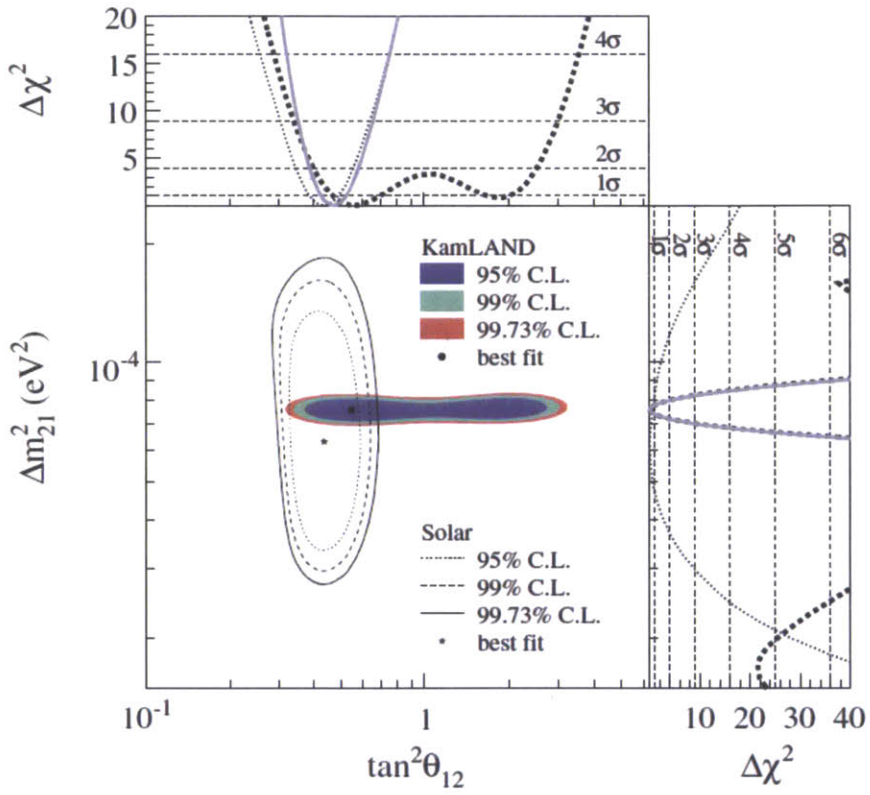

Figure 2-2: Allowed region for neutrino oscillation parameters from KamLAND and solar neutrino experiments [6, 7]. The figure is taken from Ref. [8].

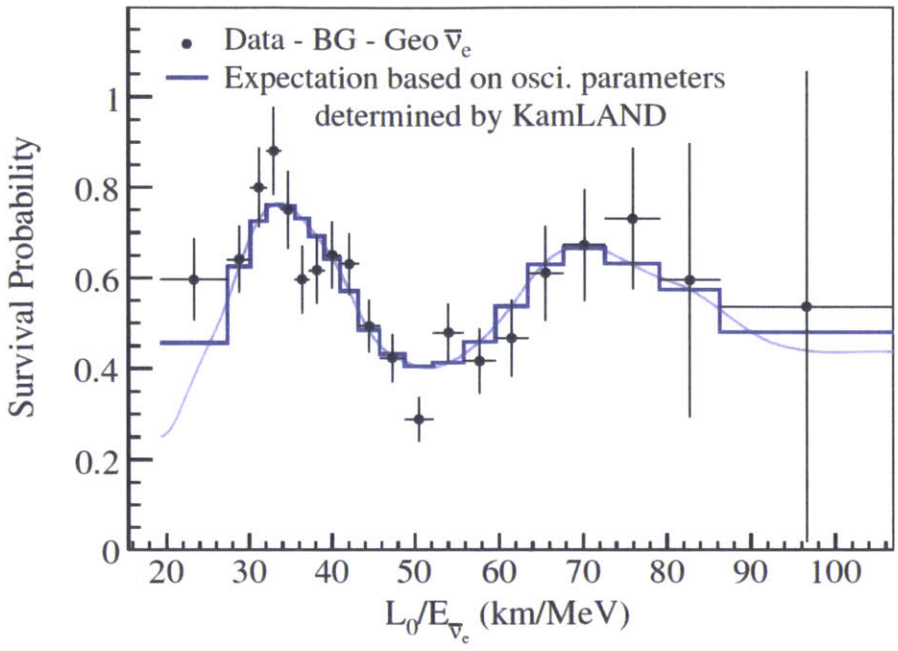

Figure 2-3: L/E oscillation pattern of $\bar{\nu}_e$ flux from nuclear reactors in Japan, taken from Ref. [8].

This measurement by KamLAND is that of a long-baseline reactor-based experiment, and is sensitive to θ_{12} and Δm^2_{21}. They measured $\Delta m^2_{21} = 7.58^{+0.14}_{-0.13}(\text{stat.})^{+0.15}_{-0.15}\text{syst.} \times 10^{-5}\text{eV}^2$ and $\tan^2\theta_{12} = 0.56^{+0.1}_{-0.07}(\text{stat.})^{0.10}_{-0.06}(\text{syst.})$ [8]. The combined fit result with solar neutrino measurements is shown in Fig. 2-2. KamLAND was able to show a beautiful oscillation shape as a function of L/E as shown in Fig. 2-3.

2.3 Measurement of θ_{13}

The last mixing parameter, θ_{13}, in the middle of the PMNS matrix in Eq. 2.2, has not been measured until very recently. While other mixing angles are found to be relatively large and hence can be measured and known well, this mixing angle is found to be small. The upper bound on this angle was long held by the CHOOZ [20] and Palo Verde [21] short-baseline reactor experiments. The measurement of θ_{13} is particularly important since a non-zero value gives rise to the possibility of measuring the CP-violating phase, δ, in the future. Two experimental approaches have been explored for the measurement of θ_{13}. The first is long-baseline accelerator-based ex-

periments, such as the MINOS and T2K experiments. Another is short-baseline reactor-based experiments, such as Daya Bay, RENO, and Double CHOOZ.

2.3.1 Past Short-Baseline Reactor Experiments

From February to July 1998, the CHOOZ experiment took place in the Chooz reactor complex, in the same location where the Double CHOOZ experiment is now taking place. It employed a single liquid scintillator detector to measure the oscillated $\bar{\nu}_e$ flux from the CHOOZ reactor core [20]. The signal detection channel was a two-fold coincidence of an IBD interaction with delayed neutron capture on gadolinium (Gd) in the liquid scintillator doped with a Gd compound. A neutron capture on Gd isotopes produces multiple gamma rays which have an energy sum of about $\approx$ 8 MeV, well above natural radioactive backgrounds. While this helps to identify the signal against backgrounds very well, at the time of the CHOOZ experiment, dissolving Gd chemical compounds in liquid scintillator was not trivial. As a result, the experiment suffered from degrading optical property of the scintillator over time. This resulted in a very short lifetime for the experiment. In addition, for a large part of the detector live-time, the Chooz reactor core was shut down. While this helped the experiment to accumulate a lot of data for background study, the signal data statistics remained low. The CHOOZ experiment result excluded the $\bar{\nu}_e$ disappearance at 90 % confidence level [20].

The Palo Verde experiment used Gd-doped liquid scintillator in a segmented detector design to observe a disappearance of $\bar{\nu}_e$ from the Palo Verde Nuclear Generating Station. For details, see Ref. [21]. The experiment ran from 1998 to 2000, and the final result also excluded the oscillation scenario at 90 % confidence level [21].

Given a small or possibly null value of θ_{13}, both experiments suffered greatly from systematic uncertainties, particularly from detector-related uncertainties and an error in the flux prediction. As described later, the modern reactor experiments use a design that greatly improves on these systematic uncertainties for their measurements of θ_{13}.

2.3.2 Hints From Long-Baseline Accelerator Experiments

More recent knowledge about the θ_{13} value comes from the MINOS and T2K experiments. These long-baseline (that is, large L/E) accelerator-based experiments search for ν_e appearance in a ν_μ beam. These experiments are somewhat complicated because the analysis requires fits to multiple fundamental parameters. Through $\nu_\mu \rightarrow \nu_e$, they are sensitive to matter effects due to forward electron scattering along the long path that the neutrinos travel. The sign of the matter effect is different for a "normal hierarchy," where the small mass splitting is at the lowest absolute mass offset, and the "inverted hierarchy," where the small mass splitting has a higher absolute mass offset than the large mass splitting. These experiments are also sensitive to the CP-violating parameter δ. The data must be fit with these additional parameters simultaneously with θ_{13}

Although the main goal of the MINOS experiment, as described above, was a measurement of θ_{23} and Δm^2_{23}, they have also analyzed the probability of $\nu_\mu \rightarrow \nu_e$ oscillation, which is sensitive to θ_{13}. In 2011, MINOS published a combined θ_{23} and θ_{13} measurement, $2\sin^2\theta_{23}\sin^2(2\theta_{13}) = 0.041^{+0.041}_{-0.031}$ $(0.079^{+0.071}_{-0.053})$ for the normal (inverted) hierarchy at $\delta = 0$. The result rejects the null θ_{13} hypothesis at 89 % C.L. [14]. The MINOS experiment has formally ended, although a future experiment, MINOS+ will begin soon.

The T2K experiment in Japan also uses protons to produce a ν_μ beam, but is designed with the primary purpose of measuring θ_{13} and CP-violating phase δ. Shortly after the MINOS result, the T2K result of $0.03(0.04) < \sin^2(2\theta_{13}) < 0.28(0.34)$ at 90 % C.L. for the normal (inverted) mass hierarchy at $\delta = 0$ [13] was published. Both T2K and MINOS results were exciting because they provide a hint for a non-zero θ_{13} value for the first time, although neither of them is yet sensitive enough to deliver a discovery. The T2K experiment continues to accumulate more data and its sensitivity will increase in the future.

2.3.3 Recent Results From Short-Baseline Reactor Experiments

A number of short-baseline reactor-based experiments have been proposed, and three of them have been built in the world: Double CHOOZ in France, Daya Bay in China, and RENO in South Korea. These experiments share a similar detector design, and the Double CHOOZ

detailed design is described in Sec. 3.2. They all use Gd-doped liquid scintillator with a greatly improved stability of optical properties compared to the CHOOZ experiment. Also, given a hint of a very small or null value for θ_{13}, these experiments are designed to improve the sensitivity from the CHOOZ experiment by employing identical detectors at two locations, the near and far, to measure oscillated and unoscillated $\bar{\nu}_e$ flux from the nearby reactor cores. The expected sensitivity of these experiments is on the order of $\sin^2(2\theta_{13}) \approx \mathcal{O}(10^{-2})$ [29].

The advantage of reactor-based experiments is that the measurement of θ_{13} from reactor-based experiments does not require a knowledge of the CP-violating phase δ, and hence their results help to constrain δ when combined with the accelerator-based experiments introduced previously. Also, due to the short baseline ($\sim$ 1 km), the mass hierarchy does not affect the interpretation of the measurement.

Despite a very similar detector design among them, it is worth noting that there are differences among the currently running reactor-based experiments. First of all, the near detector to measure an unoscillated flux is not yet running in Double CHOOZ. The results from Double CHOOZ, therefore, are based on data only from the far detector, which measures an oscillated flux. The experiment compares the measured flux with a prediction that is a combination of reactor core simulation and a measurement by the Bugey-4 experiment on the $\bar{\nu}_e$ flux at very short distance. This results in a larger systematic uncertainty in the θ_{13} measurement compared to Daya Bay and RENO, as they have both the near and far detectors under operation.

Another difference between Double CHOOZ and the other two experiments is an extraction method of θ_{13} from the prediction. Daya Bay and RENO so far perform the rate-only measurement, which simply compares the number of $\bar{\nu}_e$ event candidates observed to the prediction. Double CHOOZ, on the other hand, performs a combined rate+shape measurement. This method uses the shape information of a e^+ energy spectrum from an IBD interaction, which is close to an interacting $\bar{\nu}_e$ energy spectrum. Together with the background rate and energy spectrum prediction, rate+shape analysis is believed to provide a better understanding of backgrounds.

The first measurement of θ_{13} was from the Double CHOOZ experiment, and was presented in November 2011 at the LowNU conference in South Korea, followed by a publication in

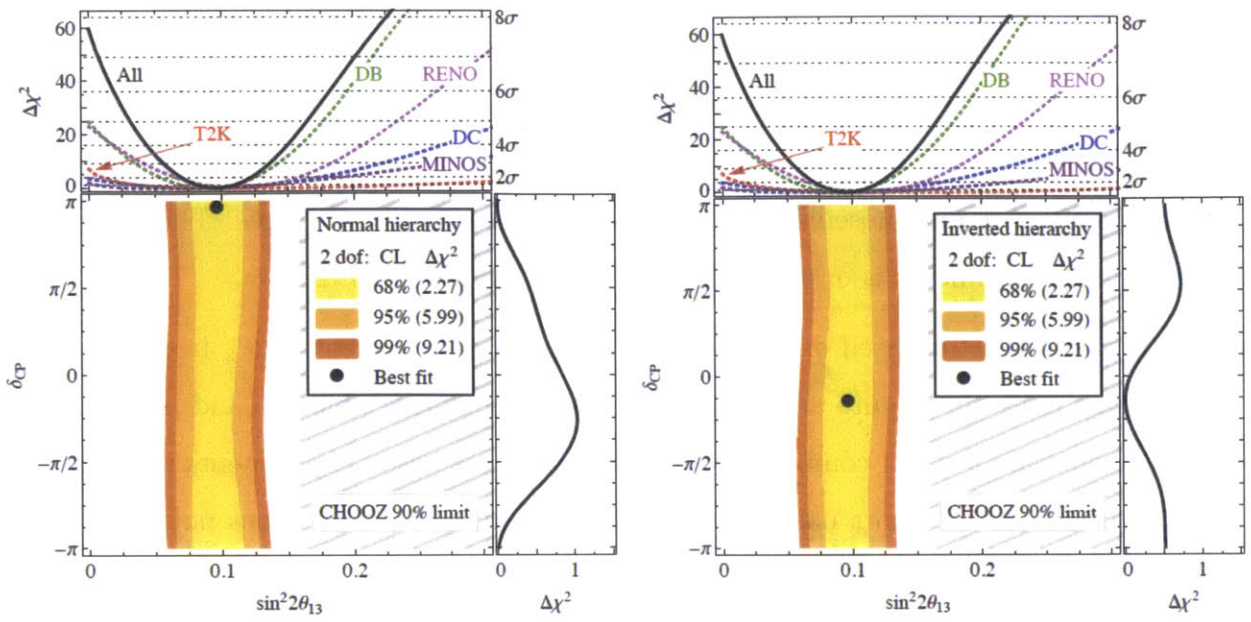

Figure 2-4: Allowed region of $\sin^2(2\theta_{13}) - \delta$ parameter space for T2K, MINOS, DC, Daya Bay (DB), and RENO combined fit from Ref. [9]. The left figure is normal and the right figure is for an inverted hierarchy.

early 2012 with the value $\sin^2(2\theta_{13}) = 0.086 \pm 0.041(\text{stat}) \pm 0.030(\text{syst.})$ [11]. This represented another hint of a non-zero θ_{13} value, which was in good agreement with preceding results from MINOS and T2K. Daya Bay and RENO published their results by summer 2012 with good agreement among all measurements [2, 3]. A null value for θ_{13} is now excluded at more than 5 σ significance level. A combined fit using the accelerator experiments, MINOS and T2K, and the first result from all reactor experiments was found to exclude the null θ_{13} hypothesis at 7.7σ (see Fig. 2-4). So far the most sensitive measurement held by one experiment is from the Daya Bay result in October 2012 with the value of $\sin^2(2\theta_{13}) = 0.089 \pm 0.010$ (stat.) ± 0.005 (syst.). A summary of all measurements is shown in Fig. 3-3.

2.3.4 Future Prospects for Double CHOOZ

The Double CHOOZ experiment is currently finishing construction of the near detector while accumulating more statistics with the far detector. The current goal in the standard analysis procedure is to reduce the statistical uncertainty in the analysis.

In parallel, the experiment sought the possibility of another detection channel: an identification of an IBD interaction using delayed neutron capture on hydrogen. As described in more detail in later sections, this new method has several attractive features. First of all, it provides a data sample statistically independent from those used in the standard assessment. A result obtained from this analysis, therefore, serves as a nice cross-check of the already published result. Secondly, with the Double CHOOZ detector, this detection channel provides double the signal statistics of the standard method. The author of this thesis made a leading contribution to the development of this analysis method, which is the focus of this manuscript.

The near future prospect of the experiment is a combined analysis of θ_{13} the main method, which uses neutron capture on Gd isotopes, and the new approach, which uses neutron capture on hydrogen. The study reported in this document proved the feasibility of the new approach, and the collaboration is currently working to improve the developed method and to combine it with the standard analysis result to improve the sensitivity of the θ_{13} measurement.

Chapter 3

Double CHOOZ Experiment

3.1 Goal of the Double CHOOZ Experiment

The Double CHOOZ experiment takes place inside the Chooz reactor complex in Ardenne (Fig. 3-1), France. The experiment measures the small mixing angle θ_{13} from the disappearance of $\bar{\nu}_e$ produced by a fission process in the reactor cores. The preceding CHOOZ experiment took place at the same site, and long held the upper limit on the value of θ_{13} [20]. The Double CHOOZ detector design is greatly improved to achieve a higher sensitivity to this small mixing angle.

3.1.1 Two Detectors Concept

The Chooz plant operates two pressurized water reactors (PWR), B1 and B2, each with 4.25 giga-watt thermal power (GW_{th}). A nuclear fission process involves multiple beta decays which emit $\bar{\nu}_e$. The reactor cores, therefore, are isotropic sources of $\bar{\nu}_e$. More details about the reactors are found in chapter 8. Emitted $\bar{\nu}_e$s travel in space following the survival probability

$$\mathrm{P}(\bar{\nu}_e \longrightarrow \bar{\nu}_e) \;=\; 1 - \sin^2 2\theta_{13} \sin^2 \left(1.267 \Delta m^2_{31}[\mathrm{eV}^2] \, \frac{L[\mathrm{km}]}{E[\mathrm{GeV}]} \right) + \text{small terms} \qquad (3.1)$$

where the small terms are neglected. The value of $\Delta m^2_{31} = 2.32 \times 10^{-3} \mathrm{eV}^2$ is used based on the result of the MINOS experiment [30]. For $\bar{\nu}_e$ with energy of 3 MeV, the oscillation probability

Figure 3-1: The Chooz reactor complex in Ardennes, France

reaches its first maximum at L=1.6 km. The expected number of signals, N_{exp}, at the detector can be expressed as

$$N_{exp}(t, E_{\bar{\nu}_e}) = \sum_{\text{Reactor}}^{\text{B1,B2}} \frac{N_{target} \cdot \epsilon}{4\pi L^2} \phi_{\bar{\nu}_e}(t, E_{\bar{\nu}_e}) \sigma(E_{\bar{\nu}_e}) \tag{3.2}$$

where N_{target}, ϵ, L, $\phi_{\bar{\nu}_e}$, σ are the number of targets in the detector, detection efficiency, distance to the reactor core, $\bar{\nu}_e$ flux, and an interaction cross section as a function of time and $\bar{\nu}_e$ energy in respective order.

In order to measure the disappearance of $\bar{\nu}_e$ with high precision, the experiment plans to employ two detectors: the near and the far detectors. The near detector is going to be located close to the reactor cores to measure an unoscillated $\bar{\nu}_e$ flux. The far detector is located at the first maximum of the oscillation probability, and hence is optimized to observe the disappearance of $\bar{\nu}_e$ flux. Both detectors take measurements underground in order to suppress cosmogenic backgrounds. The description of these detector sites is summarized in Tab. 3.1.

The two detectors share as much identical design and materials as possible. A comparison

	Ditance to B1	Distance to B2	Overburden
Far	1114.6 ± 0.1 m	997.9 ± 0.1 m	300 m.w.e.
Near	465 m	351 m	120 m.w.e.

Table 3.1: Location of the Double CHOOZ near and far detectors. Values for the near detector will be measured once the installation is finished.

of $\bar{\nu}_e$ flux at the near and far site with an identical target cancels a systematic uncertainty on ϕ and σ, and minimizes that of N_{exp} and ϵ in Eq. 3.2. This comparison technique improves the experiment's sensitivity to the value of θ_{13} greatly compared to a single-detector design. Since two detectors cannot be completely identical, however, trivially a residual difference between the two detectors remains as a systematic uncertainty.

Single-Detector Phase

Double CHOOZ finished the installation and commissioning of the far detector in April 2011 and started data taking. The most recent publication is based on data accumulated up to March 15th, 2012, from the far detector only [1]. The near detector is expected to start data taking in the middle of 2014. Until the near detector comes online, the analysis compares the observed $\bar{\nu}_e$ flux at the far detector to an expected $\bar{\nu}_e$ flux from a dedicated Monte Carlo (MC) simulation (see chapter 8) to quantify a disappearance from the expectation. This method is dominated by systematic uncertainty coming from the flux prediction, and hence affects the sensitivity to the θ_{13} measurement. The measured value of θ_{13}, however, is found to be large enough for the current sensitivity of the single detector assessment [1]. This allows Double CHOOZ to measure the value of θ_{13} by the analysis with the far detector only.

3.1.2 Experiment Signal

The signal detection channel in this experiment is an IBD interaction,

$$\bar{\nu}_e + p \longrightarrow e^+ + n. \tag{3.3}$$

The detector fiducial volume contains a proton-rich liquid scintillator which serves as a target for $\bar{\nu}_e$ and a detection medium for IBD interactions. The IBD signal is identified by a two-fold coincidence signal of e^+ annihilation with e^- followed by a delayed gamma emission from a neutron capture. A search for a coincidence signature effectively suppresses the natural radioactive background, and it has been used in many experiments since the first detection of $\bar{\nu}_e$ by Reines and Cowan [31].

The energy of the prompt event from a coincidence pair carries the energy $E_{\text{prompt}} = E_{e^+} + m_e$ where E_{e^+} is the energy of a positron and m_e is an electron mass. The interaction threshold can be computed from 4-momentum algebra,

$$E_{\bar{\nu}_e}^{thres} = \frac{(m_n + m_e)^2 - m_p^2}{2m_p} = 1.813 \text{ MeV}. \tag{3.4}$$

In the lab frame, the original $\bar{\nu}_e$ energy above $E_{\bar{\nu}_e}^{thres}$ can be expressed as

$$E_{\bar{\nu}_e} = \frac{m_n^2 - m_p^2 - m_e^2 + 2E_{e^+}m_p}{2m_p - 2E_{e^+} + 2\cos\theta\sqrt{E_{e^+}^2 - m_e^2}} \tag{3.5}$$

where m_n, m_p, θ are the neutron mass, proton mass, and angle of the e^+ momentum with respect to the $\bar{\nu}_e$ momentum. In the first order approximation, $E_{e^+}/m_p \approx o(10^{-2})$ because E_{e^+} is below $E_{\bar{\nu}_e}$ from energy conservation law, and $E_{\bar{\nu}_e}$ is typically below 10 MeV as $\bar{\nu}_e$ is produced in nuclear beta decays. This allows a simple approximation,

$$E_{\bar{\nu}_e} = E_{\text{prompt}} + 0.8 \text{ MeV}. \tag{3.6}$$

Consequently, the shape of E_{prompt} is the same as $E_{\bar{\nu}_e}$ with a constant shift of 0.8 MeV and a convolution of the detector effect. The E_{prompt} shape information is, therefore, useful to identify $\bar{\nu}_e$ from the backgrounds with a different prompt energy shape. An expected $\bar{\nu}_e$ energy spectrum is a convolution of the original $\bar{\nu}_e$ energy spectrum from nuclear decays and an IBD cross section as a function of energy (see Eq. 3.2). An example spectrum shape is shown in Fig. 3-2.

In Double CHOOZ, θ_{13} is measured in three ways: (1) the rate-only analysis uses the total

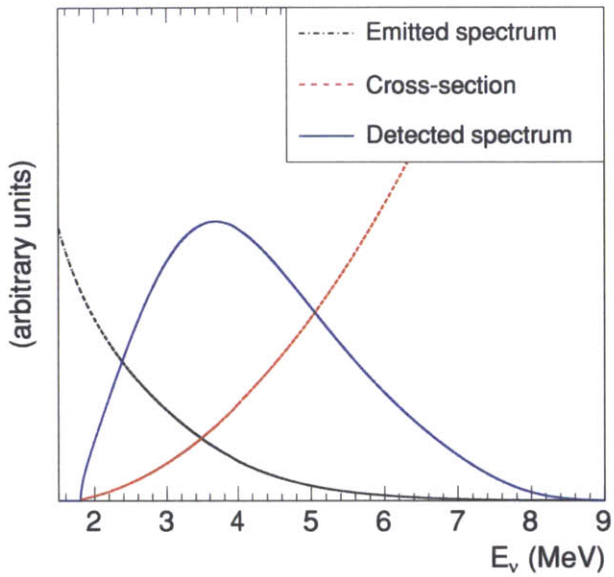

Figure 3-2: A simulated energy spectrum shape of $\bar{\nu}_e$ which is produced from a ^{235}U fission chain and undergoes an IBD interaction (blue). An IBD interaction cross section as a function of energy (red-dashed) shown with the original energy spectrum of $\bar{\nu}_e$ emitted from the fission process (black-dashed). The figure is taken from Ref. [10]

signal event statistics over energy; (2) the shape-only analysis uses the signal and predicted background E_{prompt} shape difference; (3) the rate+shape analysis uses information from both (1) and (2) to determine the best value of θ_{13}.

3.1.3 Detection of Delayed Neutron Capture

Liquid scintillators used in the detector consist of hydrogen and carbon atoms, and a thermal neutron is primarily captured by a hydrogen nucleus in a pure liquid state. In order to enhance a signal detection efficiency, however, the target liquid is doped with Gd isotopes, which have a neutron capture cross section orders of magnitude higher than hydrogen. In addition, a neutron capture on a Gd nucleus emits multiple gamma rays with total energy of $\sum E \approx 8$ MeV. Compared to a neutron capture on a hydrogen nucleus, which emits a single gamma ray with $E \approx 2.2$ MeV, capture on Gd nuclei produces a very distinct signal from natural radioactive backgrounds that arise below 4 MeV. The high-precision θ_{13} measurement to date searches for IBD signal using neutron capture on Gd nuclei, and is referred to as "Gd analysis" in this

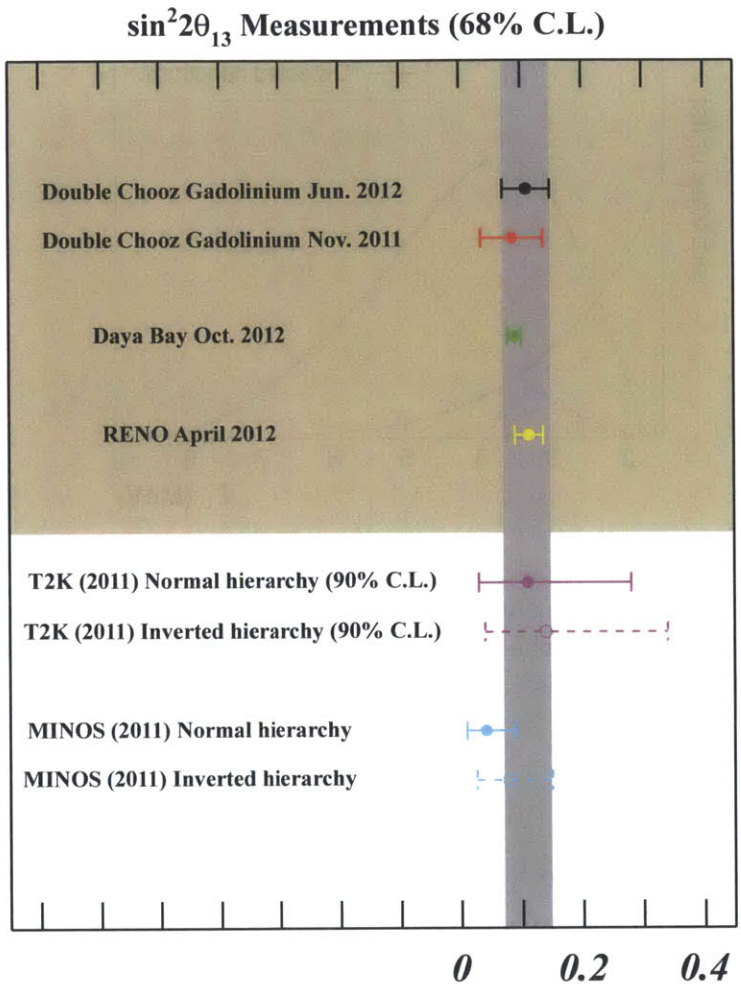

Figure 3-3: Measurements from various θ_{13} measurements [11, 1, 12, 3, 13, 14].

document. The Gd analysis method has led to a number of successful results from the reactor-based θ_{13} measurements in the last year [1, 2, 3, 11, 12]. All results, including accelerator experiments [14, 13], are found to be in a good agreement as shown in Fig. 3-3.

While the Gd analysis is used in all reactor experiments as the main analysis method, it is possible to use neutron capture on hydrogen for signal detection. This method, referred to as "hydrogen analysis," provides double the signal statistics with a data sample that is statistically

independent from that of the Gd analysis. The nature of backgrounds and detector-related systematic uncertainties are also very different from the Gd analysis. These features make the hydrogen analysis a unique, excellent cross-check to the Gd analysis. Further, the combination of the two analysis results may allow a better constraint on the value of θ_{13}. This manuscript focuses on the hydrogen analysis, which is a unique attempt only made in the Double CHOOZ experiment.

3.2 Double CHOOZ Detector

The Double CHOOZ detector consists of three sub-systems: (1) the Inner Detector (ID) for detecting IBD signal; (2) and surrounding the ID is the Inner Veto (IV) for rejection of muons and radioactive backgrounds sourced outside; (3) finally, the Outer Veto (OV) is a scintillator-strip-based muon tracking system installed above the IV system. The overview of these components is shown in Fig. 3-4. This section describes each detector component in detail.

3.2.1 Inner Detector (ID)

The ID system consists of three sub-volumes separated by cylindrical vessels. The innermost volume is the Target, an acrylic vessel of 10.3 m^3 with a radius of 1150 mm and a height of 2458 mm. The acrylic vessel of the Target volume has an extended cylindrical path on the top, called the "chimney," that reaches outside the ID. The chimney is used for filling the volume with liquid as well as for deployment of the calibration source which is described later in this chapter.

The Target is filled with gadolinium loaded liquid scintillator (Gd-LS) which consists of 80 % n-Dodecane and 20 % Phenylxylylethane (PXE). Dissolved in the Gd-LS are two fluors, PPO (7 g/L) and bis-MSB (20 mg/L), as the wavelength shifters, and Gd-β-diketonate (Gd-BDK) compound (1 g/L Gd). Primary scintillation light produced by excitation of electrons in aromatic molecules is efficiently absorbed by PPO and re-emitted at the higher wavelength (see Fig. 3-5. Then bis-MSB further shifts the re-emission photon from PPO into a higher wavelength, 430 nm to 500 nm. The attenuation length of photons at this wavelength is about

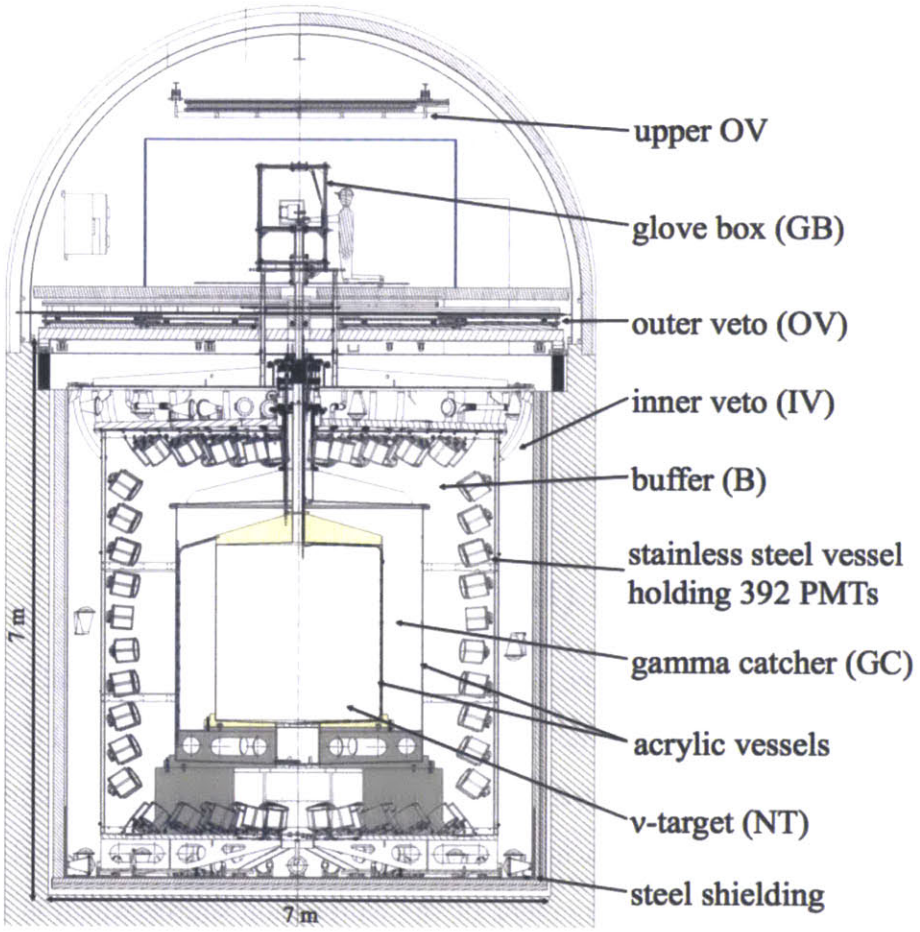

Figure 3-4: Detector overview including the ID, IV, and the OV sub-systems [1]. The ID consists of the Target, γ-catcher, and Buffer shown in the figure.

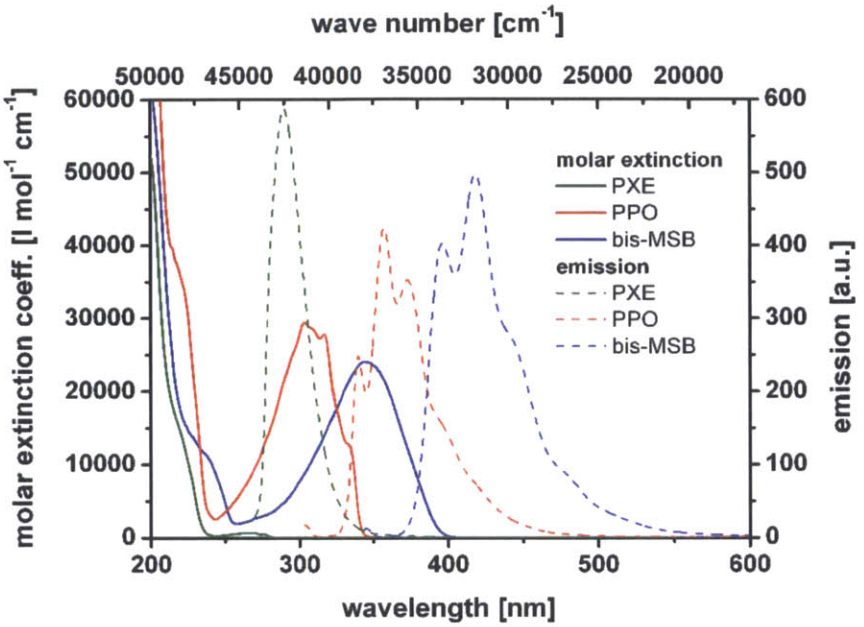

Figure 3-5: Absorption and emission spectra of aromatic components in the Target and γ-catcher liquid scintillators [15].

5 m in the Gd-LS, and hence the Target is virtually transparent for these photons while they travel toward the ID photo-detectors that are sensitive to this wavelength. Isotopes of Gd nuclei, in particular ^{155}Gd and ^{157}Gd, have a very high cross section of thermal neutron capture followed by a cascade of multiple gamma rays from a nuclear de-excitation. The emitted gamma rays carry a summed energy of $\sum E_\gamma \approx 8$ MeV, which makes it easy to distinguish them from natural radioactive backgrounds at lower energy (see Fig. 7-1 in Sec. 7.1 for an energy spectrum of natural radioactive backgrounds). The Target liquid contains about 1 g/L of Gd, and about 86 % of the neutrons in the Target are captured by Gd nuclei, while the rest of the neutrons are typically captured on hydrogen nuclei [1]. This volume is, therefore, designed to enhance a delayed neutron capture signal from IBD interactions inside with low natural radioactive backgrounds.

The γ-catcher is a larger acrylic cylindrical vessel with a radius of 1708 mm and a height of 3572 mm. This volume contains the Target vessel, and it is filled with 22.3 m^3 of liquid scintillator without Gd. The γ-catcher liquid consists of 66 % mineral oil (Shell Ondina909), 30 % n-Dodecane, and 4 % PXE, and is made with PPO (2 g/L) and bis-MSB (20 mg/L)

as wavelength shifters. The main purpose of having the γ-catcher, as its name suggests, is to efficiently convert the energy of cascade gammas from a neutron capture on Gd into scintillation light. The different composition from the Target liquid is motivated by two reasons: (1) to realize a uniform light yield through the Target and γ-catcher volumes for high photon collection efficiency homogeneity while (2) matching the density for mechanical stability of the acrylic vessel. Neutrons produced in the γ-catcher volume are primarily captured on hydrogen nuclei. A combined Target and γ-catcher region becomes, therefore, the fiducial volume for the hydrogen θ_{13} analysis.

Outside the γ-catcher is the Buffer with a radius of 2758 mm and a height of 5675 mm. The Buffer tank is made of a low-activity stainless steel with a constraint of 15 mBq/kg[29], and is filled with 98 m^3 of non-scintillating, transparent liquid mixture of 46.5 % n-alkanes (Cobersol C70) and mineral oil (Shell Ondina917). On the inner wall of the Buffer tank, 390 low-activity Hamamatsu R7081 10-inch photo-multiplier tubes (PMT) are installed [32]. The high voltage is applied such that the gain of the PMTs is set approximately to 10^7. Scintillation light produced in the Target and γ-catcher region is detected by these ID PMTs. The Buffer region, as its name suggests, shields the liquid scintillator region away from the radioactive backgrounds that are sourced from the Buffer tank and PMT materials. The stainless-steel tank optically isolates the inside from the external world. The Target, γ-catcher, and the Buffer constitute one $\bar{\nu}_e$ detector, the ID.

3.2.2 Inner Muon Veto System (IV)

The IV is a cylindrical steel-made tank and contains the Buffer tank. It is filled with 87 m^3 of liquid scintillator consisting of 50 % linear alkyl benzene (LAB) and 50 % n-alkanes (Cobersol C70). PPO (2 g/L) and bis-MSB (20 mg/L) are dissolved in it as wavelength shifters. There are 78 Hamamatsu R1408 8-inch PMTs to detect scintillation light. The volume serves as a cosmic muon veto as well as a shield toward fast neutrons sourced from outside. Outside the tank is a 15 cm thick demagnetized steel shield to suppress external gamma rays.

3.2.3 Outer Muon Veto System (OV)

The Outer Veto (OV) system consists of multi-layer scintillator strips in order to perform a precise muon track reconstruction. Modules of 64 strips cover the top of the steel shield just outside the IV. Modules are installed such that strips are aligned along the x and y axis to identify a crossing point of a muon and the modules. One purpose of the OV is to sample cosmogenic events in the ID or IV with a precisely reconstructed track of the parent muon. Also, it can be used to simply veto a stopping-muon background (see Sec. 7.4), which is a two-fold coincidence background caused by a low-energy muon that enters the ID.

The coverage of the OV modules on the horizontal plane of the detector is not complete as there is a space for the Target chimney to come outside. Modules cannot be installed in this space. However, the horizontal coverage extends to the outer circumference of the IV, and hence it can track muons that do not enter the ID or IV but pass through nearby rocks. This is expected to give an extra handle on sampling cosmogenic backgrounds produced outside the sensible volume. The upper layer of the OV is recently installed and it is attached to the ceiling. Muons that go through both the upper and the lower layer are expected to be reconstructed with great precision.

About 69 % of the DATA sample used in this analysis contains data with the lower layer OV. The upper layer was installed in summer 2012, and it is not included in this analysis. The OV has its own trigger system, and data is later synchronized based on timestamps in the data processing stream. The main use of the OV data in this analysis is to veto stopping muons, which is described in Sec. 7.4.

3.2.4 Trigger and Readout Electronics

Figure 3-6 shows the ID and IV readout system diagram. A signal waveform from the PMT is first decoupled from the high voltage (HV) DC component along a cable approximately 40 m long, and sent to the front-end electronics (FEE) where the signal is amplified. The FEE outputs summed pulses to the three trigger boards (two for ID PMTs, and one for IV PMTs) that is connected to the master trigger board. The master trigger board signals the readout electronics to store signal information when a trigger condition in any of three slave trigger

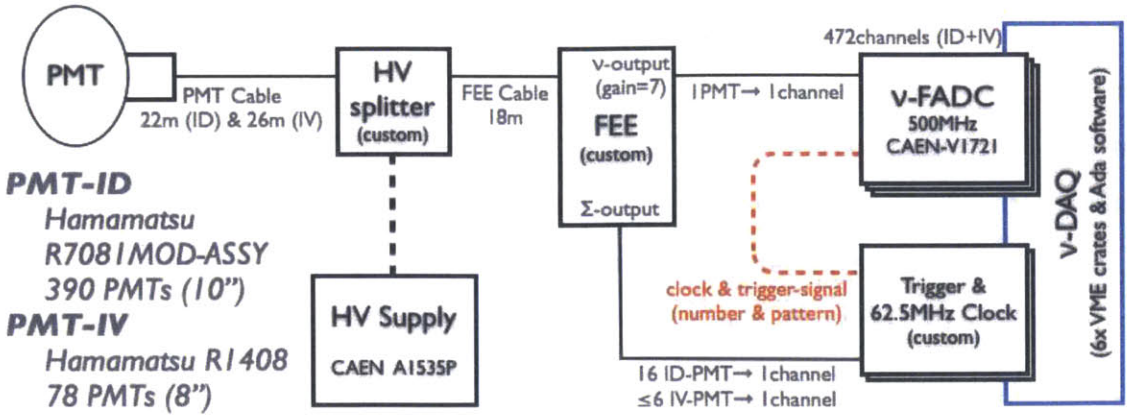

Figure 3-6: The readout diagram for the ID and IV electronics from the PMTs to the Data Acquisition (DAQ) [1].

boards is satisfied. The custom CAEN-V1721 Flash Analog to Digital Converters (FADC) continuously receive a channel-by-channel waveform from the FEEs, digitize them, and store them in 1024 rotating ring buffers. Upon receiving a signal from the master trigger board, the FADCs start writing the digitized waveform into a new rotating buffer, and the data acquisition (DAQ) stores signal information in the last buffer into a binary file on the onsite computers' disk.

Trigger System

The FEE outputs twelve groups of sixteen ID PMTs' summed pulses and one group of three ID PMTs' to each ID trigger board. One ID trigger board, therefore, handles 195 PMT channels evenly distributed over the detector. There are two trigger conditions for the ID trigger to be assigned. The first is a combination of a $\geq$2 PMT group hit multiplicity and the total signal charge exceeding the neutrino threshold. The second is a pre-scaled trigger that reads out 1/1000 of signal that has a charge sum of above $\approx$ 200 keV approximately. The IV trigger board receives five groups of three to six PMTs' summed pulses from the FEE. The IV trigger threshold is set to $\approx$ 4 MeV, which is about the end-point energy of gamma rays from natural radioactivity. A typical ID and IV combined trigger rate during the data taking was about 130 Hz, as shown in Fig. 3-7. During regular physics data taking, the master trigger board

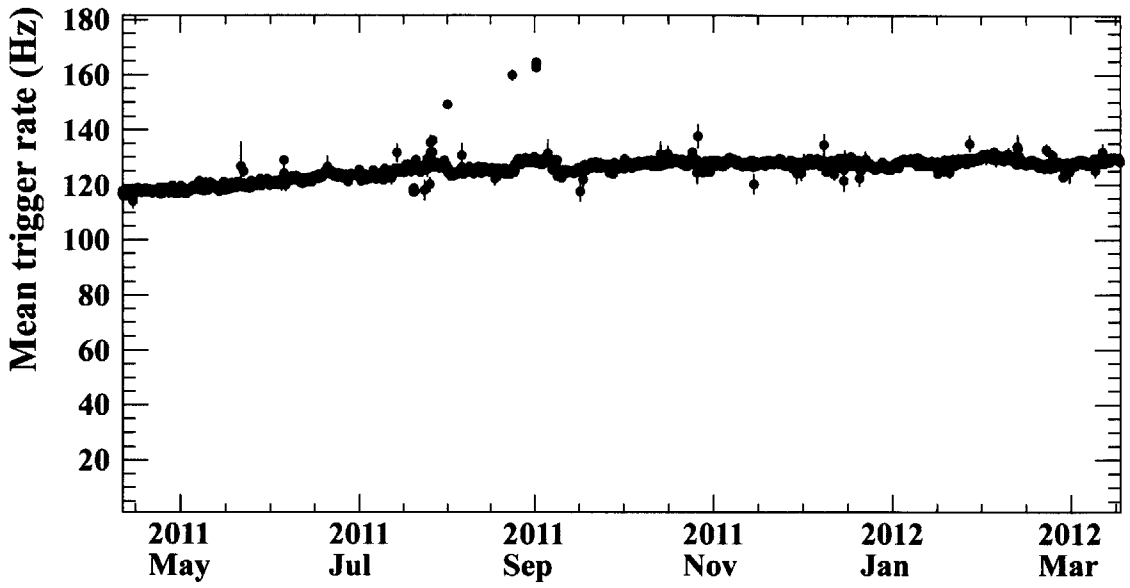

Figure 3-7: The ID and IV readout trigger rate stability during the data taking. A typical trigger rate was found to be around 130 Hz.

generates also an external trigger at a fixed rate of 1 Hz. This trigger is used to calibrate the baseline of a digitized waveform.

Signal Digitizer and Readout

The FADCs take amplified channel-by-channel PMT signal from the FEEs, and digitize at the rate of 500 MHz at 256 steps of digitized unit of current (DUI). They continuously write the digitized waveforms into rotating ring buffers. When FADCs receive a trigger signal, they changes to a new, initialized buffer while the DAQ reads out stored waveform from the last buffer. Whether the trigger signal is issued from the ID or IV trigger board, all of the ID and IV PMT waveforms are stored. The FADCs are configured to set the baseline at 210 DUI and a pulse start time of around $\approx$ 100 ns for events recognized by ID trigger boards. A typical single photo-electron (SPE) waveform has an amplitude of 6 DUI. Figure 3-8 shows an example SPE-like digitized waveform from physics data taking and a large amplitude waveform from a muon-like event.

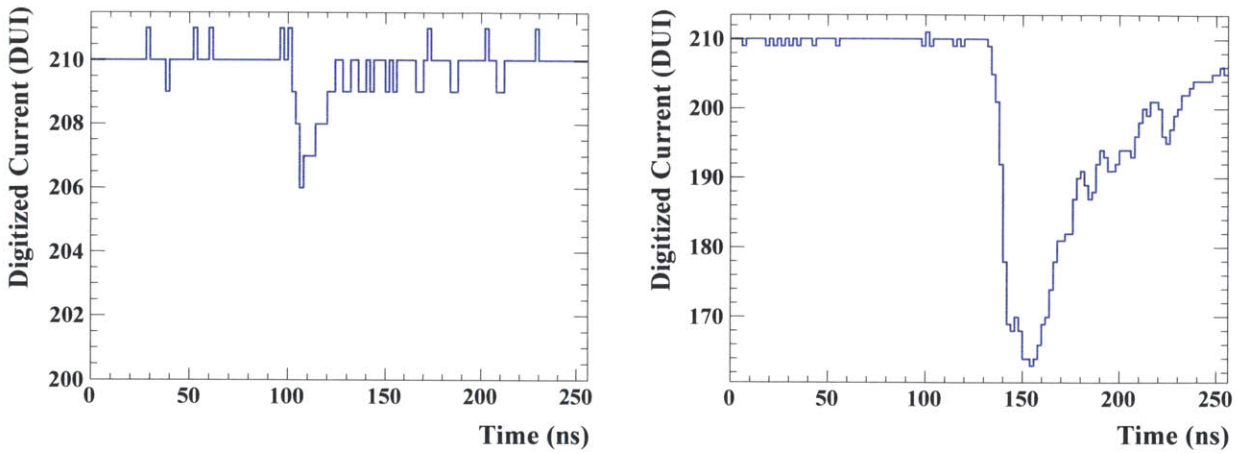

Figure 3-8: Left: a single photo-electron-like waveform from physics data. Right: a large amplitude waveform from a muon-like event. The digitization is done at 500 MHz.

Trigger Efficiency and Readout Dead Time

The trigger threshold is determined by counting the number of pre-scaled triggers that pass the neutrino threshold. Figure 3-9 shows the trigger efficiency curve. The trigger efficiency is 100 % with a negligible error at 0.7 MeV, which is the neutrino analysis energy threshold (see Sec. 5.3.1). There exists 16 ns dead time after each trigger due to the DAQ readout. Given the average trigger rate of 125 Hz, this corresponds to 2 ppm fraction of the dead time, which is a negligible correction to the detector live-time, and hence not included in the neutrino analysis.

3.2.5 Measurement of the Number of Protons

The number of protons in the Target is measured with high precision based on the liquid weight measurement during the detector filling phase and hydrocarbon fraction measurement. The uncertainty is determined to be 0.3 % [1].

For the γ-catcher volume, on the other hand, a dedicated weight measurement is not available as the region was not initially planned to be used for θ_{13} analysis. Instead, an acrylic vessel geometry survey is used to estimate the volume, and the weight is estimated in combination with the mass density measurement of the γ-catcher liquid. This method carries 0.3 % systematic uncertainty. A hydrocarbon fraction measurement carries a larger systematic uncertainty

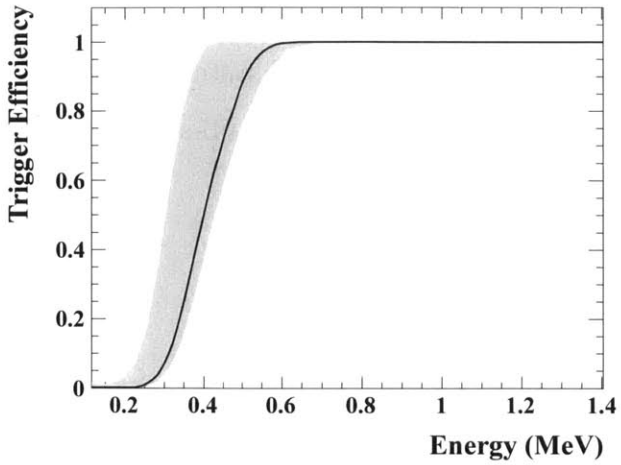

Figure 3-9: Trigger efficiency curve as a function of visible energy (for the definition of visible energy, see Sec. 4.3).

Subvolume	**# Protons**	**MC corr.** [%]
Target Liquid	$6.747 \cdot 10^{29} \pm 0.3\%$	-
GC Liquid	$1.582 \cdot 10^{30} \pm 1.0\%$	-
Buffer Liquid	$6.799 \cdot 10^{30} \pm 1.0\%$	-
Target Tank	$2.08 \cdot 10^{28} \pm 23\%$	39%
GC Tank	$6.96 \cdot 10^{28} \pm 3\%$	-

Table 3.2: A summary of number of protons estimations in each detector sub-volumes shown with the MC correction factors

of 1 % compared to the Target liquid. Combining these independent components, the systematic uncertainty on the number of target protons in the γ-catcher volume is estimated to be 1.04 %.

Table 3.2 summarizes the number of protons in Target, γ-catcher, and also other sub-volumes such as acrylic vessels and Buffer liquid. There is an MC correction factor only for the Target tank, and this is due to a simplified geometry implementation in the MC framework. Though the relative correction is as large as 39 %, this does not make a significant contribution to the overall correction as the number of protons in the Target tank geometry is much smaller compared to the LS volume.

3.3 Calibration

The detector calibration is important to achieve a small systematic uncertainty on the measurement of θ_{13}. This section briefly describes the calibration sub-systems used to better understand the Double CHOOZ far detector.

3.3.1 Calibration Sub-systems

LED Calibration System

The ID is equipped with the Inner Detector Light Injection (IDLI) system. IDLI is an LED fiber, installed at different positions on the inner Buffer wall, that generates light pulses illuminating the PMTs at multiple wavelengths (385, 425 and 470 nm) [1]. A long-wavelength LED pulse, for which the liquids are transparent, is used to calibrate the PMT timings. A high intensity short-wavelength LED pulse is used to generate scintillation light for gain calibration. Details of the usage of these light injection systems are described along with calibration methods in following sections.

Source Deployment System

In addition to the light injection systems, there are two running calibration devices, the Z-Axis and Guide-Tube. These systems can deploy a radioactive source whose properties are very well known, and allow a detailed study of the detector response. A toy drawing of these systems is shown in Fig. 3-10. Deployed radioactive sources are summarized in Tab. 3.3. The analysis described in chapter 4 and 6 uses data taken with these calibration sources/devices.

Z-Axis Calibration System

Z-Axis is a simple, automated "fish-line" system which can be used to deploy a calibration source along the vertical axis inside the Target vessel. The source is attached at the end of the wire along with a weight, and lowered by a micro-step motor which unwinds a spool of wire at the other end. The motor and the wire spool are installed inside the Glove Box, which sits above the OV (see Fig. 3-4). The inside of the Glove Box is flushed by nitrogen gas to keep

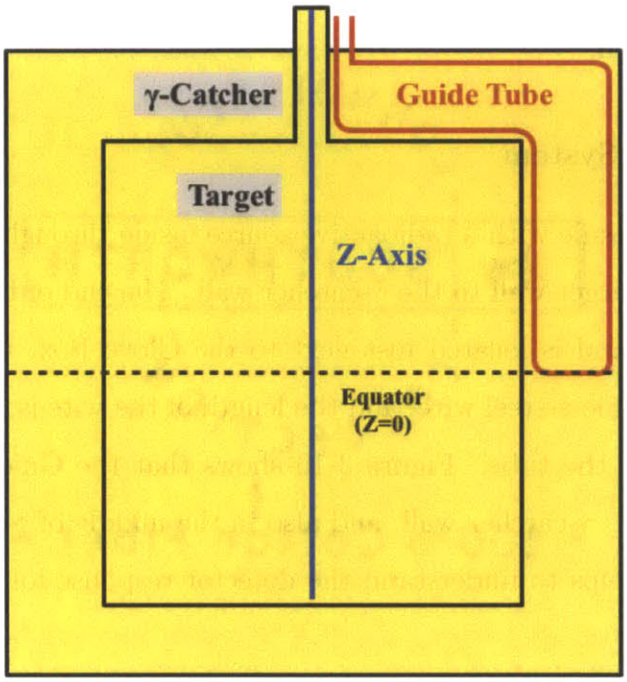

Figure 3-10: A toy drawing of Z-Axis and Guide-Tube. The Guide-Tube runs from the Chimney along the Target wall down to the equator, then comes back along the γ-catcher wall. Z-Axis is a simple fish-line which can move a radioactive source along the vertical axis inside the Target

Source	Activity	Detected Radiation	Detected Energy (MeV)
^{137}Cs		γ	0.66
^{68}Ge		γ	1.022
^{60}Co		γ	2.51
^{252}Cf		n, γ	2.22 and 7.94

Table 3.3: A summary of deployed calibration sources. ^{68}Ge goes through an electron capture to emit two 0.511 keV gamma rays. The energy of ^{252}Cf is a summed energy of gamma ray(s) from a delayed neutron capture only. The prompt events include multiple gamma rays and neutrons from the fission process.

the radio-purity of the detector, and it is also optically isolated from the outside during the calibration source deployment. Data taken with this calibration source is primarily used for the Target volume calibration.

Guide-Tube Calibration System

Guide-Tube runs a small capsule with a radioactive source inside through a stainless-steel tube that is installed along the Target wall to the γ-catcher wall. The end of the Guide-Tube comes outside along the chimney and is located just next to the Glove Box. The source capsule is attached at the end of a stainless-steel wire, and the length of the wire is controlled by a micro-step motor from the end of the tube. Figure 3-10 shows that the Guide-Tube can deploy a source along the Target wall, γ-catcher wall, and also in the middle of γ-catcher volume along the equator. This greatly helps to understand the detector response for physics events in the γ-catcher.

3.3.2 ID Photo-Multiplier Tube Gain and Timing Calibration

The most basic, and yet very important part of the calibration is the gain and timing calibration of ID PMTs. The former is a conversion factor to obtain the original number of photo-electrons produced at a photocathode from the integrated electronic charge in the waveform. The timing calibration is to correct a slight offset of hit timing among different PMTs. In order to inspect these features of PMTs, reconstructed pulse information from software called `RecoPulse` is used. Details about `RecoPulse` can be found in Sec. 4.1.

Timing Calibration

PMT timing calibration is to correct for time-offsets in the PMT readout electronics using IDLI data taken with a long-wavelength light pulse. At this wavelength, photons do not excite scintillator liquids, and they are are directly detected by ID PMT photocathodes. There are 8 IDLI light emission points used in this calibration, and the analysis uses PMTs on the other side of the detector from the emitting LED position. For each of the PMTs, an external trigger signal start time is subtracted from a reconstructed pulse time and plotted against the distance

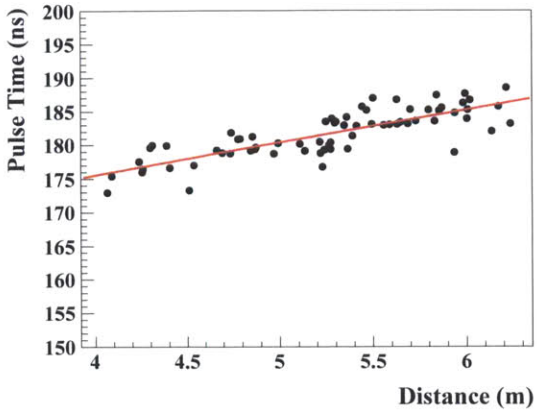

Figure 3-11: Example plot of PMT time-offset calibration using the ID Light Injection (IDLI) system. Each data point represents a PMT installed in the opposite side of the Buffer wall from the light source. The vertical axis shows a reconstructed pulse timing after subtraction of a reference time from the external trigger signal. The horizontal axis shows the distance between each PMT and the light source.

of each PMT from the light source. Figure 3-11 is an example of data points from one LED installed on the top of the Buffer tank. The data points should follow a function of a straight line with an offset. The offset comes from the reference signal. Each PMT is corrected for a time-offset which is the vertical distance from a data point to the fitted straight line.

Gain Calibration

The PMT gain calibration is performed on the `RecoPulse` output charge information. The first step is to find a SPE gain. This is done by analyzing channel-wise pulse charge distribution in the normal physics data. Figure 3-12 shows an example from one of the ID PMTs installed in the middle of the sidewall of the Buffer tank. The charge spectrum in Fig. 3-12, $Q(x)$, is interpreted by the following function form,

$$Q(x) = \sum_{n=1}^{\infty} \underbrace{\frac{\mu^n e^{-\mu}}{n!}}_{\text{Pisson}} \cdot \underbrace{\frac{1}{\sigma\sqrt{2\pi n}} \exp\left(-\frac{(x - n \cdot Q_{SPE})^2}{2n\sigma^2}\right)}_{\text{Gaussian}}, \tag{3.7}$$

which is a convolution of Poisson and Gaussian functions. The Poisson distribution is to account for Poisson statistics of electron multiplication in the dynode chain of the PMT, while

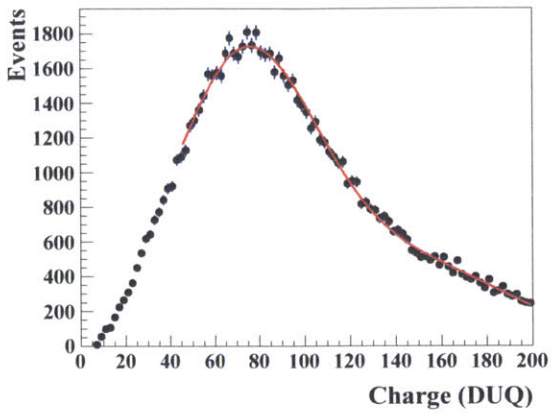

Figure 3-12: Single photo-electron (SPE) distribution from PMT 180, one of the ID PMTs installed in the middle of the Buffer side-wall. The horizontal axis shows the integrated electronic charge from reconstructed pulses. The red line shows a fitted spectrum defined in Eq. 3.7.

the Gaussian is the core function to locate the peak and resolution of the SPE spectrum. The SPE gain, g_{SPE}, is obtained by fitting this model to a low-energy physics data sample below ~ 1 MeV. The red line in Fig. 3-12 shows the fit performed up to the second order of Eq. 3.7, and shows a good agreement between the model and data. The lower range of the data spectrum is distorted by a threshold effect in the pulse reconstruction, and hence is not used for the fit. The first and second order of Eq. 3.7 are found to be sufficient to fit the low energy data sample.

The SPE gain is affected, however, by an inefficiency coming from the electronics and charge reconstruction algorithms. For multi-photo-electron (MPE) signals, therefore, the gain does not simply scale as number of photo-electrons times the SPE gain. The variance of the MPE signal is assumed to be described by the following,

$$\sigma^2_{MPE} = \sigma^2_{\text{Poisson}} + \sigma^2_{\text{SPE}} + \sigma^2_{\text{Background}} + \cdots . \tag{3.8}$$

For a large charge signal at the MPE level, the Poisson fluctuation can be assumed to be Gaussian like. That is $\sigma^2_{\text{Poisson}}$ should follow $g^2_{MPE}N$ where g_{MPE} is MPE gain and N is average number of photo-electrons. As described previously, SPE variance should follow a Gaussian distribution. Therefore, $\sigma_{SPE} = g^2_{SPE}N$. Introducing an approximation factor α such that

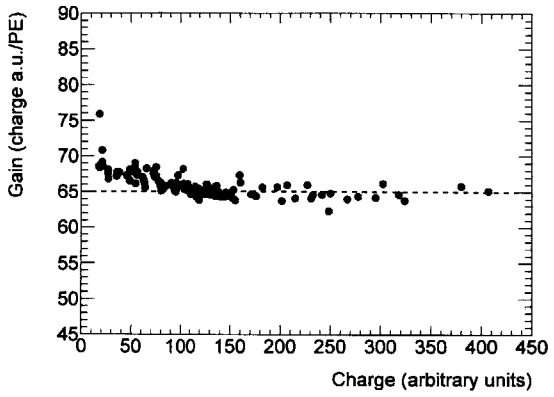

Figure 3-13: Example ID PMT gain extracted from the IDLI data points. The gain variation at low charge due to SPE inefficiency can be seen [1].

$\sigma^2_{\text{Poisson}} = \frac{g^2_{SPE}}{\alpha^2} N$, one concludes

$$g_{MPE} = \frac{\sigma^2_{MPE}}{Q_{MPE}} \frac{1}{1+\alpha^2}. \tag{3.9}$$

The IDLI calibration system is used to generate a MPE signal by generating high-intensity light at a low-wavelength which excites the liqud scintillator. The light intensity is varied at several steps to study MPE gain at different light levels.

Figure 3-13 shows an example gain calibration from one PMT using the IDLI system with varied light intensity. As described above, the low charge signal (i.e., SPE region) has a gain that is different from high charge region (i.e., MPE region). The MPE gain is extracted from the region above 200 digitized unit of charge (DUQ) where the gain is constant by fitting a constant. The parameter α in Eq. 3.9 is determined by matching the MPE gain and a linearly extrapolated line to the SPE gain, extracted from a normal physics data, from 200 DUQ. It is found that the IDLI data points in Fig. 3-13 appear nearby this extrapolated line, although it does not fit perfectly. The remaining discrepancy of the actual gain and this extrapolation method is taken into account by the energy scale systematic uncertainty (see Sec. 4.3.4).

Chapter 4

Event Reconstruction

This chapter describes how the energy and position information of an event are reconstructed from digitized waveforms from PMTs. The first step is to reconstruct a physical pulse in the waveform to obtain an electric charge and hit timing information from each PMT. The next step is to obtain spatial information, which splits into two parallel processes: a vertex reconstruction and a track reconstruction. For all events, a point-like event vertex reconstruction is applied. In addition, a track reconstruction is applied for high-energy events such as muons. In the final step, an energy reconstruction provides the energy scale, taking account of vertex dependency and the time stability of the detector response.

4.1 Low-Level Waveform Reconstruction

Pulse Reconstruction

The electric charge and the timing of the physical pulse in the PMTs' waveform is obtained from pulse information reconstructed by software called `RecoPulse`. `RecoPulse` takes three steps for pulse reconstruction: (1) waveform baseline pedestal analysis; (2) timing analysis of a pulse in waveform; (3) charge analysis by integrating a pulse.

For a given waveform, a baseline analysis is performed to find a mean and RMS of a baseline pedestal using the first eight time samples, which correspond to the first 16 ns. The mean is expected to be found around 210 DUI, as it is a configuration of FADCs, with the

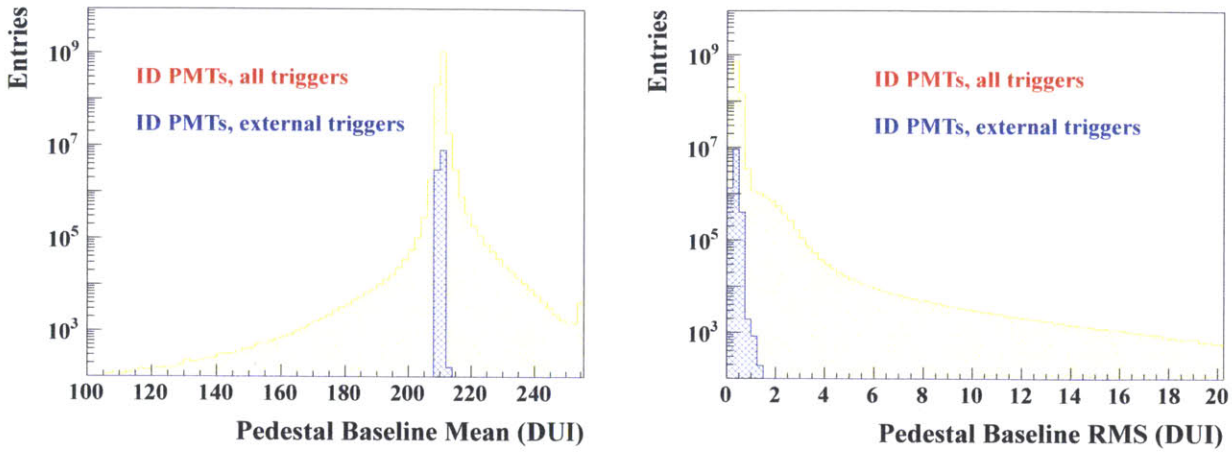

Figure 4-1: Left: a pedestal baseline mean value estimated from 8 hours' worth of physics data. The orange histogram shows the distribution from normal physics events while the red line shows that for externally triggered events. Right: a pedestal baseline RMS distribution shown in the same style as the left plot.

RMS $\approx$ 0.5 DUI due to a fluctuation of 1 DUI from the electronics. The pedestal analysis is performed, by default, for each channel in each event to determine the mean and RMS to be used for reconstructing a pulse in each waveform. This procedure fails, however, when there exists a large fluctuation or an overall mean position shift of a baseline. The former could happen with a dark noise or a pre-pulse, a signal caused by a photon hitting the first dynode in the PMT instead of a photocathode. The latter could happen on several consecutive triggers after a large signal such as a high energy muon event because of the length of time for the waveform baseline to come back to its original position. These cases can be identified by a large RMS value of the pedestal baseline samples. If the RMS is found to be large, `RecoPulse` uses a pedestal mean estimated from the previous fixed-rate trigger, which is an ideal sample for a baseline analysis with no presence of a physical pulse. Figure 4-1 shows the pedestal mean and RMS estimated from normal physics events as well as externally triggered events by a fixed-rate trigger.

Once the pedestal mean and RMS are found, `RecoPulse` searches for a physical pulse in waveform. There are two conditions in `RecoPulse` to define a pulse. The first is that the maximum amplitude of a pulse must be larger than 1 DUI. The second is that the integrated charge of a pulse, expressed in the unit of DUQ, must be above the threshold, Q_{thres}. The first

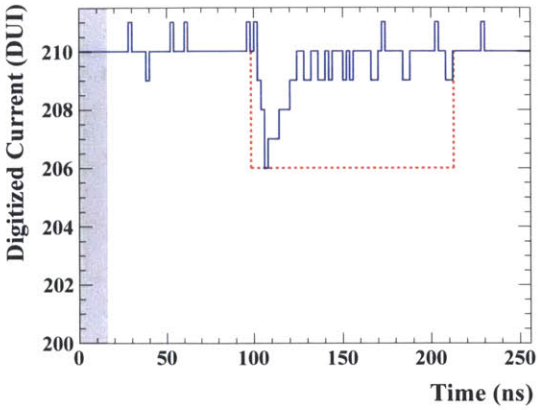

Figure 4-2: A result of applying `RecoPulse`to the waveform shown in the left of Fig. 3-8. The dotted line shows the range of reconstructed waveform in time (horizontal) and the maximum amplitude in amplitude (vertical). The shaded region on the left includes the time samples used for pedestal analysis.

condition is to distinguish a pulse from a 1 DUI fluctuation of baseline due to electronics, and the Q_{thres} in the second condition is defined as shown below

$$Q_{\text{thres}} = 5\sigma_{\text{ped}}\sqrt{\text{Time Sample}} \tag{4.1}$$

where σ_{ped} is the pedestal RMS and "Time Sample" is the time width of the reconstructed pulse. Figure 4-2 shows an example of a reconstructed pulse in the same waveform shown in Fig. 3-8. The electronic charge of a pulse is calculated by integrating the waveform from the pedestal mean within the pulse time width.

4.2 Vertex and Track Reconstruction

The localized position information of an event is a very important piece of information to understand the physics nature. Double CHOOZ employs multiple point-like vertex reconstruction and continuum track reconstruction algorithms. In this section, those algorithms used in the hydrogen and Gd analyses are described.

4.2.1 Vertex Reconstruction

A reconstruction algorithm, called `RecoBAMA`, is used for a point-like ID vertex reconstruction for all events with an ID energy deposition below 30 MeV. The `RecoBAMA` algorithm is based on both a PMT charge and hit timing likelihood function. Given a certain event vertex, a PMT charge seen by the i-th PMT can be estimated as

$$q_i = \phi(\vec{r}_i) \cdot \epsilon_i \cdot \Omega_i(\vec{r}_i) \cdot A_i(\vec{r}_i) \tag{4.2}$$

where ϕ is the source light intensity, and ϵ_i and Ω_i are the quantum efficiency and solid angle of a PMT. A_i denotes the attenuation factor, and $\vec{r}_i$ denotes the PMT position vector from the detector center. PMT hit timing can be simply estimated as

$$t_i = t_0 + \frac{|\vec{r}_i|}{c_n} \tag{4.3}$$

where t_0 is the event time stamp after correcting an offset of readout time, r_i is the distance of the event vertex to i-th PMT, and c_n is an effective speed of light in the ID liquid.

With these parameter predictions, a likelihood function for an i-th PMT is constructed and multiplied over all PMTs. Then `MINUIT`, a common minimizer package in high-energy physics, is used to reconstruct an event by minimizing a negative log-likelihood of the combined likelihood function. The performance of the algorithm is tuned using an MC sample, and benchmarked against ^{60}Co data sample. Figure 4-3 shows a reconstructed (x, y, z) position vertex of the ^{60}Co source deployed at the Target center. As shown in the figure, a vertex reconstruction resolution is about 12 to 13 cm. Any systematic error arising from `RecoBAMA` reconstructed vertex is included in the energy reconstruction and detection efficiency study (Sec. 4.3 and 6.1).

4.2.2 Track Reconstruction

For events with high-energy deposition in the IV and/or ID, a muon track reconstruction algorithm is applied. Muon track information is particularly useful to tag backgrounds that originate from muons but cannot be vetoed by the 1 ms after-muon veto period. Such back-

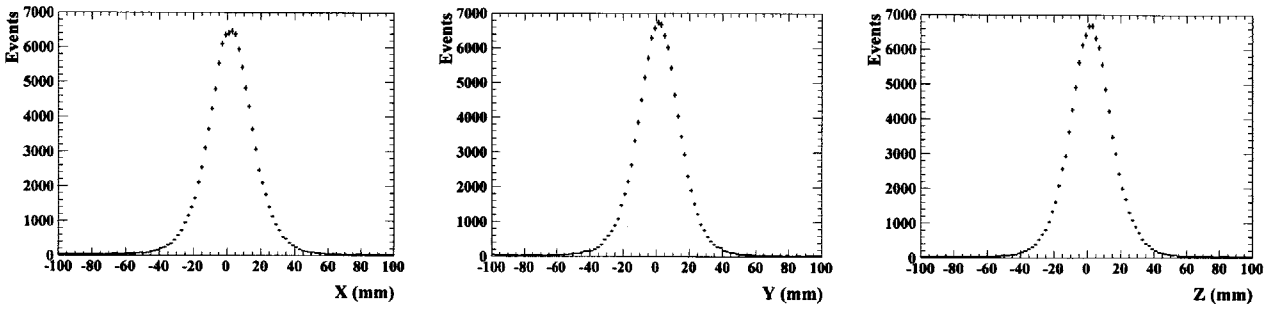

Figure 4-3: Left to right: x, y, and z reconstructed position vertex distribution by RecoBAMA using calibration data with a ^{60}Co source deployed at the Target center.

grounds are expected to stay close to a muon path in the detector.

In the hydrogen analysis, two independent muon track reconstruction algorithms are employed: (1) `RecoMuHamID` to reconstruct muons using ID PMT information; (2) `RecoJP` to reconstruct a muon track using IV PMT hit timing information. Evidently, `RecoMuHamID` is for muons that enter the ID region while `RecoJP` covers muons that go through only the IV volume. Because the fiducial volume of the hydrogen analysis extends to the boundary of the liquid scintillator and Buffer oil, `RecoMuHamID` must take into account muons crossing two liquids with very different light properties: LS and Buffer oil. In the hydrogen analysis, `RecoJP` is used as the nominal muon track reconstruction algorithm because it is expected to have a rather uniform performance on all muons. The results from both analysis methods are cross-checked (see Sec. 7.2).

4.3 Energy Reconstruction

Energy scale is a particularly important parameter to perform a θ_{13} analysis, and hence the reconstruction procedure is described in detail here. It starts from a calibrated photo-electron scale which is described in Sec. 3.3.2.

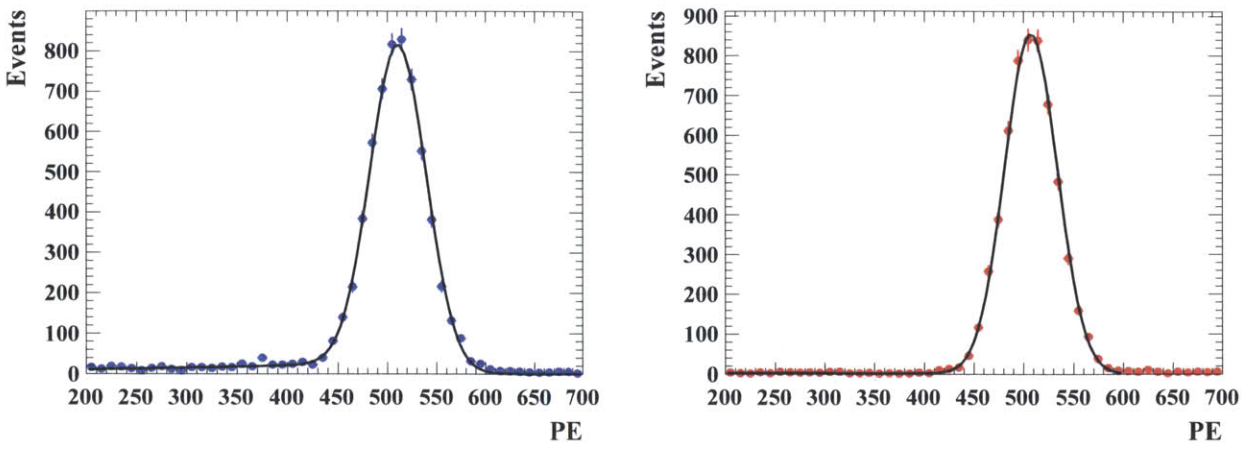

Figure 4-4: A photo-electron distribution of a gamma ray from a neutron capture on hydrogen. Left is DATA taken with a ^{252}Cf neutron source while the right represents the MC, fitted by a function defined in Eq. 4.4.

4.3.1 Energy Scale: Photo-electron to MeV Calibration

The conversion factor of the photo-electron to the MeV scale is obtained from the 2.223 MeV line of a neutron capture on hydrogen at the center of the ID. A ^{252}Cf neutron source is deployed at the Target center where light collection over PMTs is assumed to be the most uniform. A resulting photo-electron distribution is fitted using a function,

$$f(x) \;=\; N_{\text{tail}}\, e^{-\alpha x} \cdot \text{Erf}\left(\frac{x - 0.8972\mu}{\sigma}\right) + N_{peak}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{4.4}$$

where the first component describes the energy loss by Compton scattering and the second part describes the photo-peak of a released gamma ray [33]. Figure 4-4 shows a fitted photo-electron distribution for DATA and MC separately.

From the fit result shown in Fig. 4-4, we found 230 photo-electrons/MeV and 228 photo-electrons/MeV as a conversion factor for DATA and MC, respectively. This provides the MeV energy scale at the detector center.

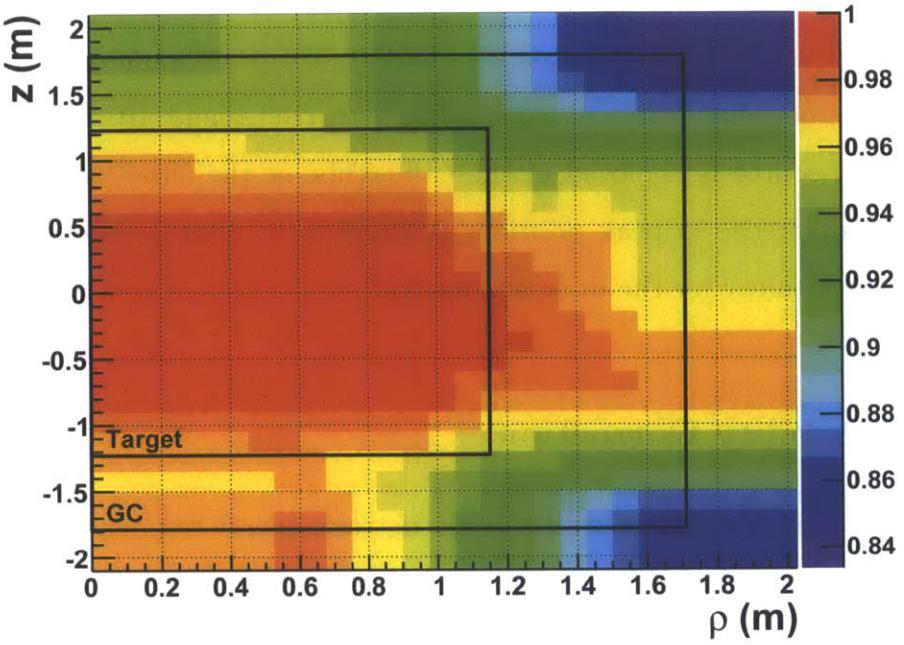

Figure 4-5: Energy uniformity map from spallation neutrons [1].

4.3.2 Energy Non-Uniformity Correction

An extension of the MeV energy scale to the whole detector is performed by using an energy uniformity map created from spallation neutrons from a muon. When a muon goes through a detector, it can produce multiple neutrons that stop and get captured by a hydrogen nucleus in the LS. Using reconstructed vertex information, post-muon photo-electron distribution is separated into 108 sub-volumes. Each photo-electron distribution is fitted using Eq. 4.4 to extract the Gaussian peak position. Finally, taking a ratio of this peak value at each sub-volume to the center of the detector, a detector map of a correction factor for the peak position is constructed. This is shown in the left of Fig. 4-5. A similar approach is taken for the MC sample using the $\bar{\nu}_e$ sample generated over the ID volume.

The systematic uncertainty of the uniformity map in the Target volume is studied using a separate map created using a neutron capture peak on the ^{157}Gd nucleus [1]. Such a map is created from spallation neutrons for the DATA and $\bar{\nu}_e$ sample for the MC. A relative difference between the DATA and MC uniformity maps is quantified as 0.43 % [1].

For the γ-catcher volume, a conservative estimate is made by taking a relative difference between DATA and MC in the peak position of a hydrogen capture gamma ray at the edge of the γ-catcher where the difference is maximal. For both DATA and MC, ^{252}Cf calibration source data is used. A source is deployed in the GuideTube at the outer-top edge of the γ-catcher volume. The difference in peak positions is found to be 1.33 %, and is used for the position uniformity uncertainty for the hydrogen analysis.

4.3.3 Time Instability Correction

Finally, there is a correction for a time variation of the detector energy response. This effect only exists in the DATA and not in the MC, and hence is a subject to be corrected. Neutron captures on hydrogen and Gd nuclei from spallation neutrons are used to monitor the detector stability. In particular, the time stability obtained from ^{157}Gd capture peak is used for the stability correction. A data sample is divided into a consecutive 5-day period during the data taking. The MeV scale with a position uniformity correction applied is fitted to extract the Gaussian peak position. An example of the fitted spectrum for both neutron capture on hydrogen and ^{157}Gd are shown in Fig. 4-6.

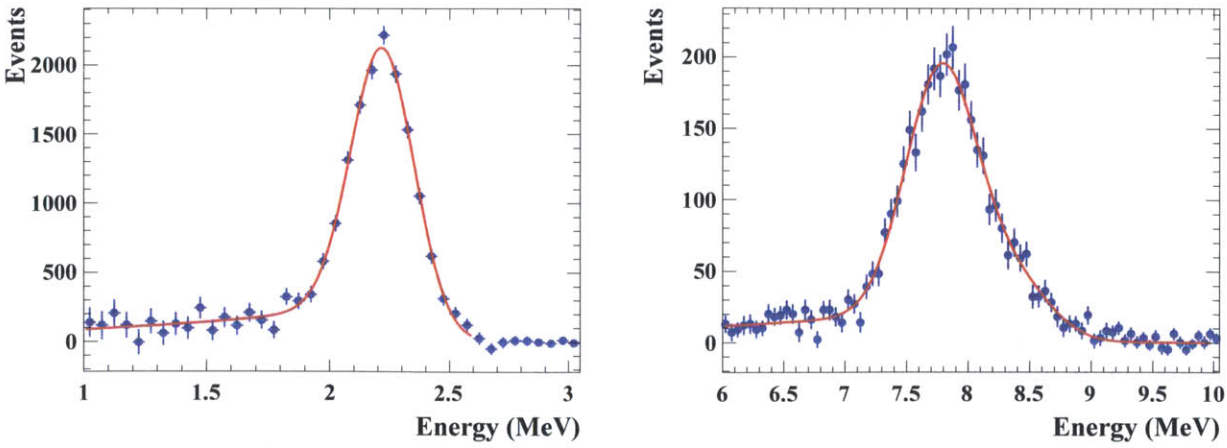

Figure 4-6: A neutron capture peak for hydrogen (left) and Gd (right) nuclei from spallation neutron data sample during the first 5 days of data taking. Extracted peak positions make up data points in Fig. 4-7.

Extracted peak position stability is shown in terms of a relative shift in Fig 4-7 for the both hydrogen and ^{157}Gd peak. Time stability correction is defined to correct a shift from the ^{157}Gd stability data points. The hydrogen capture peak stability shown in Fig. 4-7 is reprocessed with this correction factor to estimate the residual time instability to estimate a systematic uncertainty for this correction. This is found to be 0.61 %.

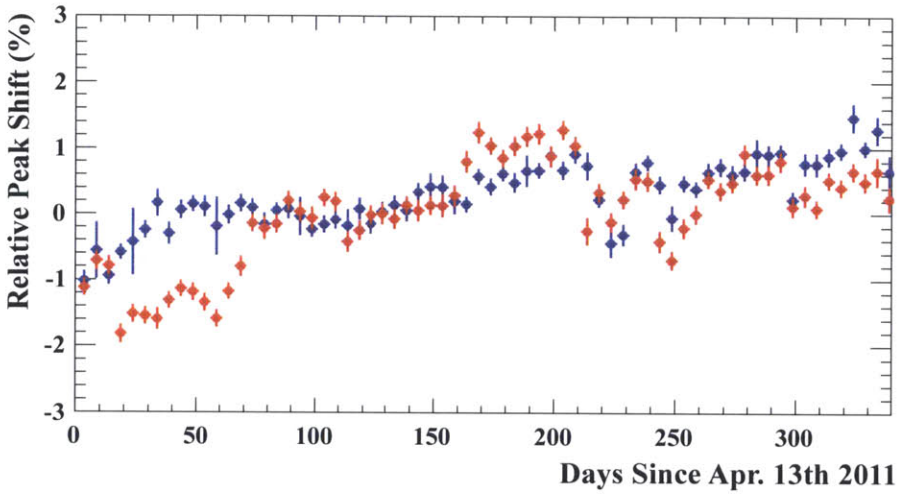

Figure 4-7: A time stability of a relative shift in neutron capture peak energy position. Red data points are from a neutron capture on hydrogen and blue for ^{157}Gd. Each data point is made from a consecutive 5 days of data taking. Error bars represent an uncertainty from the fitting of the energy distribution.

4.3.4 Residual Energy Non-Linearity

The linearity of energy scale is not necessarily the same for DATA and MC, and hence the difference needs to be quantified as a systematic uncertainty. This is quantified by comparing DATA and MC energy peak position for various calibration sources. The lowest peak is 0.662 MeV from ^{137}Cs and the highest is neutron capture on ^{157}Gd, approximately 8 MeV. This systematic uncertainty is found to be 0.85 % and is common to both Target and γ-catcher regions [1].

4.3.5 Summary of Energy Reconstruction

Table 4.3.5 summarizes the energy scale uncertainties. Overall, the hydrogen analysis carries

Syst. Error	**Gd [%]**	**hydrogen [%]**
Position non-uniformity	0.43	1.33
Time instability	0.61	0.61
Energy scale non-linearity	0.85	0.85
Total	1.13	1.69

Table 4.1: Energy-scale systematic errors [1].

a larger systematic uncertainty than the Gd analysis, and this is largely due to a conservative estimation for the position non-uniformity error. This is expected to be improved in the future with more source calibration points in the γ-catcher volume.

Chapter 5

Signal Selection

This chapter describes the data sample used in the hydrogen analysis as well as the signal candidate selection method. The selection method comes in two separate steps. The first step is a pre-selection of physics events, and is described in Sec 5.2. The second step is a coincidence event search to extract the IBD candidates, described in Sec. 5.3. Finally Sec. 5.4 presents the summary of the selected IBD candidates.

5.1 Analysis Data Sample

The data sample used in this analysis had been collected by the Double CHOOZ far detector from April 13th 2011 to March 15th 2012. It contains 251.27 days of cumulative DAQ run-time for physics analysis and several days of detector calibration data. As it is described in Sec. 5.3.1, the effective detector livetime for $\bar{\nu}_e$ analysis is 240.06 ± 0.01 days. Figure 5-1 shows both the detector run-time and livetime for analysis on each calendar day since the beginning of data taking.

5.2 Pre-Selection

The Double CHOOZ far detector trigger rate is about 140 Hz in the normal physics data taking mode. The rate consists of about 45 Hz of muons' and 10 Hz of singles' physics triggers. A large

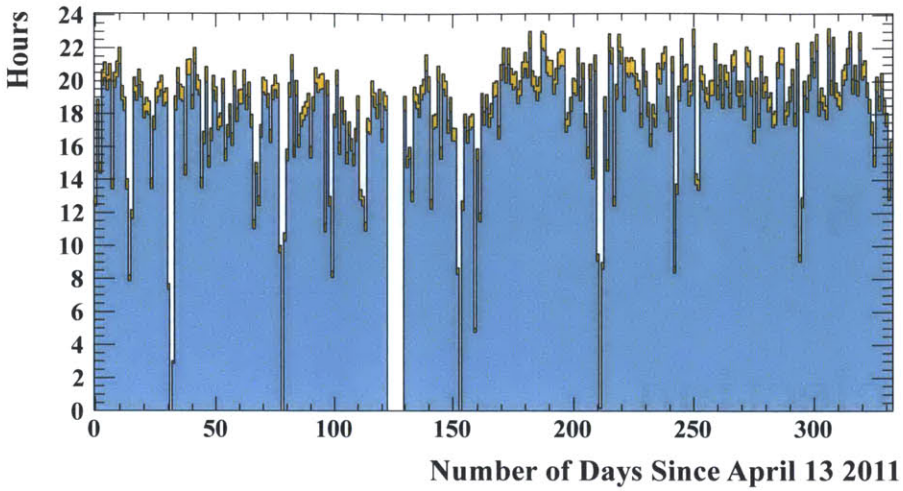

Figure 5-1: Detector run-time and livetime for analysis on each calendar day since the beginning of data taking. The cumulative livetime is 240.06 days over 335 calendar days.

fraction of the remaining 85 Hz is a non-physics, instrumental noise background called light noise (LN). The first step of the IBD signal selection is to remove such light noise backgrounds, muons, and muon induced backgrounds.

5.2.1 PMT Charge Spreadness

A light noise background event is caused by spontaneous light emission from the base of a PMT. While a detailed mechanism of light noise is still under investigation, the fact that light noise is sourced at the PMT base allows a simple and robust rejection of this background through inspecting the spread of detected optical photons over the PMTs. The MQTQ variable is defined as

$$\mathrm{MQTQ} = \frac{\mathrm{Q_{max}}}{\sum_{\mathrm{i}}^{\mathrm{All\ ID\ PMTs}} \mathrm{Q_i}} \tag{5.1}$$

where Q_i denotes an integrated charge of a reconstructed pulse from the i-th PMT, and Q_{max} is the maximum of all PMTs. This simple variable reflects the spreadness of charge over all the PMTs.

A typical physics event includes our signal arises from the liquid scintillator volume, the

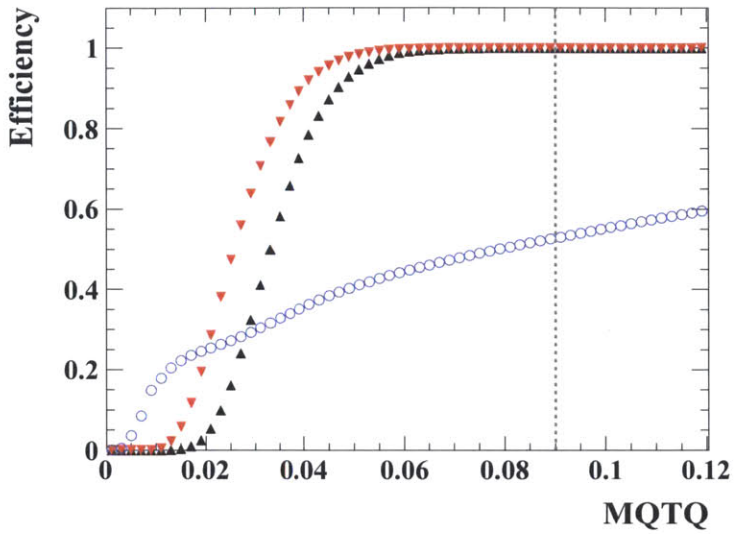

Figure 5-2: MQTQ distribution of delayed neutron capture from $\bar{\nu}_e$ MC sample (red) and ^{252}Cf data from the source deployment at γ-catcher edge (black). Blue data points show the MQTQ distribution of all events from a typical physics run. The gray dotted line indicates the default light noise rejection cut value for an IBD candidate pre-selection cut.

Target and γ-catcher region. Optical photons are produced and emitted isotropically toward all PMTs. The light yield of the Target and γ-catcher liquid is about 230 photo-electrons per MeV, and hence a physics event above the trigger threshold, 0.7 MeV, is expected to have a good spreadness of detected light over hit PMTs. Consequently, we expect a low MQTQ value for physics events. For a light noise event, on the other hand, a large fraction of light is detected by one or a few PMTs located nearby a source PMT base. Light noise backgrounds, therefore, typically come with a large value of MQTQ compared to physics events.

Figure 5-2 shows the MQTQ distribution of the $\bar{\nu}_e$ MC (red) sample that passes the default selection cuts except for MQTQ. Shown in the same figure with black data points is the MQTQ distribution of delayed neutron capture from a ^{252}Cf calibration source deployment at the γ-catcher edge point, where the spreadness of light is expected to be the worst. By applying MQTQ < 0.09, we expect negligible inefficiency on the physics data sample. Finally, in the same figure, blue data points show the MQTQ distribution of all events in one example physics run dominated by light noise background.

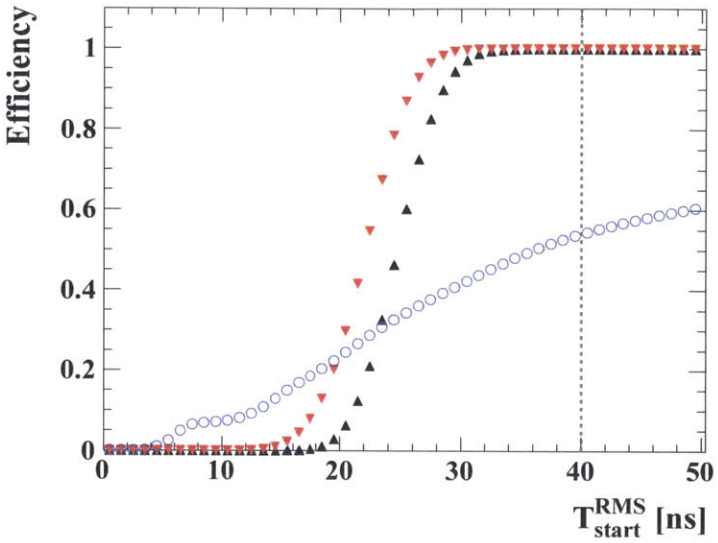

Figure 5-3: T^{RMS}_{start} distribution of delayed neutron capture from $\bar{\nu}_e$ MC sample (red) and ^{252}Cf data from the source deployment at γ-catcher edge (black). Blue data points show the T^{RMS}_{start} distribution of all events from a typical physics run. The gray dotted line indicates the default light noise rejection cut value for an IBD candidate pre-selection cut.

5.2.2 PMT Pulse Time Uniformity

Another simple and robust cut for light noise rejection is a cut on the spread of pulse start times over all hit PMTs. For a physics event, optical photons are produced from an event vertex and travel toward PMTs almost at the same time. A spread in these photons' arrival time is dominated by a scintillator de-excitation time and different distances of an event vertex to the PMTs. A light noise event does not follow the same scheme because its light emission mechanism as well as vertex of light emission point are largely different from a typical physics event. A variable T^{RMS}_{start} is defined as the root-mean-square of all hit PMTs' pulse start time to reject light noise against physics events.

Figure 5-3 shows the T^{RMS}_{start} distribution in the same manner as Fig. 5-2. The red data points are from the $\bar{\nu}_e$ MC sample, black data points from ^{252}Cf source deployment data at the γ-catcher edge, and blue data points from a typical physics run dominated by light noise. $T^{RMS}_{start} < 40\ ns$ is applied in the pre-selection to reduce light noise background. Like the case of the MQTQ cut, an inefficiency from this cut is considered to be negligible.

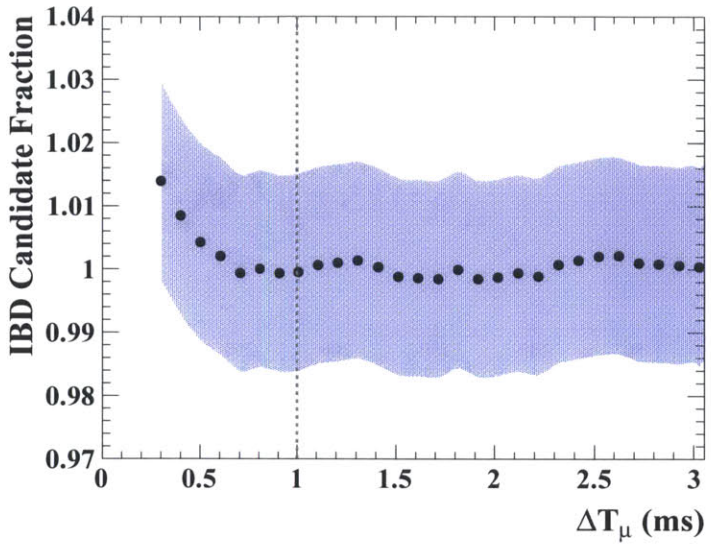

Figure 5-4: Number of IBD candidates with Gd selection cuts against the duration of after-muon veto time window. The gray line is drawn at 1 ms, indicating the default cut value used in pre-selection of candidates. A statistical uncertainty is shown in blue colored error band.

5.2.3 After-Muon Veto

An energetic muon going through or nearby the detector produces cosmogenic isotopes and multiple fast neutrons that can mimic an IBD coincidence signal. In order to avoid such backgrounds, after-muon veto is employed and any event triggered within 1 ms of muon is vetoed where a muon event is defined by either an ID energy deposition of 30 MeV or an approximate IV energy deposition of 10 MeV. Figure 5-4 shows a fluctuation in number of IBD candidates from Gd analysis after muon as a function of after-muon veto time.

After-muon veto cut causes a non-negligible analysis dead-time in data sample. Figure 5-5 shows a daily variation of muon rate. The average muon rate is about 44 Hz, corresponding to 4 % of deadtime. The live-time after taking account for this after-muon veto is 240.16 ± 0.01 days while the total run-time is 250.17 days.

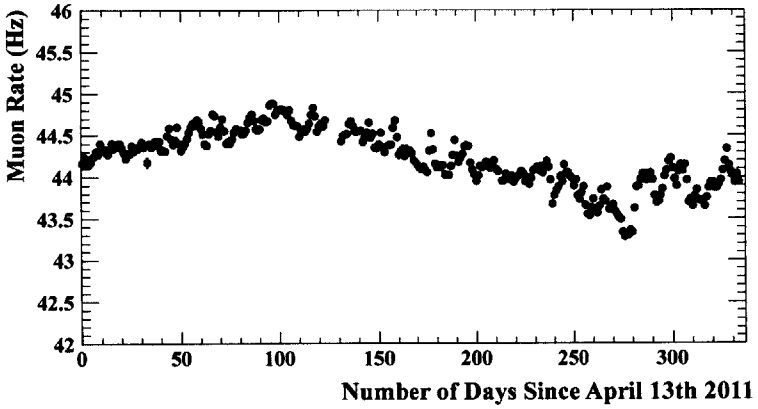

Figure 5-5: Daily muon rate stability during the data taking period.

5.3 Neutrino Event Candidate Selection

An IBD coincidence selection cut is run on the result of pre-selection described in Sec. 5.2. The first step is to search for a prompt candidate trigger, followed by the second step of search for a delayed candidate trigger. The final step is, after finding a delayed candidate trigger, to apply an isolation cut. A prompt-delayed coincidence pair that survives after these steps is counted as an IBD signal candidate.

5.3.1 Prompt Candidate Selection

A prompt candidate trigger must have a prompt energy, E_{prompt}, between 0.7 MeV to 12.2 MeV. An $\bar{\nu}_e$ event should have a minimum energy deposition of 1.011 MeV since it is an electron-positron annihilation. Given the energy resolution of the ID, 0.7 MeV is a sufficiently low energy cut to avoid an inefficiency. The upper bound of 12.2 MeV is motivated by a similar reason, to include all IBD candidates by covering up to the end point energy. It is important to note, however, that the high E_{prompt} region above 8 MeV would be particularly important to constrain background events since the $\bar{\nu}_e$ contribution is rather small in that energy region.

Another requirement for a prompt trigger candidate is not to have any triggered OV events in a coincidence within 224 ns time window. This condition is useful for rejecting stopping muon background, when OV data is available. Given the 2.7 kHz of the OV trigger rate, this

causes about 0.06 % of analysis dead-time for data taken with the OV DAQ. As a result, the analysis livetime becomes 240.06 ± 0.01 day.

5.3.2 Delayed Coincidence Candidate Selection

A delayed event candidate is required to have time and spatial correlation to the prompt event, $10 < \Delta T < 600$ μs and $\Delta R < 900$ mm, respectively. The upper bounds are motivated from the neutron thermalization study. The lower bound of the ΔT cut is set to exclude stopping muon decay backgrounds which are present in the Gd analysis. While a veto of the first 10 μs introduces a significant inefficiency to the signal in the Gd analysis, this inefficiency is small in the hydrogen analysis because of a longer neutron thermalization time in the γ-catcher. Finally, in addition, a delayed event candidate should carry energy relevant for a neutron capture on hydrogen, which is 2.223 MeV. Its energy cut is, therefore, set to be between 1.5 and 3.0 MeV.

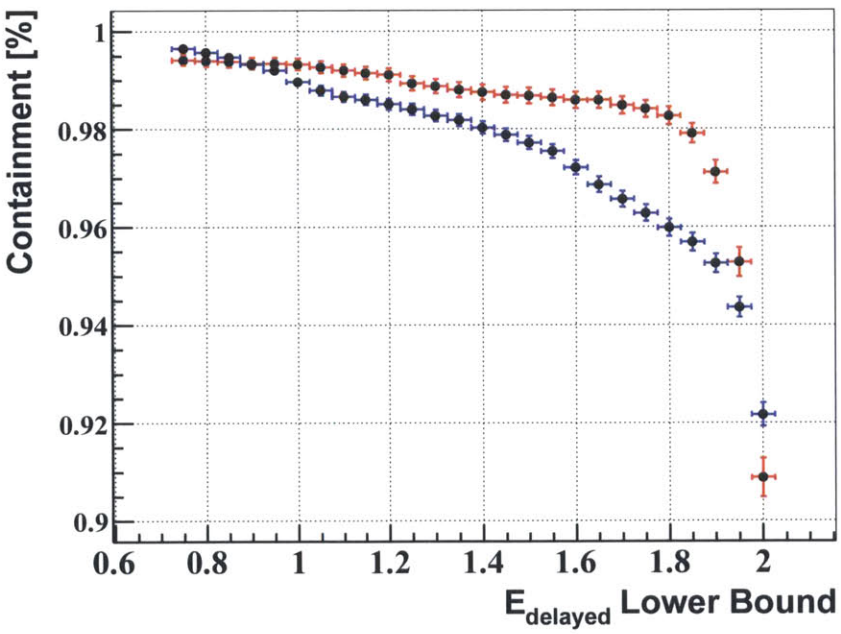

Figure 5-6: DATA and MC E_{delayed} efficiency distribution as a function of the lower boundary of *ed* cut from the ^{252}Cf calibration source in the Target center. For details of how the efficiency is studied, see Sec. 6.1. Note that the MC correction factor introduced in the Sec. 6.1 is not applied to the MC efficiency shown in this plot.

In this first attempt at the hydrogen analysis, cut values are set to keep the minimal systematic uncertainty, except for the ΔT cut lower boundary. For instance, Fig. 5-6 shows how efficiency varies for DATA and MC with a varying E_{delayed} lower cut value. The lower bound of 1.5 MeV used is where DATA and MC efficiencies start to diverge rapidly. A systematic uncertainty for a delayed coincidence selection is studied in detail using ^{252}Cf neutron source calibration data, and can be found in Sec. 6.2.

5.3.3 Isolation Cut

A muon that goes through surrounding rock without passing any active volume of the detector produces fast neutrons that can travel into the ID. Such fast neutrons are often produced in multiplicity, and those that reach the ID region are detected by means of either a recoil proton or a neutron capture event in the LS volume. Because they come in multiplicity, this background can be reduced by requiring an isolation cut to those coincidence candidate event pairs.

The Gd analysis found a sufficient rejection of fast neutrons by requiring no pre-selected trigger to exist in the [-100,400] μs period after the prompt candidate trigger except for a delayed candidate [1]. For the hydrogen analysis, an isolation cut time window is set at [-600,1000] μs after the prompt candidate trigger. The cut boundaries are changed due to neutron thermalization time in the γ-catcher volume. The expected rate of $\bar{\nu}_e$ at the far site is less than 100 events per day (see Sec. 8), and hence this enlarged isolation cut introduces a negligible inefficiency while rejecting some of the fast neutrons. A remaining fast neutron contamination is treated as one of the backgrounds in the θ_{13} analysis.

5.4 Summary

After applying the IBD candidate selection cuts described in Sec. 5.3 on the pre-selected data sample presented in Sec. 5.2, 36,284 IBD candidates were found during the livetime of 240.06 days. This corresponds to the candidate rate of 151.14 events/day. Among these candidates, as presented in the Sec. 7.6, the predicted rate of the combined background estimate is 79.11 ± 0.8

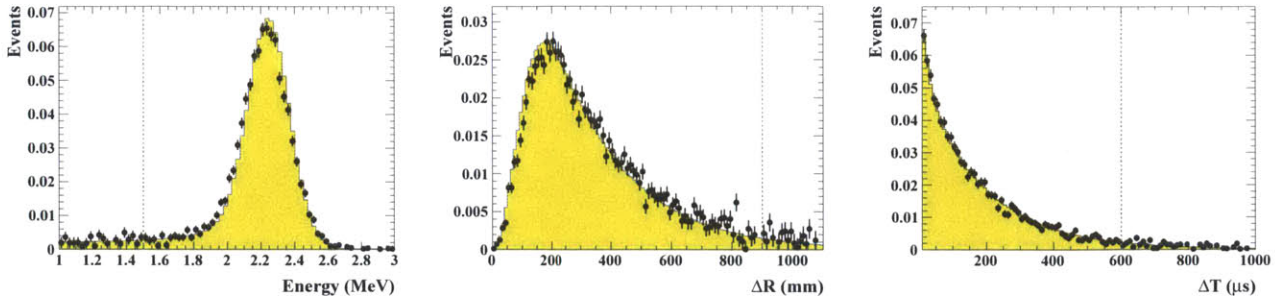

Figure 5-7: Accidentals-subtracted IBD candidates' spectrum shape. From the left to right, E_{delayed}, ΔR, and ΔT distributions are shown.

events/day, which is more than a half of the candidate rate.

A large fraction of the background is, however, accidentals background which is very well understood. This background accounts for 73.45 events/day (see Sec. 7.1 for details). The $\bar{\nu}_e$ rate prediction (see chapter 8) after applying the MC correction factor (see chapter 6) becomes 73.67±2.15 events/day. Since both the rate and spectrum shape of the accidentals background is well understood, one can easily subtract this background's spectrum from that of the candidates. The resulting spectrum is dominated by IBD candidates according to the prediction. Figure 5-7 shows ΔT, ΔR, and E_{delayed} distributions after subtracting the accidentals' spectrum shape.

Chapter 6

Detection Efficiency and Systematics

This chapter is dedicated to the IBD event detection efficiency study. Since the value of θ_{13} is obtained by a comparison of signal DATA and MC prediction, the uncertainty for the efficiency is essentially the difference between DATA and MC. In addition, the overall correction factor for the MC efficiency needs to be quantified if there is an overall offset from the detection efficiency in DATA. Section 6.1 describes the strategy of the efficiency and MC correction factor estimation in the hydrogen analysis. Section 6.2 and 6.3 present study results following the strategy presented in Sec. 6.1. Section 6.4 summarizes the overall detection efficiency and MC correction factor to be used in the θ_{13} oscillation fit.

6.1 Introduction to Efficiency Estimation

A brief description of overall strategy and inefficiencies to be accounted for are given in this section.

6.1.1 Simulation Based θ_{13} Analysis

In the current assessment, the $\bar{\nu}_e$ flux prediction is based on a MC-based approach with Bugey-4 measurement as an input. An IBD coincidence selection cut applied for DATA is also applied on the MC sample which is generated based on a null oscillation hypothesis. Consequently, we need to investigate a systematic uncertainty on common selection cuts between DATA and MC. For

some relevant cuts, a MC efficiency correction factor needs to be applied. In addition, there are some intrinsic effects that have to do with the detector itself and not with cuts. These include the neutron capture fraction to various nuclei and a mobility of neutrons between different volumes, called Spill In/Out effect. Both affect our detection efficiency of $\bar{\nu}_e$.

Selection cuts can be categorized into three groups: (1) pre-selection cuts that are applied to select a pool of candidate triggers for an IBD coincidence selection, (2) prompt candidate selection cuts, and (3) delayed candidate selection cuts. For details of cuts, see Sec. 5.2 and 5.3. These are noted as ϵ_{pre}, ϵ_{prompt}, and $\epsilon_{delayed}$ in respective order. The overall detection efficiency can be written as

$$\epsilon_{IBD} = \epsilon_{pre} \cdot \epsilon_{prompt} \cdot \epsilon_{delayed} \tag{6.1}$$

where

$$\epsilon_{pre} = \epsilon_{T_{\text{start}}^{\text{RMS}}} \cdot \epsilon_{\text{MQTQ}}, \tag{6.2}$$

$$\epsilon_{prompt} = \epsilon_{E_{\text{prompt}}} \cdot \epsilon_{trig} \cdot \epsilon_{iso}, \tag{6.3}$$

$$\epsilon_{delayed} = f_{n-H} \cdot g_{spill} \cdot \epsilon_{E_{\text{delayed}}} \cdot \epsilon_{\Delta T} \cdot \epsilon_{\Delta R}. \tag{6.4}$$

The subscripts denote a parameter on which a cut is applied. An isolation cut efficiency is noted as ϵ_{iso} and trigger efficiency on the prompt event as ϵ_{trig}. Note the E_{delayed} cut lower boundary is well above the trigger threshold, and hence ϵ_{trig} inefficiency is negligible for a delayed event. f_{n-H} denotes a fraction of neutron capture on hydrogen which is intrinsic to a liquid composition. Finally, g_{spill} represents a fractional loss of neutrons due to Spill In/Out effect.

As mentioned in Sec. 3.2, ϵ_{trig} is 100 % with a negligible uncertainty at 0.7 MeV, the lower boundary of the E_{prompt} cut. A systematic uncertainty on the $\epsilon_{E_{\text{prompt}}}$ cut is studied on the lower boundary of 0.7 MeV as the higher end is well above the end point of an expected IBD energy distribution. This study uses ^{68}Ge calibration data, which is a e^+ source. An extremely good agreement on the spectrum shape on the lower energy region is found between DATA and MC in the Gd analysis using both Z-Axis (Target) and Guide-Tube (γ-catcher) data, which yields negligible uncertainty on the DATA and MC discrepancy on the efficiency [1]. Finally,

ϵ_{iso} is quantified in Sec. 7.1 from DATA with a negligible uncertainty. The inefficiency due to an isolation cut is only relevant for DATA because the MC sample does not include any background events, and hence ϵ_{iso} simply carries no inefficiency. Accordingly, the MC correction factor of 0.98131 needs to be applied to compensate for the inefficiency in DATA.

In the efficiency study presented in this chapter, a relative efficiency is computed from the number of events that pass the default cut range and a parent set of events that passed the pre-selection cuts described above. An uncertainty is computed from the Neyman confidence belt constructed through the Clopper-Pearson frequentist approach [34].

6.1.2 ^{252}Cf Calibration Source Data

^{252}Cf calibration source data is extremely useful because it allows us to study the neutron detection efficiency of the detector in general. The source emits multiple gamma rays and neutrons in its fission process, and hence a large background rejection can be easily done by requiring a coincidence of a fission gamma ray event followed by one or multiple neutron captures.

A set of pre-selection cuts are applied on ^{252}Cf source data for which a further efficiency study is done. The pre-selection cuts include: $\Delta T_\mu > 1\ ms$ where ΔT_μ is a time after the previously detected muon; MQTQ < 0.09; and $T_{\text{start}}^{\text{RMS}} < 40\ ns$. Then the following coincidence selection cuts are applied: (1) $4 < E_{\text{prompt}} < 30$ MeV, (2) $0.7 < E_{\text{delayed}} < 30$ MeV, and (3) $0 < \Delta T < 1000\ \mu s$. Finally, in order to increase the purity of physics events, the prompt is defined only if there is no preceding trigger within 1 ms. A good visual agreement between DATA and MC data distribution can be found in Fig. 6-1. Note these distributions have no background subtracted, hence, they are showing a good purity of physics events.

6.1.3 $\bar{\nu}_e$ Sample for Calibration

IBD data is a nearly ideal data sample to study delayed neutron capture detection efficiency since its events are spread over all volumes. The first challenge is, however, to isolate the backgrounds. Since our interest is in neutron detection efficiency, fast neutrons and ^{9}Li, considered as backgrounds to the θ_{13} study, are not backgrounds here because they involve a neutron

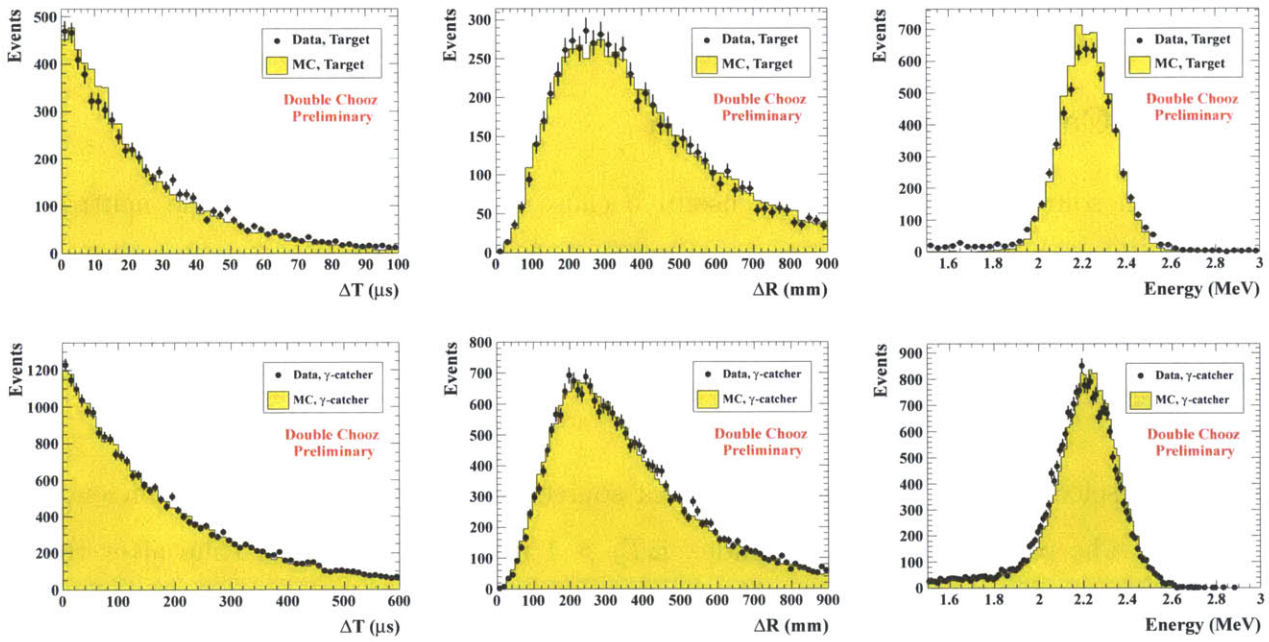

Figure 6-1: Example plots from ^{252}Cf source calibration data sample. Yellow-filled histograms are from the MC and black points are from DATA after applying selection cuts described in the main text. Top row shows distributions from Guide-Tube and the bottom from Z-Axis (Target center) source deployment. From left to right, each column shows ΔT, ΔR, and E_{delayed} distributions.

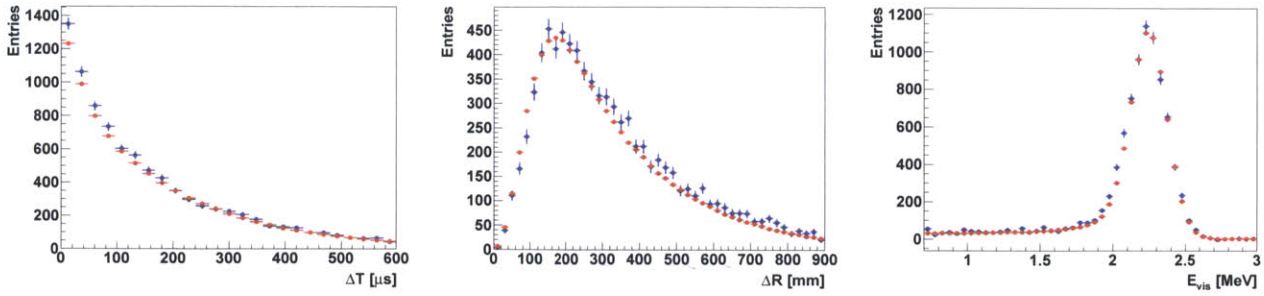

Figure 6-2: Example plots of $\bar{\nu}_e$ candidates with the modified cut of $4 < E_{\text{prompt}} < 12.2$ MeV for DATA (blue) and MC (red). Statistics in DATA and MC are not normalized, and DATA points carry backgrounds in addition to $\bar{\nu}_e$ events.

capture as their delayed event. Also, the fact that their rate is much smaller than that of IBD in the detector makes them a less of a concern. Accidental backgrounds need to be addressed as they are the most dominant backgrounds, and they involve a negligible fraction of delayed neutron capture events (see Sec. 7.1). In order to reduce accidentals backgrounds, the E_{prompt} cut range is modified to $4 < E_{\text{prompt}} < 12.2$ MeV. The DATA and MC distributions of this data sample are shown in Fig. 6-2.

6.2 Delayed Neutron Capture Detection Efficiency

The expected number of delayed neutron captures from signal events can be expressed as

$$N_{n-H} = \phi_{\bar{\nu}_e} \cdot \sigma_{IBD} \cdot N_{p,all} \cdot \epsilon_{delayed} \tag{6.5}$$

where $\phi_{\bar{\nu}_e}$, σ_{IBD}, $N_{p,all}$ represent $\bar{\nu}_e$ flux at the detector, an IBD cross section, and the total number of protons through which an IBD interaction can occur and be detected. $\epsilon_{delayed}$ is the delayed neutron capture detection efficiency. The last part of Eq. 6.5, $N_{p,all} \cdot \epsilon_{delayed}$, can be further broken down into a sum of three categories,

$$N_{p,all} \cdot \epsilon_{delayed} = N_{p,T} \cdot f_{n-H,T} \cdot \epsilon_{n-H,T} + N_{p,GC} \cdot f_{n-H,GC} \cdot \epsilon_{n-H,GC} + N_{p,Spill} \cdot f_{n-H,spill} \cdot \epsilon_{n-H,Spill}, \tag{6.6}$$

where subscripts T, GC, and *spill* denote a parameter for events originating from the Target or γ-catcher, or classified as Spill In/Out events. Accordingly, the delayed event detection efficiency is

$$\epsilon_{delayed} = \underbrace{f_{n-H,T} \cdot \epsilon_{n-H,T} \frac{N_{p,T}}{N_{p,all}}}_{\text{Target}} + \underbrace{f_{n-H,GC} \cdot \epsilon_{n-H,GC} \frac{N_{p,GC}}{N_{p,all}}}_{\gamma\text{-catcher}} + \underbrace{f_{n-H,spill} \cdot \epsilon_{n-H,spill} \frac{N_{p,spill}}{N_{p,all}}}_{\text{Spill In/Out}} \quad (6.7)$$

There is no need to understand the absolute efficiency value because the θ_{13} analysis compares the DATA to the MC sample on which the same set of candidate selection cuts are applied. The systematic uncertainties of each of the parameters in Eq. 6.7, however, need to be propagated appropriately to quantify the uncertainty on $\epsilon_{delayed}$, which is necessary information for the oscillation fit. In this section, uncertainties on the necessary parameters in Eq. 6.7 are computed, and then combined to estimate the overall $\epsilon_{delayed}$ uncertainty. Section 6.2.1 and 6.2.2 cover the Target-related and γ-catcher-related efficiencies and uncertainties. Section 6.2.4 presents how those are combined and put together in Eq. 6.7. Section 6.2.3 describes an additional uncertainty on the neutron mobility modeling in MC, and the final uncertainty value is presented in Sec. 6.2.6.

6.2.1 Fraction of Neutron Capture on Hydrogen

The neutron capture fraction on hydrogen, carbon, and Gd nuclei in the Target is studied in detail in the Gd analysis [1]. Calibration data with the ^{252}Cf neutron source is used to quantify a capture fraction of Gd isotopes. The source is deployed at the central region of the Target where negligible neutrons from ^{252}Cf fission could escape from the Target liquid. After applying coincidence selection cuts to obtain pure delayed neutron capture events, the E_{delayed} spectrum is integrated from 4 to 11 MeV, and a ratio is taken to the integral of 0.7 to 11 MeV. Two independent analyses provided 0.8657 $\pm$ 0.0011 and 0.8658 $\pm$ 0.0015 for DATA and 0.8785 $\pm$0.0007 and 0.8798 $\pm$ 0.0007 for MC [35]. The latter is taken for DATA as it carries a larger uncertainty, while a combined average is taken for MC.

The hydrogen capture fraction is quantified by subtracting these numbers from 1.0. This procedure assumes that a capture fraction on carbon nuclei is included in the integration range

of 4 to 11 MeV. This is a fair assumption as neutron capture on carbon produces a 4.95 MeV signal, and its overall fraction is in the order of 0.1 % from a capture cross section. Accordingly, neutron capture on carbon falling into energy lower than 4.0 MeV should be a negligible fraction of all neutron captures.

$$f_{n-H,T}^{DATA} = 1.0 - f_{4\ to\ 11\ MeV}^{DATA} = 13.42 \pm 0.15\% \qquad \text{(DATA)} \tag{6.8}$$

$$f_{n-H,T}^{MC} = 1.0 - f_{4\ to\ 11\ MeV}^{MC} = 12.08 \pm 0.10\% \qquad \text{(MC)} \tag{6.9}$$

where $f_{n-H,T}$ denotes the neutron capture fraction in the Target volume. Accordingly, the MC correction factor, $\mathcal{F}_{n-H}$, for a capture fraction becomes

$$\mathcal{F}_{n-H,T} = 1.1109 \pm 0.0155 \tag{6.10}$$

A hydrogen capture fraction in the γ-catcher is more important for the hydrogen analysis as the majority of signal events are expected to arise from the γ-catcher volume. Because of the geometry of γ-catcher, unlike the Target, it is hard to choose one source deployment position and claim there is negligible neutron escape to the surrounding volumes. Instead, the $\bar{\nu}_e$ delayed spectrum is used to quantify the neutron capture fraction on hydrogen in the γ-catcher. Some of the IBD selection cuts are modified for this study: $0.7 < E_{\text{delayed}} < 11$ MeV; $3.5 < E_{\text{prompt}} < 12.2$ MeV; and $90 < \Delta T < 600\ \mu s$. A modification to E_{prompt} is to reduce accidental backgrounds (see Sec. 7.1), and the increased lower boundary of the ΔT cut is to exclude neutron capture in the Target liquid. In addition, a vertex cut of $\rho_{delayed} < 1150$ mm OR $Z_{delayed} < -1229$ mm OR $Z_{delayed} > 1229$ mm is applied to further purify the sample in the γ-catcher liquid. Resulting E_{delayed} distributions from DATA and MC have an excellent agreement as shown in Fig. 6-3.

The fraction of delayed events with energy between 0.7 and 3.0 MeV is considered be from neutron capture on hydrogen. Accordingly, a capture fraction is found as

$$f_{n-H,GC}^{DATA} = 98.3 \pm 0.2\ \% \qquad \text{and} \qquad f_{n-H,GC}^{MC} = 98.3 \pm 0.2\ \%. \tag{6.11}$$

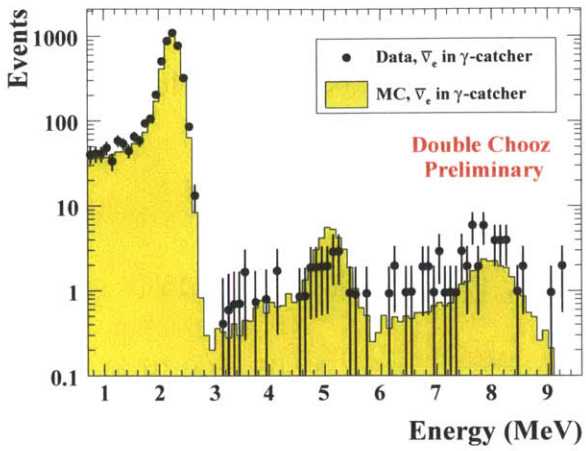

Figure 6-3: E_{delayed} distribution from an $\bar{\nu}_e$ sample selected with a modified set of cuts (see the main text). The majority of neutron captures making this distribution are expected to be from the γ-catcher volume.

The MC correction factor is, accordingly,

$$\mathcal{F}_{n-H,GC} \;=\; 1.000 \pm 0.003 \tag{6.12}$$

It should be noted that the ΔT and a vertex cut may not remove neutron capture on Gd isotopes entirely, as it is possibly seen in the energy region above 6 MeV in Fig. 6-3. However, events in this energy range could also be due to neutron capture on iron and other metal nuclei in the Guide-Tube. Therefore, it is not clear that these higher-energy capture events all come from a contamination of neutron captures on Gd. Nevertheless, we attempt to quantify a possible DATA-MC bias due to events in this energy region to validate our estimate above.

The fraction of delayed events in 6 to 11 MeV out of the whole energy range, 0.7 to 11 MeV, is 1.15 $\pm$ 0.17 % in DATA and 0.7 $\pm$ 0.14 % in MC. In the most pessimistic scenario, all events in the 6 to 11 MeV energy range come from the residual neutron capture on Gd in the Target. If this is true, there must also be a corresponding amount of neutron capture on hydrogen events from the Target contaminating the sample. Using $f_{n-H,T}$, one can estimate a bias contribution to $f_{n-H,GC}$. This yields 0.16 $\pm$ 0.02 % and 0.08 $\pm$ 0.02 % for DATA and MC, respectively. Both of these values are well covered by the assigned statistical uncertainty of 0.2 % in both

cases, and hence these pessimistic values are considered negligible.

6.2.2 ΔT, ΔR, and E_{delayed} Efficiency Study Using ^{252}Cf

The ^{252}Cf neutron source has been deployed at various positions along the Z-Axis and Guide-Tube. In order to study each volume, data points in each volume have to be somehow combined to a representative number. For this reason, source data from the Z-Axis system is used for the Target and that of the Guide-Tube is used for the γ-catcher study. Using the ^{252}Cf data sample described in Sec. 6.1.2, the following efficiency definitions are used to compare DATA and MC:

$$\epsilon_{\Delta T} = \frac{1.5 < E_{\text{delayed}} < 3.0 \text{ MeV} \otimes \Delta R < 900 \text{ mm} \otimes 10 < \Delta T < 600 \ \mu\text{s}}{1.5 < E_{\text{delayed}} < 3.0 \text{ MeV} \otimes \Delta R < 900 \text{ mm} \otimes 0 < \Delta T < 1000 \ \mu\text{s}}; \tag{6.13}$$

$$\epsilon_{\Delta R} = \frac{\Delta R < 900 \text{ mm} \otimes 1.5 < E_{\text{delayed}} < 3.0 \text{ MeV} \otimes 10 < \Delta T < 600 \ \mu\text{s}}{1.5 < E_{\text{delayed}} < 3.0 \text{ MeV} \otimes 10 < \Delta T < 600 \ \mu\text{s}}; \tag{6.14}$$

$$\epsilon_{E_{\text{delayed}}} = \frac{1.5 < E_{\text{delayed}} < 3.0 \text{ MeV} \otimes 10 < \Delta T < 600 \ \mu\text{s} \otimes \Delta R < 900 \text{ mm}}{0.7 < E_{\text{delayed}} < 3.0 \text{ MeV} \otimes 10 < \Delta T < 600 \ \mu\text{s} \otimes \Delta R < 900 \text{ mm}}. \tag{6.15}$$

Using these definitions, DATA and MC efficiencies from ^{252}Cf calibration source deployment positions in the Target and γ-catcher are computed and shown in Fig. 6-4. The data points are plotted against Z, a position along the vertical Z-Axis, for the Target and R, a distance from the detector center, for the γ-catcher.

In order to quantify the "Target and γ-catcher efficiency," data points shown in Fig. 6-4 must be somehow combined to represent the each volume. A combination of data points is done through the *sampling method.* This method proceeds as follows: (1) plot the measured efficiency data points along Z for the Target and R for the Guide-Tube deployment points (see the left column of Fig. 6-4), then interpolate linearly between these points; (2) take three thousand equally spaced sampling points along the interpolated straight lines between data points; (3) for the γ-catcher take the mean of the sampling points as the efficiency mean, weighted by the associated γ-catcher LS volume element at the given R (the shape of this weighting is shown in Fig. 6-5); (4) for the Target take a simple mean over the Z-Axis as the overall efficiency. Note the weighting by the LS volume is needed in the γ-catcher because the LS volume element is

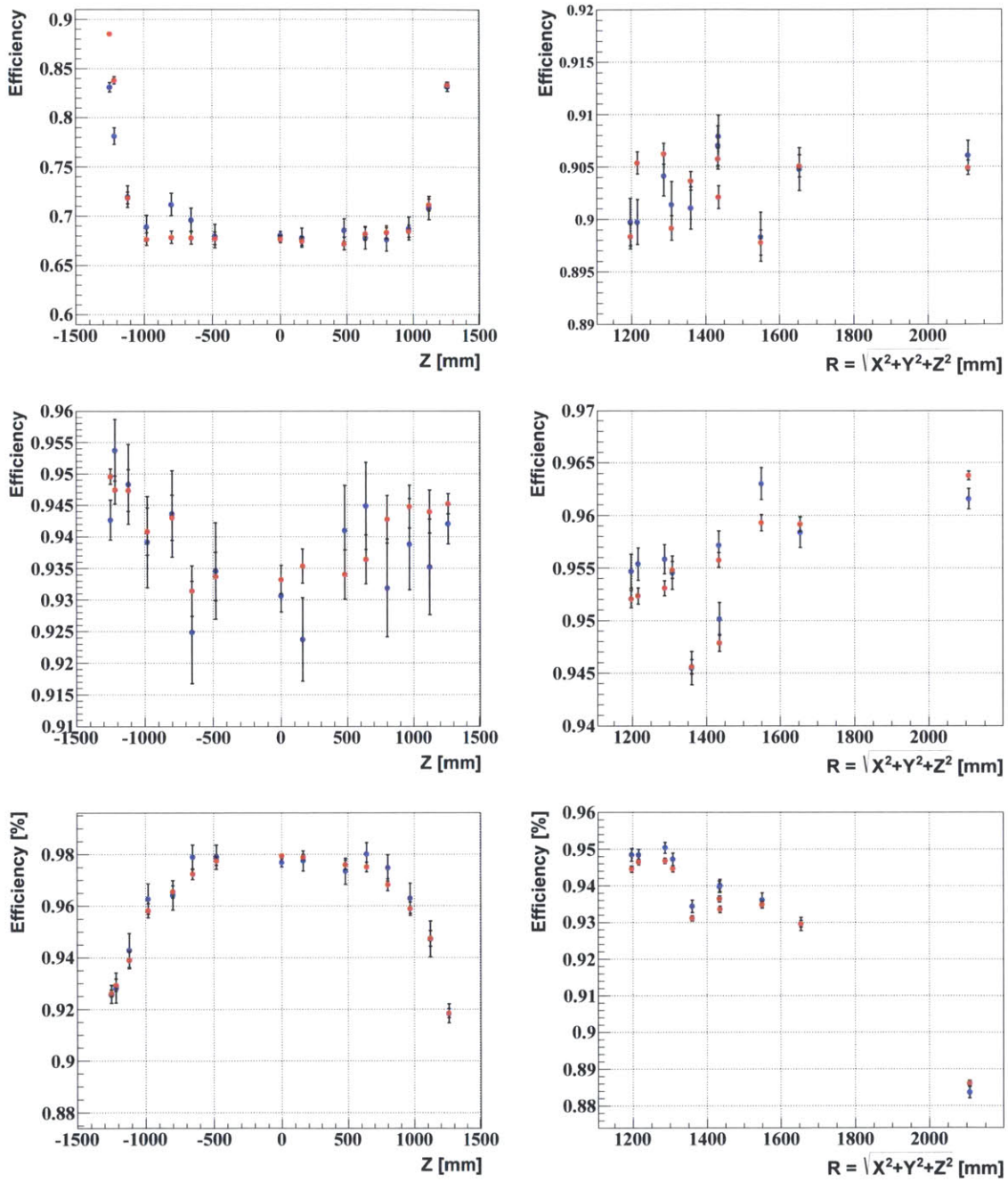

Figure 6-4: From top to bottom row: ΔT, ΔR, and E_{delayed} efficiency plotted against Z for the Target (left column) and R for the γ-catcher (right column). For each plot, the MC correction factor is applied to match the overall volume efficiency.

not uniform at a given R. A summary table of computed cut efficiency in each volume is given

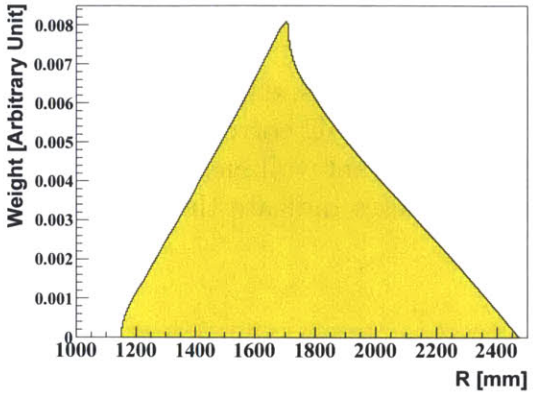

Figure 6-5: The γ-catcher liquid scintillator volume distribution as a function of $R = \sqrt{x^2 + y^2 + z^2}$. The vertical axis shows the volume in arbitrary units.

in Tab. 6.1.

Having measured both DATA and MC efficiencies using the sampling method, an overall efficiency difference between the DATA and MC can be easily quantified as the MC correction factor. This is derived from the DATA/MC efficiency ratio distribution. An example of the ΔT efficiency distribution along the Z-Axis is shown in Fig. 6-6. Simply applying the sampling

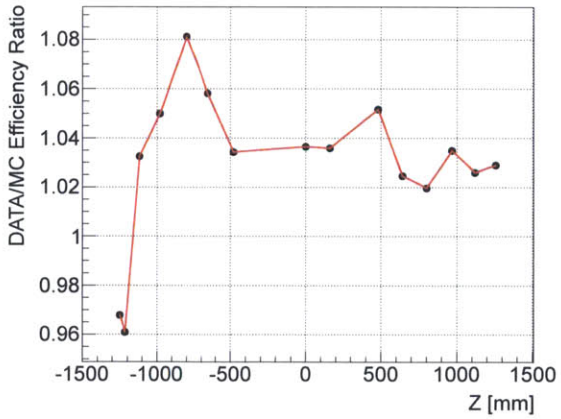

Figure 6-6: DATA/MC ΔT efficiency ratio over Z-Axis. The sampling points taken to estimate the overall MC correction factor are shown in red.

method to the ratio distribution, an overall correction factor is derived. Note that the correction

Cut	$\mathcal{F}_T$	ϵ_T (%)	$\mathcal{F}_{GC}$	ϵ_{GC} (%)
ΔT	1.0374	69.55 ± 1.78 (1.71)	0.9041	90.41 ± 0.16 (0.11)
ΔR	0.9955	93.61 ± 0.51 (0.46)	0.9920	95.84 ± 0.19 (0.17)
E_{delayed}	0.9886	96.79 ± 0.29 (0.25)	0.9920	92.20 ± 0.23 (0.20)

Table 6.1: Summary of the MC correction factors, efficiencies and uncertainties from the ^{252}Cf neutron source study. $\mathcal{F}$ and ϵ represent the MC correction factor and efficiency, respectively. The subscripts T and GC denote the relevant volume, which is either Target or γ-catcher, respectively. Numbers inside the parentheses indicate the systematic error component of the quoted total error.

factor is a constant number with no vertex dependency. The MC efficiency distributions shown in Fig. 6-4 have a relevant MC correction factor applied. MC correction factors can be found in Tab. 6.1.

Finally, a remaining DATA-MC discrepancy after applying a correction factor should be carried as a systematic uncertainty of the estimation method. This is derived, again, using the sampling method on the DATA-MC discrepancy distribution. The discrepancy between DATA and MC is defined as

$$\delta_\epsilon = \frac{\epsilon_{DATA} - \epsilon^*_{MC}}{\frac{1}{2}\left(\epsilon_{DATA} + \epsilon^*_{MC}\right)} \tag{6.16}$$

where ϵ_{DATA} and ϵ^*_{MC} represent DATA and corrected MC efficiency respectively. This is computed at each source deployment position as shown in Fig. 6-7.

The mean of sampling points from this distribution should be very close to 0 unless otherwise non-zero, since the MC overall correction factor is applied a priori. The RMS of sampling points, however, represents how detection efficiency can vary along the projected vertex in the given volume. This is taken as the uncertainty arising from an underlying vertex dependency of the MC efficiency distribution that is not modeled identically to match DATA. Uncertainties of all cuts are summarized in Tab. 6.1.

Note that, however, the discussed uncertainty is not only a systematic difference between DATA and MC but also include a statistical fluctuation. It is really a combined uncertainty. In order to ensure the statistical fluctuation is well accounted by the RMS of the DATA-MC discrepancy distribution, a cross-check analysis is carried out using a method of fake measurements. A fake measurement takes an input of measured efficiency and error at each

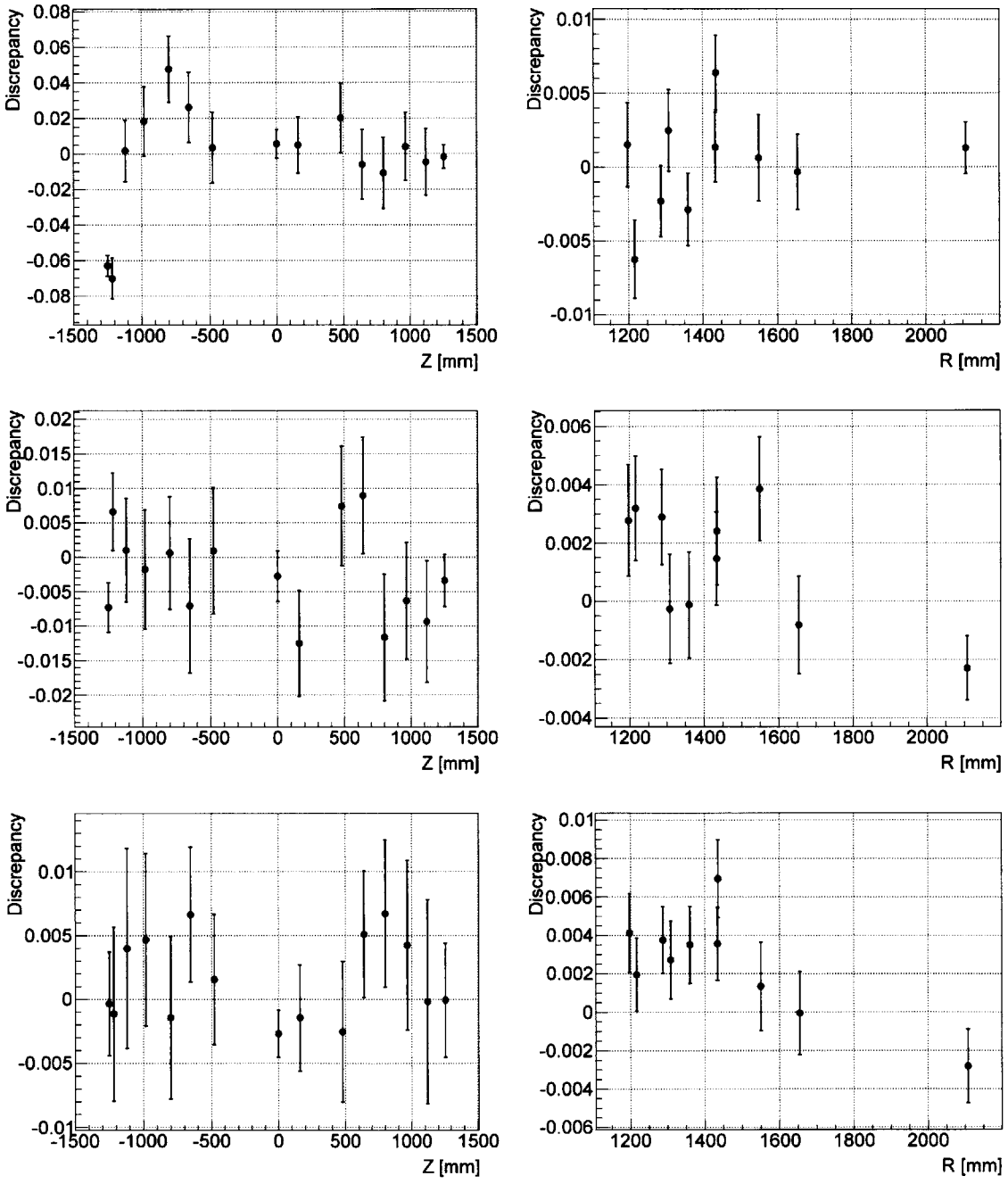

Figure 6-7: From top to bottom: ΔT, ΔR, and E_{delayed} efficiency distributions in the Target (left column) and the γ-catcher (right column), respectively.

source deployment position (see, for example, data points along the Z-Axis shown in Fig. 6-4). Using a normal distribution, which is a fine assumption for a high-statistics sample with an efficiency well away from 0 and 1, a fake data point is drawn. One set of fake measurements make a fake distribution along the Z-Axis. Applying a sampling method to this set of fake measurements, a fake efficiency mean can be obtained. This fake efficiency takes into account a statistical fluctuation at each source deployment position, and hence a fluctuation on this fake efficiency indicates a statistical fluctuation in the sampling method. For each sampling method, 10,000 fake measurements are performed, and the width of the resulting distribution is taken as a statistical fluctuation.

The uncertainty derived from the sampling method is checked to make sure the statistical fluctuation is included. An example of this fake measurement is shown in Fig 6-8. This method

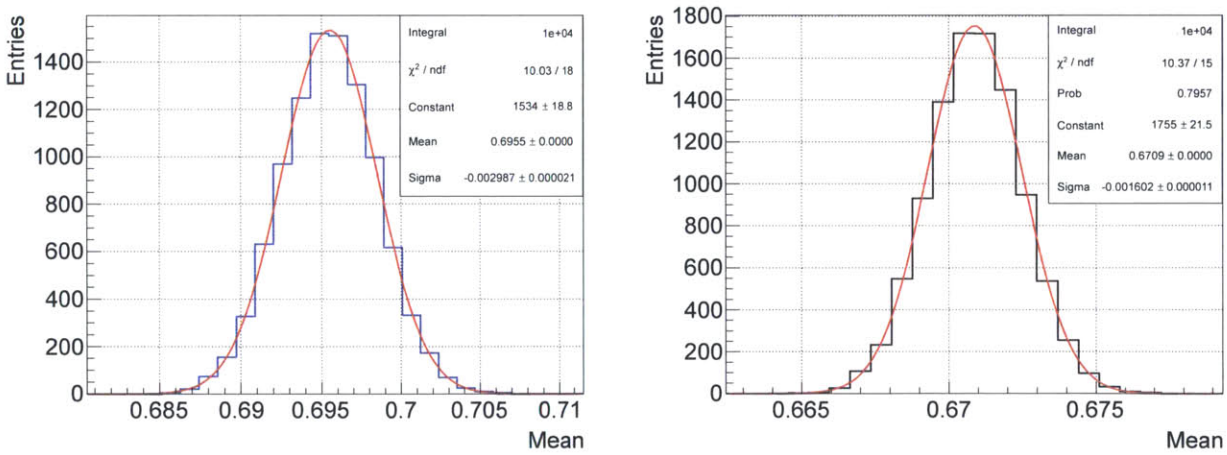

Figure 6-8: Resulting efficiency distributions from 10,000 fake experiments using a ΔT distribution shown in Fig. 6-4. Left is for DATA and the right is for MC.

of fake measurements is used to derive the statistical and systematic uncertainty breakdown in Tab. 6.1. One can see that in most cases a quoted uncertainty is dominated by the systematic component.

6.2.3 Spill In/Out

Some $\bar{\nu}_e$ interact with a proton in the Buffer just outside the γ-catcher boundary, and there is a small yet finite chance for them to be selected. There is a small fraction of IBD candidates like this, originating from neither Target nor γ-catcher volume, and they are classified in a category called Spill In/Out. The definition of this category is any event where prompt and delayed are not occurring in the same sensible volume, Target or γ-catcher. The name comes from the fact that this is due to the mobility of a neutron since its creation at the IBD vertex to the position of a capture. If a prompt is in the Target and its delayed occurs in the γ-catcher, therefore, it is also classified as Spill In/Out. Volumes that are relevant for this effect include the Target liquid, γ-catcher liquid, Target acrylic tank, γ-catcher acrylic tank, and Buffer liquid. They are expressed as "T-LS," "γ-LS," "T-Tank," "γ-Tank," and "Buffer" in this section.

Using MC IBD candidate samples for which truth event vertex information is available, the fraction of Spill In/Out events is found to be 4.59 %. "Pure" Target events where prompt and delayed both occurred in the Target liquid are found to be 5.40 %, and the "pure" γ-catcher event fraction is 90.01 %. As expected, the majority of IBD candidates are categorized in the "pure" γ-catcher events. Table 6.2 summarizes the fraction of IBD candidates that participate in Spill In/Out effect. It is worthwhile to note that the nature of IBD prefers the prompt to

Category	Volume	Event Fraction (%)
Prompt	T-LS	1.12
	γ-LS	2.18
	T-Tank	0.73
	γ-Tank	0.26
	Buffer	0.31
Delayed	T-LS	0.36
	γ-LS	1.78
	T-Tank	0.60
	γ-Tank	0.59
	Buffer	1.27

Table 6.2: Summary of the event fraction of various Spill In/Out sub-categories. The total to define a fraction is the total number of IBD candidates that are selected in the MC sample.

occur in the LS region, while the delayed could occur either in the LS volume or nearby non-LS

volume. This is because the prompt event, which is an annihilation of e^+, has relatively local energy release compared to the delayed which a compton scattering of a 2.223 MeV γ-ray from neutron capture on hydrogen. delayed neutron capture may occur in the Buffer and a capture γ-ray could jump into the LS region, while this is harder for the prompt event. As shown in Tab. 6.2, the fraction of spill events with a delayed event in the Buffer is bigger than those with its prompt occurring in the Buffer.

The hydrogen capture fraction and selection cut efficiencies in the Target and γ-catcher are described in Sec. 6.2.2 and 6.2.1, respectively. Equation 6.7 requires, however, understanding of the contribution from Spill In/Out events. A difficulty of Spill In/Out events is that they are very hard to study with DATA because there is no information about the truth event vertex, which makes it hard to separate Spill events from those they are not. Fortunately, however, the Guide-Tube runs along the wall of Target Tank as well as the γ-catcher tank where Spill events are expected to take place (see Fig. 3-10) because a considerable fraction of neutrons emitted from the ^{252}Cf source can move across the boundary of the γ-catcher. This allows an assumption to be made that the efficiency uncertainties derived from the Guide-Tube data sample may cover both the pure γ-catcher events as well as the Spill events.

This assumption can be verified nicely by comparing the ^{252}Cf study result with another independent data sample, $\bar{\nu}_e$ delayed event candidates (see Sec. 6.1.3). According to the assumption, Spill events are considered to carry the efficiency and uncertainties quantified with Guide-Tube data. From the MC event fraction noted above, the γ-catcher and Spill events constitute 94.60 % of the total, while the Target events are 5.40 %. Table 6.3 summarizes the combined volume efficiency uncertainty and MC correction factor for each cut.

For the $\bar{\nu}_e$ data sample, exactly the same definition of each cut efficiency noted in Sec. 6.2.2 is used except for the ΔT cut: the lower boundary of this cut value to define the total number of events is changed to 8 μs in order to suppress stopping-muon backgrounds. The contamination of the stopping-muon is visible in how the DATA-MC discrepancy changes as a function of the lower boundary cut value, as shown in Fig. 6-9. As a result, DATA-MC discrepancies are obtained and summarized in Tab. 6.3. The observed values are in agreement with the expected values within their statistical uncertainties. Based on this result, the assumption is considered

	$\mathcal{F}$	Expected $\delta(\epsilon)$ (%)	Observed $\delta(\epsilon)$ (%)
ΔT	0.9988	± 0.19	-0.36 ± 0.34
ΔR	0.9950	± 0.18	0.16 ± 0.27
E_{delayed}	0.9918	± 0.22	-0.33 ± 0.38

Table 6.3: MC correction factors, $\mathcal{F}$, and DATA-MC systematic uncertainty, $\delta(\epsilon)$. The expected $\delta(\epsilon)$ is derived from ^{252}Cf source data. The MC event fraction is used to determine the weight of each parameter between Target and γ-catcher values. The observed values are derived from the $\bar{\nu}_e$ data sample. The quoted uncertainty on the observed value represents the statistical uncertainty in the data set.

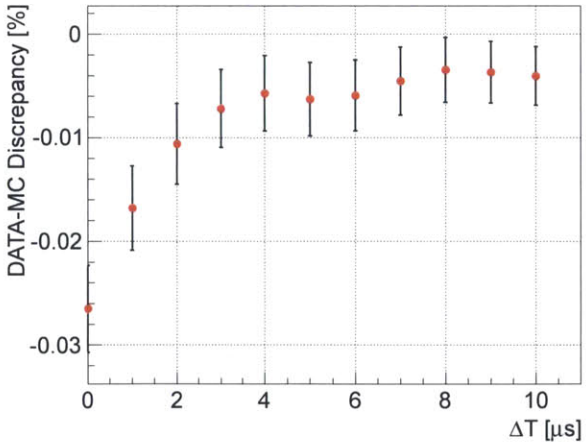

Figure 6-9: Variation of DATA-MC discrepancy as a function of the lower boundary cut value to define the full sample for the efficiency study. A fit to the distribution yields a consistent time constant with a muon lifetime of 2.2 μs.

Delayed Volume	$\mathcal{F}_{n-H}$	$\delta(f_{n-H})$ (%)	MC Event Fraction
Target	1.111	1.12	5.75
γ-catcher	1.000	0.31	91.79
T-Tank	1.000	0.45	0.60
γ-Tank	1.000	0.39	0.59
Buffer	1.000	0.31	1.27

Table 6.4: A summary table for the hydrogen capture fraction MC correction factors and uncertainties. The uncertainty is a quadrature sum of the DATA-MC discrepancy for the pure γ-catcher events and the difference in the MC hydrogen capture fraction quantified with the MC true vertex information. The Target and γ-catcher uncertainties are taken from Sec. 6.2.1.

to be valid: Spill events' detection efficiency uncertainties are considered to be covered by using the γ-catcher efficiency uncertainties extracted from the Guide-Tube data sample.

The neutron capture fraction is related to the delayed event vertex. For Spill events, therefore, one could use the MC event fraction summarized in Tab. 6.2 to assign the uncertainty of the hydrogen capture fraction. In the MC sample, the neutron capture fraction is quantified for delayed events using the MC true delayed vertex information. The pure γ-catcher events are found to carry 99.14 % of the capture fraction on hydrogen. It is found that, for delayed events occurring in the T-Tank, γ-Tank, and Buffer volumes, the hydrogen capture fractions are 98.82 %, 98.91 %, and 99.16 %. Because all of these are very close to the pure γ-catcher value, the systematic uncertainties in these volumes are computed by adding in quadrature the γ-catcher DATA-MC systematic uncertainty and the difference of the hydrogen capture fraction between the subject volume and the pure γ-catcher events. The MC correction factor is directly quoted from the γ-catcher measurement, which is 1.00 (i.e., no correction). These are summarized in Tab. 6.4.

Finally, for Spill events, the uncertainty on $N_{p,Spill}$ needs to be quantified to follow Eq. 6.7. Following the procedure taken in the past Gd analysis [11], a difference between two MC models of the neutron thermalization process is taken as a systematic uncertainty. This turns out to be 1.17 % [36] which is a large part of the total neutron detection efficiency described in the next section. This methodology needs to be improved, therefore, in order to improve the sensitivity of the hydrogen analysis.

6.2.4 Overall Delayed Selection Uncertainty

Presumably there should be no inter-volume correlation between the fraction of neutron capture on hydrogen and number of protons in each volume. From the efficiency study using the $\bar{\nu}_e$ sample in Sec. 6.2.3, however, it was found that a simple quadrature sum of an efficiency uncertainty from each volume, weighted by the MC event fraction, agrees with the observed combined uncertainty estimates. This suggests that an inter-volume correlation of cut efficiency uncertainty is negligible. Therefore, a simple quadrature sum of an uncertainty from each parameter in Eq. 6.7 is used to estimate the overall efficiency uncertainty as shown below.

$$\delta(\epsilon)^2 = \delta(f_{n-H})^2 + \delta(\epsilon_{\Delta T})^2 + \delta(\epsilon_{\Delta} R)^2 + \delta(\epsilon_{E_{\text{delayed}}})^2 + \delta(N_p)^2 \tag{6.17}$$

where each term represents the squared uncertainty of a parameter in the parentheses for all volumes combined. Table 6.3 shows the combined volume uncertainty and MC correction factor for ΔT, ΔR, and E_{delayed} cuts. Using the delayed MC event fraction in Tab. 6.2 and uncertainty values from Tab. 6.4, the hydrogen capture fraction uncertainty for the combined volume is 0.29 % with the MC correction factor of 1.0064. With the number of protons summarized in Tab. 3.2 and the prompt MC event fraction summarized in Tab. 6.2, one obtains the overall proton number uncertainty of 0.94 % with the MC correction factor of 1.0028. Altogether, the combined detection efficiency uncertainty is 1.04 % with the overall MC correction factor of 0.9948.

6.2.5 Inefficiency on Delayed Events From Muon Veto

So far the study presented in this section has focused on the DATA and MC difference in the detection efficiency that arises from either the detector effect or the applied selection cuts. There is one more inefficiency that needs to be taken into account and is specific to DATA: an inefficiency from vetoing a delayed event candidate with a muon after the prompt. If a muon hit the detector after the prompt but before the delayed event of an IBD event, this candidate pair is vetoed due to the delayed event falling in the after-muon veto period of 1 ms.

An inefficiency due to this effect can be easily calculated from DATA. A delayed neutron

ΔT distribution can be obtained from the $\bar{\nu}_e$ candidate sample, as shown in Fig. 6-2. The rate of after-muon veto window is quoted in Sec. 7.1. One can take the ΔT distribution in Fig. 6-2 as the probability distribution function (PDF), and compute a probability of a delayed event candidate to be vetoed by a muon that occurs randomly at the quoted rate. This additional MC correction factor to compensate for the inefficiency is found to be 0.9929.

6.2.6 Combining All

From Sec. 6.2.4 and 6.2.5, the overall delayed event detection efficiency uncertainty is found to be 1.04 % with the MC correction factor of 0.9877. Finally, though it is not a part of "efficiency," there is an uncertainty on the number of detected delayed events due to an uncertainty on the neutron thermalization model. This is described at the end of Sec. 6.2.3, and is 1.17 % relative to the selected number of MC IBD candidates. Adding these terms in quadrature, therefore, the overall uncertainty on the selected number of delayed event candidates becomes 1.57 %.

6.3 Pre-Selection Efficiency

The pre-selection efficiency depends on the after-muon veto, MQTQ, and $T_{\rm start}^{\rm RMS}$ cuts. The 1 ms after-muon veto is taken care of by the detector livetime calculation. The MC sample is generated for the cumulative run time of 251.27 days, while the livetime of DATA is 240.06 days. This results in a correction factor of 0.9554.

Both MQTQ and $T_{\rm start}^{\rm RMS}$ cuts are considered to carry negligible inefficiency in the Gd analysis. It turns out they are even less of a problem in γ-catcher events that constitute most of the hydrogen analysis candidate sample. The main difference between the hydrogen and the Gd analysis regarding these light noise rejection cuts is the nature of the event vertex. The IBD candidates in the hydrogen analysis are expected to come from the γ-catcher, while those in the Gd analysis come from the Target region. In order to study their vertex dependency, the same ^{252}Cf data sample used in the delayed detection efficiency is used (see Sec. 6.1.2). The same selection cuts are applied except for the MQTQ and $T_{\rm start}^{\rm RMS}$ cuts applied on the delayed events that are modified to inspect the efficiency.

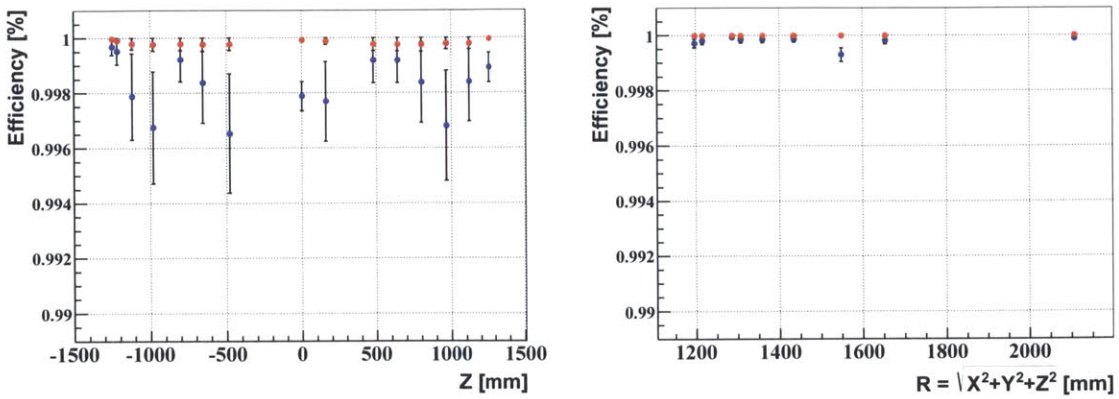

Figure 6-10: MQTQ DATA (blue) and MC (red) efficiency distribution from Z-Axis (left) and Guide-Tube (right). There is no MC correction factor applied for the plotted data points.

6.3.1 Charge Uniformity (MQTQ) Cut

In this study, the MQTQ cut efficiency is defined as

$$\epsilon_{MQTQ} = \frac{T_{\text{start}}^{\text{RMS}} < 40\ ns \otimes \text{MQTQ} < 0.09}{T_{\text{start}}^{\text{RMS}} < 40\ ns \otimes \text{MQTQ} < 0.2} \tag{6.18}$$

Both DATA and uncorrected MC efficiency distributions are shown in Fig. 6-10. Using the sampling method described in Sec. 6.2, the Target and γ-catcher efficiencies are quantified and summarized in Tab. 6.5. The quoted uncertainty in the table is purely statistical.

	Target	γ-catcher
DATA	99.80 $\pm$ 0.05	99.98 $\pm$ 0.01
MC	99.98 $\pm$ 0.01	99.999 $\pm$ 0.001

Table 6.5: DATA and uncorrected MC efficiencies for MQTQ cut. Errors are statistical only.

There are two points worth noting in Fig. 6-10. The first is a large DATA-MC difference at one source position in the Guide-Tube around R $\approx$ 1560 mm. This point is along the γ-catcher wall and, hence, expected to have the largest DATA-MC discrepancy. It is interesting to see, however, another source deployment at the γ-catcher edge ($R \approx 2110\ mm$) shows a good DATA-MC agreement. It could be a statistical fluctuation, since the point at R $\approx$ 2110 mm has been taken with much higher DATA statistics. The second point is rather surprising, and

it is the fact that the Z-axis data points in the Target seem to have a rather worse DATA-MC discrepancy. This might appear quite counter-intuitive because a better uniformity of light, which results in a smaller inefficiency in the MQTQ cut, is expected in the Target compared to the γ-catcher region.

The actual DATA and MC MQTQ distributions at the Target center (Z=0) and γ-catcher edge ($R \approx 2110\ mm$) are shown in Fig. 6-11. For the Target center data sample, there are some residual events in the region MQTQ > 0.1 that bring the mean efficiency below 1 for DATA. From the fact that the mean of the MQTQ distributions for DATA and MC are both well below 0.09 and the absence of these residual events in MC, it is quite likely these residual events with MQTQ > 0.09 are non-physics events. This may be consistent with the fact there are some LN events reconstructed at the Target center (see Sec. 7.5).

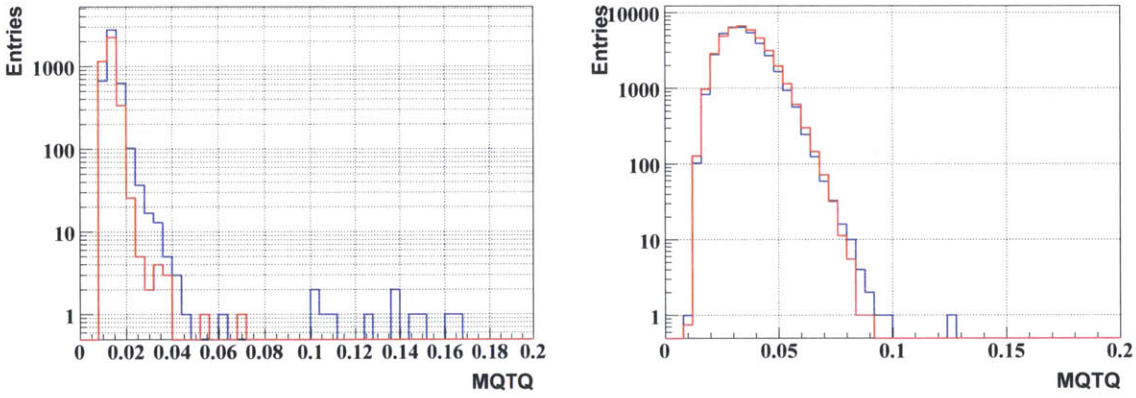

Figure 6-11: MQTQ DATA (blue) and MC (red) distributions for the Target center ^{252}Cf (left) and γ-catcher edge point (right).

The discrepancy between DATA and MC can be obtained using the same sampling method used for the delayed efficiency study (see Sec. 6.2). This discrepancy is found to be as small as 0.09 % in the Target, and is largely driven by the fact that the DATA efficiency distribution has an inefficiency around the Target center. Therefore, this could be an artifact of having the LN at the Target center. In the γ-catcher volume, the discrepancy is 0.01 %, which is negligible. Because the Target discrepancy may arise from the LN contamination, and also because it only makes up a small fraction of the hydrogen analysis IBD candidate sample, the

overall DATA-MC discrepancy in the MQTQ cut efficiency is considered to be negligible.

6.3.2 Photon Arrival Time Cut ($T_{\mathbf{start}}^{\mathbf{RMS}}$)

The definition for the $T_{\text{start}}^{\text{RMS}}$ cut efficiency is

$$\epsilon_{T_{\text{start}}^{\text{RMS}}} = \frac{T_{\text{start}}^{\text{RMS}} < 40\ ns \otimes \text{MQTQ} < 0.09}{T_{\text{start}}^{\text{RMS}} < 60\ ns \otimes \text{MQTQ} < 0.09}. \tag{6.19}$$

Following the same procedure taken for the MQTQ study in the previous section, both the DATA and MC $T_{\text{start}}^{\text{RMS}}$ efficiency along the Z-Axis and Guide-Tube are shown in Fig. 6-12. Again, the MC correction factor is not applied. The resulting efficiency and associated systematic uncertainties are summarized in Tab. 6.6. One can see that the $T_{\text{start}}^{\text{RMS}}$ cut efficiency

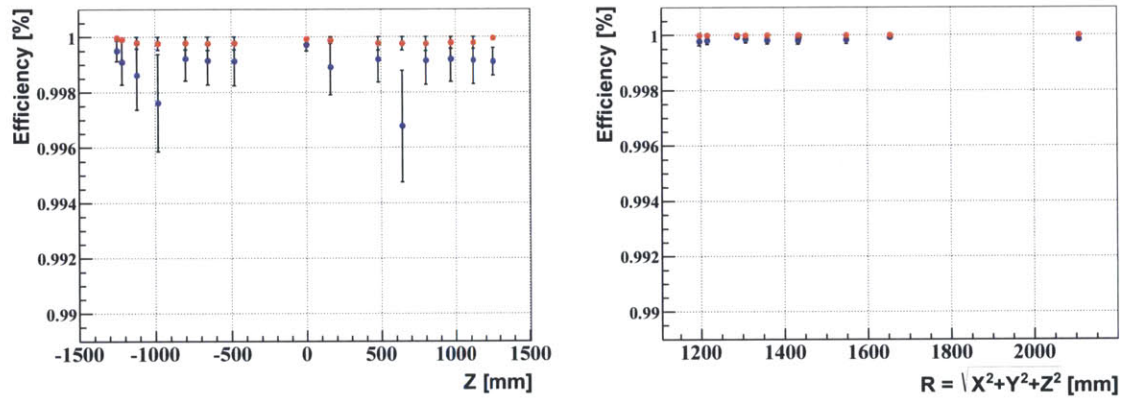

Figure 6-12: $T_{\text{start}}^{\text{RMS}}$ cut efficiency along Z-Axis (left) and Guide-Tube (right). Blue data points denote DATA and red colored dots denote MC.

	Target	γ-catcher
DATA	99.89 $\pm$ 0.04	99.98 $\pm$ 0.01
MC	99.99 $\pm$ 0.01	99.999 $\pm$ 0.001

Table 6.6: $T_{\text{start}}^{\text{RMS}}$ cut efficiencies for DATA and MC without a correction factor. Errors are statistical only.

distributions follow the same trend as the MQTQ distributions: there is a larger inefficiency along the Z-Axis in the Target compared to the γ-catcher region. Figure 6-13 compares the

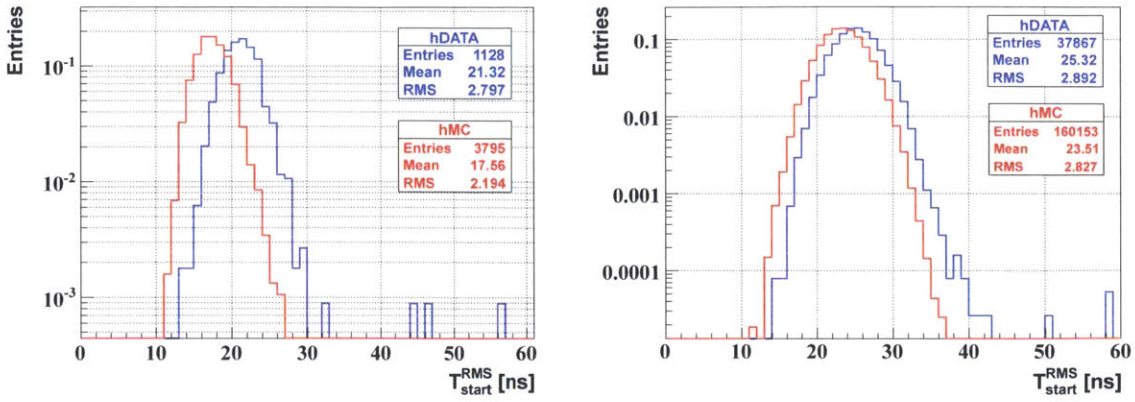

Figure 6-13: $T^{\mathrm{RMS}}_{\mathrm{start}}$ DATA (blue) and MC (red) distributions for the Target center ^{252}Cf (left) and γ-catcher edge point (right).

worst points in the Target (along the Z-Axis at Z $\approx$ 640 mm) and the γ-catcher (edge point at R $\approx$ 2110 mm). Like the MQTQ case, there are some residual events at a high value of $T^{\mathrm{RMS}}_{\mathrm{start}}$ in the left plot. Those residual events do not seem to follow the pattern of the majority of (physics) events that make a peak around 21 ns. These events are considered, therefore, a contamination of non-physics events such as LN. As concluded for the MQTQ study, the DATA-MC discrepancy and inefficiency to the signal due to $T^{\mathrm{RMS}}_{\mathrm{start}}$ cut is considered to be negligible from these results.

6.4 Summary

Referring to all preceding sections in this chapter, this section summarizes the overall uncertainty on the selected number of IBD candidates and the MC correction factor. The pre-selection process carries a correction factor of 0.9554 with negligible uncertainty, as described in Sec. 6.3. Section 6.1.1 described the MC correction factor of 0.9813 for the prompt candidate selection process with, again, a negligible uncertainty. Finally, there is 1.57 % uncertainty with the MC correction factor of 0.9877 for the number of delayed candidates selection.

Combining all these factors, the selected IBD candidates in the MC sample need to be scaled by the correction factor of 0.9260. The uncertainty on the resulting number of selected

candidates is 1.57 %.

Chapter 7

Background Prediction

This chapter presents the background prediction for the hydrogen analysis. Backgrounds in the Gd analysis are accidentals, ^{9}Li, fast neutrons, and stopping-muons [1]. These are studied in detail also for the hydrogen analysis. In addition, the correlated light noise background is found and added into the prediction. Section 7.6 summarizes the mean rate and associated uncertainty of each background considered.

The reactor OFF-OFF DATA was taken while the two reactor cores were shut down. Although the livetime of this sample is only about one week, it can be used for a cross-check of the background prediction because it contains only backgrounds and no signal. A simple cross-check is performed with the reactor OFF-OFF DATA, and presented in Sec. 7.7.

7.1 Accidentals

This section describes the analysis method used to understand an accidental coincidence background, referred to as "accidentals" in this document. Accidentals are two uncorrelated physics events occurring in the right time and space separation such that they pass our IBD coincidence selection cuts by accident. A large fraction of accidentals consist of two uncorrelated single physics events (called "singles") dominated by natural radioactivity.

Compared to the Gd analysis, the hydrogen analysis suffers from an accidentals' background rate that is orders of magnitude higher. This is because the fiducial volume of the hydrogen

analysis is the γ-catcher region, which is outside of the Target volume. A large fraction of natural radioactivity comes from either outside of the detector or a photocathode of an ID PMT. The γ-catcher liquid scintillator is closer to these singles sources, and thus carries less shielding effect compared to the Target liquid. In addition, natural radioactivity populates more toward the lower-energy region where neutron capture on hydrogen takes place. Singles events from natural radioactivity stop, at around 4 MeV. This is illustrated in the left plot in Fig. 7-1. This nature of energy dependence makes the Gd analysis quite immune to accidental backgrounds because of a delayed energy cut of 6 MeV $< E_{\text{delayed}} <$ 12 MeV. On the other hand, the 1.5 MeV $< E_{\text{delayed}} <$ 3.0 MeV cut makes the hydrogen analysis suffer from the high accidentals' background rate.

7.1.1 Off-Time Sample Selection

Since accidentals are the coincidence of two uncorrelated events, we can, in principle, collect them by using the so-called off-time selection method. The off-time selection takes the following steps: (1) search for a prompt candidate, applying an isolation cut of [-600,1000] μs around the prompt trigger; (2) from the prompt candidate's time-stamp, open 125 off-time windows to sample a delayed candidate, where the i-th window is opened from T_i + 10 μs to T_i + 600 μs with a reference time $T_i = 1s + (N-1) \times 1600\mu s$; (3) apply an isolation cut in the period of $T_i - 600\mu s$ to $T_i + 1000\mu s$ in the i-th window. This off-time search provides us with a pair of two uncorrelated singles triggers that contaminate the $\bar{\nu}_e$ sample as accidentals.

We open 125 windows in the off-time search to increase the statistics. This corresponds to searching for accidentals in the $1s$ to $1.2s$ after each IBD prompt candidate event. Accordingly, for prompt candidate events toward the end of a run, delayed candidates may not be found by this method. Since a typical run length is 3600 seconds, however, the fractional inefficiency from this effect is negligible.

Figure 7-1 shows the E_{prompt}, E_{delayed}, and ΔT spectrum shape of the data sample collected using the off-time search method. As expected, the ΔT spectrum is flat due to the unique nature of accidentals. The off-time search in the data taking period yields 2,096,445 pairs of accidental coincidence events.

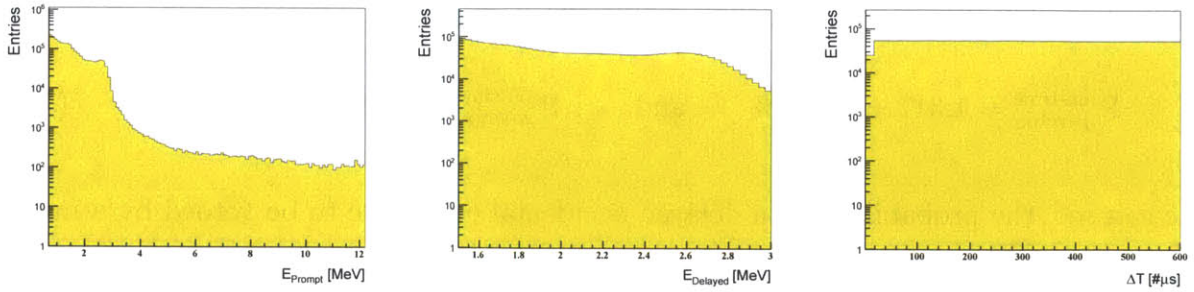

Figure 7-1: From left to right: accidentals E_{Prompt}, $E_{Delayed}$, and ΔT distributions.

7.1.2 Inefficiency Correction Factor and Rate Estimation

Unfortunately, the accidentals' rate contaminating our signal cannot be estimated directly from the sample statistics from the off-time search because the off-time search suffers from two sources of selection inefficiencies when compared to the on-time search. These inefficiencies are due to (1) a muon veto on delayed events and (2) an extra isolation cut on the delayed event in the off-time search. These are described in detail below.

Muon Veto Inefficiency When a muon comes in between a prompt and a delayed candidate, the delayed event is rejected due to the 1 ms muon veto. The probability of rejecting accidentals in the on-time search due to this effect, $\mathrm{P}^{\text{on-time}}_{\mu\text{-veto}}$, can be analytically estimated from the rate of muons, R_μ, and the following analytical formula:

$$\mathrm{P}^{\text{on-time}}_{\mu\text{-veto}} = \int_{T_0}^{T_1} t \times R_\mu \frac{dt}{(T_1 - T_0)} = R_\mu \times \frac{(T_1 + T_0)}{2} = R_\mu \times 305\mu s \tag{7.1}$$

where T_0 and T_1 are 10 μs and 600 μs, respectively. The same probability for the off-time search is much simpler because the delayed event has a 1 s time off-set. Therefore, we expect the probability for a delayed candidate in the off-time search vetoed by a muon to be:

$$\mathrm{P}^{\text{off-time}}_{\mu\text{-veto}} = 1\, ms \times R_\mu. \tag{7.2}$$

Substituting the muon rate of $R_\mu = 43.1929 \pm 0.04$ Hz, we find the numerical values as:

$$\mathrm{P}^{\text{on-time}}_{\mu\text{-veto}} = 1.317 \pm 0.001\ \% \qquad \text{and} \qquad \mathrm{P}^{\text{off-time}}_{\mu\text{-veto}} = 4.319 \pm 0.004\ \% \tag{7.3}$$

As one can see, the probability for a delayed accidental candidate to be vetoed by a muon that comes after a prompt candidate is different between the on-time and off-time selections. The correction factor for off-time search statistics to match the on-time search can be analytically estimated by the following equation:

$$f^{\text{delayed}}_{\mu\text{-veto}} = \frac{1 - \mathrm{P}^{\text{on-time}}_{\mu\text{ veto}}}{1 - \mathrm{P}^{\text{off-time}}_{\mu\text{ veto}}} = 1.03137 \pm 0.00004. \tag{7.4}$$

This correction factor must be applied to the rate yield from the off-time selection in order to account for a different delayed event selection inefficiency due to the muon veto.

Isolation Cut Inefficiency Another source of inefficiency is the isolation cut. For the on-time search, we apply the isolation cut for a prompt candidate. For the off-time search, we apply an extra isolation cut on a delayed candidate. This causes an extra inefficiency for the off-time selection. We quantified the isolation cut efficiency on a delayed event candidate for the off-time selection to be 0.98099 ± 0.00001. This corresponds to a correction factor of 1.01938 ± 0.00001 on the off-time selection statistics.

The Correction Factor Finally, combining correction factors for the two sources of inefficiencies for the off-time selection, we conclude the following correction factor to be applied on the rate estimated from the off-time selection sample:

$$\mathrm{f}^{corr}_{\text{on/off}} = \frac{1.01938 \pm 0.00001 \ \times \ 1.03137 \pm 0.00004}{N_{\text{window}}} = \frac{1.05136 \pm 0.00004}{N_{\text{window}}} \tag{7.5}$$

The correction factor can be confirmed by comparing the accidentals' sample in the off-time selection and a modified on-time search. We have performed the on-time selection by modifying the $\Delta R < 900$ mm cut to $1100 < \Delta R < 3500$ mm, where we expect only accidental coincidences to contribute in the on-time selection. Then we took a ratio of the ΔR distribution between

the on-time and off-time selection outcome. The ratio distribution is shown in the Fig. 7-2. The flat line is fitted in the figure to show that the ΔR spectrum shapes from the on-time and off-time samples are identical. The offset of the fitted flat line shows the fractional excess in the on-time sample. This is found to be $5.19 \pm 0.22\%$. This is consistent with our expectation shown in Eq. 7.5.

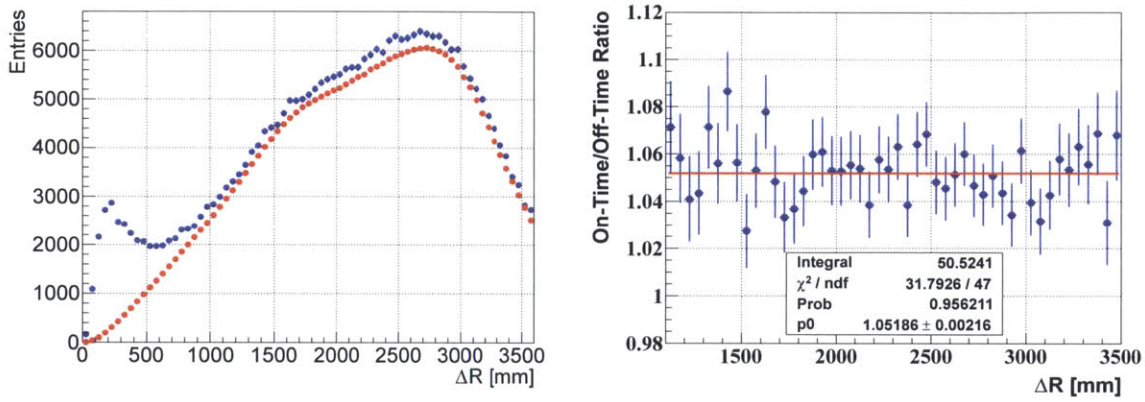

Figure 7-2: The left plot shows ΔR distributions for on-time and off-time data samples in blue and red, respectively. The total off-time data sample is scaled down by 125, the number of sampling windows. The right plot shows the bin-by-bin ratios of two distributions for $1100 < \Delta R < 3500$ mm.

The flat line fit shown in Fig. 7-2 contains 0.2% uncertainty. While this is likely due to a statistical uncertainty in each bin, we take a conservative estimate of 0.2% systematic uncertainty. This is propagated to the rate uncertainty of the accidentals background.

As described in detail above, the major component of the correction comes from a delayed event selection inefficiency due to the muon veto. We can test our understanding about this correction factor by changing the muon veto period from 1 ms to 2 ms. This yields a prediction for the correction factor to be $10.107 \pm 0.004\%$. The measured on-time excess is found to be $10.03 \pm 0.25\%$ as shown in the Fig. 7-3.

Finally, using the correction factor, the off-time search yields the estimated rate of accidental backgrounds to be:

$$\mathrm{R}_{\mathrm{accidentals}}^{\mathrm{measured}} = 73.45 \pm 0.05 \ (\mathrm{stat.}) \pm 0.15 \ (\mathrm{syst.}) \ \mathrm{events/day} \tag{7.6}$$

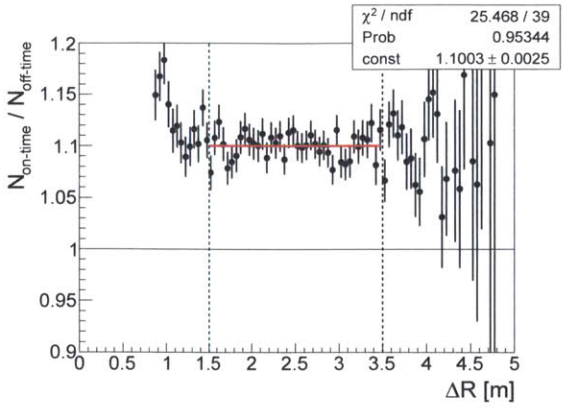

Figure 7-3: The ΔR ratio plot after increasing the muon veto period from 1 ms to 2 ms. The flat line fit indicates a $10.03 \pm 0.25\%$ excess of accidentals in the on-time sample.

where the systematic uncertainty of 0.2% is from the flat line fit shown in Fig. 7-2. As mentioned previously, we note this 0.2% is partially statistical uncertainty, and may improve with more data taking.

7.1.3 Analytical Cross-Check

The rate of an accidental coincidence can be analytically estimated based on several physics observables. This can be used as a cross check for the rate estimate of accidentals background using the off-time selection. The following equation is used for this analytical prediction.

$$\mathrm{R}_{\mathrm{accidentals}}^{\mathrm{analytical}} = \mathrm{R}_{\mathrm{prompt}} \cdot \mathrm{R}_{\mathrm{delayed}} \cdot \Delta T_{\mathrm{coinc.}} \cdot \epsilon_{\mu\text{-veto}}^{\Delta T} \cdot \epsilon_{\mathrm{isolation}}^{\mathrm{prompt}} \cdot \epsilon_{\Delta \mathrm{R}} \cdot f_{\ \mu\ \mathrm{veto}}^{\ \mathrm{delayed}} \tag{7.7}$$

The parameters in the equation above are the following: $\mathrm{R}_{\mathrm{prompt}} = 8.1974 \pm 0.0006$ Hz, the rate of prompt candidate triggers: number of events with energy between 0.7 to 12.2 MeV after OV veto; $\mathrm{R}_{\mathrm{delayed}} = 2.9909 \pm 0.0004$ Hz, the rate of delayed candidate triggers: number of events with energy between 1.5 to 3.0 MeV; $\Delta T_{\mathrm{coinc.}} = 590$ μs, the ΔT window used in the selection; $\epsilon_{\mu\text{-veto}}^{\Delta T} = \left(1 - \frac{\text{OV Corrected Muon Veto Time}}{\text{Run Time}}\right)$, the probability of a coincidence to be anti-correlated with a muon; $\epsilon_{\mathrm{isolation}}^{\mathrm{prompt}} = 0.98131 \pm 0.00001$, the efficiency of the isolation cut on prompt candidates; $\epsilon_{\Delta R} = 0.06079 \pm 0.00004$, the efficiency of the $\Delta \mathrm{R} < 900$ mm cut for accidentals; and $f_{\ \mu\text{-veto}}^{\ \mathrm{delayed}} = 1.03137 \pm 0.00004$, the correction factor for the muon veto

inefficiency for delayed candidates. Substituting all numbers, we find:

$$R_{\text{accidentals}}^{\text{analytical}} = 73.46 \pm 0.06 \text{ (stat.) events/day}, \tag{7.8}$$

which is in a great agreement with Eq. 7.6.

The most convincing feature of this cross-check can be seen in Fig. 7-4 where we compare daily accidental backgrounds' rates from the off-time selection and the analytical estimate. One can find a great agreement on all days. For your reference, the stability plot of other parameters

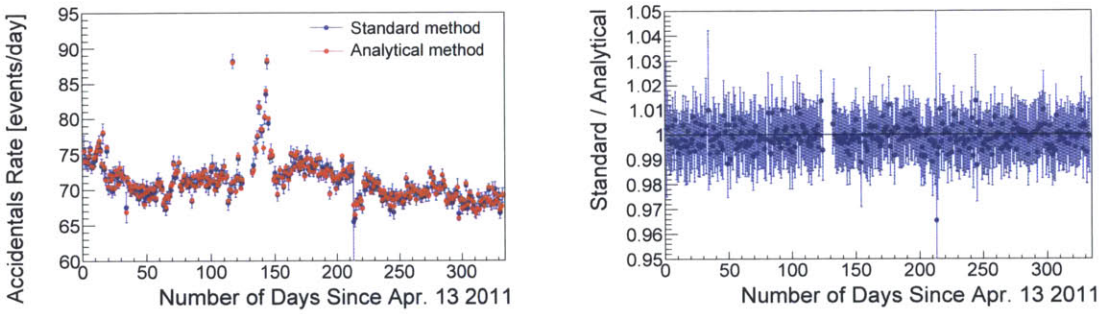

Figure 7-4: Left: A daily comparison of two accidentals rate estimation methods. The blue data points are the rate measured from the off-time selection, and the red data points are the rate derived by analytical estimation. The methods agree well over the full data-taking period. Right: Ratio of rates show in the left plot.

used in Eq. 7.7 is shown in Fig. 7-5. As each parameter shows a different time variation, the fact that the off-time search estimate matches with the analytical rate estimate using all of these parameters makes us very confident and able to claim that we understand the accidentals' rate estimation.

7.2 Lithium 9

A high energy muon going through the Target or the γ-catcher volume produces various kinds of cosmogenic isotopes through photo-nuclear interaction with carbon nuclei. Some isotopes created in this process undergo beta decay with an emission of one or more neutrons. This beta decay event followed by a delayed neutron capture can mimic an IBD coincidence signal,

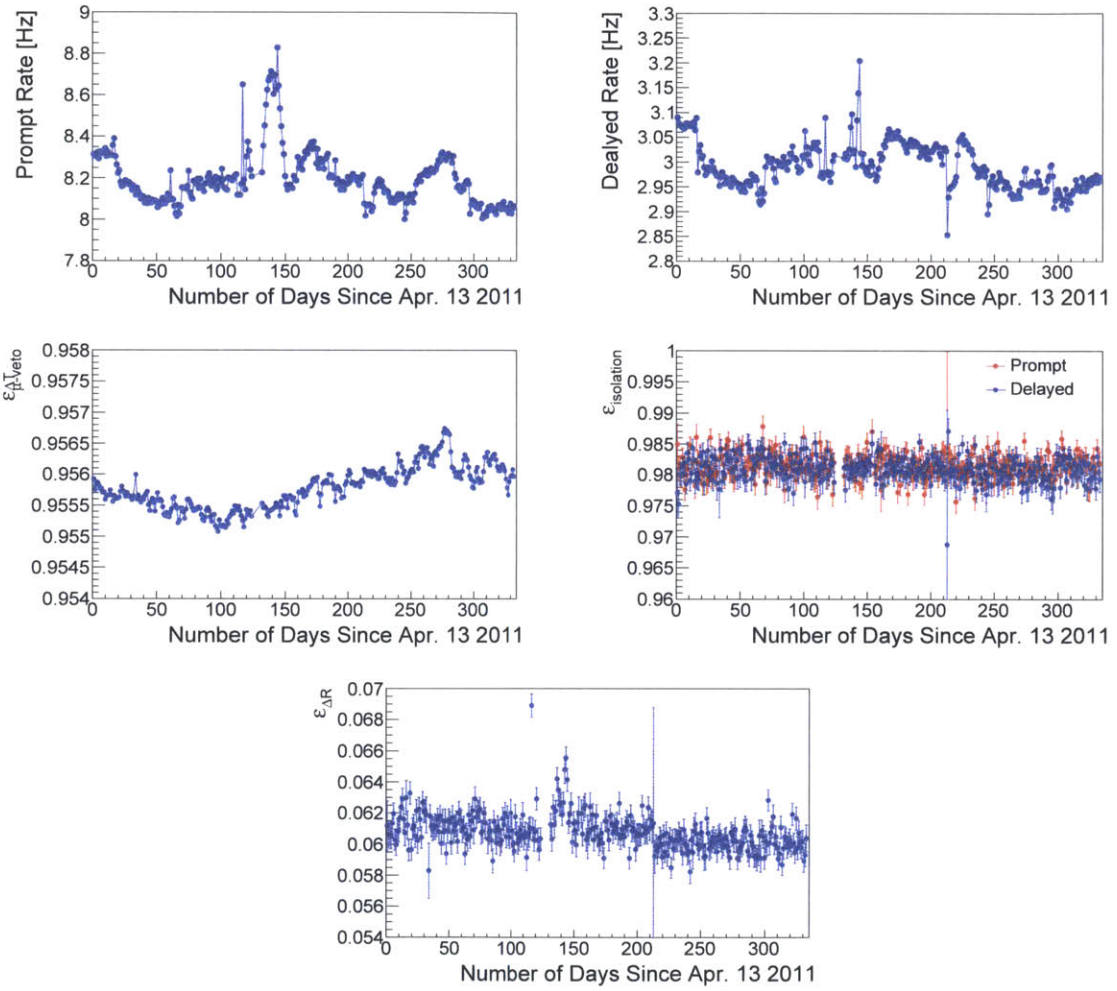

Figure 7-5: Day-to-day stability of parameters used in the analytical estimation.

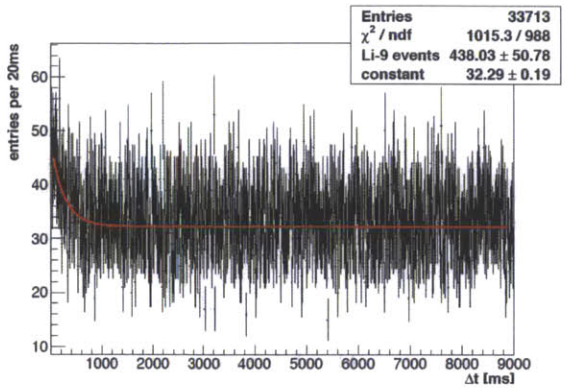

Figure 7-6: ΔT_μ distribution for IBD candidates and all preceding high-energy muons with ID energy deposition $>$ 600 MeV.

and is a background to our signal.

Among numerous cosmogenic isotopes produced by going through muons, ^{9}Li and ^{8}He are expected to be the dominant background contributors. A branching ratio of neutron emitting beta decay is 50.8 $\pm$ 0.9 % for ^{9}Li and 16 $\pm$ 1 % for ^{8}He [16]. The half-life and Q value of ^{9}Li are 178.3 ms and 13.6 MeV while those of ^{8}He are 119.0 ms and 10.7 MeV. Consequently only very small fraction of them can be removed by a 1 ms after-muon veto cut. The remaining cosmogenic backgrounds must be addressed in the final E_{prompt} spectrum fit. The next section describes an estimation of rate and E_{prompt} shape.

7.2.1 Rate Estimation

Prompt events of cosmogenic backgrounds have a time correlation to the parent muon. Figure 7.2.1 shows a time correlation between IBD prompt candidates and all high-energy muons ($>$ 600 MeV, called "showering muons") in the previous 8 seconds. The exponential fit is performed with the fixed decay time constant of a ^{9}Li isotope. From the normalization of the fitted exponential, we can estimate the rate of ^{9}Li after showering muons as 1.82$\pm$0.21 events/day. The flat component in the ΔT_μ distribution in Fig. 7.2.1 is dominated by both IBD candidates and accidental backgrounds as neither of these have any time correlation to muons.

In order to increase an efficiency of ^{9}Li event tagging, the reconstructed muon track information is used. Since the hydrogen analysis is sensitive to ^{9}Li created in non-scintillating

Buffer liquid, we employed the muon track reconstruction algorithm based on the IV PMT hit information. This is different from the Gd analysis which employed another algorithm based on ID PMT hit information. The majority of ^{9}Li backgrounds to the Gd analysis are created by muons going through the Target or γ-catcher volume, and they make a huge amount of scintillation light. The ID muon track reconstruction algorithm is expected to work fairly well for all muons going through such LS filled volumes, but the performance drops dramatically when muons go through partially in non-scintillating Buffer region. In the hydrogen analysis, however, ^{9}Li produced in the Target, γ-catcher, and Buffer region all contribute as backgrounds. For this reason, we employed IV muon track reconstruction algorithm which delivers a similar reconstructed track resolution to all muons going through the ID region.

We apply the spatial cut to require the prompt candidate to be within 435 mm of the closest reconstructed muon track. This cut carries an efficiency of 48.2 % based on the study of MC ^{9}Li sample against backgrounds taken from DATA by requiring $\Delta T_\mu > 1000$ ms. Figure 7.2.1 shows the IBD candidates' ΔT_μ distribution in black data points after applying this spatial cut. In order to further suppress the accidental backgrounds in the flat part of the ΔT_μ spectrum, we subtracted the accidental backgrounds' ΔT_μ distribution (gray histogram in the figure) taken from the off-time selection [1] with one off-time window. The subtracted spectrum shown in blue color is fitted by Eq. 7.9, a simple exponential decay function with a flat background component, using the maximum loglikelihood method.

$$N_{^9\mathrm{Li}}(\Delta T_\mu) = N_{^9\mathrm{Li}} \exp\left(-\frac{\Delta T_\mu}{\tau_{^9\mathrm{Li}}}\right) + N_{\mathrm{background}} \tag{7.9}$$

The decay time constant used, τ in Eq. 7.9, is that of ^{9}Li which is 257.2 ms. After correcting for the spatial cut efficiency, we find the ^{9}Li rate estimate of 2.84 ± 0.53 events/day. In the following sections, multiple source of systematic uncertainties are described briefly.

Monte Carlo Model of ^{9}Li Flight Distance The distance of a ^{9}Li production vertex from a reconstructed muon track depends on MC model we use. KamLAND did a sophisticated MC study using a simulation package, FLUKA [37, 38]. Figure 7.2.1 is generated using FLUKA,

[1] For a description of off-time selection, see Sec. 7.1.1

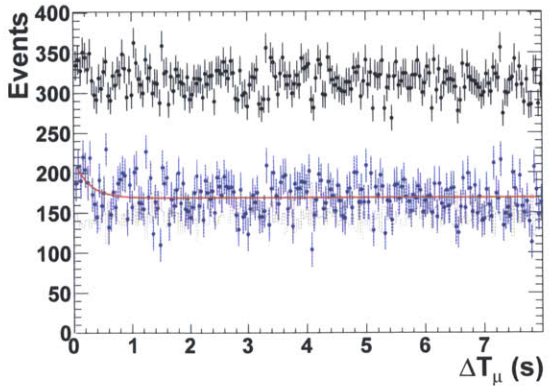

Figure 7-7: ΔT_μ distribution for ^{9}Li rate estimation. Black dots are all IBD candidates while the gray histogram shows the off-time accidentals sample for the ^{9}Li analysis. Blue data points are obtained by subtracting the off-time distribution from the IBD candidates.

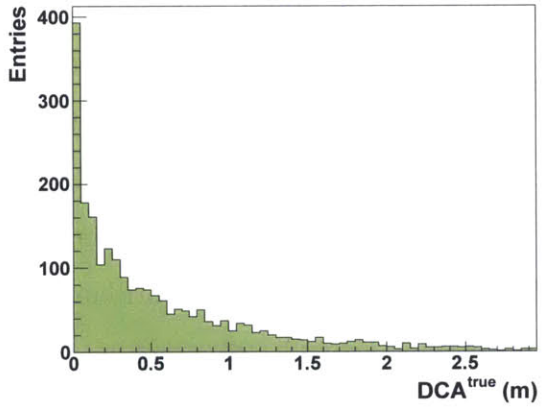

Figure 7-8: FLUKA simulation result made for the KamLAND detector [4]. The distribution shows the distance of the closest approach for ^{9}Li generation vertex from the parent muon track.

and it shows the distribution of ^{9}Li production vertex as a function of distance from the parent muon track in liquid scintillator. The distribution is found to be consistent with a double-exponential shape: one with characteristic length shorter than 10 cm, and another in the scale of one meter. Given the size of the Double CHOOZ detector, we expect to be only sensitive to the exponential component with a shorter characteristic length. The mean value of ^{9}Li production vertex is set to, therefore, 10 cm from the muon track. We have varied this distance from 0 cm to 20 cm to estimate a bias to the cut efficiency due to this MC modeling, and the systematic uncertainty due to this effect is found to be $^{+3.8}_{-13.4}$%.

Background Subtraction While the subtraction of accidentals is crucial to reduce the background rate to the ^{9}Li estimate, this could also bias the exponential fit to the ΔT_μ distribution if done incorrectly. We decided, therefore, to add a difference in the ^{9}Li mean rate between the nominal value and another one obtained without subtracting the accidentals background to the systematic uncertainty. This adds -28.4 % to the rate uncertainty.

Comparison With the Scaled Rate From Gd Analysis For the Gd analysis, the ^{9}Li rate is estimated using ID muon track reconstruction [1], which is independent from the IV muon track reconstruction used in this analysis. Since ^{9}Li production is mainly due to high energy muon hitting the carbon nuclei in the liquid scintillator, we can calculate the scaling factor to estimate the ^{9}Li event rate for this analysis using that of the Gd analysis. The main ingredients to calculate the scaling factor are the followings: (1) the difference in ^{9}Li detection efficiency between the Gd analysis and hydrogen analysis; (2) the difference in neutron capture on hydrogen and on Gd; (3) the volume difference between the Target and the γ-catcher volume; and (4) the carbon density in the Target and γ-catcher scintillator.

We can make a 1st order approximation and see how those parameters contribute. Let us express $X^{\mathrm{Gd}}_{^9\mathrm{Li}}$ as the number of ^{9}Li that may pass the Gd analysis cuts.

$$X^{\mathrm{Gd}}_{^9\mathrm{Li}} = \epsilon^{\mathrm{Gd}}_T \times N^{\mathrm{Li}}_T,$$

where ϵ^{Gd}_T and N^{Li}_T are the neutron detection efficiency and number of the ^{9}Li produced in the Target volume. Similarly, $X^{\mathrm{H}}_{^9\mathrm{Li}}$, the number of ^{9}Li that may pass the hydrogen analysis cuts, can be expressed as the following:

$$X^{\mathrm{H}}_{^9\mathrm{Li}} = \epsilon^{\mathrm{H}}_T \times N^{\mathrm{Li}}_T + \epsilon^{\mathrm{H}}_{GC} \times N^{\mathrm{Li}}_{GC}.$$

A scaling factor for ^{9}Li can be found by taking a ratio of the two:

$$\frac{X^{\mathrm{H}}_{^9\mathrm{Li}}}{X^{\mathrm{Gd}}_{^9\mathrm{Li}}} \equiv R_{\mathrm{H/Gd}} = \frac{\epsilon^{\mathrm{H}}_T}{\epsilon^{\mathrm{Gd}}_T} + \frac{\epsilon^{\mathrm{H}}_{GC}}{\epsilon^{\mathrm{Gd}}_T} \times \frac{N^{\mathrm{Li}}_{GC}}{N^{\mathrm{Li}}_T}.$$

Volume	ϵ^H	ϵ^{Gd}	Volume	f^C	ρ^{LS}
Target	0.0846	0.800	10.6910 m^3	86.1 %	804 mg/cm^3
GC	0.7853	N/A	22.6763 m^3	85.4 %	805 mg/cm^3

Table 7.1: Summary of parameter values useful for the Eq. 7.11. All of them are extracted from ^{9}Li MC sample except for ϵ^{H} and ϵ^{Gd}. For ϵ^{H}, we took numbers from Sec 6.4. For ϵ^{Gd}, numbers are taken from the Gd analysis.

To 1st order, we can estimate a ratio of $N_T^{\mathrm{Li}}/N_{GC}^{\mathrm{Li}}$ by a ratio of carbon nuclei in each volume:

$$\frac{N_{GC}^{\mathrm{Li}}}{N_T^{\mathrm{Li}}} \approx \frac{V_{GC} \times \rho_{GC}^{LS} \times f_{GC}^{C}}{V_T \times \rho_T^{LS} \times f_T^{C}},$$

where ρ_X^{LS} and f_X^C denote mass density of the LS and a carbon fraction in volume X respectively.

In order to quantify a number for $R_{\mathrm{H/Gd}}$, we used the ^{9}Li MC sample generated uniformly in all volumes of the ID.

$$\left(R_{\mathrm{H/Gd}}\right)_{^9\mathrm{Li\ MC}} = 1.86 \pm 0.01\,, \tag{7.10}$$

where the error quoted above only takes into account statistical uncertainty. As described previously, this sample is made with a uniform vertex distribution of ^{9}Li. Therefore, we expect a systematic shift coming from the fact that the carbon fraction difference in the Target and γ-catcher is not taken into account, while the volume difference is. This systematic shift can be quantified using the following expression.

$$\delta\left(R_{\mathrm{H/Gd}}\right) = \frac{\epsilon_{GC}^{H}}{\epsilon_T^{Gd}} \times \frac{V_{GC}}{V_T} \times \left(\frac{\rho_{GC}^{LS} \times f_{GC}^{C}}{\rho_T^{LS} \times f_T^{C}} - 1\right) \approx 1.4\,\%, \tag{7.11}$$

which gives an underestimation for a scaling factor obtained from ^{9}Li with uniform vertex generation. We can find some parameters in the above expression in DATA or all of them in MC. For a quick estimate, we used the MC values summarized in Tab. 7.1. Using the numbers in the table, we find $\delta\left(R_{\mathrm{H/Gd}}\right) \approx 1.4\,\%$. These systematics on the scaling factor makes a negligible contribution to the total error of ^{9}Li, which is on the order of 10 %.

We used the scaling factor in Eq. 7.10 with the quoted systematics in Eq. 7.11. The ^{9}Li rate scaled from the Gd analysis [1] is $3.81^{+1.16}_{-0.97}$ events/day, where the 1.4 % error is added to the original error of the Gd analysis in the quadrature. This result is a discrepancy of +34%

from the nominal rate in this analysis, thus, the discrepancy is included in the systematic uncertainties.

Despite the difficulties mentioned earlier about the ^{9}Li rate estimation method using ID muon track reconstruction, we have performed the analysis method [1] and found the ^{9}Li rate for the hydrogen analysis to be $3.37^{+1.13}_{-0.89}$ events/day. This is consistent with the scaled rate estimation quoted above.

Binning of ΔT_μ Distribution The maximum log-likelihood method used to fit the ΔT_μ distribution in Fig. 7.2.1 could be sensitive to the binning option we used, which is 40 ms. In order to account for a possible bias, we varied the bin size to a half and twice larger. This adds a systematic uncertainty of $^{+9.0}_{-3.0}$%.

Monte Carlo Statistical Uncertainty Muon simulation in our detector costs a huge amount of computing time, and there is a limitation on number of events generated in our MC study. We add the statistical uncertainty of $\pm 1.8\%$ from the MC sample, therefore, to our total systematic uncertainty estimate.

^{9}Li Rate With a Combined Systematic Uncertainty Finally, combining all systematic uncertainties mentioned above, we found the following as the estimated rate of ^{9}Li in this analysis.

$$N_{^9\mathrm{Li}} = 2.84 \pm 0.53(\mathrm{stat.})^{+1.02}_{-0.91}(\mathrm{syst.}) = 2.84^{+1.15}_{-1.05}\ \mathrm{events/day} \tag{7.12}$$

7.2.2 Rate Cross-Check Without Using Muon Track

It is possible that the ^{9}Li rate estimate in Eq. 7.12 could have a yet missing bias from the method of using a reconstructed muon track. This section presents a cross-check of the ^{9}Li rate estimation by a method that does not rely on reconstructed muon track information.

This alternative approach uses the candidates' E_{prompt} spectrum. After subtracting $\bar{\nu}_e$, fast neutrons, and accidentals backgrounds from the canddiates' spectrum, we obtain the remaining as the ^{9}Li candidates. We use the energy region 8.2 MeV $< E_{\mathrm{prompt}} <$ 12.2 MeV to enhance the purity of ^{9}Li in the sample. In this energy range, we expect the sample to be dominated

by fast neutrons and ^{9}Li events. Figure 7-9 shows the E_{prompt} spectrum after subtracting the accidental backgrounds. The left plot shows the full energy range of the analysis, while the right plot focuses on the energy range of our interest. The partial ^{9}Li rate is estimated by integrating

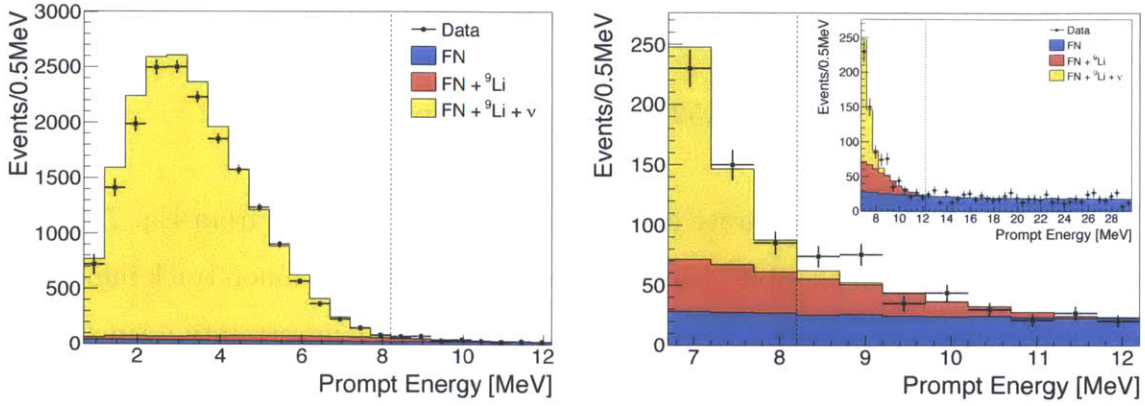

Figure 7-9: Prompt energy spectrum with expected neutrinos and backgrounds. Left plot shows the spectrum below 12.2 MeV, and right plot shows the one in [6.7-12.2] MeV. As a reference, the miniature shows the spectrum up to 29.7 MeV. In the plot, the rate of ν, ^{9}Li, and fast neutron in [0.7-12.2] MeV are scaled to 76.0 events/day, 2.84 ($^{+1.15}_{-1.05}$) events/day, and 2.50 ($\pm$ 0.47) events/day, respectively.

the E_{prompt} spectrum after subtracting non-^{9}Li events. The full ^{9}Li rate is estimated by applying an efficiency correction of the E_{prompt} energy cut, which is estimated using the E_{prompt} shape of ^{9}Li (see Sec. 7.2.3). This estimation is illustrated in the following equation

$$^9\text{Li rate [events/day]} = \frac{\text{N}_{\text{cand.}} - \text{N}_{\text{off}} - \text{N}_{\nu} - \text{N}_{\text{fn}}}{\varepsilon_{^9\text{Li,Ep}} * \text{Live time [day]}}. \tag{7.13}$$

We note that this estimation depends on the rate and shape estimation of fast neutrons and $\bar{\nu}_e$ events. The details of rate and shape estimation of $\bar{\nu}_e$ and fast neutron samples are described in Sec. 8.1 and 7.3, respectively. In addition, this method needs to take into account for energy non-linearity and ^{9}Li shape uncertainty. The Gd analysis estimated the energy non-linearity uncertainty to be about 1 % [1] by using a peak in the energy spectrum of neutron capture on Gd around 8 MeV. This is relevant to the energy region of this analysis. The ^{9}Li shape uncertainty is due to less known branching ratios of various decay modes (see Sec. 7.2.3 for

details). Taking these into account, we find the E_{prompt} cut efficiency as the following:

$$\epsilon_{^9\text{Li},E_{\text{prompt}}} = \frac{\text{Number of events in [8.2,12.2] MeV}}{\text{Number of events in [0.7,12.2] MeV}} = 14.7 \pm 0.9\ \%$$

By substituting the numbers in Eq. 7.13, we find the ^{9}Li rate to be

$$^9\text{Li} = 3.52 \pm 1.22 \text{ events/day.} \tag{7.14}$$

which is consistent with the nominal rate estimate, $2.84^{+1.15}_{-1.05}$ events/day from Eq. 7.12. Since this method does not include a possible bias due to using reconstructed muon track information, we consider the agreement is sufficient to claim that the systematic uncertainty estimate in our nominal rate estimation is conservative enough.

7.2.3 Simulation Based E_{prompt} Shape Estimation

As discussed in Sec. 7.2.1, both IBD and accidental coincidences are backgrounds for identifying ^{9}Li events. Because of the large statistics of such backgrounds, it is very hard to study the ^{9}Li spectrum using DATA only. Instead, we have done a dedicated MC study to predict the ^{9}Li E_{prompt} spectrum. This is very challenging because ^{9}Li decay branches are not very well understood [16].

Possible decay paths of ^{9}Li nuclei are illustrated in Fig. 7-10. The diagram shows the final product of ^{9}Li decay as $\alpha+\alpha+n$. ^{9}Li first decays into an intermediate state, which is one of the excited states of ^{9}Be. Then it can decay into another intermediate state of ^{5}He or ^{8}Be before the final state. A large fraction of $Q = 13.6$ MeV is shared by the α particle in the case of the ^{5}He decay branch, while the energy is taken by a neutron instead if the decay occurs through the ^{8}Be isotope. A visible energy spectrum is largely different between these two decay paths because an α particle is highly ionizing and affected by quenching much more than a neutron. In order to counter a lack of knowledge in branching ratios among these two decay modes, we decided to take one decay mode as our central shape estimation, while the difference to the other mode is assigned as the spectral uncertainty. From the references [39, 40, 41], we find the decay mode with the ^{8}Be intermediate state is more likely, and hence it is used for the central

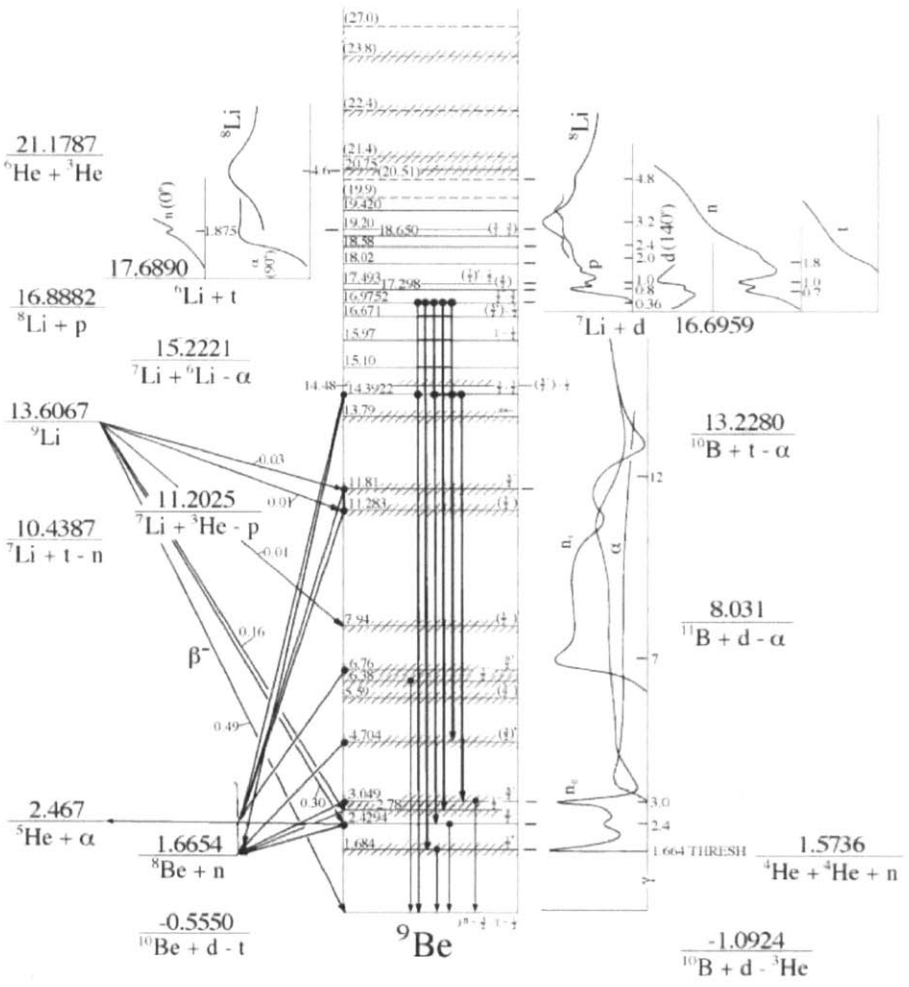

Figure 7-10: Decay branches of ^{9}Li through ^{9}Be excited states from Ref. [16].

shape estimation.

Both ^{8}Be and ^{5}He decay modes have been generated and fed into the detector MC simulation chain. The left plot in Fig. 7-11 shows the outcome of our ^{9}Li shape estimation after running the hydrogen analysis cuts on this MC sample. This MC spectrum shape is compared with DATA

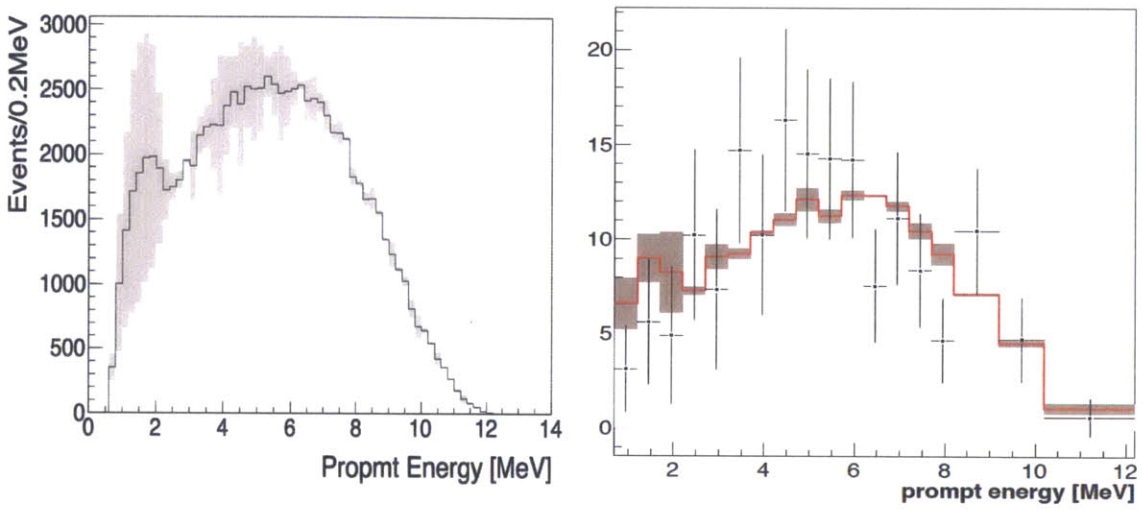

Figure 7-11: Left: Prompt energy spectrum of ^{9}Li MC sample. The central shape is generated assuming the ^{8}Be decay mode to be dominant. The shape uncertainty is taken from the difference in shape between ^{5}He and ^{8}Be decay modes. Right: a comparison between DATA in limited statistics in the Gd analysis [1].

in the Gd analysis. This is shown in the right of Fig. 7-11, and one can find a good agreement between the central shape and the DATA spectrum although the comparison is limited by low statistics.

7.3 Fast Neutrons

A high energy neutron produced by a muon going through the surrounding rock can come into the ID region and hit a proton. A recoil proton produces scintillation light and can be detected as a prompt event candidate. This event is immune to a muon veto cut because the parent muon does not go through the ID or the IV. A fast neutron, after giving up its energy to recoil protons, gets thermalized in a proton-rich liquid scintillator volume and eventually captured on a proton. This makes a coincident signal and becomes a background to the IBD signal.

Typically such fast neutrons as cosmogenic products come in a multiplicity of more than one, and they are efficiently vetoed by the multiplicity cut. In this section, we quantify the rate and E_{prompt} shape of a residual fast neutron background which passes through the multiplicity cut.

7.3.1 Low-Energy IV-Tag Analysis

As mentioned above, typically one muon going through the surrounding rock produces multiple fast neutrons. Because of less shielding effect for the IV to such fast neutrons compared to the ID, these fast neutrons have a higher probability of interacting with IV liquid scintillator than ID liquid. Consequently, we expect some fraction of the fast neutrons' prompt events to have small energy deposition in the IV, as illustrated by a drawing in Fig. 7-12. We can use this condition to tag a sample of fast neutrons in IBD candidates. In order to pass the

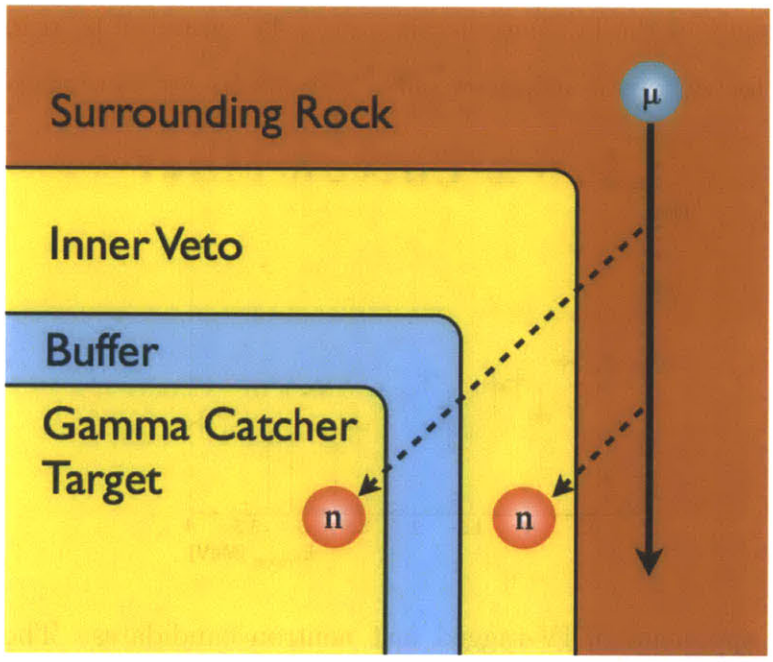

Figure 7-12: A toy model of multiple fast neutrons produced by a muon interacting separately in ID and IV liquid scintillator (yellow).

IBD selection cuts, the energy deposited in the IV must be smaller than ≈4 MeV. Unlike ID PMTs, IV PMTs are in contact with liquid scintillator directly, and they see energy deposited in the IV by radioactivity coming from themselves. This is a background to our selection of fast

neutron candidates because such radioactivity typically occurs below the 4 MeV energy region. In order to increase the purity of our fast neutron sample, therefore, we require more than two IV PMTs hit in the prompt event. This requirement is called "*IV-Tag.*" In order to cover a wide range of acceptance with small number of PMTs, the neighboring IV PMTs are typically set to face the opposite direction. By asking the IV-tag condition, and so asking for more than two IV PMTs to be hit, we dramatically reduce backgrounds coming from PMT radioactivity.

Figure 7-13 shows the IV E_{prompt} spectrum with different IV PMT hit multiplicity cuts. From the figure, it is apparent that the minimum IV-Tag condition of more than two IV PMT hits removes a large fraction of background. Figure 7-14 shows the ID E_{prompt} spectrum by the same cut condition. Because the spectrum shape is similar among different IV-Tag conditions of two, four, or six IV PMT hits, a condition of two IV PMT hits is chosen in order to increase our sample statistics. Assuming E_{prompt} >12 MeV region is dominated by fast neutrons, the efficiency of the IV-Tag selection, ϵ_{IVT}, is computed by using a sample with $E_{\text{prompt}} \in [12, 30]$ MeV for various multiplicity cuts. The results are summarized in Tab. 7.2.

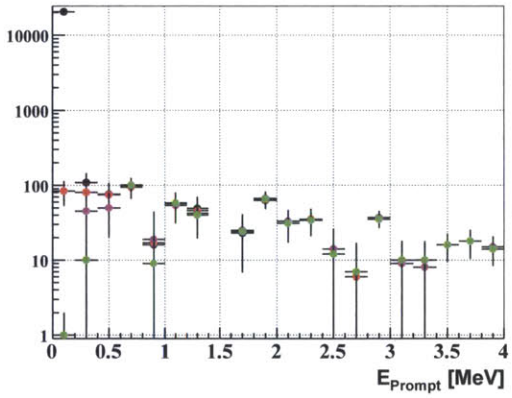

Figure 7-13: E^{IV}_{Prompt} spectrum of IV-tagged fast neutron candidates. The different colors represent different IV PMT hit conditions: 0, 2, 4, 6 minimum IV PMT hit conditions are shown in black, red, magenta, and green, respectively.

In the IV-tagged sample events, there are still a few backgrounds to our fast neutron analysis. Such backgrounds can be categorized into two classes described in the following: (1) a prompt-delayed correlated ID event occuring in an accidental coincidence with another physics event depositing energy in the IV; (2) an ID-IV correlated prompt energy deposition occuring in an

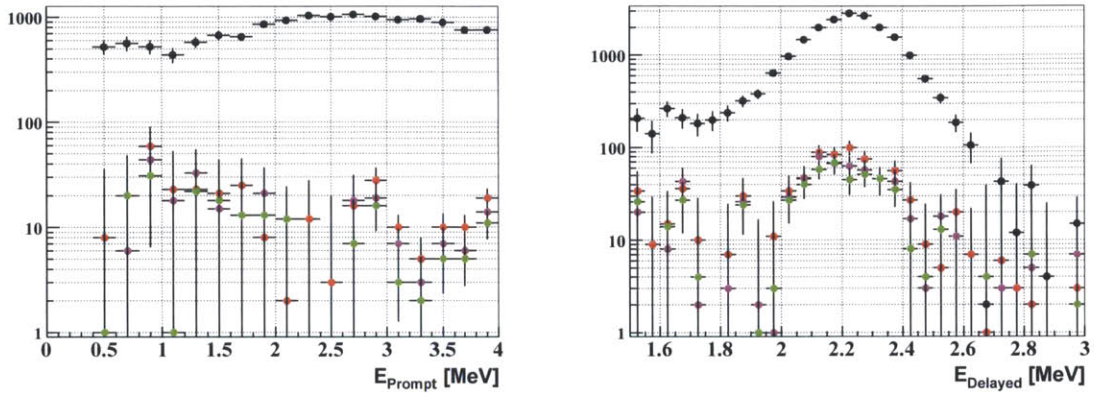

Figure 7-14: E^{ID}_{Prompt} (left) and $E^{ID}_{Delayed}$ (right) spectra of IV-tagged FN candidates. The different colors represent different IV PMT hit conditions: 0, 2, 4, 6 minimum IV PMT hit conditions are shown in black, red, magenta, and green, respectively.

Min. IV PMT Hit	$\epsilon_{IVT} \in [12, 30]$ MeV	Error
2	0.337	0.018
4	0.278	0.017
6	0.228	0.016

Table 7.2: Summary of ϵ_{IVT} for different IV-tag cut conditions.

accidental coincidence with an uncorrelated delayed physics event. An example of the first case could be a residual accidental coincidence of an IV PMT radioactivity event with an IBD event occurring in the ID.We note this class of events have no time correlation between the ID and IV PMT hit time in the prompt. An example of the second case could be a gamma ray sourced outside the detector depositing energy in both ID and IV through multiple Compton scatterings in accidental coincidence with a delayed, uncorrelated ID singles event. In this case, there is no correlation between the prompt and delayed, and in fact this is what we call accidental backgrounds to the IBD candidates. These two classes of events can be distinguished from fast neutrons because fast neutrons have a time correlation between the ID and IV PMT hit times (since they both share one parent muon) in the prompt, and also between the prompt and delayed events since they are correlated by a neutron capture time.

It is relatively simple to treat the second case of background events because events that belong to this class can be easily sampled by an off-time selection method (see Sec. 7.1.1). On the other hand, rejecting the backgrounds that belong to the first case could be tricky. To achieve this step, we define a parameter,

$$\Delta T_{start}^{ID-IV} = \text{(First ID PMT hit time)} - \text{(First IV PMT hit time)}. \tag{7.15}$$

The distribution of ΔT_{start}^{ID-IV} is shown in the left plot of Fig 7-15. The plot is made after subtracting the ΔT_{start}^{ID-IV} distribution of off-time sample [2]. We see a clear correlation in ID and IV first PMT hits around $\Delta T_{start}^{ID-IV} \approx$ 60 ns, which is consistent with the configuration for synchronizing the ID and IV signal readout. The fitted Gaussian has a width of 18.5 ns. This is due to a trigger jitter time of 16 ns. The right plot of Fig. 7-15 shows ΔT_{start}^{ID-IV} of both the IBD candidates and the off-time sample, illustrating that the IBD candidates above 12 MeV are dominated by fast neutrons that should have $\Delta T_{start}^{ID-IV} \approx$ 60 ns.

From the fitted Gaussian function in Fig. 7-15, we assumed the $\pm 3\sigma$ width from the mean includes 99.7% of the events that have correlation among ID and IV energy deposition in the prompt (the ID-IV correlated events). In this range of ΔT_{start}^{ID-IV}, however, we also expect some events with no ID-IV correlation, which appears as a flat component. We can see the profile

[2]This is done using the method described in Sec 7.1.1 with 1 off-time sampling method.

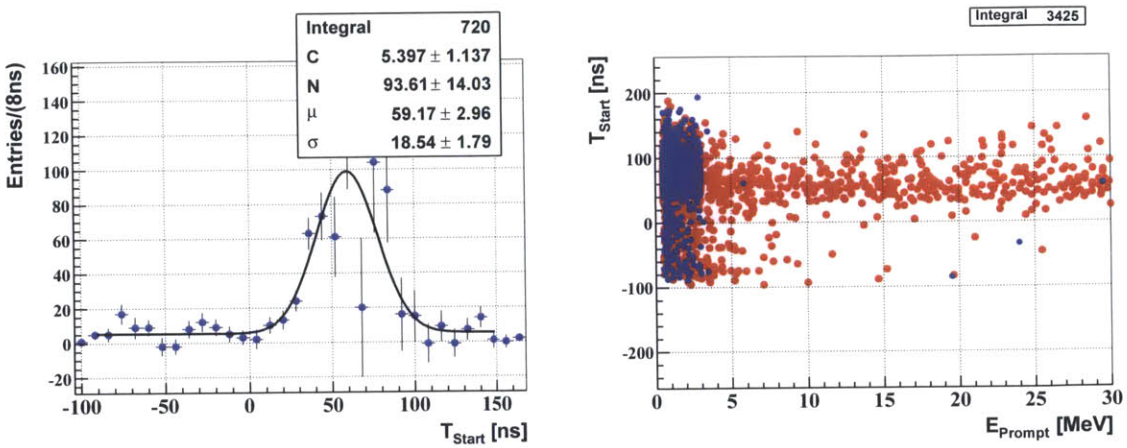

Figure 7-15: ΔT^{ID-IV}_{Start} (left) and E_{Prompt} vs. ΔT^{ID-IV}_{Start} (right) distribution. The ΔT^{ID-IV}_{Start} distribution is off-time subtracted. The right plot shows on-time sample in red and off-time sample in blue.

of $E_{\rm prompt}$ for such events by taking an event sample outside the chosen ΔT^{ID-IV}_{start} region. In order to construct the $E_{\rm prompt}$ spectrum of the ID-IV correlated sample, we first make $E_{\rm prompt}$ spectrum from events in the $\pm 3\sigma$ region of the fitted Gaussian using IBD candidates. Then we make another $E_{\rm prompt}$ spectrum with the same cut on ΔT^{ID-IV}_{start} but from the off-time selected event sample. By subtracting the latter from the former and after correcting for ϵ_{IVT}, we obtain the $E_{\rm prompt}$ spectrum of the ID-IV correlated sample with some ID-IV uncorrelated contamination. This is shown in black data points in Fig. 7-16.

The next step is to construct the $E_{\rm prompt}$ spectrum of the ID-IV uncorrelated component, so that we can subtract this spectrum from the black data points in Fig. 7-16. This is simpler because we can take the ID-IV uncorrelated sample by selecting events with $\Delta T^{ID-IV}_{start} \in [-90, 150]$ ns but outside the $\pm 3\sigma$ region of the fitted Gaussian in Fig. 7-15. In order to remove the prompt-delayed uncorrelated component, again, we subtract the $E_{\rm prompt}$ created from off-time event sample from that of the IBD candidates. Finally, assuming the ID-IV uncorrelated events make a flat distribution in ΔT^{ID-IV}_{start} space, we re-normalize the resulting $E_{\rm prompt}$ spectrum by a factor which is a ratio of 6σ divided by the width of ΔT^{ID-IV}_{start} used for the selection. This is required to match the statistics of the ID-IV uncorrelated component in the $\pm 3\sigma$ region of the fitted Gaussian. In addition, finally, we correct for ϵ_{IVT}. This re-normalized spectrum is shown in red data points in Fig. 7-16. The spectrum peaks around 3 MeV and decreases

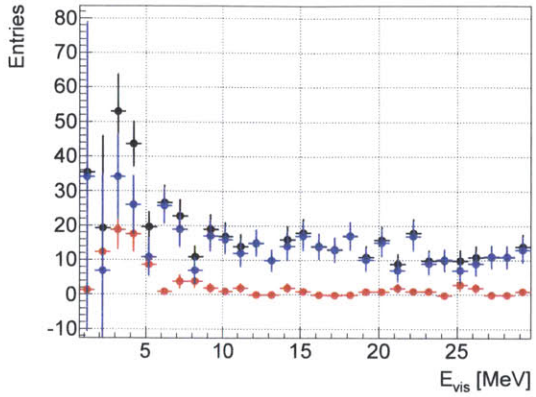

Figure 7-16: E_{Prompt} of fast neutron sample. The black data points contain the ID-IV correlated and uncorrelated component while the red data points show the ID-IV uncorrelated component. Blue data points are obtained by subtracting the red from black data points; they are the fast neutron E_{prompt} spectrum.

toward 12 MeV, which is consistent with IBD events as we expected earlier.

7.3.2 Rate and Shape of Fast Neutrons

Figure 7-16 shows blue data points as fast neutron candidate spectrum after subtracting red data points (the ID-IV uncorrelated component) from black data points (the ID-IV correlated + uncorrelated components). We fit the following function to those blue data points:

$$f_{\text{fast-n}}(E_{\text{prompt}}) \;=\; N_{\text{exp}} \cdot \exp -\frac{E_{\text{prompt}}}{C_{\text{exp}}} + C_{\text{Offset}}. \tag{7.16}$$

The result is shown in Fig. 7-17. From the fitted function and 1σ error of fit parameters, we can estimate the total number of fast neutron events by the following equation

$$N_{\text{fast-n}} = \int_{0.7MeV}^{12.2MeV} f_{\text{fast-n}}(E_{\text{prompt}})\, dE_{\text{prompt}} \tag{7.17}$$

After dividing by the detector live-time of 240.06 days, we find:

$$R_{\text{fast-n}} \;=\; 2.50 \pm 0.47 \text{ events/day} \tag{7.18}$$

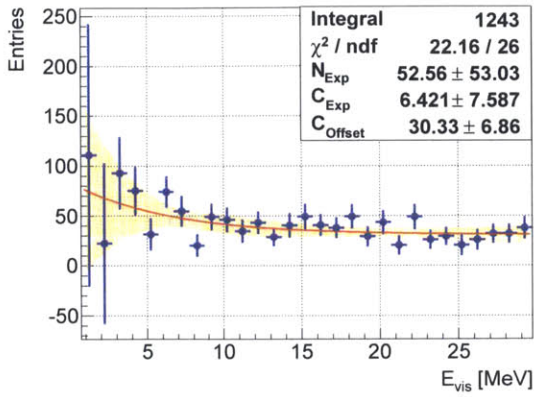

Figure 7-17: The fit result of the function in Eq. 7.16 to the fast neutron E_{prompt} spectrum.

7.3.3 Fast Neutron Analysis Cross-Check

This section describes a cross check analysis for fast neutron rate and shape estimate. An independent sample of fast neutrons from the nominal analysis method is obtained by modifying one of the IBD prompt event selection cuts: instead of $Q_{IV} < 1e4$ DUQ, we ask for $1e4 < Q_{IV} < 1e5$ DUQ. We keep all other cuts the same. This cut enhances the purity of the fast neutron sample because we are immune to radioactive backgrounds in this energy range. This means we have very low contamination from the ID-IV uncorrelated events and prompt-delayed uncorrelated events. The beauty of this analysis is the simplicity because all we require is this single modification in the selection cuts. We take the ID E_{prompt} spectrum shape from these fast neutron candidates with modified selection, and also similarly the ID E_{prompt} spectrum of accidental backgrounds using off-time selection. Subtracting the accidentals' E_{prompt} spectrum from that of the fast neutron candidates, we obtain the fast neutron E_{prompt} spectrum. Figure 7-18 shows the E_{prompt} spectrum from default IBD selection cuts (red), namely the low-Q_{IV} sample, and fast neutron candidates selected by a modified Q_{IV} prompt selection cut (blue), namely the high-Q_{IV} sample. The blue data points are normalized to match the red data points by integrating the energy range from 12 to 30 MeV, assuming this energy range is dominated by fast neutrons. We can see an excellent agreement on E_{prompt} shape between the low-Q_{IV} and high-Q_{IV} samples. Figure 7-19 shows the result of a fit using the same function defined earlier

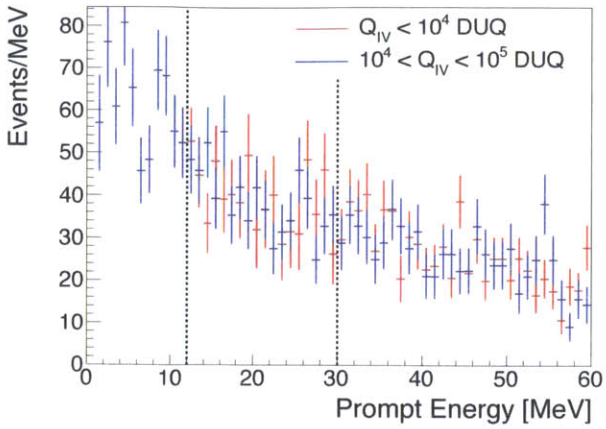

Figure 7-18: Prompt energy spectra of the low Q_{IV} (left) and high Q_{IV} cuts (right). Dotted lines show specific energy points, 12 MeV and 30 MeV from left. The high Q_{IV} spectrum is normalized to the low Q_{IV} spectrum by matching the integral of both spectra in 12 to 30 MeV.

in Eq. 7.16 to the high-Q_{IV} sample. By integrating the fitted spectrum shape in Fig. 7-19 in

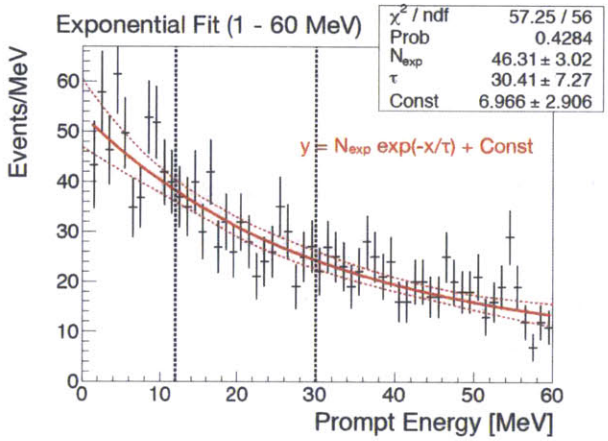

Figure 7-19: A function defined in Eq. 7.16 is fitted to the E_{prompt} spectrum of the high-Q_{IV} sample

from 0.7 to 12.2 MeV, we find

$$R_{\text{fast-n}}^{\text{High-QIV}} = 2.86 \pm 0.25 \text{ events/day}, \tag{7.19}$$

which is in a good agreement with the nominal rate estimation. The comparison of the fast neutron E_{prompt} spectrum from nominal analysis method and high-Q_{IV} method can be found in

the left plot in Fig. 7-20. The right plot in the same figure shows a comparison of fitted function shape with 1σ shape error band. Again, we find a reasonable agreement in two analyses, and hence consider this cross check analysis validates the nominal rate and shape estimation. We

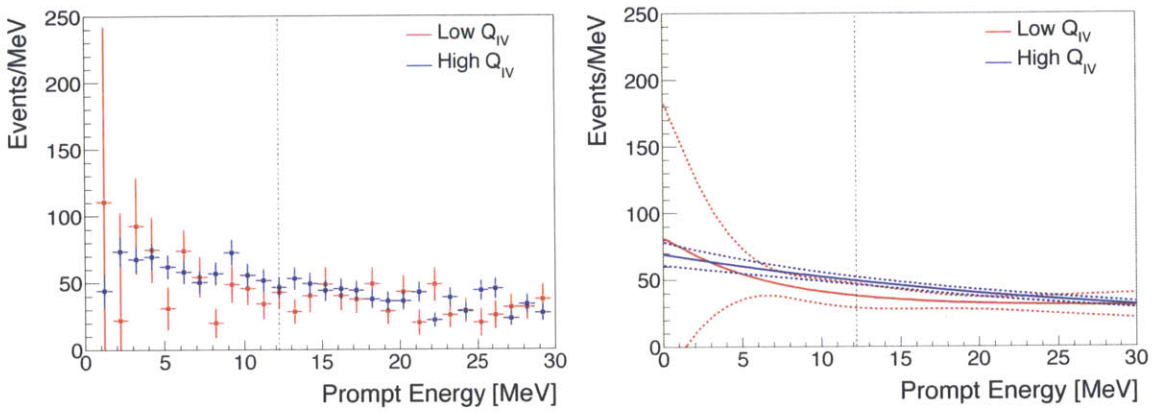

Figure 7-20: Comparison of the fast neutron sample spectrum (left) and fitted function shape (right) between the nominal fast neutron sample and high-Q_{IV} sample.

note here that, however, high-Q_{IV} sample contains fast neutrons that are completely orthogonal to fast neutrons that are contaminating our IBD signal. Despite a smaller error in both rate and shape, we did not employe high-Q_{IV} analysis outcome because we could not quantitatively justify the claim that the E_{prompt} spectrum shape is same among high-Q_{IV} and low-Q_{IV} sample in the energy range [0.7,12.2] MeV. Only handle we have had so far is to use a Monte Carlo simulation to make such a claim. Modeling muon induced fast neutrons in Monte Carlo is, however, extremely hard and not very well known. For this reason, we kept the high-Q_{IV} analysis result as a cross-check to our nominal rate and shape estimation.

7.4 Stopping-Muon Veto

A stopping-muon is a low energy muon that deposits energy in the detector as it comes to rest, and decays subsequently to produce a Michel electron. This is one of background to Gd analysis [1] but not in hydrogen analysis. This section briefly describes how stopping-muons are removed in hydrogen analysis.

7.4.1 Selection Cuts

stopping-muons are removed largely by ΔT and OV coincidene cuts. The time correlation between the prompt and delayed trigger is a typical muon lifetime of 2.2 μs. Because of our lower bound of ΔT cut, $10\mu s < \Delta T < 600\mu s$, stopping-muon background is highly suppressed in hydrogen analysis. In addition to the ΔT cut, OV coincidence cut also dramatically reduces a stopping-muon background. Because the OV detector covers a large surface area on the top of the detector, most of stopping-muons are vetoed by cutting prompt events that have a coincidence signal with OV. Only small fraction of stopping-muons that come through the chimney cannot be vetoed by the OV.

7.4.2 Verification of Zero Stopping-Muon Background

Our IBD candidates contain four kinds of events with distinct prompt-delayed correlation in terms of ΔT: (1) neutron capture time in the Target, (2) neutron capture time in the GC, (3) muon decay time, and (4) no correlation. The delayed event of an IBD, fast neutron, and ^{9}Li events are all neutron captures. It is, therefore, classified as (1) or (2) depending on where an event takes place. A stopping-muon clearly is classified to the case (3), and accidental backgrounds must be in (4) by definition. Using these facts, we can attempt to fit ΔT spectrum of IBD candidates to quantify the statistics for each of those four classes of events.

Since it is simple, we remove the accidental backgrounds by subtrcting ΔT distribution from off-time selected sample from that of IBD candidates. This leaves us a spectrum to be fitted by a triple-exponential function. In order to enhance the estimation of rates, we quantified the characteristic time for (1) to (3) before trying a combined fit. First of all, we examined ^{252}Cf source calibration data to estimate the neutron capture time in the Target. We used runs for which a ^{252}Cf source is deployed in the middle of the detector, and found $\tau_T = 26.72 \pm 0.42\mu s$. For a neutron capture time in the GC, we used the IBD candidates' ΔT spectrum in the region of $\Delta T \in [150, 600]$ μs where a neutron capture in the GC should be the only contributer to the spectrum. We found $\tau_{GC} = 194.46 \pm 8.33\mu s$. Finally, for a stopping-muon, we have run a dedicated stopping-muon selection using the following cuts: $T^{RMS}_{Start} < 40\ ns$ and $MQTQ < 0.09$ for prompt and delayed, $E_{\text{prompt}} \in [0, 60$ MeV, $E_{\text{delayed}} \in [12, 60]$ MeV, $\Delta T \in [2, 30]$ μs,

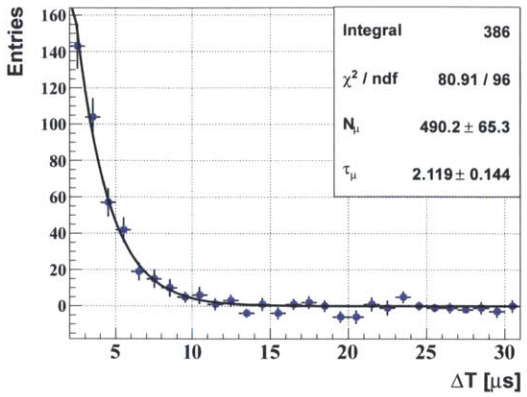

Figure 7-21: ΔT spectrum from the stopping-muon sample.

and $\Delta R < 900\ mm$. The resulting ΔT spectrum shape is shown in Fig. 7-21. The fit yields $2.119 \pm 0.144\ \mu s$ as muon lifetime, which is very reasonable[42]. A summary of all time constants obtained is shown in Tab. 7.3.

Physics	Time Constant $[\mu s]$
Muon Decay	2.119 ± 0.144
n-H in Target	26.72 ± 0.42
n-H in GC	194.46 ± 8.33

Table 7.3: Time constants used for the fit to the ΔT spectrum from the $\bar{\nu}_e$ sample.

Using the time constants summarized in Tab. 7.3, we performed a triple-exponential fit to the ΔT spectrum from the IBD candidates after subtracting the accidental backgrounds' component. This is shown in Fig. 7-22. In the fit, in order to enhance the error due to the time constants extracted from the different data sample, we defined the following χ^2 with penalty terms.

$$\chi^2 = \chi^2_{Pearson} + \sum_{i}^{\mu,NT,GC} \left(\frac{\Delta\tau_i}{\sigma_{\tau_i}} \right)^2 \tag{7.20}$$

The fit χ^2 is reasonable, and the yeilds the following as the stopping-muon component:

$$f_{\text{stop-mu}}(\Delta T) = N_{\text{stop-mu}} \cdot \exp\left(-\frac{\Delta T}{\tau_{\text{stop-mu}}} \right)$$

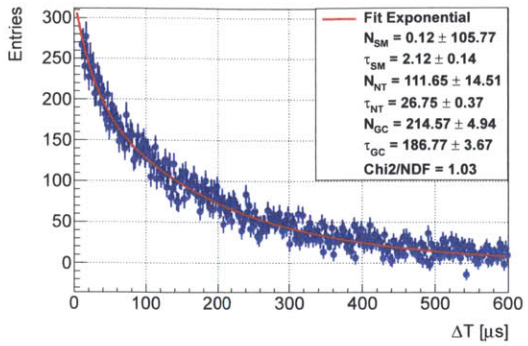

Figure 7-22: The ΔT spectrum fit with a triple exponential for ΔT spectrum of hydrogen IBD candidates after subtraction of accidental backgrounds' component.

whith $\tau_{\text{stop-mu}} = 2.119 \pm 0.144\ \mu s$ as noted earlier and $N_{\text{stop-mu}} = 0.12 \pm 105.77$ from the fit. By integrating this $f_{\text{stop-mu}}(\Delta T)$ in the range [10,600] μs, we confirm that the stopping muon rate is smaller than 0.05 events/day at 5σ level. With this result, we conclude that the stopping muon contamination to our IBD candidates is negligible.

7.5 Light Noise

This section describes a light noise background. The light noise was breifly mentioned in Sec. 5.2. The majority of this background is rejected by two pre-selection cuts, $T_{\text{start}}^{\text{RMS}}$ and MQTQ. While a large fraction of light noise event is considered to be, so far, a single uncorrelated background events, there is a small fraction that is found to make a correlated event pair.

A coincidence event of ^{214}Bi-Po and ^{212}Bi-Po arise from a natural radioactivity in the ^{238}U and ^{232}Th decay chains respectively. These coincidence events are useful to monitor the detector stability in many aspects. In both cases, a coincidence is identified as a β-decay of Bismuth isotope followed by an alpha decay of a Polonium isotope. Released energies, Q_β and Q_α, are $Q_\beta = 3.272$ MeV and $Q_\alpha = 7.687$ MeV in the case of ^{214}Bi-Po. Those of ^{212}Bi-Po are $Q_\beta = 2.254$ MeV and $Q_\alpha = 8.734$ MeV. The half-life of ^{214}Po is 164.3 μs while that of ^{212}Po is 299 ns. Neither of these serve as a background to the hydrogen analysis because of a large

quenching factor of α particle brings the delayed visible energy to be well below the E_{delayed} cut lower boundary.

However, the detector stability study using these Bi-Po coincidences showed an unexpected instability in their rates. This turned out to be correlated light noise events that are a great background to the Bi-Po analysis. Details of this study can be found in Ref. [43]. This section presents a study to quantify the possible contamination from this correlated light noise background.

7.5.1 Separation of the Correlated Light Noise Background

The possibility of a correlated LN reported in Ref. [43] directly affects the n-H analysis. The selection cuts to enhance this LN sample from the reference include well overlaps with the n-H default cuts. In the reference, three useful features of this background are reported: they are all reconstructed toward the detector center; their prompt-delayed correlation time constant is 14.1 ± 0.9 μs which is rather short; they arise in the detector since 190 days past from the beginning of the data taking. These features are used to sample and verify this background in the IBD canddiates of the hydrogen analysis. Finally, the report also presents that this backgrounds comes in a rather high multiplicity which implies they are easily vetoed by the isolation cut in the IBD selection.

Figure 7-23 shows the vertex distribution of coincidence pairs selected by the modified IBD selection cuts. The ΔT cut is modified to be $0 < \Delta T < 15$ μs and also only data taken after 190 days since the data taking is used. An anomaly at the target center is quite evident. Based on the distribution shown in Fig. 7-23, a vertex cut of $-0.8\ m < Z < 0.3\ m \otimes 0\ m < \rho^2 < 0.2\ m$ is used to sample this background. In the region selected by this vertex cut, there also exist other events such as ^{9}Li, fast neutrons and IBD. These events are all backgrounds for measuring the rate of a correalted light noise. Those backgrounds are, however, considered rather uniform. Events from the nearby volumes are used to sample those backgrounds.

To ensure this vertex cut definition collects the reported class of events in Ref. [43], ΔT distribution of this sampling method is inspected. The light noise background is collected by modified IBD selection cuts: $2 < \Delta T < 100$ μs; $-0.8 < Z < 0.3$ m; $0 < \rho^2 < 0.2$ m^2. This

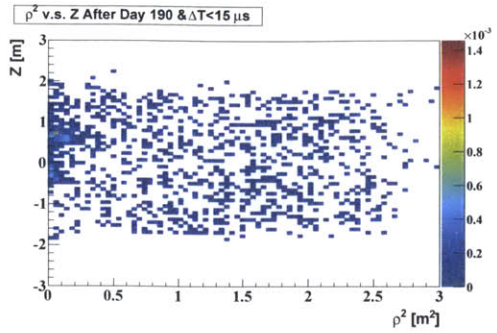

Figure 7-23: ρ^2 vs. Z vertex distribution of coincidence pairs selected with a modified IBD selection cut, $0 < \Delta T < 15$ μs. Also only data after day 190 is used to enhance the existence of the background.

volume definition is noted as V_1 in the remaining part of this section. The backgrounds for the light noise is collected from a nearby volume by modifying the vertex cut to $0.8 < Z < 0.3$ m and $0.25 < \rho^2 < 0.45$ m^2, defined as V_2. Since a neutron capture time in this region (i.e. the Target volume) is about 30 μs, the accidentals background sample in each volume is collected by using a single off-set time window at $152 < \Delta T < 250$ μs. The accidentals background's ΔT distribution is subtracted to remove the accidentals' component. The remaining ΔT distribution is that of purely correlated events, including the light noise evens and backgrounds such as IBD, fast neutrons, and ^{9}Li. Since backgrounds are considered to be uniform in space, ΔT distribution of V_1 after subtracted that of V_2 represents the ΔT distribution of the correlated light noise. This is shown on the left of Fig. 7-24. An exponential function fit yields a decaly constant of 13.88 ± 3.92 μs, which is consistent with the reported value of 14.1 ± 0.9 μs. A large error is due to a lack of statistics in the IBD sample compared to the Bi-Po sample.

In order to ensure that the volume V_1 contains all the light noise events, and also to confirm V_3 does not contain them, another volume definition V_3, $-0.8 < Z < 0.3$ m and $0.45 < \rho^2 < 0.65$ m^2, is defined and studied. The ΔT distribution of V_1 after subtracting that of V_3 is shown in the middle of Fig. 7-24. The fit yields the exponential time constant of 11.22 ± 3.49 μs, which is again consistent with the reported value. The right of Fig. 7-24 shows the ΔT distribution of V_2 after subtracting V_3. The flat distribution confirms that V_2 volume does not contain this class of background.

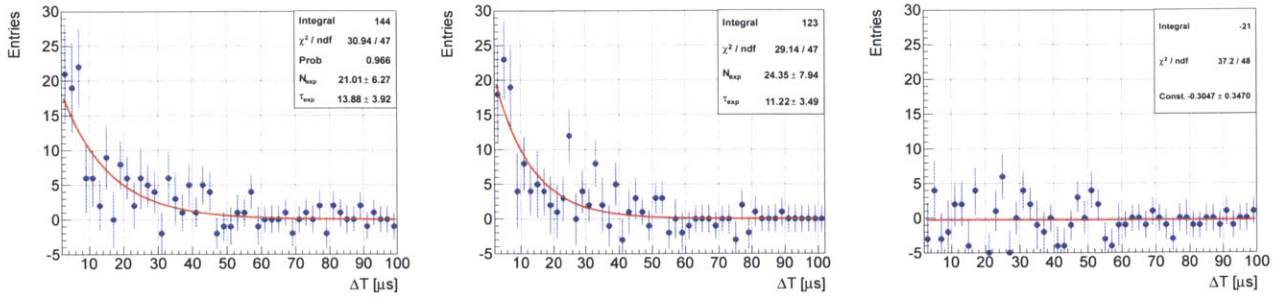

Figure 7-24: Subtracted ΔT distribution for the three volume cuts, V_1, V_2, and V_3 (see the main text for the definition). From the left to right, ΔT distribution is made by a volume-wise subtraction of $V_1 - V_2$, $V_1 - V_3$, and $V_2 - V_3$.

7.5.2 The Rate of Contamination

The rate is estimated by a simple counting method. The number of IBD candidates are counted in the volume region V_1, V_2 and V_3 (see the previous section for the definition of each volume). For the each volume definition, accidentals are counted and subtracted using a similar method described in the previous section. A subtraction of V_2 from V_1 yields the correlated light noise rate of 0.32 ± 0.07 events/day. A subtraction of V_3 from V_1 yields 0.25 ± 0.07 events/day. The $V_1 - V_2$ result has less effect from accidentals since V_3 is closer to the γ-catcher where accidentals background rate is high. Accordingly, the rate estimation from the $V_1 - V_2$ volume subtraction is used as the nominal rate estimation.

One may ask a question whether the volume V_1 fully contains the sample of light noise or not. To answer question, the size of V_1 and V_2 are changed to enlarge their volume. V_1 is enlarged to $-0.9 < Z < 0.5$ m and $0 < \rho^2 < 0.3$ m^2, and accordingly V_2 is changed to keep the same size of the volume just outside V_1. The rate estimation from these volumes yield the rate of 0.29 ± 0.08 events/day, which is completely consistent with the estimates made with the original V_1 volume definition. From this fact, it is considered that the original V_1 volume definition is sufficiently large enough to contain all of the correalted light noise events. As a reference, ΔT distribution of the modified volume is also made and shown in Fig. 7-25. The exponential time constant from the fit yields 13.34 ± 4.18 μs, which is again consistent with the report [43].

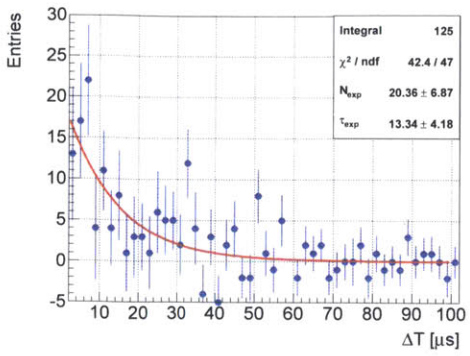

Figure 7-25: Subtracted ΔT distribution of an enlarged V_1 and V_2 definition for the cross-check.

7.5.3 The Shape of Contamination

Following the rate analysis method with V_1, V_2 and V_3 volume definitions, a similar analysis method is performed to extract the E_{prompt} spectrum shape of this background. The $V_1 - V_2$ and $V_1 - V_3$ spectra are shown in Fig. 7-26. As expected, both $V_1 - V_2$ and $V_1 - V_3$

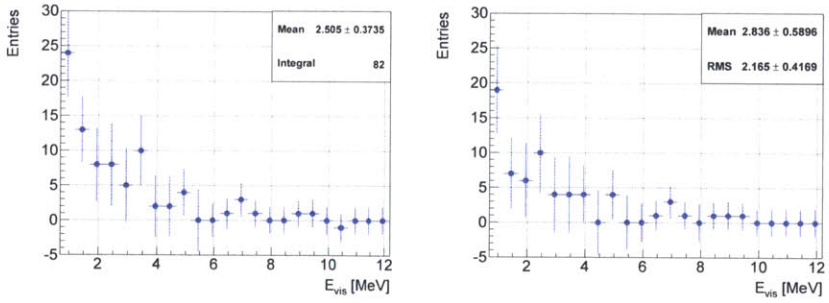

Figure 7-26: Subtracted E_{prompt} spectrum shape for $V_1 - V_2$ (left) and $V_1 - V_3$ (right). Both spectra are scaled to the corresponding rate estimate.

spectra are compatible although the shape uncertainty is large. The spectrum from $V_1 - V_2$ subtraction is used as the nominal input to the oscillation analysis with the quoted rate of 0.32 ± 0.07 events/day.

Background	hydrogen (event/day)	Gd (event/day)
Accidentals	73.45 ± 0.16	0.261 ± 0.002
^{9}Li	2.84 ± 1.15	1.25 ± 0.54
Fast Neutrons	2.50 ± 0.47	0.30 ± 0.14
Stopping-μ	N/A	0.34 ± 0.18
Light Noise	0.32 ± 0.07	N/A
Total	79.11 ± 1.25	2.18 ± 0.58
IBD Candidates	151.1 ± 0.8	36.2 ± 0.4

Table 7.4: A summary table of backgrounds in the hydrogen analysis. As a reference, values from the Gd analysis are also shown.

7.6 Summary

Combining all backgrounds' contribution, the total background rate is 79.11 ± 1.25 events/day in the hydrogen analysis. Although it is high, a large fraction of it comes from the accidentals background which is well understood. Table 7.4 summarizes the rate of each background as well as the total. The same table also holds values from the Gd analysis for shared backgrounds.

7.7 Validation with Reactor OFF-OFF

Data taken during the period in which both reactors are turned off (OFF-OFF) is very useful for us to validate our BG analysis. In this section, we show a comparison of an estimated and observed BG rate and shape.

7.7.1 Data Sample

The data sample are gathered from the all reactor OFF-OFF period in both 2011 October and 2012 June. The run-time, live-time, and OV-corrected live-time are 7.528, 7.192, and 7.188 days in respective order. The IBD candidate selection yields 599 ± 24.47 candidates during the whole period.

7.7.2 Accidentals Rate and Time Stability

The accidentals' rate estimation is the first and possibly the most important step of the OFF-OFF data analysis. We cannot use, however, the accidentals' rate extracted earlier because the accidentals' rate have a time variation as shown in Fig. 7-4. Hence, we have to extract the accidentals' rate independently using the OFF-OFF data sample.

The extraction is performed using both the off-time selection method and analytical estimation [3]. Both methods show a consistent rate estimation for each day. For the OFF-OFF sample we found the accidentals' correction factors as follows:

$$f^{\text{delayed}}_{\text{isolation}} = 1.0185 \pm 0.0001 \text{ and } f^{\text{delayed}}_{\mu\text{-veto}} = 1.032692 \pm 0.002 \tag{7.21}$$

where the error on $f^{\text{delayed}}_{\mu\text{-veto}}$ is assuming the 0.01 day error on the run/live-time. The off-time selection yields 61508 events with 125 off-time windows. With these correction factors, we find:

$$\text{R}^{\text{measured}}_{\text{accidentals}} = 72.00 \pm 0.36 \text{ events/day} \tag{7.22}$$

The input parameters for the analytical estimates are derived from the OFF-OFF data sample, and summarized in the table 7.5. We find the analytical method estiamtes the overall

R_{prompt}	$\text{R}_{\text{delayed}}$	$\epsilon^{\text{prompt}}_{\text{isolation}}$	$\epsilon_{\Delta R}$	$f^{\text{delayed}}_{\mu\text{-veto}}$
8.131 ± 0.012 Hz	2.994 ± 0.005 Hz	0.9818 ± 0.0001	0.0599 ± 0.0002	1.0327 ± 0.0021

Table 7.5: Summary of parameter values to be used for OFF-OFF analytical rate estimation.

rate to be:

$$\text{R}^{\text{analytical}}_{\text{accidentals}} = 71.95 \pm 0.29 \tag{7.23}$$

which is in a very good agreement with the selected accidentals' rate with the off-time window selection. The day-by-day stability of the accidentals' rate measurement is shown in Fig. 7-27, which proves a cross-check using the analytical method agrees with the off-time selection in daily basis.

[3] See Sec. 7.1.1 and 7.1.3

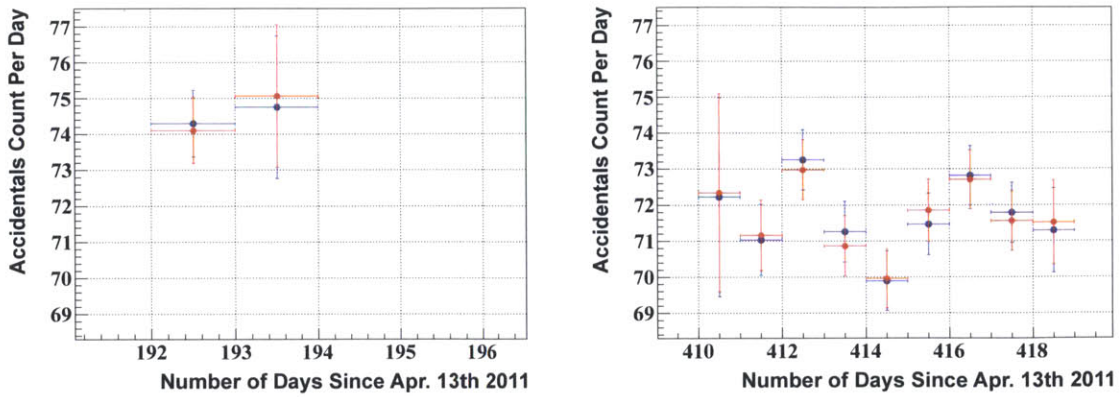

Figure 7-27: A comparison of accidentals' daily statistics estimation between the off-time window (blue) and analytical method (red). A very good agreement is found over all OFF-OFF period, as was the case for the DC2ndPub data sample.

Further, we have also checked the correction factor to the off-time selection event statistics by fitting a ratio histogram of IBD candidates to off-time selected sample at high ΔR. This is shown in Fig. 7-28. The ratio of the IBD candidates to off-time sample statistics is found to be 0.991 ± 0.012, and hence the correction factor is considered to be working well. We, therefore, use the noted accidentals' rate for the further OFF-OFF data analysis.

7.7.3 OFF-OFF Event Rate and Shape After Accidentals Subtraction

Using the Eq. 7.22, the event rate after the accidentals' subtraction is:

$$\mathrm{R}\,{}^{\text{off-off}}_{^9\text{Li+fast-n}} = 11.33 \pm 3.41 \text{ events/day} \tag{7.24}$$

The combined fast neutron and ^{9}Li event rate from the DC2ndPub estimates is:

$$\mathrm{R}\,{}^{\text{H-Analysis}}_{^9\text{Li+fast-n}} = 5.39 \pm 1.20 \text{ events/day} \tag{7.25}$$

which is roughly $\tilde{1}.74\sigma$ away from the measured rate.

The prompt energy spectrum comparison is done at a very basic level. We have used nominal

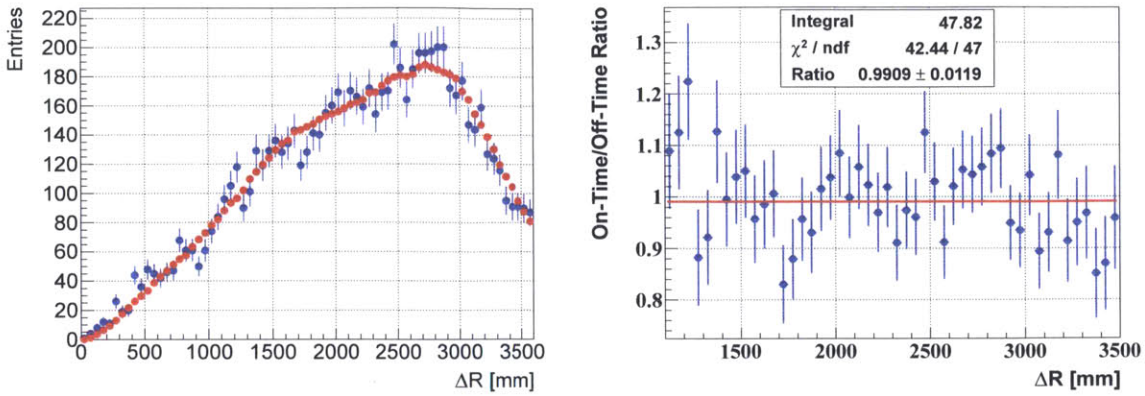

Figure 7-28: The left figure shows ΔR distribution for IBD candidates and off-time selected data sample in blue and red color, respectively. The off-time data sample is scaled using a derived scaled factor for the OFF-OFF period. The right figure shows the bin-by-bin ratio of two distributions on the left for $1100 < \Delta R < 3500$ mm.

rate and shape estimate for the ^{9}Li and fast neutron backrounds in the hydrogen analysis. We have ignored rate and shape uncertainty: the both spectrums are normalized to the central value of the estimated rate times the livetime of the detector during the OFF-OFF period. This comparison of E_{prompt} is shown in the left of Fig. 7-29. In particular, between 12.7 and

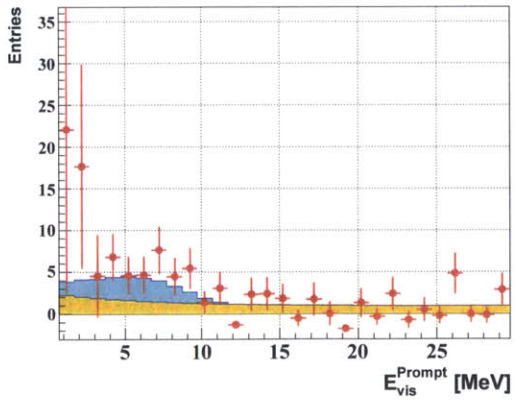

Figure 7-29: OFF-OFF data E_{prompt} is shown in the red data points. The stacked histograms represent ^{9}Li (cyan) and fast neutrons (orange).

29.7 MeV, we find the OFF-OFF data statistics of 16.02 ± 6.50 events in the histogram of Fig. 7-29. With our nominal rate and shape estimate of fast neutron backgrounds, we expect

16.74 ± 1.25 events. This is in a complete agreement.

These very minimal set of cross-check with the OFF-OFF data sample shows both the rate and the shape of accidentals, ^{9}Li and fast neutrons all seem reasonable.

Chapter 8

Signal Prediction

This chapter describes the signal prediction. Since this is a common procedure between the hydrogen and Gd analyses [1], the description is kept brief. The only difference is the signal selection method which is already described in chapter 5. The summary can be found in Sec. 8.3 where the predicted number of event candidates after applying the MC correction factor (see chapter 6) is presented.

8.1 Simulation of the Fission Rate in the Chooz Reactor Cores

Fission chains in nuclear reactor cores involve multiple beta decays that make a nuclear reactor core a rich, isotropic source of $\bar{\nu}_e$. Main fuel components include ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu. Taking the example of ^{235}U, each fission releases roughly 200 MeV energy [44] involving 6 beta decays on average [45]. As a result, approximately $\approx 10^{20}$ $\bar{\nu}_e$ $\mathrm{GW}_{\mathrm{th}}^{-1}s^{-1}$ are produced.

In order to know the number of $\bar{\nu}_e$ with precision, one has to either measure it with the near detector or extrapolate from a good understanding of fission rates in the reactor core. As the experiment does not have the near detector running, the latter treatment is taken for this assessment. The rate is, however, also constrained by DATA taken by Bugey-4 experiment [46] which helps to lower the systematic uncertainty for a precise measurement of θ_{13}. This section presents the details of a predicion of the fission rate in the reactor core, which gives a pre-

dicted number of $\bar{\nu}_e$ at the reactor core. The next section describes how the $\bar{\nu}_e$ rate is further constrained using the Bugey-4 measurement.

The flux of $\bar{\nu}_e$ at the reactor core is calculated by summing up $\bar{\nu}_e$ flux from main fuel isotopes using the following equation

$$\phi_{\bar{\nu}_e}(E,t) = \sum_{k}^{\text{Fission Isotopes}} f_k(t) \cdot \phi^k_{\bar{\nu}_e}(E) \tag{8.1}$$

where k denotes an index of four fission isotopes, f represents a fission rate of k-th isotope, and $\phi^k_{\bar{\nu}_e}$ denotes the $\bar{\nu}_e$ flux for k-th isotope. Section 8.1.1 describes a framework to make the fission rate prediction at a given time t, and Sec. 8.1.2 describes the predicted $\bar{\nu}_e$ spectrum shape at the reactor core.

8.1.1 Core Simulation

Knowing the fuel composition of the reactor core is necessary to calculate the $\bar{\nu}_e$ flux. Because the fuel composition cannot be measured while the reactor is under operation, in order to know the fission fraction at any given time, one has to extrapolate the fuel composition based on the initial condition, core assembly information, and the power history of the reactor which is relevant to the total fission rate over time. These input information are provided by EDF, the electricity company in France. This is input into two independent simulation frameworks used in Double CHOOZ, MURE and DRAGON, in order to predict the time evolution of fission fraction over time. An example thermal power history data set provided by EDF is shown in Fig. 8-1.

MURE performs a 3D, N-particle simulation of a reactor core using a neutron transport Monte Carlo simulation framework, called MCNP [47]. DRAGON [48], on the other hand, is a lattice-based 2D assembly simulatioon program which performs a deterministic calculation. Both codes have been benchmarked against Takahama-3 reactor in Japan from which three fuel rods went into a destructive assay that allowed analysis of its chemical composition. The result from this benchmark and a comparison with the prediction are used to assign the systematic uncertainty on the fractional fission rate prediction. Because MURE does a 3D simulation of

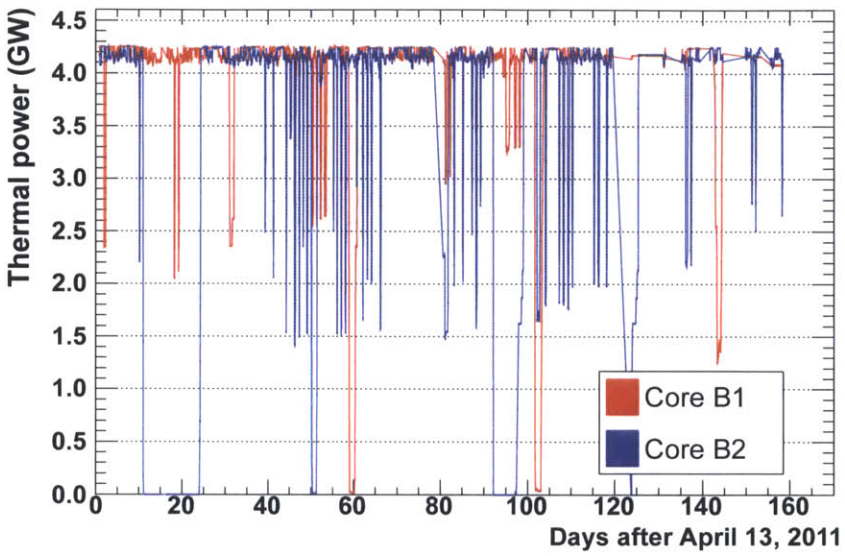

Figure 8-1: The average thermal power data monitored and provided by the electrocity company in France as a function of days since the beginning of $\bar{\nu}_e$ data taking.

the entire fuel rod, its output is used as the nominal value of the fission rate prediction. The predicted fission rates over time for main fuel isotopes are shown in Fig. 8-2.

DRAGON is used to cross-check the assembly-level fission rate predicition from MURE, and the difference is included in the systematic uncertainty. The overall uncertainty takes into account the initial fuel composition in the present fuel cycle, uncertainties in the moderator (water) density, operation of control rods, monitored temperature, and all other second order effects used in the code input parameters. The uncertainty is evaluated per fission isotope, and its time evolution is shown in Fig. 8-3.

8.1.2 Neutrino Spectrum at Reactor Core

The $\bar{\nu}_e$ flux from k^{th} isotope, $\phi^k_{\bar{\nu}_e}$ in Eq. 8.1, is estimated by a sum of beta spectrum of all daughter isotopes as shown below

$$\phi^k_{\bar{\nu}_e}(E) = \sum_i^{\text{all daughters}} A^k_i(t) \left[\sum_b^{\text{all branches}} B_j \cdot S_j(E) \right]. \tag{8.2}$$

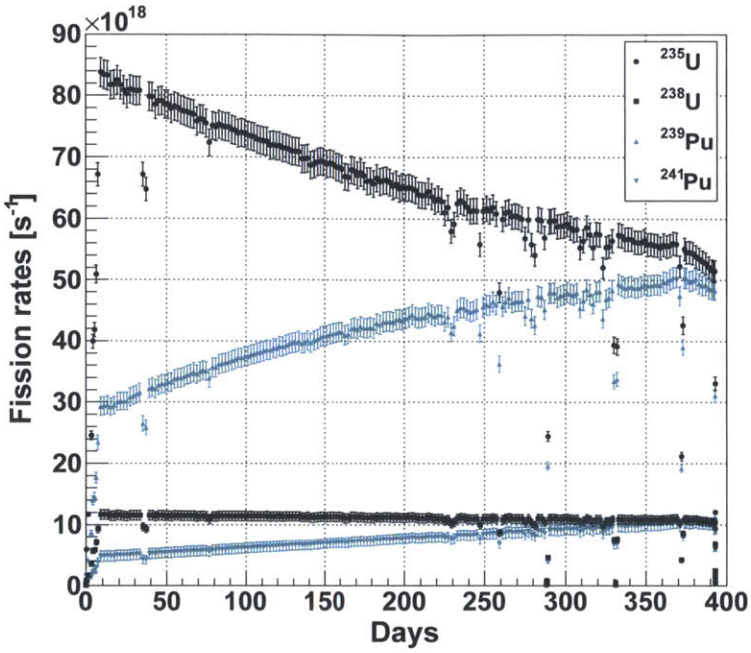

Figure 8-2: Fission rates from the main fission isotopes, ^{239}Pu, ^{241}Pu, ^{235}U, and ^{238}U, simulated by MURE with input data from the EDF.

In order to estimate $\phi^k_{\bar{\nu}_e}$, one could either take an overall beta decay spectrum measurement for the k^{th} isotope or sum up the measured beta decay spectrum from every single beta decay branch. The first approach is explored by Huber et. al [49] using the data taken at the Institute Laure Langevin (ILL) high flux reactor in Grenoble, France [50, 51, 52] for the isotope of ^{235}U, ^{239}Pu, and ^{241}. In this measurement, uranium and plutonium oxide, a typical fuel compound for nuclear reactors, are irradiated by a thermal neutron flux for 12 to 36 hours. The overall outgoing electron energy is measured by a magnetic spectrometer in steps of 50 to 100 keV. For ^{238}U isotope, the second approach, called *ab initio* [10], is recently developed and computed by Mueller et. al, and is used since there is no measurement of an overall beta spectrum available. The *ab initio*method also allows a correction for ^{235}U, ^{239}Pu, and ^{241}Pu isotopes due to long lived isotopes that pile up beyond 12 to 36 hours of the irradiation time in the ILL measurement [10]. The resulting spectrum shape and uncertainties are shown in Fig. 8-4.

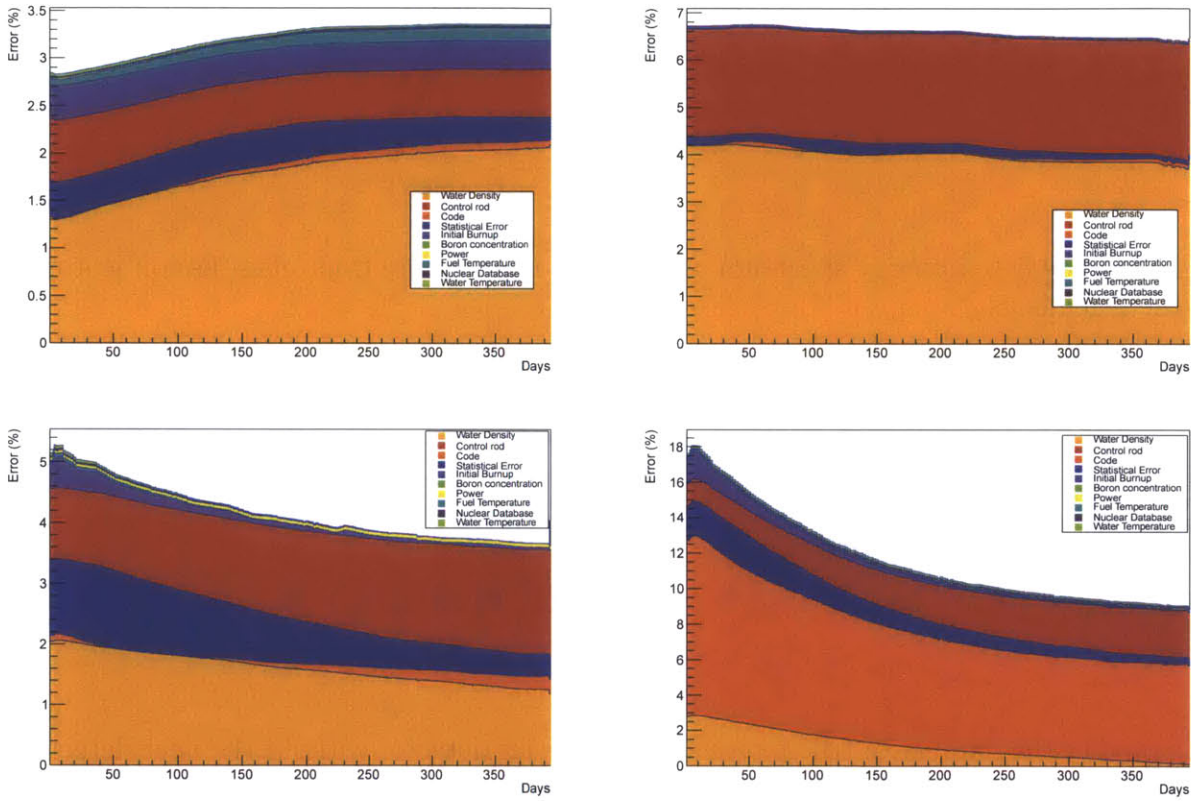

Figure 8-3: Error budgets for the main fission isotopes shown in Fig. 8-2 as a function of time based on the simulated burn-up by MURE.

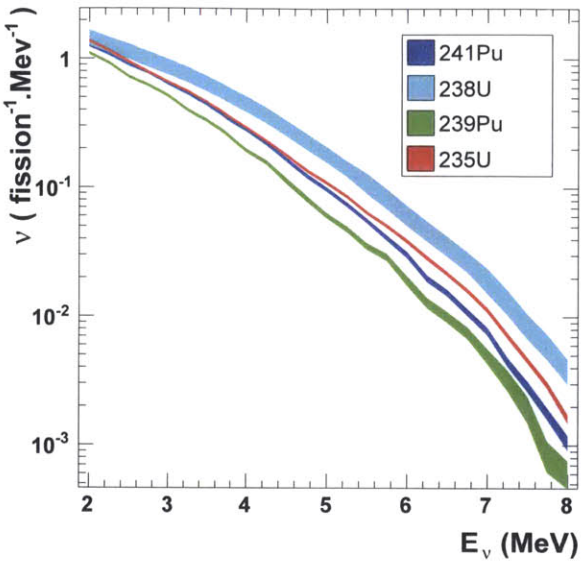

Figure 8-4: The predicted energy spectrum of $\bar{\nu}_e$ at the reactor core from main fission isotopes shown in the legends.

8.1.3 $\bar{\nu}_e$ Flux in the Detector

In order to predict the number of IBD interactions in the far detector without the near detector, one needs to predict the number of IBD interactions for each of the reactor cores. Below, α_k is defined as a fractional fission rate as opposed to the absolute fission rate, f_k,

$$\alpha_k = \frac{f_k}{\sum_\ell f_\ell} \tag{8.3}$$

Defining an average energy per fission isotope, $\langle E_f \rangle_k$, and using the total thermal power output of the reactor, $P_{th} = \sum_k f_k \langle E_f \rangle_k$, one can write

$$f_k = \alpha_k \sum_\ell \frac{P_{th}}{\alpha_\ell \langle E_f \rangle_\ell} \tag{8.4}$$

which relates the monitored total thermal power output of the reactor and fission rates. The IBD interaction rate at the detector is

$$\frac{dN_{\bar{\nu}_e}}{dt} = \sum_{\mathrm{R=B1,B2}}^{\mathrm{Reactors}} \frac{N_p}{4\pi L_R^2} \frac{P_{th,R}}{\langle E_f \rangle_R} \langle \sigma_f \rangle_R \tag{8.5}$$

where an extra subscript, R, denotes a parameter evaluated at each reactor and $\langle \sigma_f \rangle_R$ denotes the IBD cross-section weighted by $\bar{\nu}_e$ flux per fission,

$$\langle \sigma_f \rangle = \sum_k \alpha_k \int \phi^k_{\bar{\nu}_e}(E) \sigma_{IBD}(E) dE. \tag{8.6}$$

The energy spectrum prediction can be made by avoiding an integration over all energy region. The number of IBD interactions between E and $E + dE$ is

$$\frac{dN_{\bar{\nu}_e}}{dt\, dE} = \sum_{\mathrm{R=B1,B2}}^{\mathrm{Reactors}} \left[\frac{N_p}{4\pi L_R^2} \frac{P_{th,R}}{\langle E_f \rangle_R} \sum_k \alpha_{k,R} \phi^k_{\bar{\nu}_e}(E) \sigma_{IBD}(E) \right]. \tag{8.7}$$

8.1.4 BUGEY4 Anchor Point

The uncertainty in the prediction described so far is dominated by an estimation of the fractional fission rates, α_k, from the reactor core simulation. One motivation of having a near detector is to reduce this systematic uncertainty by measuring $\langle \sigma_f \rangle$ through a direct measurement with a very short baseline. While installing the near detector, the collaboration took a descision of using the measurement by BUGEY4 which measured $\langle \sigma_f \rangle^{BUGEY}$ at 15 m away from the 2.8 GW_{th} PWR BUGEY power plant. This measurement carries only 1.4 % uncertainty, and the core geometry is very similar to that of Chooz B1 and B2. One can then write

$$\langle \sigma_f \rangle_R = \langle \sigma_f \rangle^{BUGEY} + \sum_k \left(\alpha_k^R - \alpha_k^{BUGEY} \right) \langle \sigma_f \rangle_k \tag{8.8}$$

Because BUGEY and Chooz B1 and B2 reactor cores are similar, $(\alpha_k^R - \alpha_k^{BUGEY})$ the second term becomes very small, suppressing an uncertainty coming from $\langle \sigma_f \rangle_k$. By using the BUGEY4 measurement as an anchor point in this manner, a systematic uncertainty in $\langle \sigma_f \rangle_R$ is dominated by that of $\langle \sigma_f \rangle^{BUGEY}$.

Source	Uncertainty (%)
Baseline Length, L	0.038
Thermal Power, P_{th}	0.46
Average Energy per Fission, $\langle E_f \rangle$	0.16
Fractional Fission Rate, α_k	0.87
Flux Weighted Cross Section, $\langle \sigma_f \rangle$	0.17
BUGEY4 Normalization, $\langle \sigma_f \rangle^{BUGEY}$	1.42
Total	1.75

Table 8.1: Uncertainties in the $\bar{\nu}_e$ flux prediction.

8.1.5 Summary

Table 8.1 summarizes various sources of error with their size. The total uncertainty becomes 1.75 % which is the dominant error in the single detector phase of the experiment. While the BUGEY4 normalization carries the most dominant uncertainty, this results in almost a factor of two improvement in the total uncertainty by suppressing the uncertainty on $\langle \sigma_f \rangle >$ as mentioned earlier.

8.2 Detector Simulation

8.2.1 Event Generation

Each DAQ run length for a normal physics run is one hour, and corresponding MC events are generated for all DATA runs. Thermal power history, P_{th}, provided by EDF is binned into 15 minutes to twelve hour period, and an average fission rate of main isotopes are calculated for twelve to forty-eight hours long period. The fission rates and thermal power computed by a simulation, f_k^{sim} and P_{th}^{sim}, are used to obtain f_k in the following manner

$$f_k = \frac{P_{th}}{P_{th}^{sim}} f_k^{sim} \tag{8.9}$$

where P_{th} is the mean of thermal power during the longer averaging period for the simulation.

The expected number of IBD events is computed based on Eq. 8.5 with the average thermal power, the fission rates, f_k, and the length of each run. The actual number of events to be

generated in MC is determined by a random draw using a Poisson distribution with a hundred times the mean of the expected number of events. A factor of hundred is introduced to generate a high statistics MC sample to investigate the statistical effect to the sensitivity of θ_{13} analysis. The energy of $\bar{\nu}_e$ for each draw is randomly chosen from a binned, flux weighted IBD cross section. The binning size is chosen to be 250 keV from 1.625 MeV to 11.125 MeV. The detector volumes used in this simulation include the Target (liquid and tank), γ-catcher (liquid and tank), and Buffer liquid. The vertex of an IBD event is chosen randomly in each volume.

For each IBD interaction, a positron and neutron are simulated and fed into the detector MC simulation framework. Their momentum are calculated based on the incoming $\bar{\nu}_e$ energy, and istotropically distributed in their center of mass frame.

8.2.2 Geant4: Particle-Tracking Simulation

Generated e^+ and neutrons are fed into the detector and particle tracking simulation framework, `DCGLG4sim` , which is a wrapper for the particle simulation package, `Geant4` [53], and includes the Double CHOOZ specific geometry and material properties.

`Geant4` is a Monte Carlo package that simulates each step of a particle track based on provided physics libraries with data. Energy loss of e^- and e^+ are described by ionization, inelastic of nuclei, elastic scattering off another charged particle, Cherenkov and bremsstrahlung radiation. The default configuration includes a custom neutron thermalization library called `NeutronTH` which corrects for moleculer bonding effects in the neutron transport during thermalization. A gamma ray generated from a neutron capture loses its energy through Cherenkov radiation, photoelectric effects, Compton scattering, and a e^+e^- pair production.

`DCGLG4sim` implements the Double CHOOZ detector geometry and material properties such as material composition and optical property. The detector geometry include all detector volumes with calibration devices in the sensitive volume. Proton numbers of liquids are matched with the detailed measurement in the laboratory. Optical photons in the detector are simulated from either Cherenkov radiation or scintillation liht emission. For highly ionizing particles, a semi-empirical Birk's law [54] is employed to estimate the light yield. An input value, called Birk's constant, is measured for an electron and α particle from the lab measurements for

the Target and γ-catcher liquids. `DCGLG4sim` includes numbers of material dependent model parameters from the lab measurements. These include the attenuation length as a function of wavelength, scintillation light yield, resolution and spectra, and also re-emission probability. The collection and quantum efficiency of PMTs are also simulated in `DCGLG4sim` [55].

8.2.3 Read-Out Electronics Simulation

The read-out electronics simulation is done by a package called `RoSS` . This takes in the generated PE hit timing vector per PMT from `DCGLG4sim` output, and generates a waveform that mimics an output from FADC in DATA. A SPE waveform shape is tuned based on DATA following an assumption of Landau distribution fit on the baseline. `RoSS` introduces, in addition, randomly generated dark noise as well as PMT-to-PMT variation of electronics gain which mean and spread are tuned based on DATA.

8.2.4 A la DATA Reconstruction

The output of `RoSS` is meant to mimic the output of FADCs, and hence it is directly fed into the chain of reconstruction processing identical to DATA including vertex, track, and energy reconstructions.

The PMT-to-PMT gain calibration on MC sample is handled in the same way as it is done for the DATA except a sample for studying various corrections are different (see Sec. 4.3). As a result, a very good agreement of the delayed neutron capture energy spectrum can be seen for IBD delayed candidates' event among DATA and MC in Fig. 5-7.

8.3 Summary

The reconstructed MC sample is run through the same IBD candidate selection method described in chapter 5. As described in chapter 6, however, there is a certain selection inefficiency that only exists in DATA because DATA include non-$\bar{\nu}_e$ events. The selected event statistics must be corrected, therefore, using the MC correction factor presented in Sec. 6.4. Taking into account of the MC correction factor, the estimated signal statistics is 17685.6 ± 516.6 events.

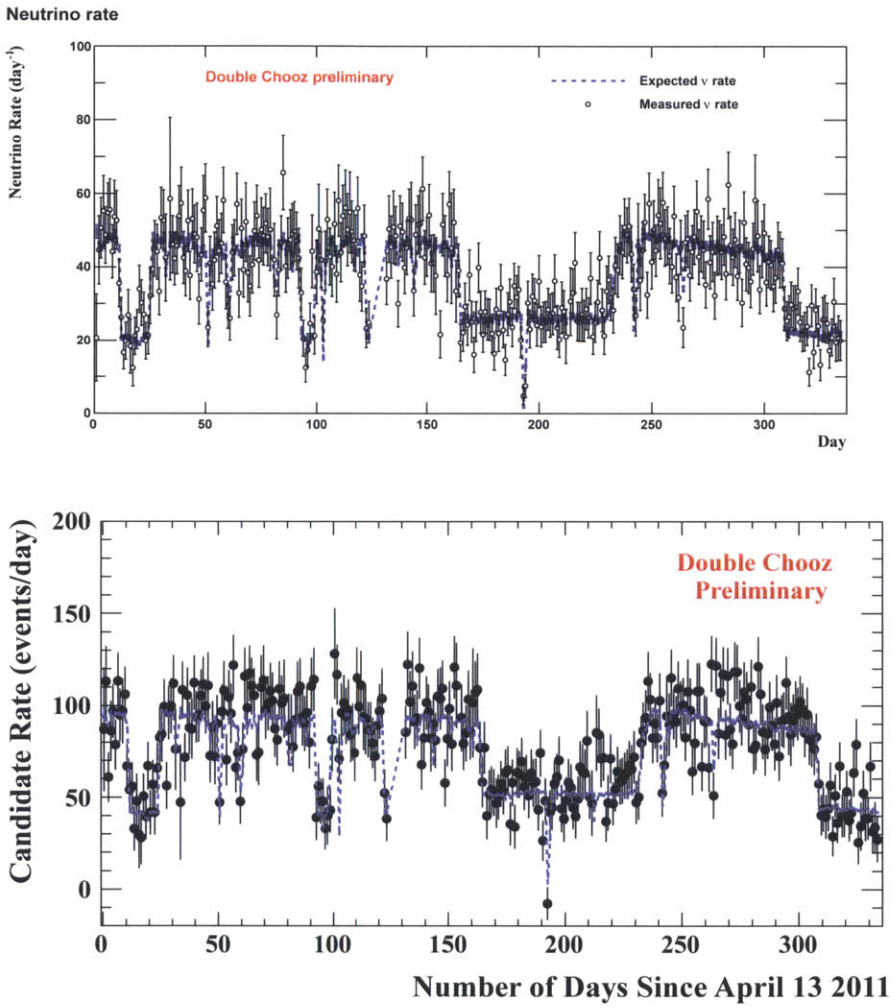

Figure 8-5: The IBD candidate rate stability plots for the Gd analysis (top) and the hydrogen analysis (bottom). The horizontal axis shows the number of elapsed days since the beginning of the data taking period. The vertical axis shows the number of events per day. The predicted rate is shown in blue while the observed rate is shown in black. Backgrounds are subtracted.

Figure 8-5 shows the expected vs. observed number of IBD candidates over the data taking period. The figure nicely shows the variation of reactor power over in the prediction followed by the observed candidates' rate from DATA.

Chapter 9

Oscillation Analysis

This chapter presents the final result of the hydrogen analysis using all inputs described previously. The procedure of the oscillation fit follows the already published method in the Gd analysis, and is briefly reviewed in Sec. 9.1. The result is available in the published manuscript which is shown in Sec. 9.2. Finally, summary and future prospects are briefly described in Sec. 9.3.

9.1 Chi-Square Approach

The oscillation fit follows the methodology used for the Gd analysis, which is based on the χ^2 statistic [1] using the prompt energy spectrum of IBD candidates. The χ^2 is defined for a comparison of DATA and MC binned histogram distributions. This method follows the generic definition of χ^2 which is

$$\chi^2 = \sum_{i,j} \underbrace{\left(X_i^{\mathrm{DATA}} - X_i^{\mathrm{MC}}\right)}_{\delta(X)_i} \mathcal{M}_{ij} \left(X_j^{\mathrm{DATA}} - X_j^{\mathrm{MC}}\right) \tag{9.1}$$

where X_i represents the i^{th} bin content of the prompt energy histogram and $\mathcal{M}_{ij}$ is an error matrix describing a bin-to-bin correlation. The index i and j run over the histogram bins. As shown above, a shorthand notation of $\delta(X)_i$ is used to represent a difference of DATA and MC prompt energy distribution in the i^{th} bin.

In addition, pull terms are added for the ^{9}Li rate, fast neutron rate, energy scale, and Δm^2_{31} value to let the fitter constrain these quantities. The input values for these pull terms are the results presented in the previous chapters. The pull terms are particularly useful for ^{9}Li and fast neutron backgrounds as their error in the prediction is large. For accidental and correlated light noise backgrounds, pull terms are not included because they are either very well understood or have a very low rate. The full expression is

$$\chi^2 = \sum_{i,j} \delta(X)_i \mathcal{M}_{ij} \delta(X)^T_j + \underbrace{\sum_b^{\text{FN},^9\text{Li}} \frac{(P_b)^2}{\sigma_b^2} + \frac{(\alpha_E - 1)^2}{\sigma^2_{\alpha_E}} + \frac{(\Delta m^2_{31} - (\Delta m^2_{31})_{\text{MINOS}})^2}{\sigma^2_{\text{MINOS}}}}_{\text{Pull Terms}}. \tag{9.2}$$

P_b is a pull term for the rate of a background "b," an index that runs over fast neutrons (FN) and Lithium-9 (^{9}Li), and σ_b is the associated uncertainty. α_E is the energy scale with an uncertainty σ_{α_E}. Finally, the Δm^2_{31} term is also added among the pull terms to let the fitter exploit a possible solution within the error obtained from the MINOS results [30]. The difference between DATA and the prediction is defined as

$$\delta(X)_i \;=\; X_i^{\text{DATA}} - \left(\sum_R^{\text{Reactors}} X_i^{\bar{\nu}_e,R} + \sum_b^{\text{Backgrounds}} X_i^b(P_b) \right). \tag{9.3}$$

where $X_i^{\bar{\nu}_e,R}$ is the predicted number of $\bar{\nu}_e$ from the reactor "R" and X_i^b is the number of predicted background class "b" in i-th histogram bin. The error matrix consists of four components:

$$\mathcal{M}_{ij} \;=\; \mathcal{M}_{ij}^{\text{Signal}} + \mathcal{M}_{ij}^{\text{Detector}} + \mathcal{M}_{ij}^{\text{Statistics}} + \sum_b^{\text{Backgrounds}} \mathcal{M}_{ij}^b \tag{9.4}$$

where the four terms in the parentheses represent covariance matrices of the signal rate and spectrum error, uncertainty in the detector response, statistical uncertainty, and background rates and spectra in respective order. The addition of covariance matrices is equivalent to an addition of errors in quadrature under an assumption of uncorrelated error components.

The covariance matrices are generated following the same approach taken in Ref. [1]. Only input parameter values have been changed. The statistical and accidental backgrounds' covariance matrices are a simple diagonal matrix with a normalization error. The spectral and

rate uncertainties are included in the ^{9}Li backgrounds' and fast neutron backgrounds' matrices. The detector and reactor covariance matrices are generated using a technique called `MultiSim` [1, 56], which is based on a large number of fake measurements by performing a random draw with input parameter values. The variance of the oscillation input parameters

Source of uncertainty	Hydrogen variance	Gd variance
Total statistical error	1.08%	1.12%
Accidentals	0.22%	0.01%
Li-9	1.56%	1.46%
Fast neutrons	0.64%	0.54%
Correlated light noise	0.10%	N/A
Energy scale	0.32%	0.32%
Detection efficiency	1.57%	1.01%
Reactor	1.75%	1.76%

Table 9.1: A comparison of uncertainties in the hydrogen and Gd analyses [1, 4]. Here, the variance is defined as: $\sqrt{\sum_{i,j} \mathcal{M}_{ij}}/N_{signal}$, where $\mathcal{M}_{ij}$ is the relevant covariance matrix, evaluated at $\sin^2 2\theta_{13} = 0$ and N_{signal} is the total number of signal events predicted at $\sin^2 2\theta_{13} = 0$.

are shown in Tab. 9.1. The table also contains those for the Gd analysis for a comparison.

9.2 Oscillation Analysis Result

First Measurement of θ_{13} From Delayed Neutron Capture on Hydrogen in the Double Chooz Experiment

Y. Abe,[30] C. Aberle,[21] J.C. dos Anjos,[5] J.C. Barriere,[15] M. Bergevin,[9] A. Bernstein,[16] T.J.C. Bezerra,[28] L. Bezrukhov,[14] E. Blucher,[6] N.S. Bowden,[16] C. Buck,[21] J. Busenitz,[2] A. Cabrera,[4] E. Caden,[10] L. Camilleri,[8] R. Carr,[8] M. Cerrada,[7] P.-J. Chang,[17] P. Chimenti,[33] T. Classen,[9,16] A.P. Collin,[15] E. Conover,[6] J.M. Conrad,[20] J.I. Crespo-Anadón,[7] K. Crum,[6] A. Cucoanes,[25] E. Damon,[10] J.V. Dawson,[4,36] S. Dazeley,[16] D. Dietrich,[32] Z. Djurcic,[3] M. Dracos,[24] V. Durand,[15,4] J. Ebert,[11] Y. Efremenko,[27] M. Elnimr,[25] A. Erickson,[16] A. Etenko,[19] M. Fallot,[25] M. Fechner,[15] F. von Feilitzsch,[26] J. Felde,[9] V. Fischer,[15] D. Franco,[4] A.J. Franke,[8] M. Franke,[26] H. Furuta,[28] R. Gama,[5] I. Gil-Botella,[7] L. Giot,[25] M. Göger-Neff,[26] L.F.G. Gonzalez,[34] L. Goodenough,[3] M.C. Goodman,[3] J.TM. Goon,[2] D. Greiner,[32] N. Haag,[26] S. Habib,[2] C. Hagner,[11] T. Hara,[18] F.X. Hartmann,[21] J. Haser,[21] A. Hatzikoutelis,[27] T. Hayakawa,[22] M. Hofmann,[26] G.A. Horton-Smith,[17] A. Hourlier,[4] M. Ishitsuka,[30] J. Jochum,[32] C. Jollet,[24] C.L. Jones,[20] F. Kaether,[21] L.N. Kalousis,[35] Y. Kamyshkov,[27] D.M. Kaplan,[13] T. Kawasaki,[22] G. Keefer,[16] E. Kemp,[34] H. de Kerret,[4,36] T. Konno,[30] D. Kryn,[4] M. Kuze,[30] T. Lachenmaier,[32] C.E. Lane,[10] C. Langbrandtner,[21] T. Lasserre,[15,4] A. Letourneau,[15] D. Lhuillier,[15] H.P. Lima Jr,[5] M. Lindner,[21] J.M. López-Castaño,[7] J.M. LoSecco,[23] B.K. Lubsandorzhiev,[14] S. Lucht,[1] D. McKee,[17] J. Maeda,[31] C.N. Maesano,[9] C. Mariani,[35] J. Maricic,[10] J. Martino,[25] T. Matsubara,[31] G. Mention,[15] A. Meregaglia,[24] M. Meyer,[11] T. Miletic,[10] R. Milincic,[10] H. Miyata,[22] Th.A. Mueller,[28] Y. Nagasaka,[12] K. Nakajima,[22] P. Novella,[7] M. Obolensky,[4] L. Oberauer,[26] A. Onillon,[25] A. Osborn,[27] I. Ostrovskiy,[2] C. Palomares,[7] I.M. Pepe,[5] S. Perasso,[10] P. Perrin,[15] P. Pfahler,[26] A. Porta,[25] W. Potzel,[26] G. Pronost,[25] J. Reichenbacher,[2] B. Reinhold,[21] A. Remoto,[25,4] M. Röhling,[32] R. Roncin,[4] S. Roth,[1] B. Rybolt,[27] Y. Sakamoto,[29] R. Santorelli,[7] F. Sato,[31] S. Schönert,[26] S. Schoppmann,[1] T. Schwetz,[21] M.H. Shaevitz,[8] S. Shimojima,[31] D. Shrestha,[17] J-L. Sida,[15] V. Sinev,[14,15] M. Skorokhvatov,[19] E. Smith,[10] J. Spitz,[20] A. Stahl,[1] I. Stancu,[2] L.F.F. Stokes,[32] M. Strait,[6] A. Stüken,[1] F. Suekane,[28] S. Sukhotin,[19] T. Sumiyoshi,[31] Y. Sun,[2] R. Svoboda,[9] K. Terao,[20] A. Tonazzo,[4] M. Toups,[8] H.H. Trinh Thi,[26] G. Valdiviesso,[5] C. Veyssiere,[15] S. Wagner,[21] H. Watanabe,[21] B. White,[27] C. Wiebusch,[1] L. Winslow,[20] M. Worcester,[6] M. Wurm,[11] F. Yermia,[25] and V. Zimmer[26]

[1] *III. Physikalisches Institut, RWTH Aachen University, 52056 Aachen, Germany*
[2] *Department of Physics and Astronomy, University of Alabama, Tuscaloosa, Alabama 35487, USA*
[3] *Argonne National Laboratory, Argonne, Illinois 60439, USA*
[4] *APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/IRFU, Observatoire de Paris, Sorbonne Paris Cité, 75205 Paris Cedex 13, France*
[5] *Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, RJ, cep 22290-180, Brazil*
[6] *The Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637, USA*
[7] *Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, CIEMAT, E-28040, Madrid, Spain*
[8] *Columbia University; New York, NY 10027, USA*
[9] *University of California, Davis, CA-95616-8677, USA*
[10] *Physics Department, Drexel University, Philadelphia, Pennsylvania 19104, USA*
[11] *Institut für Experimentalphysik, Universität Hamburg, 22761 Hamburg, Germany*
[12] *Hiroshima Institute of Technology, Hiroshima, 731-5193, Japan*
[13] *Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616, USA*
[14] *Institute of Nuclear Research of the Russian Aacademy of Science, Russia*
[15] *Commissariat à l'Energie Atomique et aux Energies Alternatives, Centre de Saclay, IRFU, 91191 Gif-sur-Yvette, France*
[16] *Lawrence Livermore National Laboratory, Livermore, CA 94550, USA*
[17] *Department of Physics, Kansas State University, Manhattan, Kansas 66506, USA*
[18] *Department of Physics, Kobe University, Kobe, 657-8501, Japan*
[19] *NRC Kurchatov Institute, 123182 Moscow, Russia*
[20] *Massachusetts Institute of Technology; Cambridge, MA 02139, USA*
[21] *Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany*
[22] *Department of Physics, Niigata University, Niigata, 950-2181, Japan*
[23] *University of Notre Dame, Notre Dame, IN 46556- 5670, USA*
[24] *IPHC, Université de Strasbourg, CNRS/IN2P3, F- 67037 Strasbourg, France*
[25] *SUBATECH, CNRS/IN2P3, Université de Nantes, Ecole des Mines de Nantes, F-44307 Nantes, France*
[26] *Physik Department, Technische Universität München, 85747 Garching, Germany*
[27] *Department of Physics and Astronomy, University of Tennessee, Knoxville, Tennessee 37996, USA*
[28] *Research Center for Neutrino Science, Tohoku University, Sendai 980-8578, Japan*
[29] *Tohoku Gakuin University, Sendai, 981-3193, Japan*
[30] *Department of Physics, Tokyo Institute of Technology, Tokyo, 152-8551, Japan*

[31] *Department of Physics, Tokyo Metropolitan University, Tokyo, 192-0397, Japan*
[32] *Kepler Center for Astro and Particle Physics, Universität Tübingen, 72076, Tübingen, Germany*
[33] *Universidade Federal do ABC, UFABC, São Paulo, Santo André, SP, Brazil*
[34] *Universidade Estadual de Campinas-UNICAMP, Campinas, SP, Brazil*
[35] *Center for Neutrino Physics, Virginia Tech, Blacksburg, VA*
[36] *Laboratoire Neutrino de Champagne Ardenne, domaine d'Aviette, 08600 Rancennes, France*

The Double Chooz experiment has determined the value of the neutrino oscillation parameter θ_{13} from an analysis of inverse beta decay interactions with neutron capture on hydrogen. This analysis uses a three times larger fiducial volume than the standard Double Chooz assessment, which is restricted to a region doped with gadolinium (Gd), yielding an exposure of 113.1 GW-ton-years. The data sample used in this analysis is distinct from that of the Gd analysis, and the systematic uncertainties are also largely independent, with some exceptions, such as the reactor neutrino flux prediction. A combined rate- and energy-dependent fit finds $\sin^2 2\theta_{13} = 0.097 \pm 0.034\,(\text{stat.}) \pm 0.034\,(\text{syst.})$, excluding the no-oscillation hypothesis at $2.0\,\sigma$. This result is consistent with previous measurements of $\sin^2 2\theta_{13}$.

PACS numbers: 14.60Pq

Neutrino oscillations are well established in the three flavor paradigm and can be described by three mixing angles (θ_{12}, θ_{23}, θ_{13}), a CP-violating phase δ, and two mass-squared differences ($\Delta m^2_{21}, \Delta m^2_{32}$). Among the three mixing angles, θ_{13} is the smallest and has recently been revealed to be non-zero [1–7]. The value of θ_{13} is a critical input for plans to measure δ and the neutrino mass hierarchy. Furthermore, it may provide important clues for physics beyond the Standard Model. The current best measurements of θ_{13} come from the reactor $\bar{\nu}_e$-disappearance experiments Double Chooz, Daya Bay, and RENO [5–7]. All three experiments rely on the detection of the inverse beta decay (IBD) interaction, $\bar{\nu}_e + p \rightarrow e^+ + n$, in Gd-doped liquid scintillator (LS). Typically these experiments search for a prompt positron signal followed by an ~ 8 MeV gamma cascade from neutron capture on Gd. Background due to natural radioactivity, which is predominantly below 4 MeV, is largely suppressed. However, in Double Chooz it is also possible to search for a prompt positron followed by a 2.2 MeV gamma ray from neutron capture on hydrogen, thanks to the low background environment in the detector. Though the latter analysis presents several challenges, it provides important benefits: a cross-check on the standard Gd analysis and improved $\bar{\nu}_e$ energy spectrum shape information which is essential to our knowledge of θ_{13}.

In this letter we present an analysis of IBD interactions with neutron capture on hydrogen in the Double Chooz far detector. Following the same approach as in previous reports [3, 6], this analysis compares the candidate event rate and prompt energy spectrum shape to the Monte Carlo (MC) prediction. This analysis, however, differs from those reported [3, 6] in two major ways. First, the definition of the delayed signal is changed from the ~ 8 MeV gamma cascade characteristic of a neutron capture on Gd to the 2.2 MeV gamma ray characteristic of a neutron capture on hydrogen. This change allows us to select a data set that is statistically independent of the Gd-based data set and has different systematic uncertainties and background characteristics. Second, because hydrogen captures occur in the undoped LS in addition to the Gd-doped region, a three times larger fiducial volume is available for analysis.

The Double Chooz far detector is located at a distance of $\sim$ 1050 m from the two 4.25 GW_{th} reactor cores of the Chooz Nuclear Power Plant, with a rock overburden of 300 meters water equivalent. The central region of the detector consists of three concentric cylinders, collectively called the inner detector (ID). The innermost cylinder is the 10.3 m^3 target. This is surrounded by a γ-catcher (22.5 m^3). The target liquid is a PXE-based LS doped with Gd at a concentration of 1 g/l [8], while the γ-catcher liquid is an undoped LS. Outside the γ-catcher is the buffer, a 105 cm thick layer of non-scintillating mineral oil contained in a stainless steel tank. Light from the target and γ-catcher volumes is collected by 390 low-background 10-inch PMTs installed on the inner wall of the buffer tank [9–11]. Outside the buffer tank, and optically isolated from it, is the inner veto (IV), a 50 cm thick layer of liquid scintillator in a steel tank. The IV is equipped with 78 8-inch PMTs and serves as a veto for cosmic rays and fast neutrons entering the detector. The IV is surrounded by a 15 cm thick layer of demagnetized steel which suppresses γ-rays from radioactivity in the surrounding rock. Above the IV is the outer veto (OV) detector, a scintillator-strip-based muon tracking system. The OV system was installed during the data taking period, and about 2/3 of the data in this analysis benefit from OV use. A more detailed description of the entire detector can be found in Ref. [6].

The number of protons is estimated to be $(6.747 \pm 0.020) \times 10^{29}$ in the target [6] and $(1.582 \pm 0.016) \times 10^{30}$ in the γ-catcher volume, the latter being based on a geometrical survey and measurements of the scintillator hydrogen fraction.

The IBD signal is a twofold coincidence of a prompt positron energy deposition, E_{prompt}, and a delayed gamma energy deposition, E_{delay}, resulting from a neu-

tron capture on hydrogen or Gd. The separation in time and space, Δt and Δr, of the coincident events are determined by neutron capture physics. Neutron capture times are 200 μs in the γ-catcher and 30 μs in the target, where the presence of Gd greatly increases the neutron capture probability. In this analysis, where we search for $E_{delay} \approx 2.2$ MeV without any fiducial volume cuts, we expect to detect candidates in both the target and γ-catcher. Given that only 13 % of the IBD interactions in the target volume are followed by neutron capture on hydrogen [6], 95 % of the signal events used in this analysis are located in the γ-catcher.

Vertex reconstruction is based on a likelihood maximization of the charge and timing of the pulses detected at each PMT [6]. It allows the spatial correlation of prompt and delayed events, effectively removing accidental backgrounds.

We reconstruct the energy of all events via two steps: (1) a total charge (Q_{tot}) to photoelectron (PE_{tot}) conversion; and (2) a PE_{tot} to visible energy (E_{vis}) conversion as done in the Gd analysis [6]. The first step takes into account a channel-by-channel, non-linear gain calibration. The second step uses a light yield of ~ 230 PE/MeV, defined by the neutron capture peak on hydrogen in ^{252}Cf calibration source data. By applying correction factors derived from spallation neutron data, this step also corrects for the time variation and vertex dependence of the detector response. The same method is used to determine E_{vis} for the MC sample.

This analysis uses data collected by the Double Chooz far detector between April 13, 2011 and March 15, 2012, which is the same time-period used in the latest Double Chooz Gd analysis [6]. The total live time is 240.1 days, which is different from 227.9 days used in the Gd analysis [6] because of different analysis cuts.

The IBD candidate selection is performed via the following procedure. To reduce muon-induced backgrounds, we reject all events that occur less than 1 ms after a cosmic muon crosses the IV or the ID. We use PMT charge isotropy and PMT pulse simultaneity cuts to reduce backgrounds caused by light emitted from PMT bases ("light noise") [3]. We apply the following coincident selection cuts to the remaining events: 0.7 MeV $< E_{prompt} <$ 12.2 MeV; 1.5 MeV $< E_{delay} <$ 3.0 MeV; 10 μs $< \Delta t <$ 600 μs; $\Delta r <$ 90 cm. Furthermore, we reject prompt candidates that are coincident with a signal detected in the OV. This veto, along with the 10 μs lower bound of the Δt cut, renders backgrounds due to stopped muons negligible. Finally, we apply a multiplicity cut to reduce fast neutron backgrounds. This cut demands that no trigger occur in the 600 μs preceding the prompt candidate and that no trigger other than the delayed candidate occur in the 1000 μs following the prompt candidate.

The selection cuts yield a total of 36284 events. Among these IBD candidates are backgrounds due to uncorrelated accidental coincidences, fast neutrons produced by muons traversing the nearby rock, long-lived cosmogenic isotopes (mainly ^{9}Li), and a small contribution from light noise. Accidentals are the dominant background, comprising almost half the IBD candidate sample.

We measure the rate and energy spectrum of accidentals by analyzing a sample of off-time coincidences. We collect this sample by looking for a delayed trigger between 1 s + 10 μs and 1 s + 600 μs after a prompt candidate event and applying a multiplicity cut for a period of 1 s − 600 μs to 1 s + 1000 μs. To increase sample statistics, we open 124 consecutive windows after this first window, thus sampling accidentals between 1 s and 1.2 s after each prompt candidate. After correcting for inefficiencies associated with this selection method, we obtain an accidentals rate of 73.45 $\pm$ 0.16 events/day. The result is cross-checked among multiple independent methods, and the quoted value includes the largest systematic uncertainty among them.

The fast neutron background consists of a proton recoil, the prompt event, in coincidence with the capture of the neutron, the delayed event. A single muon passing close to the detector may generate one or more fast neutrons which traverse the IV and ID. We tag the number of IBD candidates in which fast neutrons are recorded simultaneously in the IV and ID by requiring $\geq$ 2 IV PMT hits and an ID-IV pulse-timing correlation. We estimate the tagging efficiency from an event sample with $E_{prompt} >$ 12 MeV, following the same method as used for the Gd analysis [6]. From this sample we obtain a spectrum shape and, using the tagging efficiency and sample purity, we calculate the fast neutron rate to be 2.50 $\pm$ 0.47 events/day.

Muon-induced radioactive isotopes which emit a neutron immediately following β-decay, such as ^{9}Li, can be a background to IBD reactions. As the lifetime of ^{9}Li is 257 ms, we use the correlation of the ^{9}Li decay events to previously detected muons to estimate the ^{9}Li background rate. To increase the purity of ^{9}Li in our sample, we consider only the subset of IBD candidates for which the spatial separation between the prompt event and the reconstructed muon track is within a defined distance. While ID PMTs are used to reconstruct the muon tracks in the Gd analysis [6], IV PMTs are used in this analysis to account for muons going through non-scintillating buffer liquid. To estimate the ^{9}Li rate in this subsample, we fit the time difference Δt_μ between the IBD candidate prompt events and preceding muons with an exponential function characterized by the ^{9}Li lifetime, plus a flat function to accommodate remaining accidentals and IBD candidates. The estimated rate is found to be consistent with that in the Gd analysis [6], accounting for the different fiducial volumes and selection efficiencies, and the difference is included in the systematic uncertainty. We find a ^{9}Li rate of 2.8 $\pm$ 1.2 events/day. Muon track reconstruction efficiency is evaluated by a MC study and

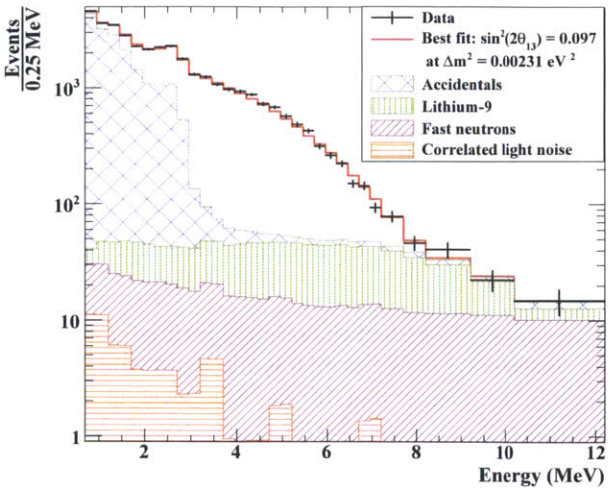

FIG. 1. (Color online) Stacked histogram showing the prompt energy spectrum of neutrino candidates without background subtraction (black data points with statistical error bars). The red (grey) line is the best fit oscillation hypothesis. Also shown are contributions from accidentals (blue cross-hatched), ^{9}Li (green vertical lines), fast neutrons (purple diagonal lines), and correlated light noise (orange horizontal lines).

TABLE I. Summary of the number of observed IBD candidates and the predictions for the signal and background contributions used as input for the oscillation fit analysis.

Source	Predicted/Observed events
$\bar{\nu}_e$ prediction (no osc.)	17690
Accidentals	17630
Cosmogenic isotopes	680
Fast neutrons	600
Light noise	80
Total prediction	36680
Observed IBD candidates	36284

added into the systematic uncertainty. We estimate the shape and associated systematic uncertainty from MC, as was done in the Gd analysis [6].

Finally, we found a small number of light noise events creating two consecutive triggers that are identified as IBD candidates. A volume cut on the reconstructed vertex is used to quantify the rate and E_{prompt} spectrum shape for this type of background. We estimate this background rate as 0.32 ± 0.07 events/day.

In this analysis, neutron detection efficiency ϵ_n includes both the efficiency of the IBD selection and the fraction of neutron captures which occur on hydrogen. We evaluate ϵ_n from ^{252}Cf neutron source calibration data and find it to be $\epsilon_n = 0.0846 \pm 0.0018$ in the target and $\epsilon_n = 0.7853 \pm 0.0036$ in the γ-catcher. Weighting by the fraction of predicted IBD candidates in each region, we estimate the uncertainty in the detection efficiency over the entire fiducial volume as 1.0 %. Finally, we find that an uncertainty of 1.2 % accounts for the MC modeling of neutron migration, called spill-in/out [6], between detector subvolumes. Adding these factors in quadrature, we obtain a total detection efficiency uncertainty of 1.6 %.

TABLE II. Summary of signal and background normalization uncertainties relative to the predicted signal.

Source	Uncertainty [%]
Reactor Flux	1.8
Statistics	1.1
Accidental background	0.2
Cosmogenic isotope background	1.6
Fast neutrons	0.6
Light noise	0.1
Energy Scale	0.3
Efficiency	1.6
Total	3.1

Energy scale uncertainty arises from three sources: time variation, non-linearity, and non-uniformity in the detector response. We treat the first two effects exactly as in Gd analysis [6]. The third effect has a larger impact on the hydrogen analysis because of its extended fiducial volume. We estimate it by comparing data and MC from calibration source deployments in the γ-catcher. In total, we find an energy scale uncertainty of 1.7 %, as compared to 1.1 % used in the Gd analysis [6].

The reference E_{prompt} spectrum is selected from the same reactor power-based $\bar{\nu}_e$ MC sample generated for the Gd analysis [6]. Systematic uncertainties on the reference spectrum are the same as for the Gd analysis. We use the Bugey4 measurement to minimize the systematic uncertainty on the reactor neutrino flux prediction [6, 12], which is the dominant uncertainty in this analysis. The no-oscillation expectation for the number of neutrino candidates is 36680 ± 520, including background. The predicted number of events for both signal and backgrounds are summarized in Tab. I, and uncertainties relative to the predicted signal statistics are shown in Tab. II.

To extract $\sin^2 2\theta_{13}$ we compare both the rate and shape of the data to the reference E_{prompt} spectrum in 31 variably sized energy bins from 0.7 to 12.2 MeV. The fit procedure is identical to that used in the Gd analysis [3, 6], except that we use a single integration period and include the Δr cut efficiency as an additional source of uncertainty. We use the MINOS value of $\Delta m^2 = (2.32 \pm 0.12) \times 10^{-3}\text{eV}^2$ as input for the fit [13]. We find a best fit of

$$\sin^2 2\theta_{13} = 0.097 \pm 0.034\,(\text{stat.}) \pm 0.034\,(\text{syst.})$$

with χ^2/DOF of 38.9/30. As in the Gd analysis [6], we define statistical error as the portion of the 1 σ error which can be improved by collecting more data. This includes uncertainty from our current statistics (see Tab. II) and uncertainty on background shapes. We de-

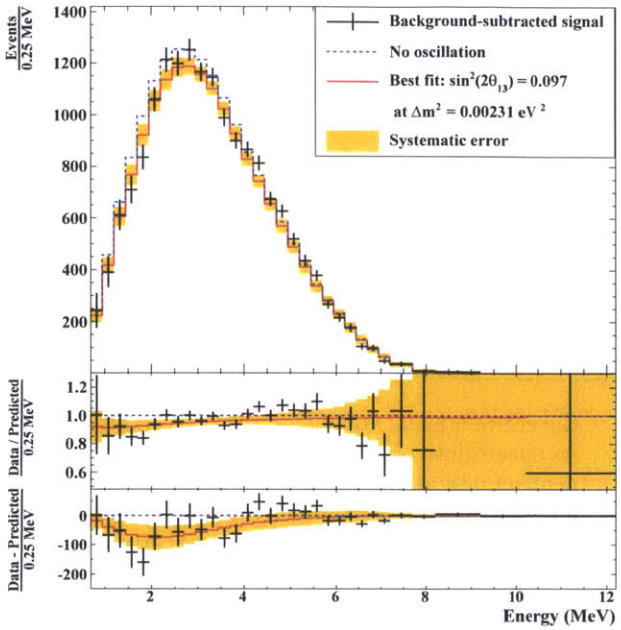

FIG. 2. (Color online) Top: Background-subtracted data (black points with statistical error bars) are superimposed on the prompt energy spectra expected in the case of no oscillations (dashed blue line) and for our best fit $\sin^2 2\theta_{13}$ (solid red line). The best fit has χ^2/DOF of 38.9/30. Solid gold bands indicate systematic errors in each bin. Middle: The ratio of data to the no-oscillation prediction (black points with statistical error bars) is superimposed on the expected ratio in the case of no oscillations (blue dashed line) and for our best fit $\sin^2 2\theta_{13}$ (solid red line). Gold bands indicate systematic errors in each bin. Bottom: The difference between data and the no-oscillation prediction is shown in the same style as the ratio (above).

fine systematic error as the uncertainty which cannot be reduced simply by collecting more data. Figure 1 shows the complete spectrum of IBD candidates with the fitted background contributions, while Fig. 2 shows the background-subtracted E_{prompt} spectrum along with the best fit. The pull parameters from the fit are summarized in Tab. III together with the input values. We have performed a frequentist study to determine the compatibility of the data and the no-oscillation hypothesis. Based on a $\Delta\chi^2$ statistic, defined as the difference between the χ^2 at the best fit and at $\sin^2 2\theta_{13} = 0$, the data exclude the no-oscillation hypothesis at 97.4 % ($2.0\,\sigma$). A fit incorporating only the rate information yields $\sin^2 2\theta_{13} = 0.044 \pm 0.022\,(\text{stat.}) \pm 0.056\,(\text{syst.})$. A simple ratio of observed to expected signal statistics yields $R = 0.978 \pm 0.011\,(\text{stat.}) \pm 0.029\,(\text{syst.})$ at the far site.

The smaller best-fit value of $\sin^2 2\theta_{13}$ by the rate-only analysis can be explained by the ^{9}Li background. The fit to the energy spectrum indicates a larger ^{9}Li background

TABLE III. Summary of pull parameters in the oscillation fit. The input values are determined by measurements, and the best-fit values are outcome of oscillation fit.

Pull parameter	Initial value	Best-fit value
Cosmogenic Isotope [day^{-1}]	2.8 ± 1.2	3.9 ± 0.6
Fast neutrons [day^{-1}]	2.5 ± 0.5	2.6 ± 0.4
Energy scale	1.00 ± 0.02	0.99 ± 0.01
$\Delta m^2(10^{-3}\text{eV}^2)$	2.32 ± 0.12	2.31 ± 0.12

contamination than the original estimate, although it is consistent within the systematic uncertainty.

In summary, due to the low level of backgrounds achieved in the Double Chooz detector, we have made the first measurement of $\sin^2 2\theta_{13}$ using the capture of IBD neutrons on hydrogen. This technique enabled us to use a different data set with partially different systematic uncertainties than that used in the standard Gd analysis [6]. An analysis based on rate and spectral shape information yields $\sin^2 2\theta_{13} = 0.097 \pm 0.034\,(\text{stat.}) \pm 0.034\,(\text{syst.})$, which is in good agreement with the result of the Gd analysis $\sin^2 2\theta_{13} = 0.109 \pm 0.030\,(\text{stat.}) \pm 0.025\,(\text{syst.})$ [6]. With increased statistics and a precise evaluation of the correlation of the systematic uncertainties, a combination of the two results is foreseen for the future.

We thank the French electricity company EDF; the European fund FEDER; the Région de Champagne Ardenne; the Département des Ardennes; and the Communauté des Communes Ardennes Rives de Meuse. We acknowledge the support of the CEA, CNRS/IN2P3, the computer center CCIN2P3, and LabEx UnivEarthS in France; the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) and the Japan Society for the Promotion of Science (JSPS); the Department of Energy and the National Science Foundation of the United States; the Ministerio de Ciencia e Innovación (MICINN) of Spain; the Max Planck Gesellschaft, and the Deutsche Forschungsgemeinschaft DFG (SBH WI 2152), the Transregional Collaborative Research Center TR27, the excellence cluster "Origin and Structure of the Universe", and the Maier-Leibnitz-Laboratorium Garching in Germany; the Russian Academy of Science, the Kurchatov Institute and RFBR (the Russian Foundation for Basic Research); the Brazilian Ministry of Science, Technology and Innovation (MCTI), the Financiadora de Estudos e Projetos (FINEP), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the São Paulo Research Foundation (FAPESP), and the Brazilian Network for High Energy Physics (RENAFAE) in Brazil.

[1] K. Abe *et al. Phys. Rev. Lett.*, 107 041801 (2011).
[2] P. Adamson *et al. Phys. Rev. Lett.*, 107, 181802 (2011).

[3] Y. Abe *et al. Phys. Rev. Lett.*, 108, 131801 (2012).
[4] F.P. An *et al. Phys. Rev. Lett.*, 108, 171803 (2012).
[5] J.K. Ahn *et al. Phys. Rev. Lett.*, 108, 191802 (2012).
[6] Y. Abe *et al. Phys. Rev. D*, 86, 052008 (2012).
[7] F.P. An *et al. accepted by Chinese Physics C*, arXiv:1210.6327 (2012).
[8] C. Aberle *et al. JINST*, 7, P06008 (2012).
[9] E. Calvo *et al. NIM A*, 621, 222 (2010).
[10] C. Bauer *et al. JINST*, 6, P06008 (2011).
[11] T. Matsubara *et al. NIM A*, 661, 16 (2012).
[12] Y. Declais *et al. Phys. Lett. B*, 338, 383 (1994).
[13] P. Adamson *et al. Phys. Rev. Lett.*, 106, 181801 (2011).

9.3 Summary and Future Prospect

The Double CHOOZ experiment has developed an alternative analysis method of using delayed neutron capture on hydrogen to measure the value of θ_{13} in PMNS matrix. The result is found to be $\sin^2(2\theta_{13}) = 0.097 \pm 0.034$ (stat) ± 0.034 (syst). This is the first result from the hydrogen analysis lead by the author of this manuscript and several colleagues in Double CHOOZ collaboration.

The outcome of the hydrogen analysis is not only the measurement of θ_{13} but also a better understanding of the detector response and backgrounds. The number of protons in the γ-catcher is now understood to the level of 1 % error, which will be reduced once the near detector is available for the near-far comparison. A beautiful agreement of the E_{delayed} distribution between the DATA and MC shows a success of the energy calibration in the γ-catcher region, and hence expand the high level physics analysis possibility using this volume. The study of accidental backgrounds revealed numbers of inefficiency corrections that were not accounted before in the Gd analysis. Good understanding of this backround was a key to perform the hydrogen analysis, and its rate error is now understood much better than other backgroudns. Although it is not presented as a part of this thesis, the high statistics ^{9}Li DATA sample from the hydrogen analysis greatly helps to improve the E_{prompt} spectrum shape uncertainty. This also benefits the Gd analysis in future.

Despite the achievements described above, the hydrogen analysis needs be improved to achieve a higher sensitivity to measure θ_{13}. This is an on-going effort in the collaboration. The next goal of the Double CHOOZ before starting data taking with the near detector is a combined θ_{13} measurement of the hydrogen and Gd analyses, which is considered to be the best θ_{13} measurement in the single detector phase of the experiment.

Appendices

Appendix A

Further Contributions

A.1 Generic Production Framework

The author of this thesis led the development, implementation, and management of `DCProd`, the Double CHOOZ data production software framework. The framework is written in `C++` and `Python` computer languages as a part of `DOGS`, the Double CHOOZ offline software framework. In this section a brief description of `DCProd` and one specific implementation of the `DCProd` are introduced as the author's main contribution.

A.1.1 `DCProd`

To process a large data sample, one can divide a data sample into many and process them in parallel to save time. This is typically what is done for analyzing physics data samples at a computing cluster with many servers for a batch computation, which is called a *batch server* below. At a large computing cluster, there exist many batch servers managed by a job scheduler software such as `GridEngine`, `Condor`, etc. [57, 58]. A job scheduler receives requests from users to send jobs to batch servers, and it distributes these jobs to batch servers that match the jobs' specification as they become available. A user needs to provide a script which describes the contents of a job and learn how to use a particular batch scheduler software.

The `DCProd` package provides basic utility needs for a large data production with the kind of job scheduler described above. It can be considered as a wrapper program for a job scheduler

to run analysis using DOGS. In particular the first version is equipped with GridEngine and Condor. In this section, a brief overview is given before going into specific example usages.

Readability of a User Script

A user needs to provide a so-called *job macro*, which is a formatted text file to define what process needs to be run by DCProd. A job macro follows the format very close to CINTand PyROOT script, which is what a user usually uses to run an analysis under the DOGS framework. Because it follows a format of analysis code itself, anyone who is familiar with an analysis script using DOGS can intuitively understand the contents of a job macro.

Job Submission

DCProd has an interface to GridEngine and Condor job schedulers. It is very simple to implement another job scheduler as long as it is a job scheduler of a similar type. DCProd takes care of generating an executable script for the chosen batch job scheduler and submission of jobs. It provides the flexibility to submit multiple jobs with a single job macro, which is an advantage since a user needs to maintain a single job macro for a record later.

Storage Interface

A computing cluster typically comes with a special storage system for large data files. The use of such a special storage system may require storage specific protocols for transferring files to the storage server. DCProd is equipped to interact with HPSS and iRODS data storage servers that are both used at CCIN2P3 computing cluster where the processing of data takes place for the experiment. One can add another easily if necessary. For DOGS output files from a batch server, DCProd follows the protocol of the chosen storage system to transfer files. This avoids having users to learn storage specific protocols.

Monitoring and Tracking

Upon a user's request, DCProd provides tracking and monitoring of submitted jobs and their output through the MySQL database server. In particular, two types of information are mon-

itored: the job status and an output file path in a storage server. The first is useful to keep track of a long-lasting jobs. It also records the length of operation, which may be useful to profile the performance from many jobs later. The second is extremely useful to keep output files organized and easily trackable. It allows users to share the output files of a job with three identity keys: a string and two numbers. A string is associated with a job macro, and the two identity keys are associated with each submitted job (recall one job macro is typically used to submit multiple jobs). There is a dedicated database interface which takes in these three identity keys and locates the output files. This allows to keep a simple account of hundreds of output files from many jobs, and to share them with others by passing three identity keys.

A.1.2 Implementation Case: Common Trunk

One specific example of an implementation of `DCProd` is another framework called the Common Trunk (CT). The CT framework is responsible for running an automated data reconstruction as it arrives from the experiment on-site. Since a reconstruction can be treated just like an analysis process in `DOGS`, the CT uses the `DCProd` framework for submission and monitoring of batch jobs as well as the tracking of output files that are used for physics analysis.

The author of this thesis led the development of the CT framework. Figure A-1 shows a data flow diagram in the CT. There are four daemon processes always running to monitor the status of different production stages, which is recorded to the MySQL database for both monitoring and decision-making purposes.

Storage Monitoring Daemon

The storage monitoring daemon keeps track of new files transferred from the on-site computers to the `iRODS` server at `CCIN2P3`. This triggers the whole sequence of operation on those new files. Once they arrive, files are automatically transferred to `HPSS` server, which is a permanent storage system using a tape. Once the daemon confirms a copy of files on `HPSS`, it triggers staging of those files on the `XROOTD` disk cache server, which is an interactive disk for data processing. The actual staging of files is done by the storage manager daemon described below. The storage monitoring daemon keeps track of the status of these copying processes

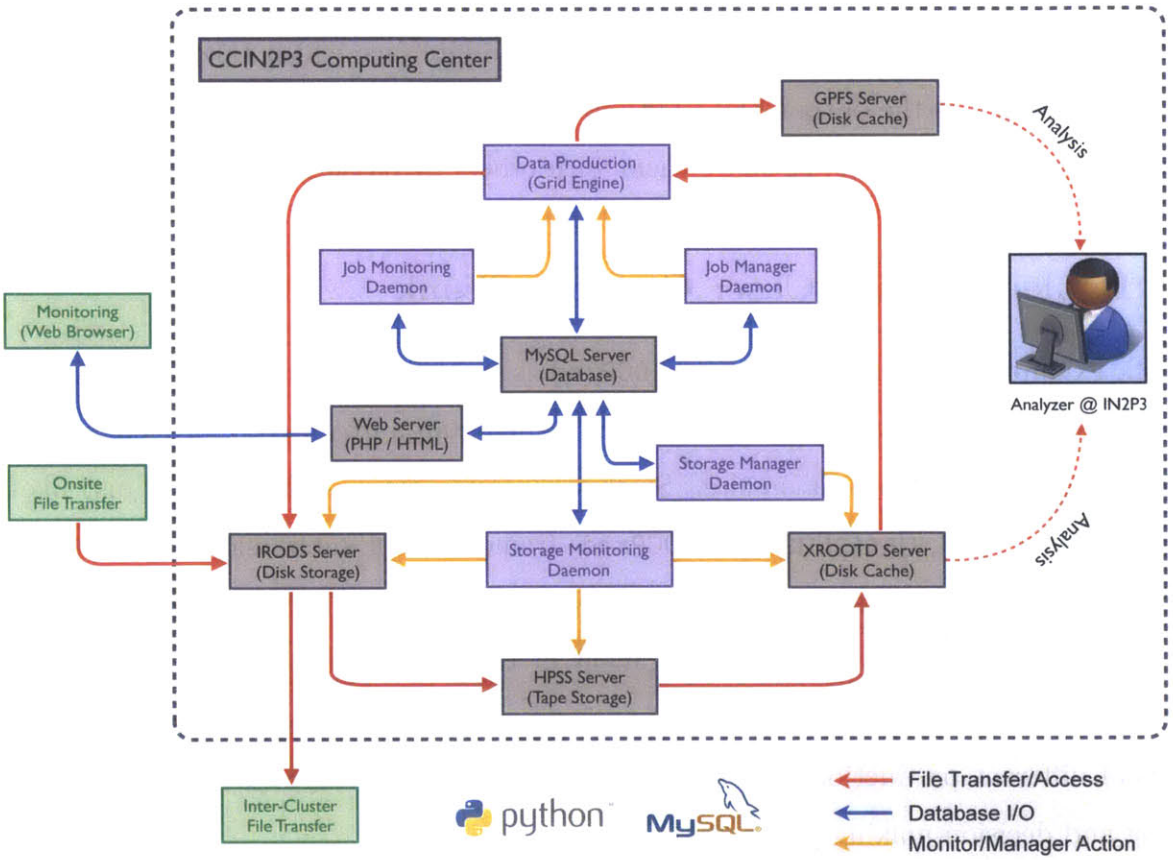

Figure A-1: CT data processing flow chart at `CCIN2P3`. File storage servers and batch servers are monitored and managed by multiple daemon processes constantly as described in the main text. Upon successful ending of the CT processing, output files become accessible to the analyzers and also are ready for inter-cluster data transfer.

in the database which can be tracked from a web server. When the staging of files on the XROOTD server is completed, the files are ready to be processed by the CT reconstruction process. The storage monitoring daemon signals this through the database.

Storage Management Daemon

The storage management daemon constantly checks the database for signals that are sent from the storage monitoring daemon. Given a signal from the storage monitoring daemon, the storage management daemon is responsible for two actions: deletion and creation of files on the iRODS and XROOTD servers. In principle, the deletion is never performed on DATA unless the signal is explicitly confirmed by an expert. This occurs when some files are confirmed to be corrupted, and are removed from the disk. The most important task is to create files on the XROOTD server which is needed for a further processing of the CT.

Job Manager Daemon

Provided with a signal in the database from the storage monitoring daemon, the job manager daemon performs the final check and submits a job to process the CT reconstruction on the data files. The final check involves configurations that are specified separately in the database by the experts. For instance, an expert can specify the total number of jobs to exist in the batch farm, avoid or prioritize a job submission for certain files such as calibration data files, and so on. Once a job is submitted, it logs the submitted status in the database. Another responsibility of the job manager daemon is to clean up failed jobs. For instance, some jobs may crash while writing a part of the output files on the storage server. Such files need to be removed before making another attempt of submitting the same job. A job manager daemon, upon a signal given from the job monitoring daemon described below, takes the necessary steps before attempting another job submission to process the same files.

Job Monitoring Daemon

Once a job is submitted by the job manager daemon, the job is constantly monitored by the job monitoring daemon. Either a success or a failure of a job is recorded in the database with a

corresponding status flag. A job failure may be classified into multiple cases with corresponding flags.

Summary

The CT framework is automated with the daemon processes described above. The output files are easily accessible by the collaborators using a standard procedure of using three identification keys to access `DCProd` output files. It has successfully processed data files from the Double CHOOZ far detector at `CCIN2P3` computing cluster, which corresponds to more than 40,000 DAQ runs with approximately 900,000 files of about 1 giga-byte size each.

A.2 Detector Commissioning

The author has been involved in many aspects of the detector integration and commissioning work. Among those, this section briefly describes two major contributions: monitoring of on-site data processing and quality check.

A.2.1 On-Site Data Quality Check

After the DAQ finishes running, binary data files from a run are immediately processed on the on-site computers into the `ROOT` data file format. In this process, a monitoring framework is run which monitors a very basic set of parameters and creates plots to be monitored by on-site shifters for identifying possible problems in the data. From his experience, the author is an expert in identifying many detector-related issues. His analysis code to identify and debug problems during the commissioning has become a large part of algorithms run in the monitoring framework.

In addition, the author has contributed to making a `PHP`-based framework for shifters to access monitoring plots through a web browser. Monitoring of the data quality with this framework has become a part of the shifters' task, and it served many times in the past to identify problems. A list of problems can be found in Ref. [59, 60]. Upon validation, the status flag and shifters' comments are saved in the database and accessible for later reference.

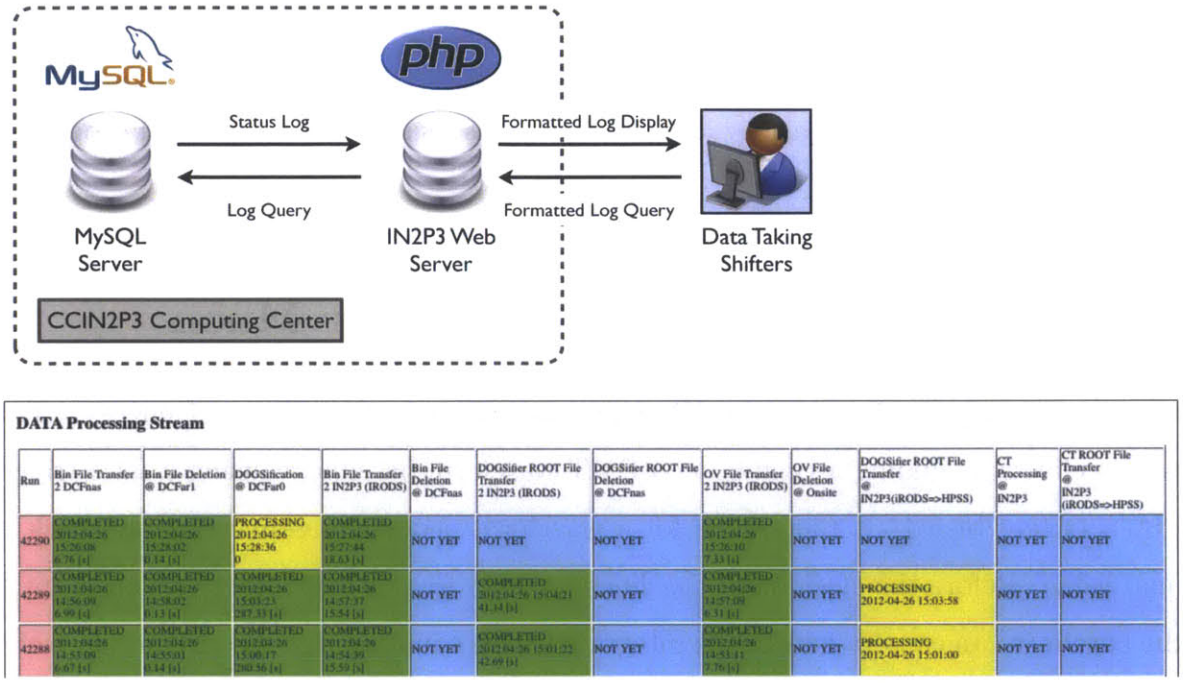

Figure A-2: Data stream monitoring scheme and example viewer.

A.2.2 Data Stream Monitoring

As described in the main body of this thesis, the DAQ runs for one hour for a typical physics data taking. Corresponding data files are roughly 15 giga-bytes of ROOT files and nearly 20 giga-bytes of binary files. Given the limited space on the on-site computers, a smooth processing of data files and transferring is extremely important. The framework called `DataMigro`is developed to do this operation, and the author has contributed to make a `PHP`-based framework to monitor steps of operation taken by the `DataMigro`.

A.3 Monitoring Data Processing and Quality Control

Figure A-2 shows the part of a web viewer that shows the status of each operations taken by the `DataMigro`. It has become the shifters' responsibility to make sure that the processing and transfer of data files are on-going smoothly through this web viewer. It has served an important task of alerting the experts in the past by identifying various issues such as failure in processing and network.

A.4 Analysis of Detector Response After Muon

High-energy muons passing through the ID region deposit a large amount of energy. This causes various effects, including a ringing on a PMT waveform baseline, causing a false trigger assigned by a trigger board. It takes time for a waveform pedestal baseline to come back to the nominal value. Before the electronics baselines recovers from a high energy muon event, a shift in the baselines can cause a bias in the integrated charge in the pulse reconstruction which turns into a bias in the estimated energy. The bias on the pedestal mean estimation depends on the deposited muon energy and time since the muon went through the detector. Figure A-3 shows the pedestal mean bias as a function of these parameters. The author contributed to set the minimum time separation from a previous muon of any energy above 30 MeV to be 150 μs during the commissioning period [61].

Spallation neutrons from muons going through the ID are the unique calibration source as they spread more or less uniformly in the detector. The uniformity and time stability correction in the energy reconstruction both rely on the spallation neutron data sample. It is very important, therefore, to quantify how this affects the energy reconstruction. In addition, the set boundary of 150 μs during the commissioning period was a conservative decision, and it greatly reduces the sample of neutron capture on Gd isotopes since a neutron capture time in the Target is about 30 μs. The author has led the investigation of the bias and a possibility of using a lower boundary value, and the result has contributed to the study of uniformity and time stability correction in the energy reconstruction. The details of the study can be found in the Ref. [62].

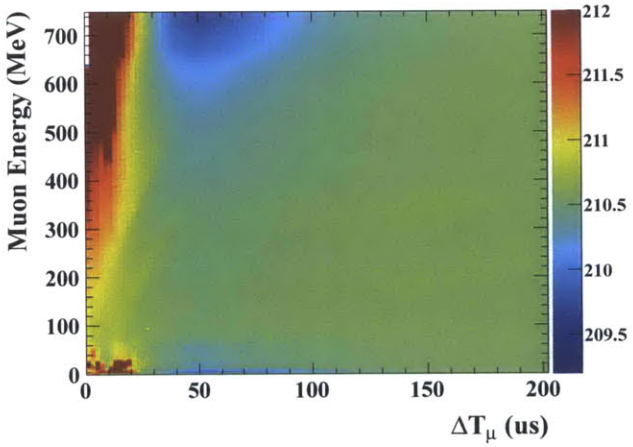

Figure A-3: The averaged pedestal mean value over all channels (Z-axis) for different muon energy (Y-axis) and time after muon (X-axis). One can see pedestal bias changes as a function of these two parameters.

Appendix B

Time Projection Chamber for Detecting Fast Neutrons

Nuclear Instruments and Methods in Physics Research A 673 (2012) 22–31

Contents lists available at SciVerse ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

A prototype detector for directional measurement of the cosmogenic neutron flux

J. Lopez, K. Terao, J.M. Conrad, D. Dujmic, L. Winslow *

Physics Department, Massachusetts Institute of Technology, Cambridge, MA 02139, United States

ARTICLE INFO

Article history:
Received 25 August 2011
Received in revised form
3 January 2012
Accepted 11 January 2012
Available online 21 January 2012

Keywords:
Neutron detector
Neutron production
Muon interactions
Neutrino

ABSTRACT

This paper describes a novel directional neutron detector. The low pressure time projection chamber uses a mix of helium and CF_4 gases. The detector reconstructs the energy and angular distribution of fast neutron recoils. This paper reports results of energy calibration using an α source and angular reconstruction studies using a collimated neutron source. The best performance is obtained with a 12.5% CF_4–87.5% He gas mixture. At low energies the target for fast neutrons transitions is primarily helium, while at higher energies, the fluorine contributes as a target. The reconstruction efficiency is both energy and target dependent. For neutrons with energies less than 20 MeV, the reconstruction efficiency is $\sim$40% for fluorine recoils and $\sim$60% for helium recoils.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Precise measurement of the energy and direction of cosmogenic-muon-induced fast neutrons, with various levels of underground shielding, will be valuable for future neutrino and dark matter experiments [1]. In this paper, we present a first-generation detector for this purpose. The detector design is based on that for the directional dark matter search DMTPC [2]. We introduce modest modifications for directional neutron detection [3]. Our first application will be in the Double Chooz (DC) neutrino experiment [4], and it is, therefore, called DCTPC.

DCTPC is a low pressure time projection chamber (TPC) filled with two gasses: (1) He, the fast neutron target, and (2) CF_4, a scintillant and a quencher for the electron avalanche. This detector is blind to x-rays and minimum ionizing particles (MIP), like muons, because the density of primary ionization is too small to be detected. The ionization electrons from a non-MIP particle track drift down to a stainless steel ground mesh. Between the ground mesh and the anode, an avalanche of electrons occurs. A charge-coupled device (CCD) camera installed in the detector images the visible scintillation light from the avalanche. The charge created in the avalanche is readout from the anode plate and ground mesh. The track direction and energy can be reconstructed using the CCD image and the charge information. The reconstructed track from the recoiling nucleus can then be correlated to the incoming neutron direction and energy. In this paper we present on-surface measurements from a small 2.8 l, first-generation detector for an α-source run and a ^{252}Cf neutron source run as well as an on-surface cosmogenic run.

2. Motivation for DCTPC

DCTPC provides a benchmark for the Double Chooz fast neutron simulation at the specific sites of the two Double Chooz neutrino detectors [4]. The detectors sites have 114 and 300 meter water equivalent (m.w.e) shielding, respectively. The final DCTPC design calls for a large detector located at each site. The small first-generation detector described here will be installed in the 300 m.w.e site for initial studies. The neutron flux energy and angle measurements from the rock can be used to tune the Monte Carlo and demonstrate a clear understanding of this background in the Double Chooz neutrino oscillation analysis. However, the results will provide wider benefits as backgrounds from fast neutrons produced in rock are an issue for all low background experiments.

Low energy neutrons arise from (α,n) interactions due to natural radioactivity in the shielding rock. At locations with low shielding, this neutron source is overwhelmed by neutron production from cosmic ray muons. However, muons are attenuated with depth and the (α,n) background becomes significant at 300 m.w.e. [5]. DCTPC, with its very low energy threshold, can sample these (α,n) events. The comparison of the 114 m.w.e. to 300 m.w.e. depths at Double Chooz will be useful in isolating the (α,n) from the muon-induced neutrons.

* Corresponding author.
E-mail address: lwinslow@MIT.EDU (L. Winslow).

0168-9002/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2012.01.021

The remaining two broad categories of neutron production are caused by cosmic ray muon interactions with rock: stopped muon capture and deep inelastic interactions. The rates depend upon the depth of the final experiment because low energy muons are attenuated with shielding. DCTPC sensitivity will reach to >20 MeV in fast neutron reconstruction, overlapping with the KARMEN2 (surface) and LVD (3200 m.w.e.) data sets. As we show below, the maximum reconstructed recoil energy depends upon the size of the detector.

The FLUKA [6,7] and GEANT4 [8,9] simulations are generally used to predict >10 MeV fast neutrons that are produced by muons. DCTPC adds data points at two depths which will complement the existing measurements. Information on the energy spectra of fast neutrons is scarce, coming only from KARMEN2 experiment on the surface [10], KamLAND at 2700 m.w.e. [11] and the LVD experiment at 3200 m.w.e. [12]. DCTPC data will fill the gap in depth between KARMEN2 and deeper experiments. DCTPC also provides an angular distribution of events. In this paper, we show the quality of this angular reconstruction. The unique directional information can help disentangle different neutron sources.

3. The first generation detector

We have developed a first generation of DCTPC. The TPC has a drift volume above an amplification plane that has two active regions. The amplification plane is imaged with one CCD, and the charge is read out for each active region. The redundancy in the readout of the amplification plane is powerful for event reconstruction and background rejection. The detector schematic is shown in Fig. 1 and a photograph of the detector vessel is shown in Fig. 2.

3.1. The time projection chamber

The detector follows the classic design of TPCs. We have chosen a mixture of CF_4 and He as our operating gas. We study several gas ratios below, but most data were taken with the gasses mixed at 75 Torr and 525 Torr respectively. Helium and fluorine both have large cross-sections for fast neutron scattering [13]. The ^{4}He nucleus is closer to the neutron mass, and therefore the recoils will produce longer tracks than the fluorine nucleus.

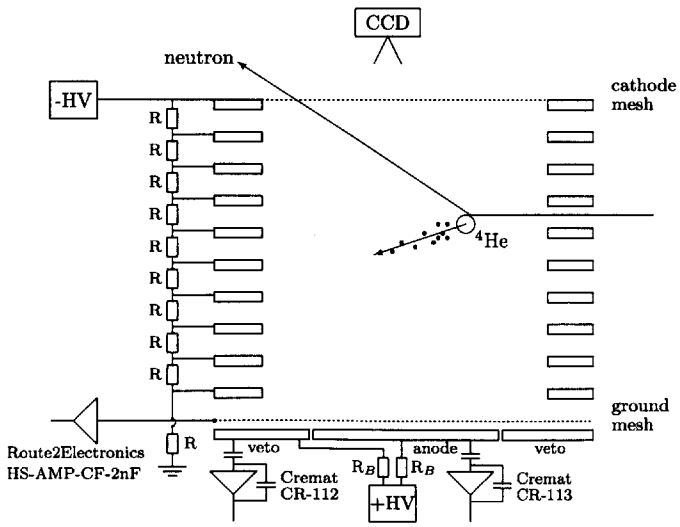

Fig. 1. A schematic of the detector: the drift field is created by a cathode mesh, field-shaping rings attached to a resistor chain, and a ground mesh. Primary ionization from a recoiling nucleus drifts down to the ground mesh. The high-field amplification region is formed by the ground mesh and the anode plane. The ground mesh is read out with a fast amplifier and the veto and anode are readout with charge-sensitive preamplifiers. Scintillation light from the amplification region is recorded with the CCD camera.

Within the vessel, a uniform drift field is molded by a series of copper rings with an inner diameter of 26.9 cm. The rings are connected with resistors, see Fig. 1. The top ring carries a high transparency stainless-steel mesh with wire pitch of 512 μm and wire diameter of 31 μm. This top mesh is maintained at a negative high voltage. The total height of the field cage is 10 cm. Primary ionization electrons are drifted down to a grounded stainless-steel mesh with wire pitch of 256 μm and wire diameter of 30 μm.

The amplification region is created by the ground mesh and the copper anode that are separated with insulating tubes, called spacers, 440 μm in diameter, placed 2.56 cm apart [14]. This high-field amplification gap multiplies the primary ionization charge by a factor of approximately 10^5, as we discuss below.

The anode bias scheme is shown in Fig. 1, where resistor R_B is 210 MΩ. The anode of the amplification plane is divided into two regions. The inner part of the anode, diameter 24.7 cm, is used for the detection of tracks. A veto ring, with inner diameter 24.8 cm and outer diameter 26.8 cm, surrounds the central plane and is used to veto charged particles originating from decays from within the detector components. A central square region, 16.7 cm × 16.7 cm, of the amplification plane is read out optically using the CCD camera. Using this geometry, we calculate that the active volume of the detector that has a charge readout is 5.6 L, while the volume with both CCD and charge readout is 2.8 L.

3.2. The gas mixture and associated operating voltages

In order to select the gas mixture, we studied three gases: 100%CF_4, 12.5% CF_4–87.5% He, and 6.25% CF_4–93.75% He. Operating voltages, listed in Table 1, were optimized for each gas. However, we had difficulty finding suitable operating conditions for the 6.25% mixture. Small variations in voltage caused very large gain differences. We also found that the absolute and relative energy calibrations were inconsistent over several gas fills with this mixture, likely due to the much steeper increase of gain with respect to anode voltage. Therefore, we rejected this mixture and do not report the results of the 6.25% CF_4 below.

3.3. CCD readout

The CCD camera records scintillation light created in the amplification region, see Fig. 1. The camera is the U6 model made by Apogee Instruments Inc., with the Kodak KAF-1001E CCD chip. The chip is kept at -20 °C to minimize the dark current. An image from the amplification plane is focused to the CCD chip using a Nikon lens, of focal length 55 mm and set to $f/1.2$. The CCD chip consists of 1024 by 1024 square pixels where the side length of a pixel is 24 μm. The active area on the anode plate is a square of 16.7 cm × 16.7 cm, see Fig. 2. In order to maximize the signal-to-noise ratio, we merge the CCD pixels into 4×4 bins during readout, so that the effective spatial resolution is 652 μm. The unit for one pixel count returned by the CCD readout is an ADU (Analog-to-Digital Unit).

The camera does not use a shutter and the CCD sensor is always live, even during the CCD readout. This leads to two types of exposures: a nominal exposure lasting 1 s, and a shorter (typically 250 ms) 'parasitic' exposure taken during the CCD readout and event processing. Track images recorded during the short parasitic exposure can be shifted in the direction of readout of CCD rows, and do not have corresponding signal in the charge readout channels. Therefore, they are easy to remove during analysis.

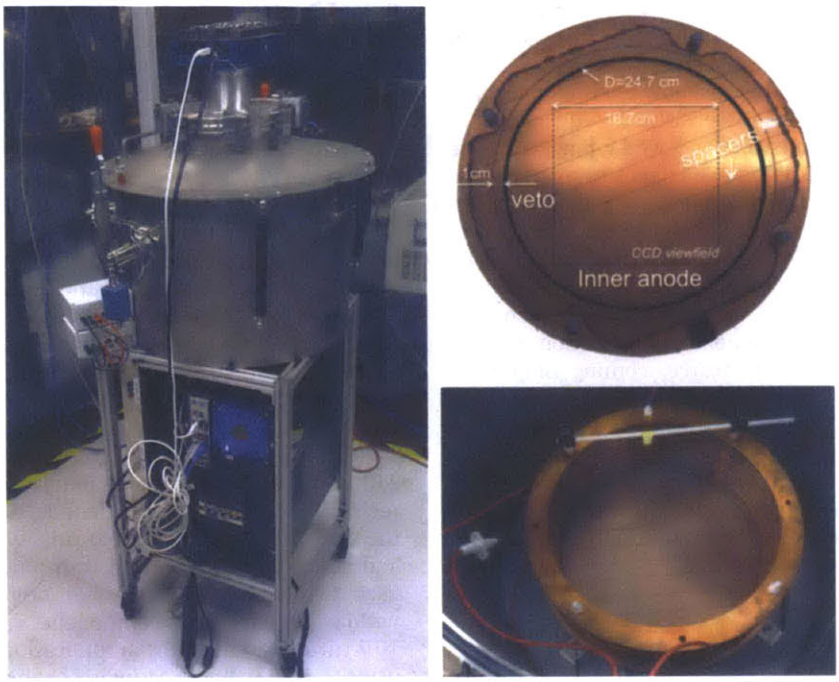

Fig. 2. Left: the first generation detector. The detector vessel is sitting on top of the cart which holds the readout electronics. The CCD camera is mounted on the top flange. Right-top: a diagram of the anode plane. Right-bottom: a photograph of the field cage and amplification plane of the TPC. The ground mesh is difficult to see, however, the spacers separating it from the anode plane are visible as horizontal lines.

Table 1
Summary of detector calibrations for running in the various gas mixtures. We report the run conditions for the 6.25% mixture here, however, due to instability, we do not report on this mixture further (see text). The trigger threshold is that used for neutron and background running. Higher thresholds are needed for the energy calibration runs.

CF_4 %	P (Torr)	(V)		G_{CCD} (ADU/keV$_\alpha$)	G_{anode} (mV/keV$_\alpha$)	Gas gain	Threshold (mV)
		V_{anode}	V_{drift}				
100	75	680	−1200	16.3	0.346	1.1×10^5	20.3
12.5	600	700	−1600	9.18	0.187	$> 6.0 \times 10^4$	18.8
6.25	600	600	−1600	15.3	0.518	$> 1.6 \times 10^5$	20.3

3.4. Charge readout

The charge readout is very powerful in suppressing CCD-specific backgrounds (*e.g.*, direct hits in CCD chip and residual bulk images). This is the major improvement over the CCD-only approach described in Ref. [3]. The anode plate contains three regions separated by electrically insulated gaps, as shown in Fig. 2. The outermost circular region is connected to a common ground, onto which the ground mesh is attached. The inner disk region is the actual anode, maintained at 680 V. The circular-ring region between is used as a veto and is kept at the same potential as the anode.

When a particle enters from outside, its path must cross the veto region. Therefore, we can detect this type of event by inspecting the charge signal from the veto region. The veto is not applied at the trigger level. The veto signal is stored for use in the analysis.

The anode signal is collected using a charge-sensitive preamplifier (CSP), the Cremat CR-113 with a gain of 1.5 mV/pC. The pulse height of the output of the CSP corresponds to the total charge collected on the anode. Therefore, the pulse height is used to reconstruct the total energy of the event. The veto electrode is read out through the more sensitive CSP, the Cremat CR-112 with a gain of 15 mV/pC. The ground mesh is read out through a fast current amplifier, the Route2Electronics HS-AMP-CF-2nF with 1 ns rise time and gain of 80. The fast amplifier has a built-in protection that guards against discharges from this high-capacitance detector. The CSPs are protected with 300 Ω resistors. All three signals are digitized using the ATS860, a 250 MS/s 8-bit flash ADC made by AlazarTech. The raw unit for charge readout in the following discussion is mV. The digitization of all three channels is triggered by a signal on the anode and stored with the current CCD image. The trigger threshold used is gas dependent.

3.5. Event readout, reconstruction and correlation

The combination of the four measurements, CCD image, charge deposited on the inner anode and veto anode and the current recorded at the ground mesh, is used for reconstruction. The CCD track reconstruction starts with subtracting the bias image that is obtained at the beginning of every run. At this time, any unusually high-valued pixels are removed. The image is then smoothed. With the smoothed image, a search is conducted for clusters of adjacent pixels with values 3.7σ larger than the mean value of the image [2]. If enough such pixels are found, the cluster is determined to be a possible track. An additional ring of pixels around the main cluster is added to account for diffusion, and the track is the corresponding pixels from the original un-smoothed

image. For each of these tracks, the energy, range, angle, maximum value, and several moments are calculated.

We require that the track be contained within the TPC. The veto removes signals from radiation emitted from the drift rings or the vessel wall. However, this also vetoes long tracks produced within the active volume that enter the veto. Additional inefficiency comes from the requirement that the track be fully contained in the CCD view-field. We show below that these are not a significant constraints, even in the first generation detector.

Tracks imaged by the CCD camera are matched to charge waveforms using the reconstructed energy. The rate of events in the detector, even with the neutron and α sources, is relatively low, and most exposures contain no tracks.

4. Calibration

Calibration of the detector was performed using a ^{241}Am α source. A window on the source degrades the energy of the α particles leading to a mean energy of 4.40 ± 0.04 MeV [15]. Example α events in two different gas mixtures are shown in Fig. 3.

4.1. Spatial calibration

The spatial calibration is obtained by measuring the positions of the wire spacers separating the anode plane from the ground mesh, shown in Fig. 2. The spacers create a region of low gain. A ^{137}Cs source is used to generate electronic recoils so that a small but diffuse amount of light is seen in each image. By integrating many images, the spacers become visible as regions with low light levels. The spacers are placed 2.5 cm apart, and the fit gives a calibration of 163 μm per pixel. The total area of the 1024 by 1024 pixel region imaged by the CCD is then 16.7 cm $\times$ 16.7 cm, as discussed in Section 3.3. The uncertainty on the spatial resolution is believed to be dominated by systematics from misplacement of spacers, optical distortions, and deviations in the spacer shape from a line.

4.2. Energy calibration

The total ionization left by the track is measured in three ways: (1) integrating the light in the CCD track, (2) measuring the pulse peak height of the central anode signal and (3) integrating the current signal from the mesh channel. At the drift voltage used in these data, the α tracks saturate the digitizer of the ground mesh and possibly the anode preamp. Using an 8-bit ADC, it is not possible to have both a fine energy granularity and a large range. The gain is optimized to have good energy resolution for low-momentum recoils created in the neutron scattering rather than the α studies. For this reason only methods (1) and (2) are used for determining track energy. Electron attachment was measured to be negligible at our drift fields [16].

The energy distributions from the anode and CCD are shown for two gas mixtures in Fig. 4. We have fit the peak position for the energy calibration. The results of the calibration are given in Table 1. The energy is quoted in units of alpha equivalent energy to denote that the detector response is calibrated with an alpha source. We do not apply corrections due to ionization quenching at lower recoil momenta due to lack of experimental data.

The gas gain shown in Table 1 is calculated from the charge energy calibration, the preamplifier gain of 1.5 mV/pC, and the gas work function. The work function of CF_4 gas was measured to be 33.8 eV [17]. This is used for calculating the gain of all mixtures.

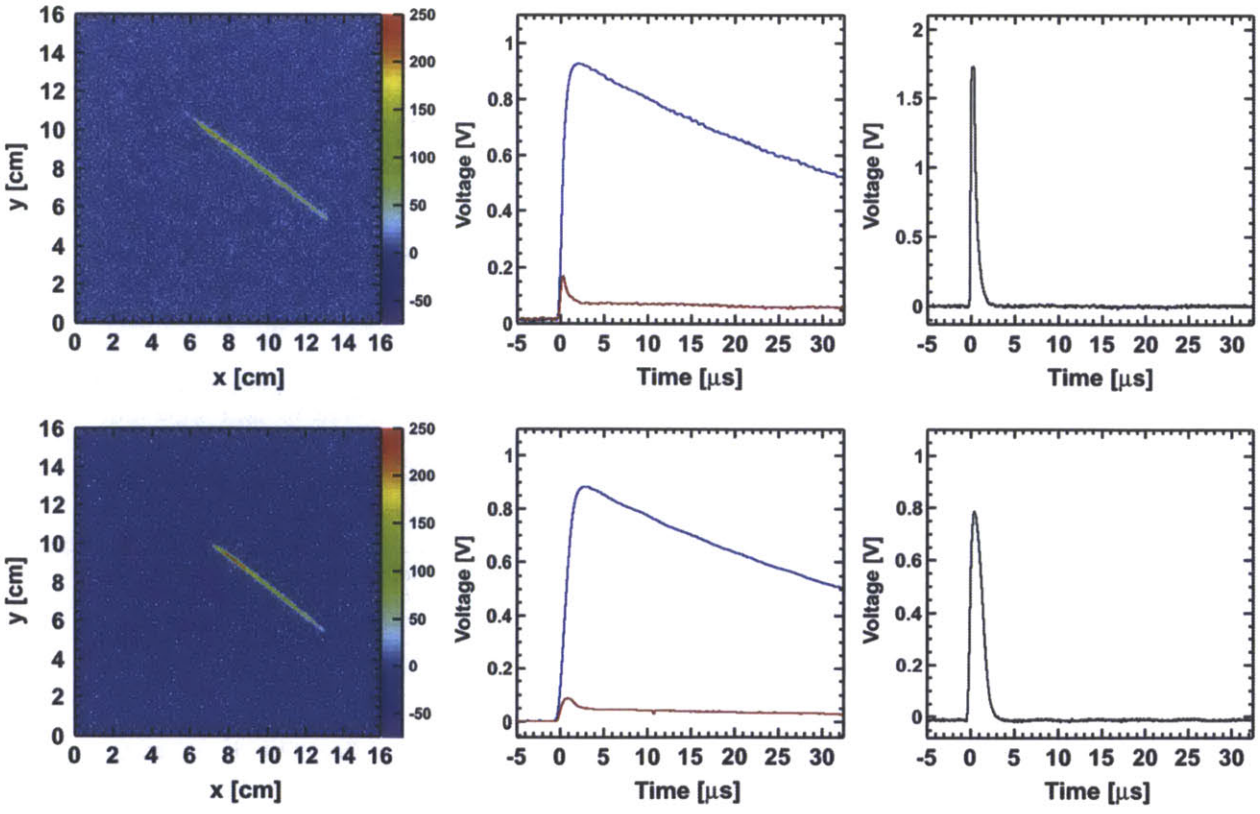

Fig. 3. For α particle tracks from the ^{241}Am source, images of the CCD and charge read-out system are shown. Left: CCD image of α particle tracks. Middle: waveforms from the Veto and Anode charge read-out. Right: waveforms from the Ground-Mesh charge read-out. Top: 100% CF_4; bottom: 12.5% CF_4 + Helium.

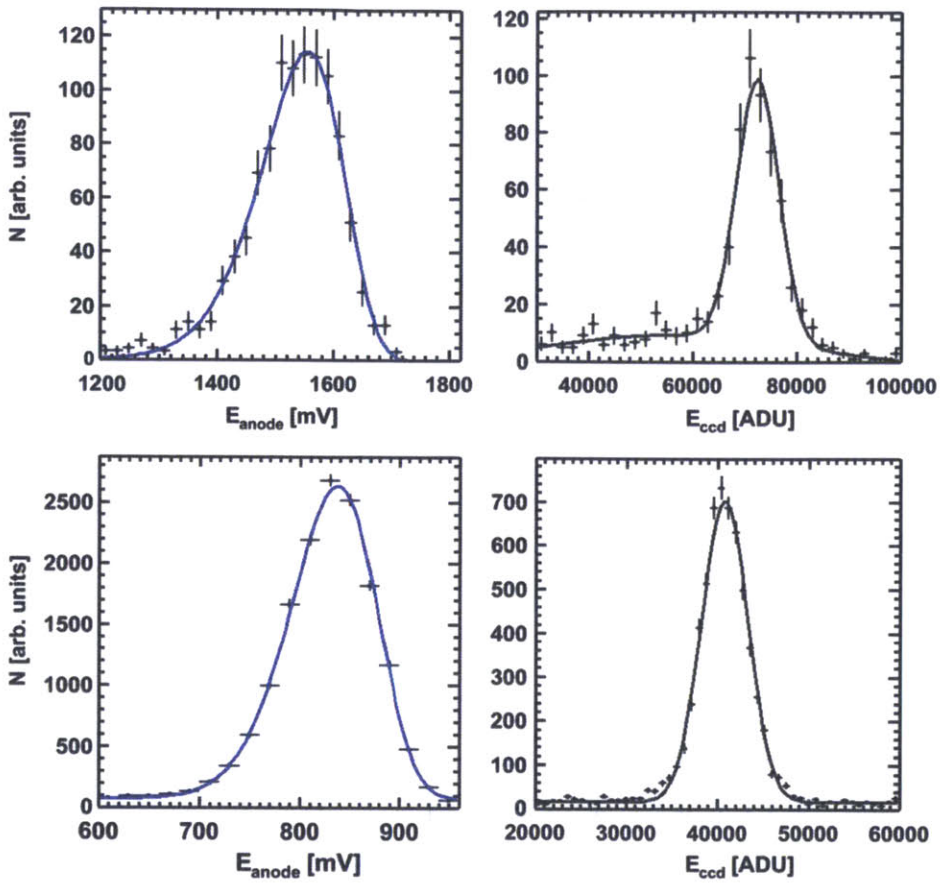

Fig. 4. Energy gain calibration using the charge read-out (left) and CCD (right). For 100% (top) and 12.5% (bottom) CF_4 mixtures.

5. The neutron source run

The response of the detector to neutrons was characterized using a 1.24 mCi ^{252}Cf source. ^{252}Cf decays via spontaneous fission 3% of the time emitting multiple neutrons with a mean energy of 2.35 MeV [18]. Given the activity of this source, we expect 5.7×10^6 neutrons per second in all directions. The source is placed in a collimator made of borated plastic, 2.1 m from the detector. The high flux of both neutrons and x-rays causes an increase in the rate of sparks in the TPC from ~ 0.01 Hz to 0.1 Hz. A spark is identified as an image with a mean number of ADU 1% greater than the previous image. Events within six exposures of a spark are excluded from analysis to allow sufficient time for the high voltage to recover.

5.1. CCD, charge and matching cuts

The cuts to isolate good events are applied to the CCD image, the charge-readout signal and the charge-light matching. For each run condition described in Table 1, we identified a set of optimized cuts. Table 2 describes the cuts placed on CCD images. Only cuts on track-length are gas-specific. The other cuts are designed to restrict tracks to the fiducial volume and remove backgrounds from ionization within the CCD chip, noise artifacts, and hot pixels.

The charge-readout cuts are more gas dependent than the CCD cuts. They can be divided into rise time cuts and pulse height cuts, summarized in Tables 3 and 4, respectively. For the veto we define two variables: the time is takes the pulse to rise from 10% to 90% of the peak pulse height, t_{veto}, and the ratio of the mesh peak pulse height to the veto peak pulse height, R_{veto}. For the mesh and anode, we define pulse rise times, t_{mesh} and t_{anode}. The pulse height range over which t_{mesh} and t_{anode} are defined is tuned for each gas mixture. For the mesh we also define two ratios for the pulse height: R_{elec} the ratio of the electron peak mesh pulse height to the height of the anode pulse and R_{ion} the ratio of the ion

Table 2
Cuts applied to CCD images for the study.

% of CF_4	Requirement
All cases	Two adjacent pixels with counts $>3.7\sigma$ above mean [2]
	Tracks are contained within $24 \le x,y \le 1000$ pixels
	Maximum pixel <250 ADU and $<25\%$ of total signal
	Only one track in the image
	No two tracks within 12 pixels in 1 run
100%	Tracks shorter than 80 pixels (1.3 cm)
12.5%	Tracks shorter than 160 pixels (2.6 cm)

Table 3
Rise time requirements. Variables are defined in the text. Events are accepted if they pass the requirements listed. The pulse height range over which t_{veto} is measured is always 10%–90%. The pulse height range for the definition of t_{mesh} and t_{anode} vary as listed in column 4.

% of CF_4	t_{veto} (ns)	t_{anode} and t_{mesh}	Pulse height range (%)
100%	<800	$t_{mesh} < 30$ ns for pulse height <100 mV	25–75
		$t_{mesh} < 25$ ns for pulse height >100 mV	25–75
12.5%	<800	$t_{anode} < 1250$ ns	10–90

peak mesh pulse height to the anode pulse height. R_{mesh} is a simple ratio of the anode pulse height to the mesh pulse height.

The CCD and charge readout run autonomously. Matching the signals proceeds by comparing the CCD track energy to the charge energy, where the best energy match is identified as the true signal from the track. The requirements for matching are shown in Table 5. Optical effects are estimated by measuring the mean ratio of the energies of the light and charge signals in neutron calibration running as a function of the distance from the image center. This is used to correct the energy of the light signal to obtain a linear relationship between the two energy measurements. The energy from the anode channel is more accurate than the CCD energy, so it is used as the final track energy. Neutron source data shown in Fig. 5 confirms a linear relationship between the summed anode energy and the CCD energy, demonstrating a good match between a CCD track and an anode waveform can be found for most of the events.

Table 4
Pulse height ratio requirements. The ratios are defined in the text. Events are accepted if they pass the requirements listed in this table. N/A indicates that this variable is not used.

% of CF_4	R_{veto}	R_{elec}	R_{ion}	R_{mesh}
100%	>4	$1.7<R_{elec}<2.7$	$1.9<R_{ion}<3.5$	N/A
12.5%	>3	N/A	N/A	$1.3<R_{mesh}<3$

Table 5
Requirements for CCD and charge event matching in this study, where $|\Delta E| = |E_{CCD} - E_{anode}|$.

% of CF_4	Requirement
100%	$\lvert\Delta E\rvert < 40$ keV$_\alpha$ if $E_{anode} < 150$ keV$_\alpha$; $\lvert\Delta E\rvert < 75$ keV$_\alpha$ if $E_{anode} > 150$ keV$_\alpha$;
12.5%	$\lvert\Delta E\rvert < 75$ keV$_\alpha$ for $E_{anode} < 250$ keV$_\alpha$; $\lvert\Delta E\rvert < 125$ keV$_\alpha$ for $250 < E_{anode} < 500$ keV$_\alpha$; $\lvert\Delta E\rvert < 150$ keV$_\alpha$ for $E_{anode} > 500$ keV$_\alpha$

The efficiency of the cuts, including charge-light matching for neutron events is given in Fig. 6. These data are dominated by nuclear recoils rather than CCD artifacts, so is ideal for the efficiency determination. The efficiency is defined as the ratio of the number of tracks passing all CCD cuts, charge cuts and charge-light matching compared to the number of CCD tracks found without using the charge signals.

5.2. Discussion of neutron source run

The neutron source generates a large number of nuclear recoils, with $\mathcal{O}(1)$ neutron-induced nuclear recoils detected per 1 s exposure. CCD images and charge waveforms of neutron-induced recoils in two events with different gas mixtures are shown in Fig. 7. A large number of electronic recoils from the high

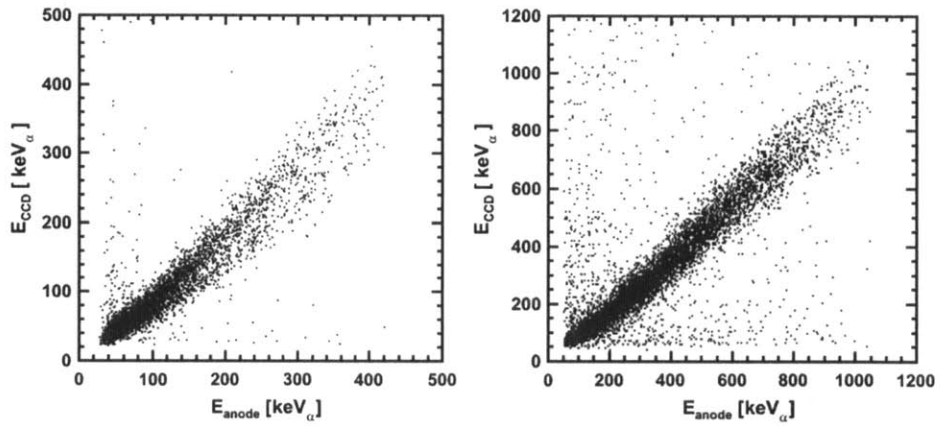

Fig. 5. To validate the α source calibration, we use CCD light versus charge read-out signal for neutron scatters from a ^{252}Cf source. Left: 100% CF_4. Right: 12.5% CF_4 + He.

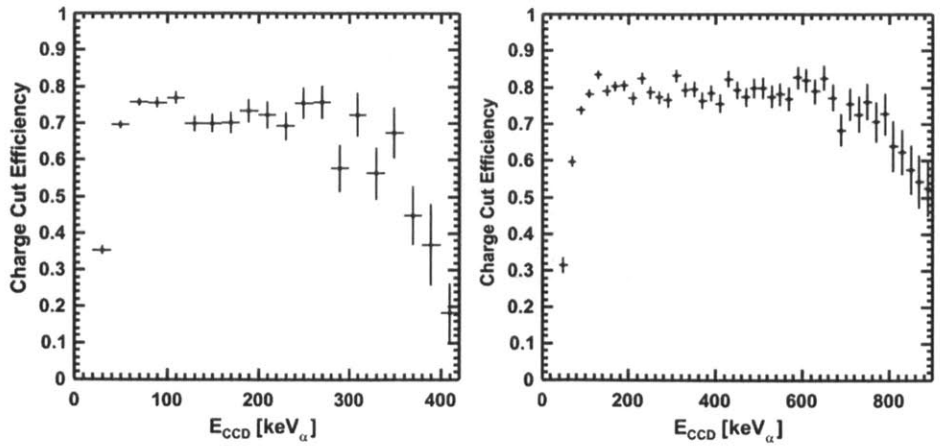

Fig. 6. Efficiency of charge analysis in accepting tracks identified by the CCD analysis. The efficiency plateaus at 70–80% are due to exposure difference between the charge and light readout and depend on the event readout time. The decrease in efficiency at higher energies is from the mesh pulse reaching the maximum value read out by the digitizer. Left: 100% CF_4 gas. Right: 12.5% CF_4 + He mixture. The errors on the data points reflect the statistics of the calibration run.

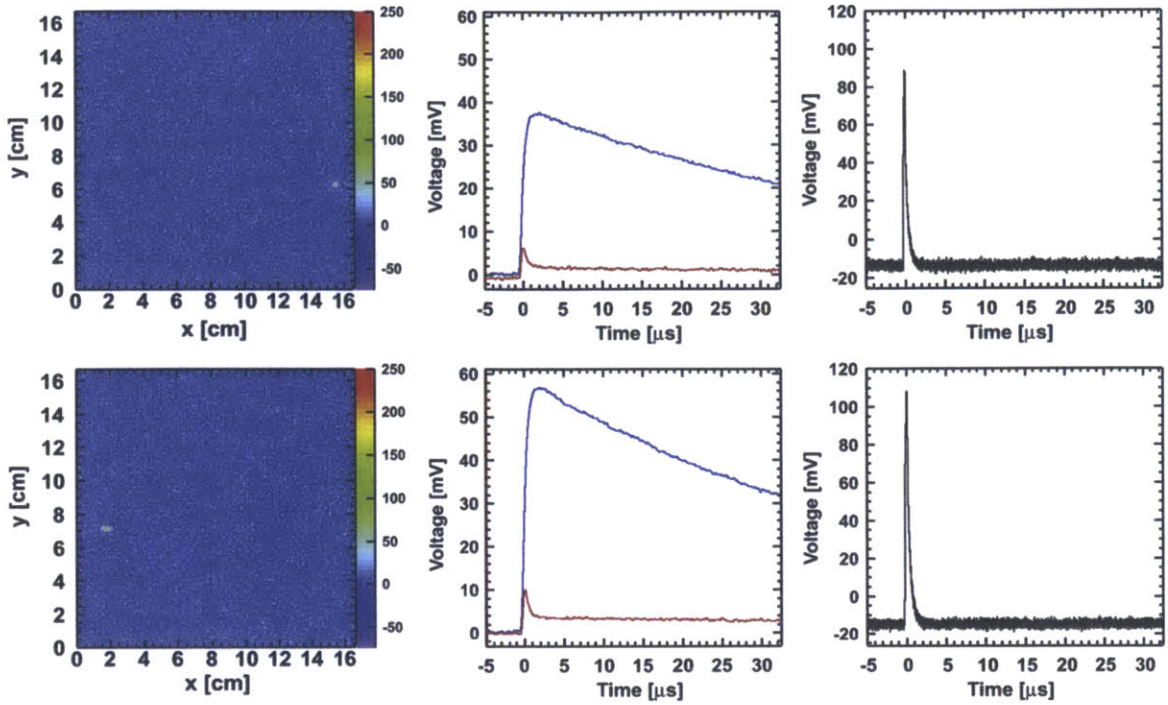

Fig. 7. Example figures of neutron recoils from ^{252}Cf source showing graphical images of CCD and charge read-out system. Left: CCD image of ^{4}He recoil. Middle: waveform from the Veto and Anode charge read-out. Right: waveform from the Ground-Mesh charge read-out. Top row is 100% CF_4 and bottom is 12.5% CF_4.

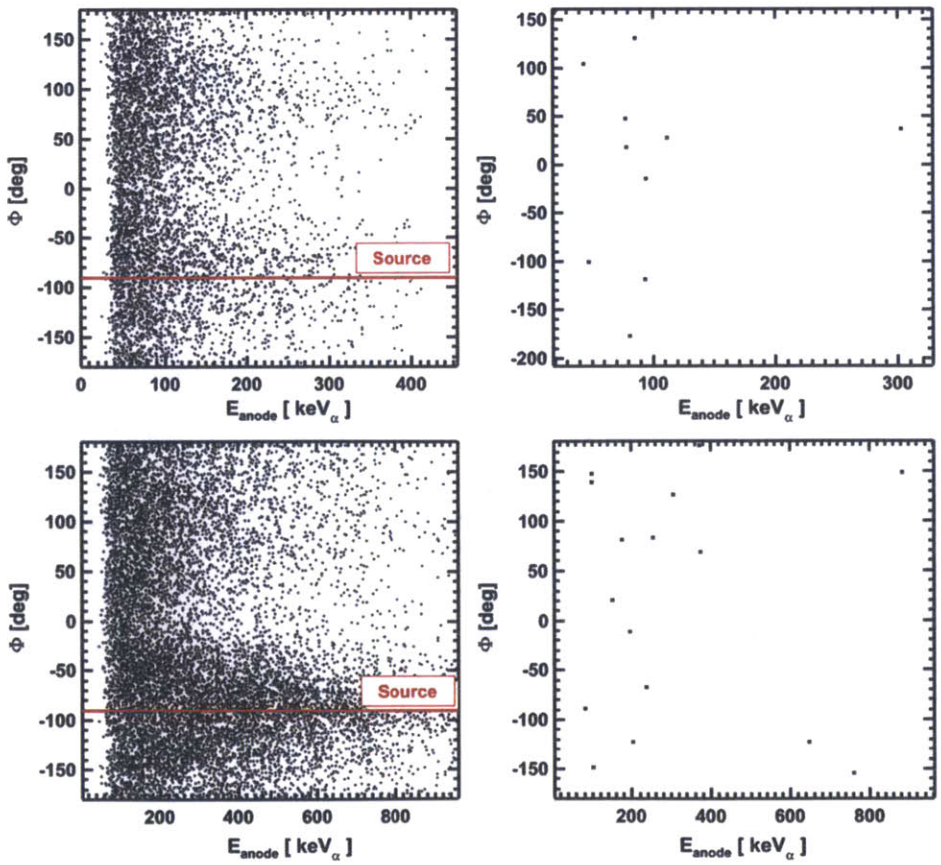

Fig. 8. Energy-angular distribution for calibration and background runs for two gas mixtures. Left: calibration run with ^{252}Cf source. Right: background run. Top: 100% CF_4 gas. Bottom: 12.5% CF_4. The direction of the neutrons from the source is indicated by a red solid line at -90°. The population at $+90^{\circ}$ is due to incorrect reconstruction of the track direction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

γ flux are observed in the charge signals. However, the CCD remains blind to these events due to their low primary ionization density. The source creates a much higher number of CCD-induced artifacts, likely from electronic recoils inside the active volume of the CCD chip. Even with these large source activity issues, the cuts are sufficient for this analysis.

After applying all cuts, we use the measured track angles and ranges, Figs. 8 and 9, to evaluate the performance of the detector and the analysis. The measured two-dimensional track range generally falls just below the expected mean three-dimensional range calculated with the publicly available program SRIM [19]. For the 12.5% CF_4 mixture, the measured two-dimensional track range falls close to the expected range for helium tracks. This follows from the fact that helium recoils are kinematically favored across most of the energy range measured here. At lower energies, where both helium and fluorine recoils are expected, the track ranges appear to be consistent with either nucleus.

We use the track angles to reconstruct the location of the neutron source. Tracks along the direction of the incoming neutrons should have a decreasing light profile, and this "head-tail effect" can be used to reconstruct the sense of track direction [20]. Two peaks are seen in the angular distribution of Fig. 8. The larger peak at -90° is along the true mean direction of the nuclear recoils. The second peak at $+90^\circ$ is primarily from events where the track direction was mis-reconstructed. The large width of the distribution is expected from elastic scattering kinematics. Additionally, some nuclear recoils are from neutrons that have scattered several times within the laboratory, and no longer appear to originate at the source position.

A model of helium and fluorine recoils is fitted to the measured recoil spectrum in Fig. 10. The fit uses the likelihood approach with the probability density functions for helium and fluorine based on simulation, and a ^{252}Cf spectrum based on ENDF tables [13]. We find the fraction of helium recoil events to be $61 \pm 1\%$. Another way to present this result is in terms of the unfolded neutron spectrum. We use the recoil data and the unfolding matrix, based on the simulation, to obtain the neutron spectrum as shown in Fig. 10. There is a reasonable agreement with the neutron spectrum from a ^{252}Cf source that is based on ENDF tables [13]. In order to improve the agreement between the data and the simulation of the recoil and neutron spectra, we need to improve the simulation to include the efficiency loss due to sparks. This requires additional measurements, and a full simulation of the detector material which modifies the energy spectrum of neutrons entering the sensitive region.

6. Surface run

With the DCTPC prototype, we have also taken data with no sources to measure cosmogenic, environmental and detector backgrounds in a surface lab on the MIT campus. The lab is on the ground floor of a two-story building. The lab has approximately 1 m concrete walls on all sides. After removing spark events, we obtain 112 713 exposures in pure CF_4 and 103 822 exposures in the 12.5% CF_4 mixture. As in the other runs, each exposure is 1 s, so this corresponds to 1.3 and 1.2 days of exposure, respectively.

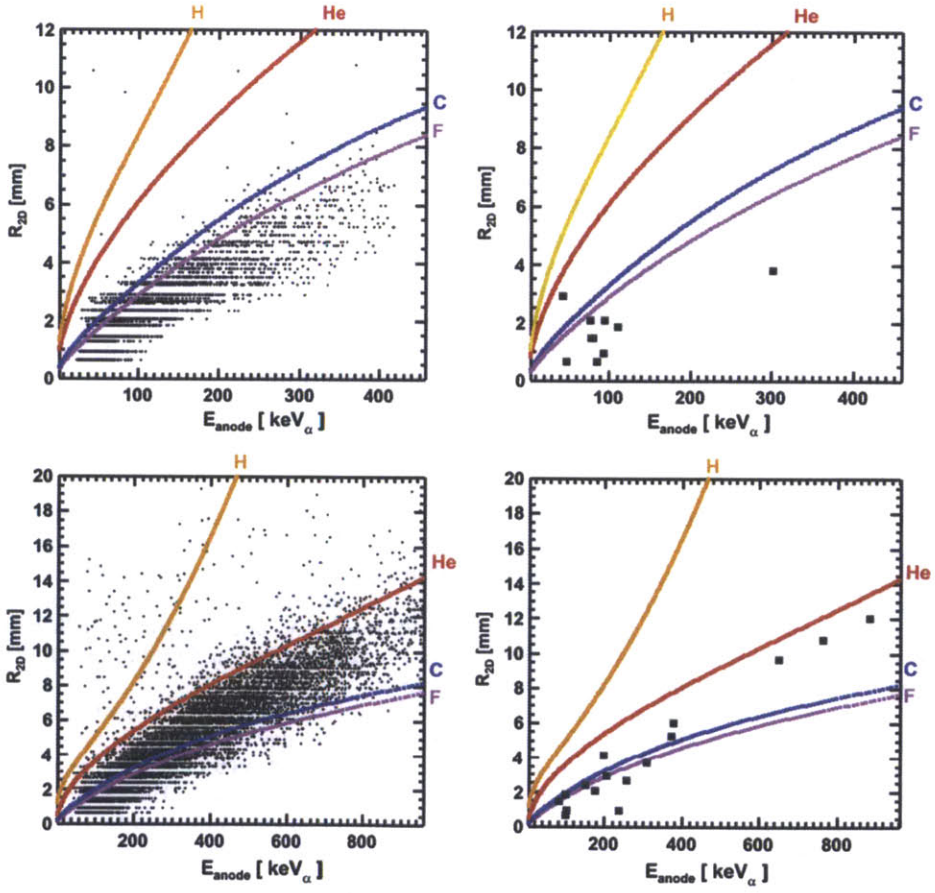

Fig. 9. Energy-range distribution, calculated in two dimensions, for calibration and background runs with two different gas mixtures. Left: calibration run with ^{252}Cf source. Right: background run. Top: 100% CF_4 gas. Bottom: 12.5% CF_4 mixture. Three-dimensional analytical prediction of the dE/dx curve for different nuclear elements of the gas mixture is shown in colored lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

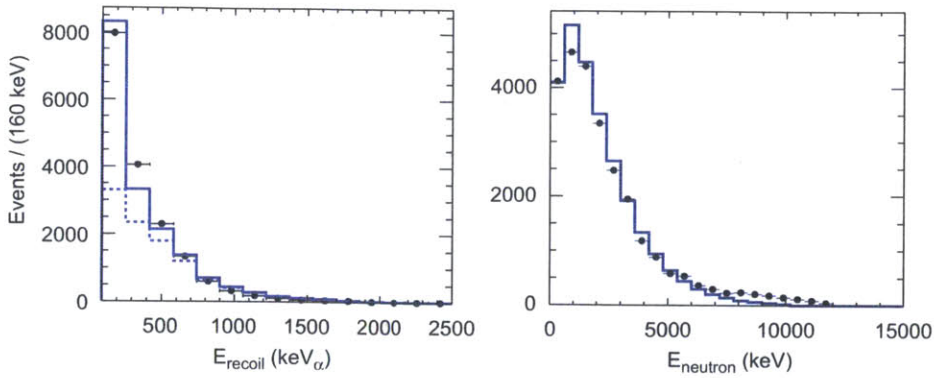

Fig. 10. Left: a fit to a recoil spectrum (markers) using a likelihood model based on helium and fluorine recoils. The helium contribution (dashed) and total recoil distribution (full line) are shown as histograms. Right: unfolded neutron spectrum from a calibration run with ^{252}Cf source (markers), and a spectrum based on the ^{252}Cf table [13] (histogram).

In the pure CF_4 run, we identified 128 possible tracks in the CCD analysis. Most of these tracks appear to be from CCD backgrounds such as noise artifacts or ionization in the CCD. Of these 128 tracks, only 10 were matched to a charge trigger. Several additional tracks had too much energy to be accurately measured by the charge channels, although the pulses were collected. These tracks are mostly low energy ($E < 100\,keV_\alpha$) and the ranges generally appear to be consistent with the expected values from SRIM for fluorine recoils. A single event has a range more consistent with helium or hydrogen tracks.

In the 12.5% mixture run, we identified 71 potential tracks in the CCD analysis, with 16 having a matching charge signal. As with the pure CF_4 run, most of the potential tracks seen in the CCD are likely noise artifacts or tracks left directly in the silicon of the CCD chip. As expected, we see more tracks in this run due to the presence of helium. The tracks seen here are also generally at higher energies than were seen in pure CF_4. This is to be expected if most of the tracks come from elastic scattering from neutrons due to the favorable kinematics of neutron–helium scattering compared to neutron–fluorine scattering. Fewer events at lower energies are seen due to the differing energy thresholds of the two gas mixtures. There does not appear to be a favored direction of these tracks, although much higher statistics would be necessary to confirm this. The events with $E > 300\,keV_\alpha$ here have ranges consistent with helium recoils but not carbon or fluorine, while most of the other tracks have ranges consistent with either carbon or fluorine nuclei or near vertical helium nuclei.

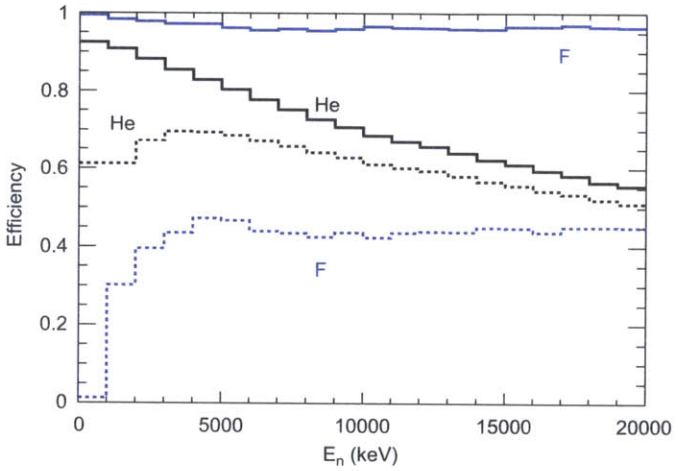

Fig. 11. The efficiency for reconstructing helium and fluorine recoils as a function of neutron energy. The geometric efficiency (full line) requires full track containment in the CCD view-field. The full efficiency (dashed line) assumes a step-function for the recoil reconstruction efficiency with thresholds given in the text.

7. Future underground measurement

The first generation detector is being installed in the far-hall of Double Chooz at 300 m.w.e. We can use the 13 neutrons per day measured in the surface run to make an estimate for our event rate in the far-hall. For the surface run, we measure a muon rate of 100 muons per $m^2\,s^1$. We assume the mean muon energy is ~ 3 GeV. The muon rate in the far-hall is 0.4 muons per $m^2\,s^1$ with a mean muon energy of 60 GeV [4]. Using the power law scaling from Ref. [11], we estimate ~ 0.5 events per day.

A study of the detector efficiency, shown in Fig. 11, indicates that events collected in this run is expected to sensitive to neutrons from (α,n) and muon interactions. Fluorine and helium recoils are efficiently reconstructed above 150 and 100 keV of recoil energy, respectively. Since fluorine recoils are shorter, they have a high probability to fit within CCD view-field. Geometric acceptance efficiencies are shown as full lines in Fig. 11. Helium recoils receive more energy from a collision with neutrons and leave longer tracks—a combination that is favored by reconstruction algorithms. Helium recoils are more efficient at energies $\lesssim 20$ MeV. Fluorine appears more efficient if the detector is being used for neutrons with the energy of tens of MeV's due to more compact tracks.

7.1. Conclusion

We have modified the DMTPC dark matter detector design to produce a 2.8 l, first-generation directional neutron detector. The primary change to the DMTPC design was to add helium as a target for the neutrons. We have demonstrated that a gas mixture of 75 Torr of CF_4 and 525 Torr of helium is capable of extracting the energy and angular distributions of the neutrons from a ^{252}Cf source.

The sensitivity to higher energy neutrons and an estimate of 0.5 events per day at the Double Chooz far site indicates that the first generation detector could make an interesting measurement at the far site on its own. The success of this detector motivates the construction of a full-sized DCTPC system. This will consist of two 60 l detectors, located at the near and far halls of Double Chooz, which can be interchanged for comparison of systematics. The goal is to measure the neutron flux, with information on both energy and direction, as a function of time, in order to tune the fast neutron simulation.

Acknowledgments

The authors thank Prof. Peter Fisher, of MIT, and the Double Chooz collaboration for valuable input. JMC, KT and LW are supported by the National Science Foundation. DD is supported by the Department of Energy. JL is supported by a Massachusetts Institute of Technology Lyons Fellowship and the Institutè for Soldier Nanotechnology.

References

[1] J.A. Formaggio, C.J. Martoff, Annual Review of Nuclear and Particle Science 54 (2004) 361.
[2] S. Ahlen, et al., Physics Letters B 695 (2011) 124.
[3] A. Roccaro, et al., Nuclear Instruments and Methods in Physics Research Section A 608 (2009) 305.
[4] F. Ardellier, et al., Double Chooz: A Search for the Neutrino Mixing Angle θ_{13}, 2006.
[5] D. Mei, A. Hime, Physics Review D 73 (2006) 053004.
[6] A. Ferrari, P.R. Sala, A. Fasso, J. Ranft, FLUKA: A Multi-Particle Transport Code (Program version 2005), CERN-2005-010.
[7] A. Fasso, et al., The Physics Models of FLUKA: Status and Recent Development, 2003.
[8] J. Allison, IEEE Transactions on Nuclear Science NS-53 (2006) 270.
[9] S. Agostinelli, et al., Nuclear Instruments and Methods in Physics Research Section A 506 (2003) 250.
[10] B. Armbruster, et al., Physics Review D 65 (2002) 112001.
[11] S. Abe, et al., Physics Review C 81 (2010) 025807.
[12] M. Aglietta, et al., Measurement of the Neutron Flux Produced by Cosmic-Ray Muons with LVD at Gran Sasso, 1999.
[13] M.B. Chadwick, et al., Nuclear Data Sheets 107 (2006) 2931.
[14] D. Dujmic, et al., Astroparticle Physics 30 (2008) 58.
[15] H. Yegoryan, Study of Alpha Background in a Dark Matter Detector, Senior Thesis, Massachusetts Institute of Technology, 2010.
[16] T. Caldwell, et al., Transport Properties of Electrons in CF(4), 2009.
[17] I.C. Wolfe, Measurement of Work Function in CF_4 Gas, Senior Thesis, Massachusetts Institute of Technology, 2010.
[18] K. Nakamura, et al., Journal of Physics G 37 (2010) 075021.
[19] J.F. Ziegler, J.P. Biersack, M.D. Ziegler, SRIM: The Stopping and Range of Ions in Matter, ⟨http://www.srim.org/⟩, 2011.
[20] D. Dujmic, et al., Nuclear Instruments and Methods in Physics Research Section A 584 (2008) 327.

Bibliography

[1] Y. Abe et al. Reactor electron antineutrino disappearance in the Double Chooz experiment. *Phys. Rev. D*, 86, 052008, 2012.

[2] F.P. An et al. Observation of electron-antineutrino disappearance at Daya Bay. *Phys. Rev. Lett.*, 108, 171803, 2012.

[3] J.K. Ahn et al. Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment. *Phys. Rev. Lett.*, 108, 191802, 2012.

[4] Y. Abe et al. First measurement of θ_{13} from delayed neutron capture on hydrogen in the double chooz experiment. *ArXiv*, 1301.2948, 2013.

[5] T. Schwetz, M. Tortola, and J. W. F. Valle. Global neutrino data and recent reactor fluxes: status of three-flavour oscillation parameters. *ArXiv*, 1103.0734, 2011.

[6] B. Aharmim et al. Electron energy spectra, fluxes, and day-night asymmetries of ^{8}b solar neutrinos from measurements with nacl dissolved in the heavy-water detector at the sudbury neutrino observatory. *Phys. Rev. C*, 72, 055502, 2005.

[7] J. N. Bahcall and A. M. Serenelli. New solar opacities, abundances, helioseismology, and neutrino fluxes. *Astrophys. J.*, 621, L85, 2005.

[8] S. Abe et al. Precision measurement of neutrino oscillation parameters with kamland. *Phys. Rev. Lett.*, 100, 221803, 2008.

[9] P. A. N. Machado et al. Combining accelerator and reactor measurements of θ_{13}: the first result. *ArXiv*, 1111.3330, 2012.

[10] Th. A. Mueller et al. Improved predictions of reactor antineutrino spectra. *Phys. Rev. C*, 83, 054615, 2011.

[11] Y. Abe et al. Indication of Reactor $\bar{\nu}_e$ Disappearance in the Double Chooz Experiment. *Pysical Review Letters*, 108, 131801, 2012.

[12] F.P. An et al. Improved measurement of electron antineutrino disappearance at daya bay. *ArXiv*, 2012.

[13] K. Abe et al. Indication of Electron Neutrino Appearance from an Accelerator-produced Off-axis Muon Neutrino Beam. *Phys. Rev. Lett.*, 107 041801, 2011.

[14] P. Adamson et al. Improved search for muon-neutrino to electron-neutrino oscillations in MINOS. *Phys. Rev. Lett.*, 107, 181802, 2011.

[15] C. Aberle. Ph.D Thesis: Optimization, simulation and analysis of the scintillation signals in Double Chooz experiment. *Ph.D Thesis*, 2011.

[16] D. R. Tilley et al. Energy levels of light nuclei a=8,9,10. *Nuclear Physics A*, 745:155-362, 2004.

[17] R. Davis et al. Search for neutrinos from sun. *Phys. Rev. Lett.*, 20:1205-1209, 1968.

[18] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1572-1567, 1998.

[19] P. Adamson et al. Measurement of the neutrino mass splitting and flavor mixing by MINOS. *Phys. Rev. Lett.*, 106, 181801, 2011.

[20] M. Appolonio et al. Limits on neutrino oscillations from the chooz experiment. *Phys. Lett. B*, B466, 415, 1999.

[21] F. Boehm et al. Search for neutrino oscillatins at the palo verde nuclear reactors. *Phys. Rev. Lett.*, 84 17, 1999.

[22] Ch. Kraus et al. Final results from phase ii of the mainz neutrino mass search in tritium beta decay. *Eur. Phys. J.*, C40:447-468, 2005.

[23] B. Kayser. On the quantum mechanics of neutrino oscillation. *Phys. Rev. D*, 24, 110, 1981.

[24] W. Hampel et al. Gallex solar neutrino observations: Results for gallex iv. *Phys. Lett. B*, 447:127-133, 1999.

[25] J. N. Abdurashitov et al. Solar neutrino flux measurements by the soviet-american gallium experiment (sage) for half the 22-year solar cycle. *Journal of Experimental And Theoretical Physics*, 95:181, 2002.

[26] L. Wolfenstein. Neutrino oscillations in matter. *Phys. Rev. D*, 17, 2369, 1978.

[27] B. Kayser, F. Gibrat-Debu, and F. Perrier. Physics of massive neutrinos. World Scientific Lecture Notes in Physics. World Scientific Pub Co. Inc., 1989.

[28] P. Vogel and F. Boehm. Physics of massive neutrinos. Cambridge University Press, 1992.

[29] F. Ardellier et al. Double chooz: A search for the neutrino mixing angle θ_{13}. *ArXiv*, 0606025v4, 2006.

[30] P. Adamson et al. Measurement of the Neutrino Mass Splitting and Flavor Mixing by MINOS. *Phys. Rev. Lett.*, 106, 181801, 2011.

[31] C. L Cowan Jr., F. Reines, F. B. Harrison, H. W. Kruse, A. D McGuire. Detection of the Free Neutrino: a Confirmation. *Science*, 124(3212), 103-104, 1956.

[32] T. Matsubara et al. Evaluation of 400 low background 10-in. photo-multiplier tubes for the Double Chooz experiment. *Nucl. Instrum. Methods Phys. Res.*, Sect. A 661, 16, 2011.

[33] Andi S. Cucoanes. A Toy Model for the Gamma Spectrum Generated in n-Captures on Gd. *Double Chooz Internal*, DocDB 3139.

[34] C. J. Clopper and E. S. Pearson. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, Vol. 26, No. 4, pp. 404-413, 1934.

[35] Z. Djurcic et al. Neutron Efficiency Systematics for 2nd Publication. *Double Chooz Internal*, DocDB-4023, 2012.

[36] B. Reinhold et al. Spill In/Out for n-H analysis. *Double Chooz Internal*, DocDB-4156, 2012.

[37] A. Ferrari, P.R. Sala, A. Fasso, and J. Ranft. Fluka: A multi-particle transport code (program version 2005). CERN-2005-010, 2005.

[38] A. Fasso et al. The Physics Models of FLUKA: Status and Recent Development. 2003.

[39] Y. Prezado et al. Large asymmetry in the strongest β-transition for A=9. *Physics Letters B*, 576:55-61, 2003.

[40] Y. Prezado et al. Low-lying resonance states in the ^{9}Be continuum. *Physics Letters B*, 618(1-4):43-50, 2005.

[41] T. A. D. Brown et al. Decay studies for states in ^{9}Be up to 11 MeV: Insights into the n+^{8}Be and α+^{5}He cluster structure. *Physical Review C*, 76:054605, 2007.

[42] J. Beringer et al. The Review of Particle Physics. *Physical Review D*, 86:010001, 2012.

[43] K. Nakajima. Cluster Noise Events and ^{214}Bi-Po Increase. *Double Chooz Internal*, DocDB-4104, 2011.

[44] V. Kopeikin et al. Reactor as a source of antineutrinos: Thermal fission energy. *Physics of Atomic Nuclei*, 67:1892-1899, 2004.

[45] C. Bemporad et al. Reactor-based neutrino oscillation experiments. *Rev. Mod. Phys.*, 74:297-328, 2002.

[46] Y. Declais et al. Study of reactor antineutrino interaction with proton at bugey nuclear power plant. *Physics Letters B*, 338:383-389, 1994.

[47] O. Meplan. Mure. mcnp utility for reactor evolution. user guide. version 1.0. *Tech. Rep. LPSC 0912 and IPNO-09-01 2009.*

[48] G. Marleau, R. Roy, and A. Herbert. Dragon: A collision probability transport code for cell and supercell calculations. *Report IGE-157*, 1994.

[49] P. Huber. Determination of antineutrino spectra from nuclear reactors. *Phys. Rev. C*, 84:024617, 2011.

[50] F. von Felitzsch et al. Experimental beta-spectra from ^{239}pu and ^{235}u thermal neutron fission products and their correlated antineutrino spectra. *Physics Letters B*, 118:162-166, 1982.

[51] K. Schreckenbach et al. Determination of the antineutrino spectrum from ^{235}u thermal neutron fission products up to 9.5 mev. *Physics Letters B*, 160:325-330, 1985.

[52] A. A. Hahn et al. Antineutrino spectra from ^{241}pu and ^{239}pu thermal neutron fission products. *Physics Letters B*, 218:365-368, 1989.

[53] S. Agostinelli et al. Geant4: A simulation toolki. *Nucl. Instrum. Meth.*, A506:250-303, 2003.

[54] J. B. Birks. Scintillations from organic crystals: Specific fluorescence and relative response to different radiations. *Pceedings of the Physical Society. Section A*, A539:217-235, 2005.

[55] Dario Motta and Stefan Schonert. Optical properties of bialkali photocathodes. *Nucl. Instrum. Meth.*, A539:217-235, 2005.

[56] U.S. Double Chooz Group. Technical Report for CUFits Framework. *Double Chooz Internal*, DocDB 3222.

[57] Open grid scheduler/grid engine.

[58] Htcondor: High throughput computing.

[59] Kazuhiro Terao. Waveform/trigger level flagging and exploring light noise. *Double Chooz Internal*, DocDB-2195, 2011.

[60] Kazuhiro Terao. Re-newed post-dogsifier validation tool. *Double Chooz Internal*, DocDB-2961, 2011.

[61] Kazuhiro Terao. n-h capture analysis for muon spallation neutrons. *Double Chooz Internal*, DocDB-3429, 2011.

[62] Kazuhiro Terao. Detector effects on spallation neutron analysis. *Double Chooz Internal*, DocDB-3865, 2012.