Gran Sasso Science Institute

ASTROPARTICLE PHYSICS
DOCTORAL PROGRAMME

Cycle XXXI - AY 2015/2018

High Precision Solar Neutrino Spectroscopy with Borexino and JUNO

PHD CANDIDATE
Xuefeng Ding

PhD Thesis submitted
April 22nd, 2019

ADVISORS

Nicola Rossi
Laboratori Nazionali
del Gran Sasso

Matteo Agostini
Technical University Munich

Gioacchino Ranucci
Sezione di Milano, INFN

GRAN SASSO
SCIENCE INSTITUTE
SCHOOL OF ADVANCED STUDIES
Scuola Universitaria Superiore

ma per seguir virtute e canoscenza
SISSA
Scuola Internazionale Superiore
di Studi Avanzati

Abstract

This PhD thesis focuses on the study of solar neutrinos with Borexino and JUNO.

The first simultaneous analysis of all solar neutrino fluxes has been performed using Borexino Phase-II data. For this purpose, an innovative multivariate analysis has been designed and implemented. A software framework is created to perform such an analysis. Thanks to parallel computing techniques and graphics processing units (GPU), this software framework reduces the computing time by a factor of 200 compared with the existing tools. New analytical models of the detector response have been developed and benchmarked against simulations. A comprehensive study of the sensitivity and systematic uncertainties has also been performed. The results obtained from this analysis assuming the MSW-LMA model are:

- $\nu(pp)$: $(6.1 \pm 0.5\ ^{+0.3}_{-0.5}) \times 10^{10}$ [$\mathrm{cm^{-2}s^{-1}}$]
- $\nu(^7\mathrm{Be})$: $(4.99 \pm 0.11\ ^{+0.06}_{-0.08}) \times 10^{9}$ [$\mathrm{cm^{-2}s^{-1}}$]
- $\nu(pep)$: $(1.27 \pm 0.19\ ^{+0.08}_{-0.12}) \times 10^{8}$ [$\mathrm{cm^{-2}s^{-1}}$] (Assuming HZ ν(CNO)),
  and $(1.39 \pm 0.19\ ^{+0.08}_{-0.13}) \times 10^{8}$ [$\mathrm{cm^{-2}s^{-1}}$] (Assuming LZ ν(CNO))
- ν(CNO): $< 7.9 \times 10^{8}$ (95% C.L.) [$\mathrm{cm^{-2}s^{-1}}$]

The estimation of the fluxes has been combined with an independent analysis of the ^{8}B solar neutrino flux based on Borexino data in order to perform phenomenological studies on the solar models and neutrino oscillation models. It is found that, assuming the MSW-LMA model, the Low Metallicity model (B16-LZ) is rejected with a confidence level of 96.6% (2.2 σ); it is also found that, assuming the High Metallicity model (B16-HZ), the vacuum oscillation model is rejected with a confidence level of 98.2% (2.4 σ).

The sensitivity of Borexino to CNO solar neutrinos has been studied assuming a constraint on the its major background ^{210}Bi. Different methods to measure the ^{210}Bi decay rate with its daughter ^{210}Po have been considered and characterized. In order to reduce the systematic uncertainties, we have created a special dataset whose events are reconstructed using only the PMTs that have been operating stably over the considered data-taking period. This dataset is proved to be be useful for various studies, including testing fundamental assumptions made in the method to measure the ^{210}Bi decay rate and improving the light yield monitoring in the Monte Carlo simulation.

In the context of the JUNO experiment, fundamental aspects of the analysis and simulation framework needed for studying the solar neutrinos have been developed, including the event generator used in simulation, the algorithms used to suppress event pile-up effect and contribution of dark noise of PMTs, and energy/vertex reconstruction softwares optimized for sub-MeV events.

The signal and backgrounds rates have been evaluated, as well as the efficiencies of various event selection for detecting the ^{8}B solar neutrinos with JUNO. It is demonstrated that JUNO can achieve around 2 to 1 signal-to-background ratio in the observed energy range from 2 to 3 MeV, making it possible to test new physics models. A comprehensive study of the sensitivities of JUNO to different solar neutrino components has also been performed. According to our estimation, the evaluated sensitivities are

- $\nu(^7\mathrm{Be})$: 7% statistical + systematic uncertainty (assuming 0.5% light yield accuracy)
- ν(CNO): 13% statistical + systematic uncertainty (assuming 0.1% light yield accuracy)

Preface

In 2015 when I started my PhD project, the answers to two questions in solar neutrino physics were still at large:

- Which of the two solar metallicity models is correct?
- What is the solar neutrino survival probability in the transition region?

One year later, my colleague and former room-mate demonstrated a method to reach the lowest ^{8}B solar neutrino detection threshold in the world. Due to the lack of statistics, no conclusions could be drawn. I instead decided to work on the precise measurement of solar neutrino fluxes to answer the first question and joined the Borexino collaboration. In 2018 when I started writing this thesis, Borexino has just put its result of Phase-II data analysis on arXiv, a moderate preference towards the High Metallicity model.

My work in Borexino has been a journey full of excitements. With Dr. Nicola Rossi as my supervisor, I started to study the detector response model. Finally I found the correct formulae and performed the measurement of ^{7}Be solar neutrino interaction rate with 3% accuracy. At the same time, I was unsatisfied with the hour-scale fitting speed of our fitting code, which prevented our progress in understanding our spectrum and killed our passion, so I reformed our fitting code to accelerate it with parallel computing, and the fitting time finally reached a few seconds. In 2017 Summer, the collaboration decided to announce the results on solar neutrinos based on 3.5 years of new data. My tutor Dr. Matteo Agostini has encouraged my classmate to use a multi-dimensional probability density function based fitting method to extract the solar neutrino interaction rate. I joined this project by estimating its shape-related systematic uncertainty using a new method just before the deadline for submitting the analyses to the collaboration, and after a few weeks of overtime working the three of us succeeded to finalize the analysis in time. Later on I also deployed this method for other fitting approaches. A new challenge emerged when it was found we had problems modeling our data after 2016 summer. At the same time our main computing farm "centro nazionale dell’Istituto Nazionale di Fisica Nucleare" (CNAF) was flooded on November 9th. 2017 and had been unavailable for 148 days till April 20th. 2018. This dark period became a good chance to slow down and review weak points of our analysis. After a few weeks of meditation, a few hours of discussing in the `Whatsapp` group and a few more hours of checking, I found that the temporal stability of the light yield was worse than estimated. I proposed to produce a new dataset using only the 1000 stablest high quality PMTs to get rid of such instability. Moreover, this dataset provided also a way to track the temporal stability of detector for Monte Carlo simulation.

At the same time, significant progress has been made towards the detection of CNO solar neutrinos. We re-evaluated Borexino's sensitivity on CNO solar neutrino flux and it was found that we could reach a 3 σ discovery of CNO solar neutrino given a 10% constraint of the ^{210}Bi decay rate under the high metallicity hypothesis. The main background of CNO solar neutrinos, ^{210}Bi, can be estimated through its daughter ^{210}Po. In the last days, after the coldest winter in my life, I started to look at the distribution of the ^{210}Po event vertexes, and I found a beautiful vortex. The vortex bottom has a stable position, so I was convinced that this was not due to fluctuations. The rate at the vortex bottom was low enough to claim a CNO solar neutrino discovery. I was so excited that night that I texted all my friends that I saw the most beautiful vortex in the world.

In the second part of my thesis, I worked with Dr. Gioacchino Ranucci as my supervisor on JUNO. We formed a little group with a few people from Italy and started studying the sensitivity of JUNO to ^{7}Be solar neutrinos, hoping to study all solar neutrinos one-by-one. However we found that the impact of the dark noise was critical assuming the baseline design of the PMTs, and thus I started developing reconstruction algorithms to suppress dark noise. At the same time, Chinese colleagues evaluated the background level and drew the conclusion that the background level for ^{8}B solar neutrino through elastic scattering channel was too high. With my experience in Borexino, I found they had underestimated the muon-neutron veto efficiency for two important cosmogenic backgrounds, and they didn't treat the ^{238}U chain backgrounds properly. After correcting these two factors, I found JUNO to be able to reach 2 MeV energy threshold, which would be even as important as determining the mass ordering.

At last, I would like to dedicate this thesis to my father. It is he who planted the seeds of eagerness to explore in my heart, and he supported me all the way. 谨以此书献给我的父亲志民。是他在我心里种下渴望探索的种子，也是他支持我一路前行。

Introduction

In 2018, neutrino astronomy was surely the hottest topic in the physics community. On 2018 July the large ice based Cherenkov detector IceCube announced the discovery of the first extra-galactic neutrino source, a blazer visible also by the Gamma-ray telescopes Fermi and MAGIC, and several other experiments[1].

However neutrino astronomy is not a fresh concept. Neutrinos have been proposed to be used to study the Sun since 1964[2,3]. Although at the beginning people were confused by the missing neutrinos due to neutrino oscillation phenomena[4], in 2001 the solar model was confirmed with the measured solar neutrino flux[5]. In 2009, a more realistic model was proposed[6], however it is in tension with helioseismology data. The new model and old model can be distinguished through the neutrino fluxes. The SuperK experiment, with 50 kilo-tons of water, did measure the ^{8}B solar neutrino flux with 2% precision[7]. However the result is exactly at the middle of two model predictions and no conclusions could be drawn.

Borexino, on the other hand, had been working hard to improve its precision of ^{7}Be solar neutrino flux. Because the ^{7}Be solar neutrino flux is three orders of magnitude higher than the ^{8}B solar neutrino flux, the precisions are almost the same even though Borexino is only around 1/100 of SuperK. The main challenge is to reduce the radioactive contaminations in the detector by more than ten orders of magnitude. After 3 years of data taking from 2007 May till 2010 May, Borexino succeeded to make a 5% ^{7}Be solar neutrino flux measurement. In order to improve the signal to background ratio of the previous measured solar neutrino fluxes, Borexino undertook a dedicated purification campaign and started another round of data taking in 2011 December. In 2017 summer at the TAUP conference, the collaboration announced a simultaneous measurement of pp, ^{7}Be and pep solar neutrino fluxes and the precision on ^{7}Be solar neutrino flux reached 3% while the significance on pep solar neutrino detection reached more than 5 σ. Combined with Borexino's ^{8}B solar neutrino measurement, a moderate preference towards the old model was found.

Another exciting challenge is the detection of CNO solar neutrinos. The CNO-cycle is another mechanism of proton-proton fusion and is more important compared to the pp-chain mechanism for stars that are heavier or older than the Sun, and the neutrinos from this process are not detected yet. The predicted CNO solar neutrino flux

---

[1] T. IceCube et al., "Multi-messenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A", *Science*, eaat1378, July 2018, DOI: `10.1126/science.aat1378`, arXiv: `1807.08816`

[2] R. Davis, "Solar Neutrinos. II. Experimental", *Physical Review Letters*, vol. 12, no. 11, pp. 303–305, Mar. 1964, DOI: `10.1103/PhysRevLett.12.303`

[3] J. N. Bahcall, "Solar neutrinos. I. theoretical", *Physical Review Letters*, vol. 12, no. 11, pp. 300–302, 1964, DOI: `10.1103/PhysRevLett.12.300`

[4] R. Davis, D. S. Harmer, and K. C. Hoffman, "Search for neutrinos from the sun", *Physical Review Letters*, vol. 20, no. 21, pp. 1205–1209, 1968, DOI: `10.1103/PhysRevLett.20.1205`, arXiv: `arXiv:1011.1669v3`

[5] The SNO collaboration, "Measurement of the Rate of $\nu_e + d\top + p + e^-$ Interactions Produced by ^{8}B Solar Neutrinos at the Sudbury Neutrino Observatory", *Physical Review Letters*, vol. 87, no. 7, p. 071301, July 2001, DOI: `10.1103/PhysRevLett.87.071301`

[6] M. Asplund et al., "The Chemical Composition of the Sun", *Annual Review of Astronomy and Astrophysics*, vol. 47, no. 1, pp. 481–522, Sept. 2009, DOI: `10.1146/annurev.astro.46.060407.145222`, arXiv: `0909.0948`

[7] The Super-Kamiokande collaboration, "Solar neutrino measurements in Super-Kamiokande-IV", *Physical Review D*, vol. 94, no. 5, pp. 1–33, 2016, DOI: `10.1103/PhysRevD.94.052010`, arXiv: `1606.07538`

differs by almost a factor 2 between the two solar models and a precise measurement could provide a conclusive evidence in favor of one of them. Borexino might claim a 3 σ discovery of CNO solar neutrino if its main background ^{210}Bi is measured within 10%. To achieve such a precision, the convective motions need to be reduced and a reliable model describing the ^{210}Po background is needed. The collaboration has been working on it.

Besides studying the Sun, solar neutrinos are also an excellent source to study the neutrino oscillation properties, especially those related to forward scattering on electrons (i.e. the matter effect), because solar neutrinos need to propagate through a large distance of ultra high density region before leaving the Sun. For *pp* and ^{7}Be solar neutrinos the matter effect is negligible while for ^{8}B solar neutrino the matter effect dominates because of the so-called MSW resonance. The non-trivial energy dependence of the neutrino survival probability in the transition region[8] is not observed yet and interesting for the physicists. Besides, the magnitude of the matter effect is determined by Δm^2_{12}[9]. There is tension between its value determined by solar neutrinos[10] and that determined by reactor neutrinos[11]. New physics models, like non-standard interactions have been proposed to reconcile the two results, and a precise measurement of the ^{8}B solar neutrino flux between 3 to 5 MeV neutrino energy range would rule out or confirm such proposals[12]. Currently the measurement with the lowest average neutrino energy is 7.9 MeV and the precision is around 8%, given by Borexino Phase-II analysis[13,14].

JUNO as the next generation experiment is designed to determine the mass ordering of neutrinos. It will be located not as deep underground as Borexino and the requirement on purity is less stringent, so the signal to background level for solar neutrinos will be worse as compared with Borexino. These disadvantages, however, would be compensated by its huge mass assuming good knowledge of the theoretical energy spectra and the detector response. For ^{8}B solar neutrinos, because we would like to measure the energy distribution, it is essential to suppress cosmogenic and natural radioactivity backgrounds. The cosmogenic backgrounds can be suppressed with algorithms using the reconstructed muon tracks and the vertexes of cosmogenic neutrons. The natural radioactive background events in the critical visible energy range (2–3 MeV) can be suppressed with the analysis of the scintillation light arrival time of each event. After applying all these techniques the detection threshold can reach 2 MeV and the neutrino energy range can reach 7 MeV.

Personal contributions

The work presented in this thesis was conducted at Laboratori Nazionali del Gran Sasso (INFN) in the context of the Borexino collaboration and at Gran Sasso Science Institute in the context of the JUNO collaboration. My contributions can be summarized as the following:

[8] See Definition of the transition region see Section 1.1.5.

[9] C. Giunti and C. Kim, *Fundamentals of neutrino physics and astrophysics*, vol. 54, 2, Oxford University Press, USA, 2007, p. 720, DOI: `10.1093/acprof:oso/9780198508717.001.0001` Section 9.3

[10] The Super-Kamiokande collaboration, "Solar neutrino measurements in Super-Kamiokande-IV"

[11] The KamLAND collaboration, "Reactor on-off antineutrino measurement with KamLAND", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 88, no. 3, pp. 1–10, 2013, DOI: `10.1103/PhysRevD.88.033001`, arXiv: `1303.4667 [hep-ex]`

[12] M. Maltoni and A. Yu. Smirnov, "Solar neutrinos and neutrino physics", *The European Physical Journal A*, vol. 52, no. 4, p. 87, Apr. 2016, DOI: `10.1140/epja/i2016-16087-0`, arXiv: `1507.05287`

[13] The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure", vol. 016, pp. 1–13, Sept. 2017, arXiv: `1709.00756`

[14] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

1. I played a leading role in the development of the fit method for Borexino using the energy observable `charge` and the analytical response function, in order to match the required precision of the Phase-II analysis. The main advantage of this energy observable is that its resolution is better than for other variables in the region of interest of ^{7}Be, *pep* and CNO solar neutrinos. The ^{7}Be solar neutrino interaction rate is correlated with resolution parameters. After improving the resolution model, I could use a common model for decaying ^{210}Po and constant-rate components, and the ^{210}Po peak can pin down one degree of freedom of the model. After applying non-uniformity corrections, I almost removed the corresponding resolution terms and eliminated the other degree of freedom, and the sensitivity on ^{7}Be solar neutrino interaction rate reached 3%.

2. I developed a new procedure in Borexino to estimate the energy-distribution related systematic uncertainties. In Borexino the signal and backgrounds are highly entangled, and thus a careful estimation of the energy-distribution related systematic uncertainties is needed to guarantee the reliability of the fitting result. I have proposed to extract the systematic uncertainty by testing the fit on a large ensemble of pseudo-experiment parameterized with a toy Monte Carlo including all systematic uncertainties. This method takes into account bin-to-bin correlations introduced by each systematic source properly, and it is not limited to the χ^2 regime. I also studied the systematic uncertainties related to the stability of detector light yield, which were previously neglected. I developed procedures to track the detector light yield using ^{210}Po and ^{14}C backgrounds. The tracked light yield is also used as a new input for Monte Carlo simulation.

3. In JUNO I played a leading role in studying JUNO's physics potential regarding solar neutrinos. I have developed the event generator of solar neutrinos used in simulation, which provides the energy and momentum of recoil electrons, and a specialized reconstruction algorithms to suppress dark noise and ^{14}C pile-ups, which affects the precision on ^{7}Be and *pp* solar neutrinos, respectively. I also reviewed backgrounds estimations and proposed new methods to suppress backgrounds for ^{8}B solar neutrinos. According to my estimation, we could reach 2 MeV detection threshold for ^{8}B solar neutrino. JUNO could hence be a unique experiment and determine Δm^2_{12} from both reactor anti-neutrinos and solar neutrinos. It could also measure the survival probability in the transition region, confirming or rejecting non-standard interactions as the cause of the 2 σ discrepancy between KamLAND and the solar neutrino global fit results.

4. I also created and developed *GooStats*. *GooStats* is a package to run multivariate fits on graphics processing units. The computation of likelihoods is parallelized and the fitting time is shortened

from days to seconds. The flexible design makes it easy to implement new statistical analysis modules. Modules used to extract solar neutrino rates and magnetic moments have been developed and used in recent Borexino publications[15,16]. Modules used to differentiate neutrino mass ordering through the three neutrino vacuum oscillation pattern has also been developed. The package is described in [17].

Thesis layout

The thesis is organized in the following way:

1. Chapter 1 introduces the solar models, the neutrino oscillation theory, and the role of Borexino and JUNO in these two fields. Future experiments are also discussed.

2. Chapters 2–8 cover the analyses methods and results of Borexino. Chapter 2 describes the main features of the Borexino detector and reviews the analyses at the beginning of 2016, i.e. the starting point of this thesis. Chapter 3 discusses the new fitting procedure using a multivariate fit analytical response function accelerated by graphics processing units. Chapter 4 discusses an improved detector response models. Chapter 5 discusses various sources of systematic uncertainty, especially temporal stability of detector light yield, and the method to estimate the magnitude of these systematic uncertainties. Chapter 6 describes the results and discusses their physics impact. Chapter 7 is devoted to the estimation of sensitivity of Borexino on CNO solar neutrinos assuming a measured ^{210}Bi decay rate. Chapter 8 discusses the ^{210}Bi meausurement using its daughter ^{210}Po.

3. Chapter 9–11 cover the physics potential of JUNO for solar neutrinos. Chapter 9 introduces the JUNO detector and its physics program. Chapter 10 presents the estimation of signal and background rates after event selection for the ^{7}Be solar neutrinos. Chapter 11 describes the work on reconstruction for sub-MeV events and the sensitivity estimation for ^{7}Bi and CNO solar neutrinos, respectively.

4. Chapter 12 summarizes the conclusions of this thesis.

Related publications

The collaboration has published the Phase-II results as

- The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

The content of this *Nature* paper is strongly based on the results presented in this thesis. The software used to obtain them has been published separately as

[15] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos"

[16] The Borexino collaboration, "Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data", *Physical Review D*, vol. 96, no. 9, p. 091103, Nov. 2017, DOI: `10.1103/PhysRevD.96.091103`, arXiv: `1707.09355`

[17] X. Ding, "GooStats: A GPU-based framework for multi-variate analysis in particle physics", *Journal of Instrumentation*, vol. 13, no. 12, P12018–P12018, Dec. 2018, DOI: `10.1088/1748-0221/13/12/P12018`

- X. Ding, "GooStats: A GPU-based framework for multi-variate analysis in particle physics", *Journal of Instrumentation*, vol. 13, no. 12, P12018–P12018, Dec. 2018, DOI: `10.1088/1748-0221/13/12/P12018`

Besides, additional papers describing the analysis details are currently under preparation[18].

---

[18] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II", pp. 1–8, July 2017, arXiv: `1707.09279`

Contents

List of Figures

List of Tables

1

Introduction

In this chapter, an introduction to the solar physics and neutrino oscillations is given. Experimental searches on solar neutrinos are also introduced.

In 1968, results from Homestake[1], that the measured solar neutrino flux was well below the predicted value[2], known as the *Solar Neutrino Problem*[3], has triggered numerous experimental and theoretical development on neutrino physics. The *Solar Neutrino Problem* was solved in 2001 by the SNO experiment discovering the neutrino flavor transformation of solar neutrinos[4,5]. Neutrino flavor transformation is governed by the neutrino oscillation. It is a quantum mechanical phenomenon, and is described with Pontecorvo-Maki-Nakagawa-Sakata matrix[6]. For solar neutrinos, the oscillation is amplified due to the neutrino coherent scattering on matter and is described by the Mikheyev–Smirnov–Wolfenstein effect[7]. Nowadays, open questions remains, such as the absolute scale of neutrino masses, octant of θ_{23}, value of CP violation phase angle δ_{CP}[8].

Neutrinos, due to their low cross sections of interaction with matter, can easily travel astrophysical distance, and thus can be used as a probe to study the astronomy. Solar neutrinos have been used as one of the most powerful tools to test the model of the Sun, the *Standard Solar Model*. The *Standard Solar Model* plays a fundamental role for stellar models. It is used to calibrate stellar models, and is also a benchmark against which we can test additional physical processes in stars. A solar model was proposed in 2009[9] and it was believed to be improved significantly in several aspects. However, it lost the agreement with high precision helioseismic data, and this is known as the *Solar Metallicity Problem*. Solar neutrinos may again provide a key to this mystery.

This chapter is organized as in the following: Section 1.1 introduces the neutrino oscillation phenomena, including vacuum oscillation and oscillation in matter. Section 1.2 introduces the neutrino astronomy, the *Standard Solar Model* and neutrinos as a probe to study the *Standard Solar Model*. Section 1.3 introduces the physics results achieved by the Borexino experiment and the physics program and potential of the JUNO experiment. At last, Section 1.4 introduces future experiments that have potentials on solar neutrino physics.

---

[1] R. Davis, D. S. Harmer, and K. C. Hoffman, "Search for neutrinos from the sun", *Physical Review Letters*, vol. 20, no. 21, pp. 1205–1209, 1968, doi: `10.1103/PhysRevLett.20.1205`, arXiv: `arXiv:1011.1669v3`

[2] J. Bahcall, N. Bahcall, and G. Shaviv, "Present Status of the Theoretical Predictions for the ^{37}Cl Solar-Neutrino Experiment", *Physical Review Letters*, vol. 20, no. 21, pp. 1209–1212, 1968, doi: `10.1103/PhysRevLett.20.1209`

[3] For a review of the problem, see J. N. Bahcall and R. K. Ulrich, "Solar models, neutrino experiments, and helioseismology", *Reviews of Modern Physics*, vol. 60, no. 2, pp. 297–372, 1988, doi: `10.1103/RevModPhys.60.297`

[4] The SNO collaboration, "Measurement of the Rate of $\nu_e + d\top + p + e^-$ Interactions Produced by ^{8}B Solar Neutrinos at the Sudbury Neutrino Observatory", *Physical Review Letters*, vol. 87, no. 7, p. 071301, July 2001, doi: `10.1103/PhysRevLett.87.071301`

[5] The SNO collaboration, "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory", *Physical Review Letters*, vol. 89, no. 1, pp. 1–6, 2002, doi: `10.1103/PhysRevLett.89.011301`, arXiv: `0204008 [nucl-ex]`

[6] B. Pontecorvo, "Mesonium and anti-mesonium", *Sov.Phys.JETP*, vol. 6, p. 429, 1957; B. Pontecorvo, "Inverse beta processes and nonconservation of lepton charge", *Sov.Phys.JETP*, vol. 7, pp. 172–173, 1958; Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the Unified Model of Elementary Particles", *Progress of Theoretical Physics*, vol. 28, no. 5, pp. 870–880, Nov. 1962, doi: `10.1143/PTP.28.870`

[7] L. Wolfenstein, "Neutrino oscillations in matter", *Physical Review D*, vol. 17, no. 9, pp. 2369–2374, May 1978, doi: `10.1103/PhysRevD.17.2369`; S. P. Mikheyev and A. Y. Smirnov, "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos", *Sov. J. Nucl. Phys.* Vol. 42, pp. 913–917, 1985

[8] P. F. de Salas et al., "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 782, pp. 633–640, 2018, doi: `10.1016/j.physletb.2018.06.019`, arXiv: `1708.01186`

[9] M. Asplund et al., "The Chemical Composition of the Sun", *Annual Review of Astronomy and Astrophysics*, vol. 47, no. 1, pp. 481–522, Sept. 2009, doi: `10.1146/annurev.astro.46.060407.145222`, arXiv: `0909.0948`

1.1 Neutrino oscillations

In this section, the neutrino oscillation phenomena is introduced. We start from a simple explanation using a 2-neutrino model, then we review the history of its discovery, after that its current description in vacuum and in matter, and at last the open questions. Discussions are related but not limited to solar neutrinos. More general discussions, including the discovery of neutrinos, historical and current experimental and theoretical studies can be found in[10,11,12]. For even more introductory materials, I recommend reading the textbook written by Griffith[13].

1.1.1 Intro: oscillation with 2ν model

Neutrino oscillations can be understood with basic quantum mechanics[14]. Consider neutrinos have two mass eigenstates, and they constitute the base of the Hilbert space. In the Schrödinger picture, they evolve with time

$$|m_i(t)\rangle = e^{-iHt}\,|m_i(0)\rangle \tag{1.1}$$

$$H\,|m_i\rangle = E_i\,|m_i\rangle \tag{1.2}$$

where H is the hamiltonian operator, E_i is the energy of i-th eigenstate.

When neutrinos are produced or detected, they are always in another set of eigenstates, the flavor eigenstates, which also constitute the base of the Hilbert space. The mixing matrix can be described with a mixing angle θ:

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_x\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} |m_1\rangle \\ |m_2\rangle \end{pmatrix} \tag{1.3}$$

$$\begin{pmatrix} |m_1\rangle \\ |m_2\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} |\nu_e\rangle \\ |\nu_x\rangle \end{pmatrix} \tag{1.4}$$

Consider an electron neutrino $|\nu_e\rangle$ is produced, then the probability that it is detected as the state of electron neutrino, according to quantum mechanics, is

$$P\left(|\nu_e\rangle \to |\nu_e\rangle\right) = |\,\langle \nu_e|\mathcal{U}(t,0)|\nu_e\rangle|^2 \tag{1.5}$$

where $\mathcal{U}(t,0)$ is the time evolution operator and is e^{-iHt}.

After some calculation,

$$P\left(|\nu_e\rangle \to |\nu_e\rangle\right) \simeq 1 - \sin^2 2\theta \cdot \sin^2\left[1.27 \times 10^3 \cdot \Delta m^2 \ (\text{eV}^2) \cdot \frac{L}{E}\ (\text{km/MeV})\right]$$

where Δm^2 is the mass square difference of two neutrino mass eigenstate, and the assumption that neutrino are ultra-relativistic particles is used, i.e. $E \simeq p + m^2/(2p)$, where p is the momentum. From this we can see the survival probability oscillate periodically with the wavelength

$$L_{\text{osc}} = 2\pi \cdot \frac{4E}{\Delta m^2} \tag{1.6}$$

[10] S. F. King, "Models of neutrino mass, mixing and CP violation", *Journal of Physics G: Nuclear and Particle Physics*, vol. 42, no. 12 2015, DOI: `10.1088/0954-3899/42/12/123001`, arXiv: `1510.02091`

[11] Particle Data Group, "Review of Particle Physics", *Physical Review D*, vol. 98, no. 3, p. 030001, Aug. 2018, DOI: `10.1103/PhysRevD.98.030001` Section 14

[12] C. Giunti and C. Kim, *Fundamentals of neutrino physics and astrophysics*, vol. 54, 2, Oxford University Press, USA, 2007, p. 720, DOI: `10.1093/acprof:oso/9780198508717.001.0001`

[13] D. J. Griffiths, *Introduction to Elementary Particles*, 2008, DOI: `10.3769/radioisotopes.60.527`

[14] Many excellent introductory books to Quantum Mechanics exist. Such as J. J. Sakurai, *Modern quantum mechanics; rev. ed.* Reading, MA: Addison-Wesley, 1994

The survival probability of 2 MeV neutrinos, assuming $\sin^2\theta = 0.32$ and $\Delta m^2 = 7.6 \times 10^{-5}\,\mathrm{eV}^2$, is shown in Figure 1.1.

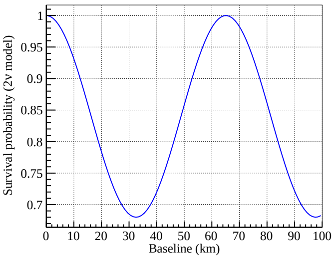


Figure 1.1: Survival probability of 2 MeV neutrinos assuming 2ν model. $\sin^2\theta = 0.32$ and $\Delta m^2 = 7.6 \times 10^{-5}\,\mathrm{eV}^2$

From the figure we can see that if a detector is placed at ~30 km away, the detected neutrino flux will be 30% less than expected.

Now we consider the average neutrino survival probability when the size of the source (such as the Sun) is much larger than the periods of neutrino oscillation. The neutrino oscillation is the interference of two mass eigenstates. If the size of the region where neutrinos are produced or detected is too large (much larger than the wavelength), or if the energy resolution is not good enough, the oscillation in the survival probability is averaged out. Consider the distance follows a Gaussian distribution with expectation μ_L and standard deviation σ_L

$$\begin{aligned} f(l;\,\mu_L,\,\sigma_L) &= \frac{1}{\sqrt{2\pi\sigma_L^2}}\exp\left(-\frac{(l-\mu_L)^2}{\sigma_L^2}\right) \\ \text{average}\{P_{\nu_e\to\nu_e}\} &= \lim_{n\to\infty}\int_{\mu_L-n\sigma_L}^{\mu_l+n\sigma_L} f(l;\,\mu_L,\,\sigma_L)\cdot P_{\nu_e\to\nu_e}(l)\mathrm{d}l \\ &= 1-\frac{1}{2}\sin^2 2\theta\left[1-\exp\left(-2\pi^2\cdot\frac{\sigma_L^2}{L_{\text{osc}}^2}\right)\right] \qquad (1.7) \end{aligned}$$

where l is the distance between the neutrino source and the detector, $f(l)$ is the distribution of l, and here the Gaussian distribution is used. When $\sigma_L \gg L_{\text{osc}}$, the last term is 0, then the survival probability is given by $1-\frac{1}{2}\sin^2 2\theta$.

1.1.2 *History of the* Solar Neutrino Problem

In 1964, Raymond Davis Jr. and John N. Bahcall proposed an experiment to test whether converting hydrogen nuclei to helium nuclei in the Sun is indeed the source of sunlight[15,16]. In 1968, Raymond and his colleagues published their results[17] of the Homestake experiment that only about one third as many radioactive argon atoms as were predicted[18]. This discrepancy soon became known as the *Solar Neutrino Problem*. At that time, people were more worried about the possible mistakes in the theoretical predictions and calibration of the measurement, while the correct idea[19] that had been anticipated by 10 years, was not taken seriously.

Figure 1.2: Raymond David Jr. (left) and John Bahcall in miner's clothing and protective hats. The photograph was taken in 1967 about a mile underground in the Homestake Gold Mine in Lead, South Dakota, USA.

[15] J. N. Bahcall, "Solar neutrinos. I. theoretical", *Physical Review Letters*, vol. 12, no. 11, pp. 300–302, 1964, DOI: `10.1103/PhysRevLett.12.300`

[16] R. Davis, "Solar Neutrinos. II. Experimental", *Physical Review Letters*, vol. 12, no. 11, pp. 303–305, Mar. 1964, DOI: `10.1103/PhysRevLett.12.303`

[17] Davis, Harmer, and Hoffman, "Search for neutrinos from the sun"

[18] Bahcall, Bahcall, and Shaviv, "Present Status of the Theoretical Predictions for the ^{37}Cl Solar-Neutrino Experiment"

[19] Pontecorvo, "Mesonium and antimesonium"

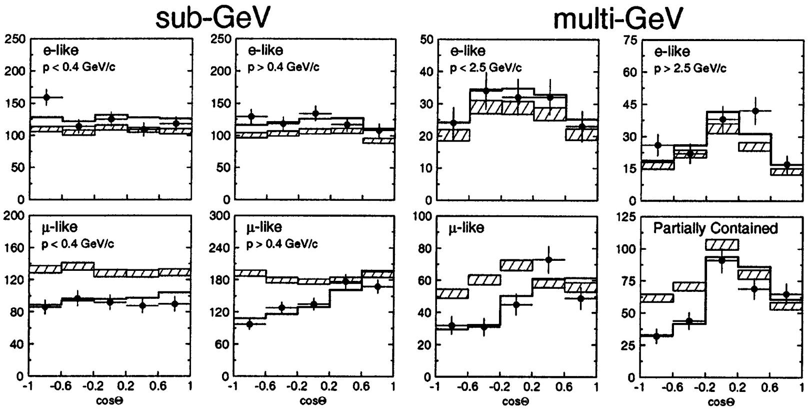


Figure 1.3: Zenith angle distributions of μ-like and e-like events for sub-GeV and multi-GeV data sets of the Super-Kamiokande experiment. Reproduced from The Super-Kamiokande collaboration, "Evidence for oscillation of atmospheric neutrinos", *Physical Review Letters*, vol. 81, no. 8, pp. 1562–1567, 1998, DOI: `10.1103/PhysRevLett.81.1562`, arXiv: `9807003 [hep-ex]`.

In 1989, the deficit reported by Homestake is confirmed by the real time differential detection experiment[20] Kamiokande-II in 1989[21], and later as Kamiokande-III in 1996[22]. Even worse, the ratio of the measured to the predicted flux did not agree with that of Homestake. Physicists started to consider that maybe they did not understand how neutrinos behave when they travel astronomical distances. In the following decade (1990s), two new solar neutrino detectors were constructed, i.e. GALLEX[23], SAGE[24], and they also confirmed the deficit. Another experiment is Super-Kamiokande. It is aimed at searching for proton-decays, while it is also sensitive to solar neutrinos, and it again confirmed the solar neutrino deficit in 1998[25]. What's more, it gave the first definite evidence of neutrino oscillation using atmospheric neutrinos through the zenith angle dependent deficit of muon neutrinos in the same year[26,27]. At the same time, apart from the clear evidence of neutrino oscillations, the *Standard Solar Model* was also being improved. One groundbreaking achievement is the measurement of sound speeds in the Sun in 1996[28] using periodic fluctuations observed in ordinary light from the surface of the Sun, and the measurement matches the predicted sound speeds using the *Standard Solar Model* at that time within 0.1% precision .

[20] real time: the interaction is recorded at the time when the interaction happened. differential detection: the distribution of the recoil electrons are measured. These are different from the radio-chemical experiments.

[21] K. Hirata et al., "Observation of 8B solar neutrinos in the Kamiokande-II detector.", *Physical review letters*, vol. 63, no. 1, pp. 16–19, 1989, DOI: `10.1103/PhysRevLett.63.16`

[22] The Kamiokande collaboration, "Solar neutrino data covering solar cycle 22", *Physical Review Letters*, vol. 77, no. 9, pp. 1683–1686, 1996, DOI: `10.1103/PhysRevLett.77.1683`

[23] The GALLEX collaboration, "Solar neutrinos observed by GALLEX at Gran Sasso", *Physics Letters B*, vol. 285, no. 4, pp. 376–389, 1992, DOI: `10.1016/0370-2693(92)91521-A`

[24] The SAGE collaboration, "Measurement of the solar neutrino capture rate with gallium metal", *Physical Review C - Nuclear Physics*, vol. 60, no. 5, p. 32, 1999, DOI: `10.1103/PhysRevC.60.055801`, arXiv: `0901.2200`

[25] Y. Fukuda et al., "Measurements of the solar neutrino flux from super-kamiokande's first 300 days", *Physical Review Letters*, vol. 81, no. 6, pp. 1158–1162, 1998, DOI: `10.1103/PhysRevLett.81.1158`, arXiv: `9805021 [hep-ex]`

[26] The Super-Kamiokande collaboration, "Evidence for oscillation of atmospheric neutrinos", *Physical Review Letters*, vol. 81, no. 8, pp. 1562–1567, 1998, DOI: `10.1103/PhysRevLett.81.1562`, arXiv: `9807003 [hep-ex]`

[27] The MARCO collaboration, "Measurement of the atmospheric neutrino-induced upgoing muon flux using MACRO", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 434, no. 3-4, pp. 451–457, 1998, DOI: `10.1016/S0370-2693(98)00885-5`

[28] J. Christensen-Dalsgaard et al., "The Current State of Solar Modeling", *Science*, vol. 272, no. 5266, pp. 1286–1292, May 1996, DOI: `10.1126/science.272.5266.1286`

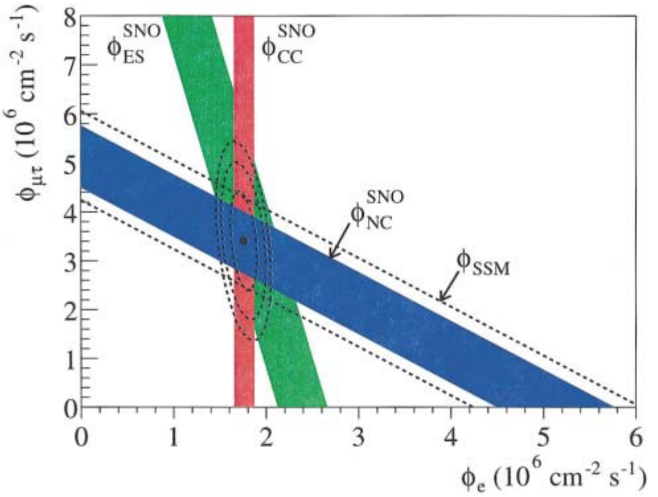


Figure 1.4: Flux of ^{8}B solar neutrinos which are μ or τ flavor vs flux of electron neutrinos deduced from the three neutrino reactions in SNO. The diagonal bands show the total ^{8}B flux as predicted by the *Standard Solar Model* (dashed lines) and that measured with the NC reaction in SNO (solid band). The intercepts of these bands with the axes represent the $\pm 1\sigma$ errors. The bands intersect at the fit values for Φ_e and $\Phi_{\mu\tau}$, indicating that the combined flux results are consistent with neutrino flavor transformation assuming no distortion in the ^{8}B neutrino energy spectrum. Reproduced from "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory".

Another essential piece of the jigsaw was also found in 1980s. Although physicists knew the neutrino oscillations can be described with the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix[29] as early as 1960s, the predicted flavor transformation was not enough to explain the deficit. In 1985, based on the work of Wolfenstein[30], it was proposed a new mechanism through which full flavor transformation can be achieved even with small mixing angle[31]. It soon became famous and known as the Mikheyev–Smirnov–Wolfenstein (MSW) effect.

Although physicists were already almost sure that the *Solar Neutrino Problem* could be explained by neutrino oscillations and the MSW effect, they still need a smoking gun. The transformed neutrinos need to be found and the sum need to be proved to match the prediction. To find the smoking gun, in 1985 Chen proposed a detector based on deuteron in heavy water[32]. It was constructed at the Sudbury Neutrino Observatory (SNO), Canada. Three years later another detector based on Boron called BOREX was proposed by Sandip Pakvasa and Raju Raghavan[33], but that was never materialized. Instead, a smaller version was made and that is the BOREXino experiment[34]. On June 18, 2001 at 12:15 PM (eastern daylight time) the SNO collaboration made a dramatic announcement: they had solved the solar neutrino mystery[35]. The fluxes of electron neutrinos and μ/τ neutrinos can be extracted by combining the results using neutral current interaction and elastic scattering interaction channels, and it was found that the sum of three neutrinos matched the *Standard Solar Model* prediction[36].

The complete picture of the *Solar Neutrino Problem* is obtained including the Kamland results. After SNO result, multiple choices of answers to the *Solar Neutrino Problem* are consistent with experimental results, and they differ in the magnitude of neutrino mass splitting and mixing angles. In 2003, the KamLAND published their result on the evidence of reactor antineutrino disappearance and eliminated all other choices except the Large-Mixing-Angle (LMA) solution.

[29] Pontecorvo, "Mesonium and antimesonium"; Pontecorvo, "Inverse beta processes and nonconservation of lepton charge"; Maki, Nakagawa, and Sakata, "Remarks on the Unified Model of Elementary Particles"

[30] Wolfenstein, "Neutrino oscillations in matter"

[31] Mikheyev and Smirnov, "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos"

[32] H. H. Chen, "Direct Approach to Resolve the Solar-Neutrino Problem", *Physical Review Letters*, vol. 55, no. 14, pp. 1534–1536, Sept. 1985, DOI: `10.1103/PhysRevLett.55.1534`

[33] R. S. Raghavan and S. Pakvasa, "Probing the nature of the neutrino: The boron solar-neutrino experiment", *Physical Review D*, vol. 37, no. 4, pp. 849–857, Feb. 1988, DOI: `10.1103/PhysRevD.37.849`

[34] In Italian -ino means little

[35] See Queens University's webpage "SNO First Results" for the announcement.

[36] The SNO collaboration, "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory"

1.1.3 Neutrino oscillation in Vacuum

As written in Section 1.1.1, the neutrino oscillation is a quantum mechanics phenomena. Following the notation used in PDG[37], considering 3 active neutrino scenario, denote the mixing matrix as

$$|\nu_l\rangle = \sum_j U_{lj}^* \left|m_j\right\rangle, \, l = e, \, \mu, \, \tau \tag{1.8}$$

where $|\nu_l\rangle$ are the flavor eigenstates and $\left|m_j\right\rangle$ are the mass eigenstates, then the transition probability from flavor l to l' is

$$\begin{aligned} P_{ll'}(L) &= \left| \langle \nu_{l'} | e^{-iHt} | \nu_l \rangle \right|^2 &(1.9)\\ &= \left| \sum_k U_{lk}^* U_{l'k} e^{-iE_k t} \right|^2 &(1.10)\\ &= \sum_{j,k} U_{lk}^* U_{l'k} U_{lj} U_{l'j}^* \exp\left(-i2\pi \frac{L}{L_{kj}^{\text{osc}}} \right) &(1.11) \end{aligned}$$

where $L_{kj}^{\text{osc}} = 4\pi E/\Delta m_{kj}^2$ is the oscillation length, U is the PMNS matrix, and it is factorized as

$$U = \begin{pmatrix} 1 & & \\ & c_{23} & s_{23} \\ & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & & s_{13}e^{-i\delta_{CP}} \\ & 1 & \\ -s_{13}e^{i\delta_{CP}} & & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & \\ -s_{12} & c_{12} & \\ & & 1 \end{pmatrix}$$

where s_{ij} and c_{ij} are used to denote $\sin\theta_{ij}$ and $\cos\theta_{ij}$, respectively, and δ_{CP} is the CP-violation phase angle.

1.1.4 Neutrino oscillation in Matter

Here for introduction purpose we give results in the 2ν framework. For 3ν framework, readers can refer to Section 10.1, page 160.

Neutrino mass eigenstates are modified when matter is present:

$$\begin{aligned} \Delta m_m^2 &= \Delta m^2 \cdot \sqrt{(\cos 2\theta - \varepsilon)^2 + \sin^2 2\theta} &(1.12)\\ \tan 2\theta^m &= \frac{\tan 2\theta}{1 - \dfrac{\varepsilon}{\cos 2\theta}} &(1.13)\\ \varepsilon &= \frac{2 \cdot E_\nu \cdot V_{CC}}{\Delta m^2} &(1.14)\\ V_{CC} &= \sqrt{2} G_F \cdot n_e &(1.15) \end{aligned}$$

where V_{CC} is the matter potential, G_F is the Fermi factor, n_e is the electron density, and Δm_m^2 and θ^m are the mass splitting and mixing angle in matter, respectively.

When $\varepsilon = \cos 2\theta$, or when

$$n_e = \frac{\Delta m^2 \cos 2\theta}{2\sqrt{2} G_F \cdot E_\nu},$$

, the mixing angle reaches the maximum, and the mass splitting reaches the minimum. This is called the MSW resonance:

$$\begin{aligned} \sin 2\theta_R^m &= 1 &(1.16)\\ \Delta \left(m_{mR}\right)^2 &= \Delta m^2 \sin 2\theta &(1.17) \end{aligned}$$

[37] K. Nakamura and S. T. Petcov*Section 14.7 Neutrino Masses, Mixing, and Oscillations* Particle Data Group, "Review of Particle Physics" 2017, page 261

At the resonance, there is possibility that the state of the neutrino transits between $|\nu_1^m\rangle$ and $|\nu_2^m\rangle$, and the crossing probability is given by[38]

$$P_c = \frac{\exp\left(-\frac{\pi}{2}\gamma_R F\right) - \exp\left(-\frac{\pi}{2}\gamma_R \frac{F}{\sin^2\theta}\right)}{1 - \exp\left(-\frac{\pi}{2}\gamma_R \frac{F}{\sin^2\theta}\right)} \tag{1.18}$$

where γ_R is the adiabaticity parameter at the resonance

$$\gamma_R = \frac{\Delta m^2 \sin^2 2\theta}{2E\cos 2\theta\,|\mathrm{d}\ln n_e/\mathrm{d}x|_R},$$

parameter F depend on the electron density profile. For the Sun, it is calculated to be[39]

$$P_c = (10^{-9}\text{–}10^{-7})\left(\frac{E}{10\ \mathrm{MeV}}\right)^2 \tag{1.19}$$

so it's negligible for solar neutrinos.

Ignoring the crossing transit, which is called the adiabatic approximation, the survival probability is given by[40]

$$P^{\text{adiabatic}}_{\nu_e\to\nu_e}(L) = \frac{1}{2} + \frac{1}{2}\cos 2\theta^m_{\mathrm{p}}\cos 2\theta^m_{\mathrm{d}} + \frac{1}{2}\sin 2\theta^m_{\mathrm{p}}\sin 2\theta^m_{\mathrm{d}}\cos\left(\int_0^L \frac{\Delta m^2_m(x)}{2E}\mathrm{d}x\right) \tag{1.20}$$

where the subscript `p` denotes the production and `d` denotes the detection.

If the production or the detection region is large enough, or if the detector energy resolution is not good enough to see the oscillation patter, the integration of cos in Equation 1.20 is averaged out:

$$P^{\text{adiabatic}}_{\nu_e\to\nu_e}(L) = \frac{1}{2} + \frac{1}{2}\cos 2\theta^m_{\mathrm{p}}\cos 2\theta^m_{\mathrm{d}} \tag{1.21}$$

1.1.5 MSW transition region problem

Let us consider the situation where neutrinos are produced in dense environment and detected in vacuum, such as in the Sun. The mixing angle at the detection site is the same as the mixing angle in vacuum. For the mixing angle at the production site, it follows

$$\cos 2\theta^m = \frac{\cos 2\theta - \epsilon}{\sqrt{(\cos 2\theta - \epsilon)^2 + (\sin 2\theta)^2}} \tag{1.22}$$

$$\epsilon = \frac{2\cdot E_\nu \cdot \sqrt{2}G_F n_e^{\mathrm{p}}}{\Delta m^2} \tag{1.23}$$

where E_ν is the neutrino energy, G_F is the Fermi factor and n_e^{p} is the average electron density in the production region. From the above

[38] Giunti and Kim, *Fundamentals of neutrino physics and astrophysics* Equation (9.96), page 338

[39] P. C. de Holanda, W. Liao, and A. Y. Smirnov, "Toward precision measurements in solar neutrinos", *Nuclear Physics B*, vol. 702, no. 1-2, pp. 307–332, 2004, DOI: `10.1016/j.nuclphysb.2004.09.027`, arXiv: `0404042 [hep-ph]` Equation (27)

[40] Giunti and Kim, *Fundamentals of neutrino physics and astrophysics* Equation (9.83), page 335

equations we can see the two extreme situation where the

$$E_\nu^c = \frac{\cos 2\theta \Delta m^2}{2\sqrt{2} G_F n_e^{\mathrm{p}}} \tag{1.24}$$

$$\text{Vacuum regime } E_\nu \ll E_\nu^c: \quad \cos 2\theta^m \to \cos 2\theta \tag{1.25}$$

$$P_{ee}^{\mathrm{day}} \to c_{13}^4 \left(1 - \frac{1}{2}\sin^2 2\theta\right) \tag{1.26}$$

$$\text{Matter regime } E_\nu \gg E_\nu^c: \quad \cos 2\theta^m \to -1 \tag{1.27}$$

$$P_{ee}^{\mathrm{day}} \to c_{13}^4 s^2 \tag{1.28}$$

For ^{8}B solar neutrinos, the average electron density weighted according to the production location is $n_e^{\mathrm{p}} = 89.6\ \mathrm{mol/cm}^3$[41], so $E_\nu^c = 5.5$ MeV ($\Delta m^2 = 7.53 \times 10^{-5}\ \mathrm{eV}^2$). The region around this energy is called the MSW transition region.

Currently the value of Δm^2 determined from KamLAND reactor anti-neutrino results[42] and that determined from global fit of solar neutrino experiments[43] are inconsistent with 2σ level, and the comparison of the predicted solar neutrino survival probability based on two values is shown in Figure 1.5.

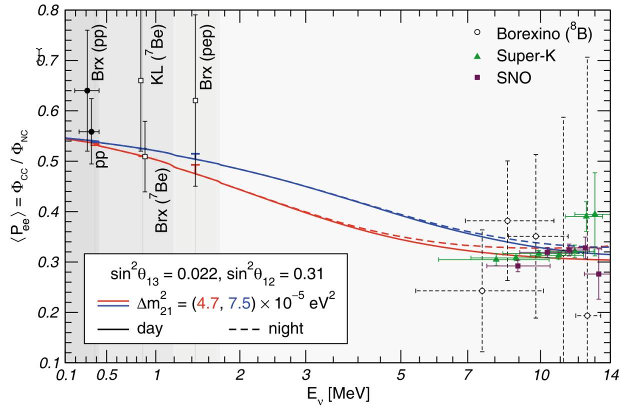


Figure 1.5: Comparison of solar neutrino survival probability based on Δm^2 determined from KamLAND and from solar neutrino experiments. Reproduced from M. Maltoni and A. Yu. Smirnov, "Solar neutrinos and neutrino physics", *The European Physical Journal A*, vol. 52, no. 4, p. 87, Apr. 2016, DOI: `10.1140/epja/i2016-16087-0`, arXiv: `1507.05287`.

The discrepancy can be reconciled with new physics. Various models have been proposed and are shown in Figure 1.6. More details can be found in[44].

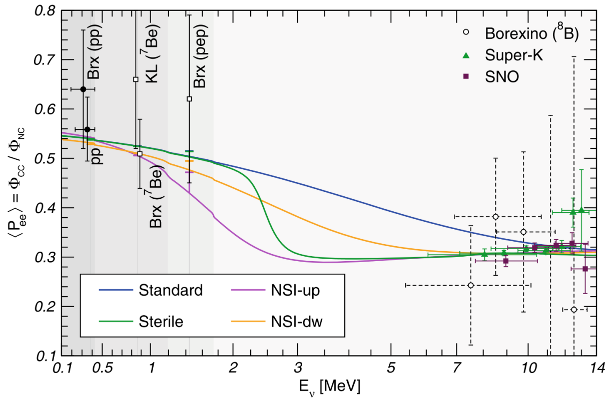


Figure 1.6: Comparison of solar neutrino survival probability based on Δm^2 determined from KamLAND and from solar neutrino experiments. Reproduced from M. Maltoni and A. Yu. Smirnov, "Solar neutrinos and neutrino physics", *The European Physical Journal A*, vol. 52, no. 4, p. 87, Apr. 2016, DOI: `10.1140/epja/i2016-16087-0`, arXiv: `1507.05287`.

---

[41] F. Vissani, "Joint analysis of Borexino and SNO solar neutrino data and reconstruction of the survival probability", pp. 1–12, Sept. 2017, arXiv: `1709.05813`

[42] The KamLAND collaboration, "Reactor on-off antineutrino measurement with KamLAND", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 88, no. 3, pp. 1–10, 2013, DOI: `10.1103/PhysRevD.88.033001`, arXiv: `1303.4667 [hep-ex]`

[43] The Super-Kamiokande collaboration, "Solar neutrino measurements in Super-Kamiokande-IV", *Physical Review D*, vol. 94, no. 5, pp. 1–33, 2016, DOI: `10.1103/PhysRevD.94.052010`, arXiv: `1606.07538`

[44] M. Maltoni and A. Yu. Smirnov, "Solar neutrinos and neutrino physics", *The European Physical Journal A*, vol. 52, no. 4, p. 87, Apr. 2016, DOI: `10.1140/epja/i2016-16087-0`, arXiv: `1507.05287`

1.2 Neutrino astrophysics and solar physics

In this section, the neutrino astrophysics is introduced. We briefly reviewed the neutrino astrophysics, then discuss the *Standard Solar Model*, and at last using neutrino as the probe to study the Sun. More contents can be found in[45,46,47,48]

1.2.1 Neutrino astrophysics

Astrophysics is the branch of astronomy that studies the nature of astronomical objects with the principle of physics and chemistry. Neutrino astrophysics uses neutrinos as the probe. Neutrinos are produced in the core region of the Sun and other starts, they are produced by the super novae explosion, and also produced during the Big Bang. Because the cross-section of neutrino interactions is relatively low, neutrinos is an ideal probe to study astronomy distant in space and in time, because it conducts little interactions during propagation. Recently the IceCube collaboration announced that correlation of the neutrino with the flare of TXS 0506+056 is statistically significant at the level of 3 standard deviations)[49], and this indicates that blazers may be one of the long-sought sources of very-high-energy cosmic rays.

1.2.2 *The* Standard Solar Model

The *Standard Solar Model*s are the model of the Sun calibrated to match present day surface properties of the Sun, including the present-day solar luminosity $L_\odot$, radius $R_\odot$, and surface metal [50] to hydrogen abundance ratio $(Z/X)_\odot$. The calibration is done by adjusting the mixing length parameter (α_{MLT}) and the initial helium and metal mass fractions (Y_{ini} and Z_{ini} respectively).

The *Standard Solar Model*s have played a fundamental role in discovering the neutrino flavor oscillations. Neutrinos on the other hand also confirmed the *Standard Solar Model* by probing the core region of the Sun. After several decades, development of the helioseismology made it possible to construct an accurate picture of the solar interior. The agreement between *Standard Solar Model*s and sub percent precision helioseismic results was astonishingly good, and this provide both strong support to stellar evolution theory and to neutrino oscillations[51].

Later on a new generation of *Standard Solar Models*[52] was proposed. The new models are believed to be improved in several aspects. For these new models, all determinations of the solar metallicity based on the new generation of spectroscopic studies yield a solar metallicity that is substantially lower than the older spectroscopic results, in particular, for the volatile and most abundant C, N, and O. However, they have brought about a series of problems because they fail to reproduce all helioseismic probes of solar properties and this is called the Solar Abundance Problem.

[45] A. Gallo Rosso et al., "Introduction to neutrino astronomy", *The European Physical Journal Plus*, vol. 133, no. 7, p. 267, 2018, DOI: `10.1140/epjp/i2018-12143-6`

[46] N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: `10.3847/1538-4357/835/2/202`, arXiv: `1611.09867`

[47] A. Serenelli, "Alive and well: A short review about standard solar models", *The European Physical Journal A*, vol. 52, no. 4, p. 78, Apr. 2016, DOI: `10.1140/epja/i2016-16078-1`, arXiv: `1601.07179`

[48] J. N. Bahcall and C. Peña-Garay, "Solar models and solar neutrino oscillations", *New Journal of Physics*, vol. 6, no. 04, pp. 1–19, 2004, DOI: `10.1088/1367-2630/6/1/063`, arXiv: `0310030 [astro-ph]`

[49] T. IceCube et al., "Multi-messenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A", *Science*, eaat1378, July 2018, DOI: `10.1126/science.aat1378`, arXiv: `1807.08816`

[50] Astronomers refer to all elements other than hydrogen and helium as 'metals'.

[51] Christensen-Dalsgaard et al., "The Current State of Solar Modeling"

[52] Asplund et al., "The Chemical Composition of the Sun"

1.2.3 *Neutrino as probe of Solar Physics*

Neutrinos can be used to solve the Solar Abundance Problem. Neutrinos are produced when protons are converted to heliums in the core region of the Sun. They can escape the Sun without interactions and thus preserving the information of the interior, such as the interaction rate. The High Metallicity models and Low Metallicity models predict different fluxes of neutrinos, especially those produced from the CNO-cycle process. A 5% precision CNO solar neutrino flux measurement would provide a moderate evidence in favor of one of two models[53].

The predicted solar neutrino fluxes are summarized in Table 1.1, and the energy spectra are shown in Figure 1.7.

Table 1.1: Predicted fluxes of *pp*, ^{7}Be (862 +384 keV), *pep* and CNO solar neutrinos assuming the MSW-LMA oscillation parameters. Unit: [$\text{cm}^{-2}\text{s}^{-1}$]. Reference: N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: `10.3847/1538-4357/835/2/202`, arXiv: `1611.09867`

Solar ν	B16(GS98)-HZ	B16(AGSS09)-LZ
pp	$5.98\,(1 \pm 0.006) \times 10^{10}$	$6.03\,(1 \pm 0.005) \times 10^{10}$
^{7}Be	$4.93\,(1 \pm 0.06) \times 10^{9}$	$4.50\,(1 \pm 0.06) \times 10^{9}$
pep	$1.44\,(1 \pm 0.009) \times 10^{8}$	$1.46\,(1 \pm 0.009) \times 10^{8}$
CNO	$4.88\,(1 \pm 0.11) \times 10^{8}$	$3.51\,(1 \pm 0.10) \times 10^{8}$
^{8}B	$5.46\,(1 \pm 0.12) \times 10^{6}$	$4.50\,(1 \pm 0.12) \times 10^{6}$

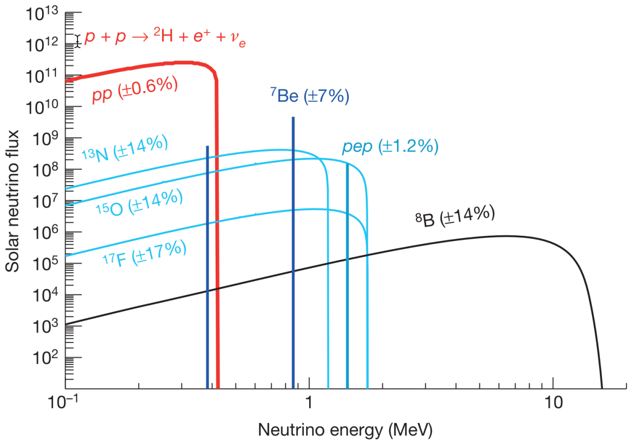


Figure 1.7: Solar neutrino spectra. Reproduced from The Borexino collaboration, "Neutrinos from the primary proton–proton fusion process in the Sun", *Nature*, vol. 512, no. 7515, pp. 383–386, Aug. 2014, DOI: `10.1038/nature13702`, arXiv: `1508.05379`.

[53] J. Bergström et al., "Updated determination of the solar neutrino fluxes from solar neutrino data", *Journal of High Energy Physics*, vol. 2016, no. 3, p. 132, Mar. 2016, DOI: `10.1007/JHEP03(2016)132`, arXiv: `1601.00972`

1.3 The Borexino and JUNO experiments

In this section, the Borexino and JUNO experiments are introduced. The physics results achieved by Borexino are reviewed, and the physics programs and their significance of JUNO are presented. The design of the Borexino and JUNO detectors are described dedicatedly in Section 2.1 and 9.1, respectively.

1.3.1 The Borexino experiment

Borexino, the little (-ino) BORon solar neutrino EXperiment, started data taking on May 16, 2007. In 2010 May, the data taking was stopped and Borexino conduct a purification and calibration campaign. After that, Borexino started another round of data taking since 2011 December till now. The period between 2007 May and 2010 May is called Phase-I and the period between 2011 December and now is called Phase-II.

In Phase-I, Borexino achieved vast number of groundbreaking results. In 2011, it reported the first precision measurement of ^{7}Be solar neutrino interaction rate[54]. One year later it reported the first evidence of *pep* solar neutrinos[55]. In 2014 Borexino reported the first real time detection of *pp* solar neutrinos[56]. Besides, Borexino also reported results on various topics. In 2010, it reported 3 MeV detection threshold measurement of ^{8}B solar neutrinos[57] and detection of geo-neutrinos[58], etc. In Phase-II data, Borexino continued to contribute. For example, recently it published the comprehensive measurement of *pp*-chain solar neutrinos[59]. This is the first time we use a unique detector and dataset to measure all *pp*-chain solar neutrinos.

1.3.2 The JUNO experiment

JUNO, the Jiangmen Underground Neutrino Observatory, is an under-construction reactor anti-neutrino experiment and is expected to start data-taking in 2021. It is aimed at determining the neutrino mass ordering and at same time has potential in vast number of physics goals, such as percent-level precision measurement of neutrino oscillation mixing parameters, detection of supernova neutrinos, diffusive supernova neutrino backgrounds, solar neutrinos, atmospheric neutrinos, geo-neutrinos, nucleon decays etc[60].

1.4 Future searches for solar neutrinos

In this section, future experiments that are sensitive to solar neutrinos are introduced, together with the prospects of solar neutrino physics. SNO+, Hyper-K, DUNE and Argo are discussed.

1.4.1 The SNO+ experiment

SNO+ uses the refurbished facility of the SNO experiment[61]. The

---

[54] The Borexino collaboration, "Precision measurement of the Be7 solar neutrino interaction rate in Borexino", *Physical Review Letters*, vol. 107, no. 14, pp. 1–5, 2011, DOI: `10.1103/PhysRevLett.107.141302`, arXiv: `1104.2150`

[55] The Borexino collaboration, "First Evidence of *pep* Solar Neutrinos by Direct Detection in Borexino", *Physical Review Letters*, vol. 108, no. 5, p. 051302, Feb. 2012, DOI: `10.1103/PhysRevLett.108.051302`, arXiv: `1110.3230`

[56] The Borexino collaboration, "Neutrinos from the primary proton–proton fusion process in the Sun", *Nature*, vol. 512, no. 7515, pp. 383–386, Aug. 2014, DOI: `10.1038/nature13702`, arXiv: `1508.05379`

[57] The Borexino collaboration, "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector", *Physical Review D*, vol. 82, no. 3, p. 033006, Aug. 2010, DOI: `10.1103/PhysRevD.82.033006`, arXiv: `0808.2868`

[58] The Borexino collaboration, "Observation of geo-neutrinos", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 687, no. 4-5, pp. 299–304, 2010, DOI: `10.1016/j.physletb.2010.03.051`, arXiv: `1003.0284`

[59] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

[60] F. An et al., "Neutrino physics with JUNO", *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, no. 3, p. 030401, 2016, DOI: `10.1088/0954-3899/43/3/030401`, arXiv: `1507.05613`

[61] V. f. t. S. c. Lozza, "Scintillator phase of the SNO+ experiment" Jan. 2012, DOI: `10.1088/1742-6596/375/1/042050`, arXiv: `1201.6599`

main goal of the experiment is to test if a neutrino is its own antiparticle by measuring the half life of the neutrino-less double beta decay process of the ^{130}Te isotope. SNO+ filled the center detector with liquid scintillator without ^{130}Te first, then replace it with the liquid scintillator doped with ^{130}Te. During this period, the SNO+ experiment has sensitivity to *pep* and CNO solar neutrinos.

1.4.2 The Hyper-Kamioka experiment

The Hyper-Kamioka experiment is a giant water based Cherenkov detector. Similar to the Super-Kamioka experiment, it has the potential to measure ^{8}B solar neutrinos. It is expect to be able to claim observation of day-night asymmetry for more than 4 σ significance with 2 years of data and to claim evidence of MSW upturn with 10 years of data for 3(5) σ significance with 4.5(3.5) MeV detection threshold[62].

1.4.3 The DUNE experiment

The Deep Underground Neutrino Experiment (DUNE) will use the liquid Argon time projection chamber technique, and it has potential to measure the ^{8}B solar neutrino flux. Considering that the target mass of DUNE is designed to be 40 kilo-tons, within a reasonable time it can collect an exposure of 100 kton-year and thus more than 10^5 signal events above 5 MeV. With such exposure, DUNE is expected to be able to claim the day-night asymmetry with more than 10 σ significance, and might confirm the discrepancy on Δm^2_{21} between solar experiments and reactor experiments at more than 5 σ assuming current best fit values. It will be able to measure the ^{8}B solar neutrino flux within 2.5% precision and *hep* solar neutrinos within 11% precision[63].

1.4.4 The ARGO experiment

The Argon Observatory (ARGO) experiment is the next generation of liquid Argon based dark matter experiment. Its target mass is 300 ton. ARGO has the advantage in detecting CNO solar neutrinos that it will be without carbon and thus cosmogenic ^{11}C background, and it will be optimized for detecting low energy events. It is expected that with 400 ton × years exposure, precisions at the level of around 2%, 10% and 10% for the ^{7}Be, *pep* and CNO solar neutrino interaction rates, respectively, can be achieved[64].

[62] S.-h. Seo, "Physics Potentials of the Hyper-Kamiokande Second Detector in Korea", pp. 2–6, 2018, arXiv: `1811.06682v1`

[63] F. Capozzi et al., "DUNE as the Next-Generation Solar Neutrino Experiment", pp. 1–19, Aug. 2018, arXiv: `1808.08232`

[64] D. Franco et al., "Solar neutrino detection in a large volume double-phase liquid argon experiment", *Journal of Cosmology and Astroparticle Physics*, vol. 2016, no. 8 2016, DOI: `10.1088/1475-7516/2016/08/017`, arXiv: `1510.04196`

2

Solar neutrinos with Borexino

In this chapter, an introduction to the Borexino experiment is given. Solar neutrinos released with fusion reactions in the core of the Sun, unlike the high energy gamma rays, can pass through the solar plasma immediately and make it possible to study the core of the Sun, and Borexino has demonstrated that it is possible to measure the individual branches of the *pp*-chain above 190 keV[1,2,3,4]. Moreover, the determination of the abundances of heavy elements in the core region is the key to solve the solar abundance problem[5]. Using new and old abundances leads to different predictions for the neutrino flux[6]. This is especially true for the neutrinos from the CNO cycle, that are to date not yet measured.

Solar neutrinos also provide a unique probe to study the matter effect of neutrino oscillations thanks to the high electron density in the Sun[7]. A wide energy range of solar neutrino spectra have been observed covering both the adiabatic (the *pp*, ^{7}Be and *pep*) and vacuum regime (high energy part of ^{8}B) of MSW effect. Borexino can even uniquely reach the transition region by the detection of ^{8}B neutrino with the lowest threshold as low as 3 MeV[8,9].

After Phase-I, more data has been collected between December 14th, 2011 and May 21st, 2016, which corresponds to an exposure of 1291.51 days × 71.3t (Phase-II). With larger exposure and better radio-purity levels, as well as improved detector response function and improved Monte Carlo[10], we are expected to improve the precision of fluxes of *pp*, ^{7}Be and *pep* solar neutrinos fluxes and to reach the same level of upper limit for CNO neutrinos with weaker assumptions. In this chapter, we report the analysis techniques adopted to suppress the background and to determine the response functions, the methods used for the estimation of the systematic uncertainties related to detector response and the expected uncertainties on neutrino interaction rate measurements.

This chapter is organized as in the following: Section 2.1 introduces the components of the Borexino detector. Section 2.2 summarizes the unprecedented radio-purities achieved by Borexino. Section 2.3 reviews the signal and backgrounds of Borexino Phase-I and Phase-II data. Section 2.4 presents the list of backgrounds. Section 2.5 reviews the background reduction analysis methods. At last, Section 2.6 presents the offline softwares.

[1] The Borexino collaboration, "Precision measurement of the Be7 solar neutrino interaction rate in Borexino", *Physical Review Letters*, vol. 107, no. 14, pp. 1–5, 2011, DOI: 10.1103/PhysRevLett.107.141302, arXiv: 1104.2150

[2] The Borexino collaboration, "First Evidence of *pep* Solar Neutrinos by Direct Detection in Borexino", *Physical Review Letters*, vol. 108, no. 5, p. 051302, Feb. 2012, DOI: 10.1103/PhysRevLett.108.051302, arXiv: 1110.3230

[3] The Borexino collaboration, "Neutrinos from the primary proton–proton fusion process in the Sun", *Nature*, vol. 512, no. 7515, pp. 383–386, Aug. 2014, DOI: 10.1038/nature13702, arXiv: 1508.05379

[4] The Borexino collaboration, "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector", *Physical Review D*, vol. 82, no. 3, p. 033006, Aug. 2010, DOI: 10.1103/PhysRevD.82.033006, arXiv: 0808.2868

[5] N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: 10.3847/1538-4357/835/2/202, arXiv: 1611.09867

[6] ibid.

[7] F. Vissani, "Joint analysis of borexino and sno solar neutrino data and reconstruction of the survival probability", *Nuclear Physics and Atomic Energy*, vol. 18, no. 4, pp. 303–312, 2017, DOI: 10.1207/s15327868ms0402_3, arXiv: 1709.05813

[8] The Borexino collaboration, "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector"

[9] The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure", vol. 016, pp. 1–13, Sept. 2017, arXiv: 1709.00756

[10] The Borexino collaboration, "The Monte Carlo simulation of the Borexino detector", *Astroparticle Physics*, vol. 97, pp. 136–159, Jan. 2018, DOI: 10.1016/j.astropartphys.2017.10.003, arXiv: 1704.02291

2.1 *The Borexino detector*

The Borexino detector is a liquid scintillator based calorimeter. Its structure is shown schematically in Figure 2.1[11]. In the most inner part there is the scintillator target, an organic solution. Its solvent is PC (pseudocumene, 1,2,4-trimethylbenzene $C_6H_3(CH_3)_3$), and the solute is PPO (2,5-diphenyloxazole, $C_{15}H_{11}NO$) at a concentration of 1.5 g/L (0.17% by weight). Its mass is around 278 tons. The neutrinos can be detected via the elastic scattering channel on electrons, and the anti-neutrinos can be detected via the inverse beta decay on the hydrogen atoms. The scintillator target is surrounded by quenched liquid scintillator which shields it from the gammas rays from the natural radioactivity in the PMTs and stainless steel sphere (SSS), and they are separated by the inner vessel (IV), a 125 μm specially treated ultra-low-radioactivity nylon vessel. Another nylon vessel, the outer vessel, together with the inner vessel works both as the barriers against radon atoms diffusing inward from outer part of the detector. The radius of the inner vessel, outer vessel and the stainless steel sphere are 4.25 m, 5.5 m and 6 m, respectively. Outside the stainless steel sphere there is the water Cherenkov detector playing the role of active muon veto and passive shield of the environmental radiation. 3800 meters water equivalent rock provides passive shield to further suppress muon and cosmogenic backgrounds.

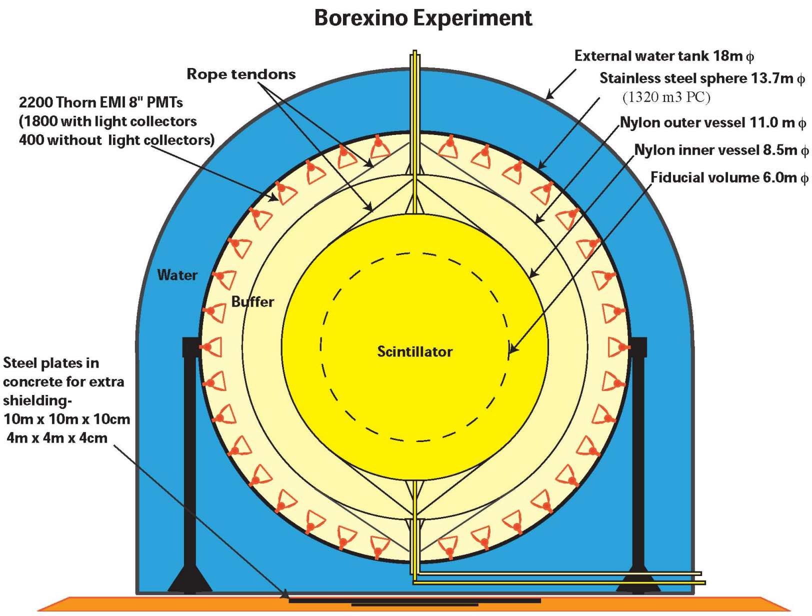


Figure 2.1: Schematic drawing of Borexino detector.

The scintillation light is collected by 2212 PMTs installed on the inner surface of the stainless steel sphere. The number of photonelectrons collected by Borexino detector is around 500 p.e./MeV/2000 PMTs and the region of interests extends from tens of keV up to a few MeV, so PMTs work in single p.e. regime mostly. The multiple hit probability is of the order of 10% for a 1 MeV energy deposition event in the detector center. For each hit, two signals, one for the energy

[11] The Borexino collaboration, "The Borexino detector at the Laboratori Nazionali del Gran Sasso", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 600, no. 3, pp. 568–593, 2009, DOI: `10.1016/j.nima.2008.11.076`, arXiv: `0806.2400`

and one for the time measurements, are produced by the front-end circuit[12].

2.2 *Unprecedented Radio-purity*

Detection of solar neutrinos and measurement of their fluxes are not easy tasks. Solar neutrinos are detected via the elastic scattering channel. To achieve the expected sensitivity and precision, it is of prime importance to purify the liquid scintillator and remove the natural radio-active elements in the liquid scintillator. Borexino as the pioneer of low background control has achieved unprecedented radioactive-purity level. In Phase-I from 2008 to 2010, the concentration of ^{238}U and ^{232}Th in the liquid scintillator are $(5.3 \pm 0.5) \times 10^{-18}$ g/g and $(3.8 \pm 0.8) \times 10^{-18}$ g/g, respectively[13]. In Phase-II, their concentrations are even reduced to $< 9.4 \times 10^{-20}$ g/g (95% C.L.) and $< 5.7 \times 10^{-19}$ g/g (95% C.L.). Such radio-purity levels were achieved by the extensive purification campaign in 2011.

2.3 *Comparison of signal and backgrounds*

In this section we compare the signals of the low and high energy branches of the solar neutrino chain to the background in the conditions of the Phase-I and Phase-II data taking periods. The backgrounds of low energy branches in Phase-II will be explained in more detail in Section 2.4 on page 38. In this section we compare

Borexino is a liquid scintillator detector. It can detect the energy deposition of the recoil electron in the elastic scattering process of a neutrino on an electron. By counting the number of times of detection, it measures the solar neutrino fluxes. Decay of natural radioactive nuclei in the liquid scintillator or in the outer components of the detector, or that of short lived cosmogenic isotopes, produce α, β or γ particles, and the energy deposition of these decay products can mimic the recoil electron and thus forms the backgrounds. Borexino differentiates the signal and background using various techniques, and they will be explained in the next section.

2.3.1 *pp, ^{7}Be, pep and CNO analysis*

The *pp*, ^{7}Be, *pep* and CNO solar neutrino interaction rates are extracted by fitting the energy spectrum with the help of various background reduction techniques. The comparison of energy spectra obtained in Phase-I[14] and Phase-II[15] is shown in Figure 2.2. Unlike the ^{8}B solar neutrino analysis, the type of major backgrounds are the same between Phase-I and Phase-II. For the signals, due to around loss of around 50 PMTs per year, the energy resolution in Phase-I is better than that in Phase-II. For the backgrounds in Phase-II, the rate of the main background of ^{7}Be, *pep* and CNO solar neutrinos, the ^{210}Bi decay events, is reduced by almost 60%; the rate of the main background of ^{7}Be solar neutrinos, the ^{85}Kr decay events, is reduced

[12] V. Lagomarsino and G. Testera, "A gateless charge integrator for Borexino energy measurement", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 430, no. 2-3, pp. 435–446, July 1999, DOI: `10.1016/S0168-9002(99)00170-9`

[13] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 89, no. 11, pp. 1–68, 2014, DOI: `10.1103/PhysRevD.89.112007`, arXiv: `1308.0443`

[14] ibid.

[15] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

by more than 80%. The average rate of ^{210}Po decay events, which is also a main background of ^{7}Be solar neutrinos, is one orders of magnitude lower than that of Phase-I.

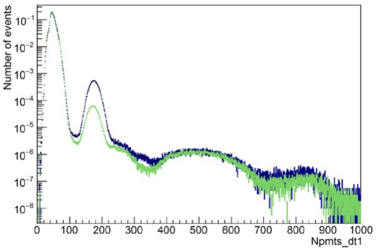


Figure 2.2: Comparison of energy spectra obtained in Phase-I (black) and Phase-II (green).

2.3.2 ^{8}*B analysis*

The ^{8}B solar neutrino flux is extracted by counting the number of events passing all cuts and subtracting the estimated number of residual background events. In Phase-I analysis[16], events within 3 meters radius are selected. In Phase-II analysis[17] all events fallen in the inner vessel are considered and the exposure is almost 10 times larger. In both analyses, the major backgrounds include the muon events, cosmogenic neutrons, decay of other cosmogenic isotopes, decay of ^{208}Tl inside the liquid scintillator or on the inner vessel, and γs from neutron capture on outside components. Muon events and short lived isotopes are removed by muon veto. Long lived isotope, such as ^{11}Be and ^{10}C are removed by Three-Fold-Coincidence cut[18]. ^{214}Bi-Po cut is applied to remove ^{214}Bi decay events. External backgrounds are subtracted according to results of radial fit. In Phase-II, because the fiducial volume is enlarged and include peripheral events, a new component called "emanation ^{208}Tl" is added. This component consists of two types of events. Firstly, nuclei in the ^{232}Th decay chain may recoil into the liquid. Alternatively, ^{220}Rn, an volatile progenitor of ^{208}Tl can diffuse into the liquid scintillator. Its radial distribution is derived based on ^{212}Bi events.

2.4 *Backgrounds*

In this section we present the list of background components. Fit results of an Asimov dataset are shown in Figure 2.3. All important background components are shown.

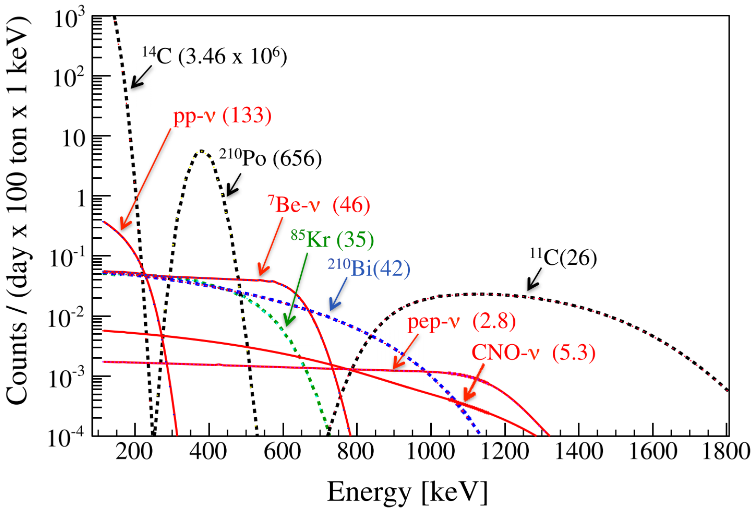


Figure 2.3: Fit results of an Asimov dataset. All signals and backgrounds are shown. Reproduced from The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 89, no. 11, pp. 1–68, 2014, DOI: `10.1103/PhysRevD.89.112007`, arXiv: `1308.0443`.

^{14}*C* The ^{14}C isotope accompanies ^{12}C in the liquid scintillator. It is a β-emitter with 156 keV end point. Its mean life is 8270 years

[16] The Borexino collaboration, "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector"

[17] The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure"

[18] See also Section 2.5.4 page 41

and is geologically short, but it is constantly being replenished by the cosmic-ray flux. In order to lower the contribution of this background, the Borexino scintillator is derived from petroleum from deep underground. The ^{14}C rate is measured to be 40 ± 1 cpd/100 ton independently from fitting.

pile-up events When two or more physics events occur within short time, they can be identified as one single cluster. It is the major background for *pp* solar neutrinos. The major components of pile-up events in the sensitive region of *pp* solar neutrinos are the pile-up of two ^{14}C events or the pile-up of one ^{14}C and one external background event. The pile-up events can be included by adding an independent component or include the pile-up effects to each componnet

^{85}***Kr*** The isotope ^{85}Kr is present in the air and it has an average concentration of around 1 Bq/m^3. It is a β-emitter with 687 keV end point (99.57% branching ratio). Its mean life is 15.4 years. Its contribution is lowered due to use of N_2 with low Kr content during the scintillator manipulations. With a branching ratio of 0.43%, ^{85}Kr decays by emitting a β particle then a γ ray with 2.06 μs mean life. This fast β–γ sequence can be used to measure a ^{85}Kr concentration independent from the one resulting from the spectral fit.

^{210}***Bi*** The isotope ^{210}Bi is a β-emitting daughter of ^{210}Pb with 7.23 days mean life and Q value of 1160 keV. Its energy distribution is similar to that of electrons recoiling after the interaction of CNO neutrinos in the scintillator.

^{210}***Po*** The isotope ^{210}Po is a mono-energetic α emitter with 200 days mean life, but with strong ionization quenching of the scintillator. The observed energy of ^{210}Po is similar to an electron of 0.4 MeV and is in the energy range of the recoiled electrons after the interaction of ^{7}Be solar neutrinos.

^{11}***C*** The isotope ^{11}C is a β^+ emitter and is a muon-induced cosmogenic background. About 95% of this nuclide is produced by muons accompanied with a free neutron:

$$\mu +^{12} \mathrm{C} \rightarrow \mu +^{11} \mathrm{C} + n \tag{2.1}$$

The mean life of ^{11}C is 29.4 minutes. The total energy released in the detector is between 1020 and 1980 keV and lies in the energy region of interest of the *pep* and CNO neutrinos. Its concentration is enhanced in the TFC-enhanced spectrum and suppressed in the TFC-subtracted spectrum[19].

External $\boldsymbol{\gamma}$***s*** γs from ^{40}K, ^{214}Bi and ^{208}Tl decays occurred in the detector components outside the liquid scintillator form backgrounds for the *pep* and CNO neutrinos.

[19] TFC: Three Fold Coincidence. See Section 2.5.4, page 41.

^{10}C and ^{6}He ^{10}C and ^{6}He are cosmogenic backgrounds. Their contribution is negligible in the TFC-subtracted spectrum and thus is only included in the TFC-enhanced spectrum.

2.5 Background reduction analysis techniques

In this section we reviewed the analysis techniques developed to reduce the background contribution. The background events can be rejected according to single or multiple quantities. They can also be disentangled from the signal events through spectral fitting.

2.5.1 Fiducial volume cut

Events reconstructed outside a pre-selected volume are rejected. In this way the contribution of γ rays produced in the outer region of the detector is suppressed. These γs, especially those from decay of ^{40}K, deposit energy in the fiducial volume and form backgrounds for *pep* and CNO solar neutrinos. In this analysis, the so-called *pep* fiducial volume is used:

- $z > -1.8$ m and $z < 2.2$ m
- $R < 2.8$ m

2.5.2 Muon cut and muon daughter cut

Muons may deposit significant amount of energy in the liquid scintillator and saturate Borexino electronics, and thus events adjacent to muons are rejected. Muon events are identified using the outer water Cherenkov veto detector trigger or using the time and space distribution of hits collected by the inner detector. Decay of short lived cosmogenic isotopes are rejected based on their time with respect to the most recent muons. The muon and muon daughter cut is defined as

- $\Delta t > 0.3$ s

where Δt is the time to the last tagged μ event.

2.5.3 Radon filter

If two events are adjacent in space and time, then it is unlikely they are uncorrelated. Both events will be removed and this cut is called the coincidence cut or the radon filter. It can remove events such as decays of ^{214}Bi which will be followed by the decay of ^{214}Po. The coincidence criteria is defined as the following:

- $\Delta t < 2$ micro-seconds
- $\Delta r < 1.5$ meters

2.5.4 *Three Fold Coincidence*

In part of production processes of cosmogenic isotopes, one or more neutrons are also present. When the production is not accompanied with neutrons, it is possible that at the production position some neutrons are produced due to the production of other type of cosmogenic isotopes. This is further discussed in Section 10.3.4 page 177.

This information can be used to suppress decay events of long lived cosmogenic isotopes such as ^{11}C, ^{10}C, ^{6}He and ^{11}Be. Neutron events close to a muon track are more likely cosmogenic neutrons. Events that are close to a cosmogenic neutron in space and time are more likely the decay of cosmogenic isotopes, and tagged as "TFC selected", or otherwise "TFC suppressed". The TFC is short for three-fold-coincidence, which consists of a muon track, a neutron event and a physics event. The tagging is not simply based on a cut of the spatial and time distance to the cosmogenic neutron, but based on a projected likelihood function taking into account all useful information except the event energy, which otherwise will introduce distortions to the event energy distribution.

The comparison of TFC selected and TFC suppressed is shown in Figure 2.4. It is obvious that the neutrino signals are not affected while decays of cosmogenic ^{11}C are strongly suppressed. This is critical for extracting the *pep* solar neutrino flux and the CNO solar neutrino flux.

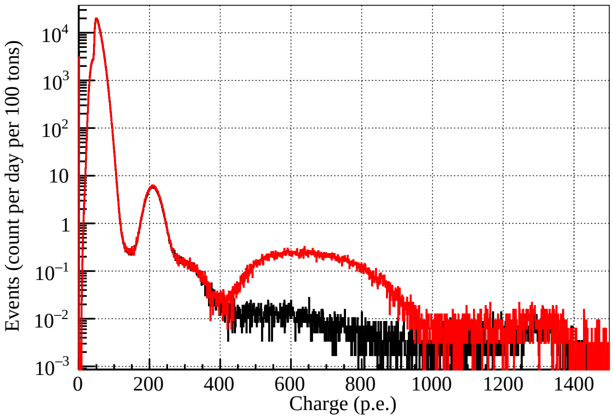


Figure 2.4: Comparison of the energy spectrum of events selected by TFC and that of events surviving the TFC cut. Both spectra are normalized by the exposure.

2.5.5 *Complementary fit*

The complementary fit technique allows to improve the precision of the *pp* and ^{7}Be solar neutrino fluxes by increasing the effective exposure. Instead of abandoning TFC-selected events, both spectra are utilized and the product of the likelihood terms built upon two spectra is minimized:

$$\mathcal{L} = \mathcal{L}^{\text{TFC vetoed}} \cdot \mathcal{L}^{\text{TFC tagged}} \tag{2.2}$$

Considering that the spatial temporal distribution of TFC selected events is not uniform and that ^{210}Po decay events are not uniform in space and in time, the ^{210}Po decay rates in two spectra are independent. For other species, except cosmogenic ^{11}C, ^{10}C and ^{6}He, their rates are required to be the same in both spectra. In the analytical fit, the detector response function parameters, including the light yield and resolution parameters, are required to be the same in two spectra.

2.5.6 *Multivariate fit*

The multivariate fitting is used to further disentangle the signals from the backgrounds.

The "radius", which is the distance between the event vertex to the detector center, is used to suppress the impact of external gammas. The neutrino elastic scattering and cosmogenic ^{11}C event vertexes are uniform distributed, while the flux of gammas penetrating the buffer into the fiducial volume originated from the radioactivities on the PMT glasses and stainless steel sphere reduces exponentially along the radial axis toward the center, thus can be distinguished through different radial distribution. For ^{210}Bi, dedicated studies are performed and no strong non-uniformity is observed. The involved energy range is carefully chosen to reduce the systematic uncertainties from the light yield as well as to avoid ^{210}Po, which is strongly non-uniform.

e^+/e^- pulse shape discriminator $\mathcal{L}(\mathbf{r})$ (PS-$\mathcal{L}_{\mathrm{PR}}$), which is the minimized likelihood used by the position reconstruction algorithm[20], is used to suppress the ^{11}C events. Its distribution is shown in Figure 2.5. The position reconstruction uses the time-of-flight (TOF), or the time that the photons spend moving from the production position to the PMT photocathodes where the photon are detected. The reconstruction algorithm assumes the event is an electron. Electrons deposit energy in a volume of millimeters, while gammas and positrons deposit energy in a volume of centimeters, so they have additional variation in the distribution of the TOF of hits, and the likelihood is on average larger compared with an electron event of the same energy. Its distributions for electrons are validated against the ^{214}Bi events selected through fast coincidence and those for positrons are validated against special TFC selected ^{11}C events.

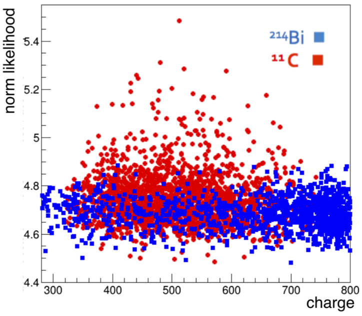


Figure 2.5: Distribution of $\mathcal{L}(\mathbf{r})$ and `charge` for ^{214}Bi (blue) and ^{11}C (red) events. Reproduced from S. Marcocci, "Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with Xe -loaded Liquid Scintillators", PhD thesis, Gran Sasso Science Insitute, 2016.

The observables radius and the e^+/e^- pulse shape discriminator are included by multiplying additional terms of likelihood[21]:

$$\mathcal{L}^{\mathrm{MV}} = \mathcal{L}^{\mathrm{TFC\ vetoed}} \times \mathcal{L}^{\mathrm{TFC\ tagged}} \times \prod_i \mathcal{L}_i^{\mathrm{radial}} \times \prod_j \mathcal{L}_j^{\mathrm{pulse\text{-}shape}} \tag{2.3}$$

where i and j indicate different pieces of likelihoods selected based on the energy of the events.

[20] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy"

[21] S. Davini, "Measurement of the pep and CNO solar neutrino interaction rates in Borexino-I", *The European Physical Journal Plus*, Springer Theses, vol. 128, no. 8, p. 89, Aug. 2013, DOI: `10.1140/epjp/i2013-13089-9`

2.5.7 α/β discrimination

The ^{210}Po α-decay is the major background of ^{7}Be solar neutrinos. It sits in the sensitive region for ^{7}Be solar neutrino where ^{7}Be solar neutrino interactions are correlated with ^{210}Bi decays and ^{85}Kr decays. In Phase-I analysis they were suppressed using the "Gatti" parameter[22]. In Phase-II analysis they are kept and modeled. Due to the high statistics and the asymmetry of the shape of ^{210}Po α-decay events, symmetric Gaussian response function will introduce biases. There has been studies on doing fitting with ^{210}Po α-decay events removed using a Multi-Layer Perceptron algorithm[23]. In this method, because the efficiency depends on the energy exponentially, the rejected α-like events are also integrated in the final likelihood terms using the so-called complementary fit methods, in order to extract the corresponding constants.

2.6 Offline softwares

In this section, we introduced the overview of the offline softwares used in the Borexino analysis.

The first step is processed by `Echidna`. Data collected are first packed in binaries. They are passed to `Echidna`. Functional units in `Echidna` are organized as regulated modules. Each module is assigned a number representing its priority level, and the modules with smaller number, or higher priority level, are processed first. These levels are specified in the configuration file `echidna.cfg`. First raw data are decoded by the `reader` module. The hit time and charge of each hit collected on each PMT packed as events is extracted. Then they are passed to the `decoder` module. Corrections are applied to the hit time and hit charge. Then they are passed to the `find cluster` module. Hits are grouped into clusters for each event. After that, the hits are passed to reconstruction algorithms, and the vertex and energy estimators are calculated for each cluster. In summary, `Echidna` takes raw data in binaries as inputs and produce reconstructed vertex and energy of clusters in `CERN ROOT TTree` as outputs.

The second step is processed by `bxfilter`. The software loops over each cluster, and selects clusters that fulfills criteria of each analysis. Different type of selections are wrapped as different modules. In the low energy branches solar neutrino analysis, the selection is defined in the `bx_nusol` module.

The third step is processed by `spectral-fitter`, `bx-GooStats`[24] or `bx-stats`. In the last step, all inputs, including the energy spectrum of TFC tagged and suppressed events, the exposure, the distribution of the number of live PMTs, the Monte Carlo probability density functions, etc. are produced and saved in `ROOT` files. In this step, the above mentioned softwares, or the fitters, perform minimizations and extract the best fit of rates and the corresponding HESSIAN statistical uncertainties.

The last step is processed by `bx-sampler`. This software is used

[22] E. Gatti *et al.*, Energia Nucleare (Milan) 17, 34 (1970). http://www-3.unipv.it/donati/papers/6d.pdf.

[23] I. Drachnev, "New Spectral Analysis of Solar B Neutrino with the Borexino Detector", PhD thesis, Gran Sasso Science Institute, 2016

[24] X. Ding, "GooStats: A GPU-based framework for multi-variate analysis in particle physics", *Journal of Instrumentation*, vol. 13, no. 12, P12018–P12018, Dec. 2018, DOI: `10.1088/1748-0221/13/12/P12018`

to study the impact of various fitting model related systematic uncertainties. It produces large amount of pseudo-experiment spectra, then they are passed to the softwares mentioned in the third step. The width of the distribution of best fit of the results are used to estimate the systematic uncertainties.

3

Development of Analytical multivariate fitting

In this chapter, the development of the analytical multivariate fitting method is introduced. It is a combination of the analytical response functions and multivariate method. The analytical response function provides an alternative way to extract the detector response information, such as the light yield and resolutions. It is useful to crosscheck results and understand the systematic uncertainties from the detector modeling. The multivariate fitting method provides a way to suppress background. Beside the energy, the radius and pulse shape parameters can be used to suppress backgrounds, especially external ^{40}K decay γs and cosmogenic ^{11}C β^+ decays. These two are critical backgrounds for *pep* solar neutrinos.

It is difficult to construct an analytical model of the joint distribution of energy and radius or pulse shape discriminators. To solve this problem, Monte Carlo, or `g4bx2`, is used to predict the joint distributions. Although in this way two models of energy response are used, the systematic uncertainty introduced is acceptable. Besides, in previous studies, development of the analytical multivariate fitting method suffers from its unacceptable long converging time — it takes more than one week to converge with the CPU fitter, and we have to fix the light yield and several species and fit only part of the energy range to get results in acceptable time. Thanks to the development of `GooStats`[1], it takes only a few minutes to finish an analytical multivariate fit, making this study feasible.

This chapter is organized as in the following: Section 3.1 introduces the concepts related to the spectrum fitting methods. Section 3.2 discusses the new fitting method: the analytical multivariate fitting method. At last, section 3.3 explains how parallelization speeds up the fitting process.

---

[1] X. Ding, "GooStats: A GPU-based framework for multi-variate analysis in particle physics", *Journal of Instrumentation*, vol. 13, no. 12, P12018–P12018, Dec. 2018, DOI: `10.1088/1748-0221/13/12/P12018`

3.1 Spectrum fitting method in Borexino

In this section, a pedagogical introduction to the understanding of the spectrum fitting method based on the experience of Borexino analysis is given, then the energy estimators and fitting methods used in Borexino spectral analyses are introduced. In Borexino, the signals and backgrounds are highly entangled, and they are differentiated through the difference of their energy distributions. It is crucial to understand what will affect the precision of the extracted rates of signals and how to control them. In this section, I will tell you the details of Borexino spectrum fitting.

3.1.1 Concept of correlation in spectrum fitting

Fitting is the activity that allows you to extract the counts of signal events and background events based on two information: the distribution of collected events and the models of components. In the fitting process multiple pieces of information are extracted, and the main obstacle of fitting is the entanglements among these pieces of information, which are called correlations. Thus we can also say that fitting is just the process of breaking the correlations among components, or more precisely, the correlation among the extracted counts of components. Correlation means you know the sum of two, while you know little about their difference. It also means two things are similar. Counting analysis is the extreme case of spectrum fitting: all components are completely correlated and indistinguishable.

The correlation can be measured as the correlation coefficient between two fit parameters. It is a statistical concepts and fit parameters are considered as random variables, or test statistics. When response function parameters are also included as free parameters, it is possible that correlation exists between counts of some components and the detector response parameters, such as the light yield, or the resolution parameters. In this case, correlation cannot be simply explained as similarity between the shape of distribution of two components.

The type of correlation can be defined according to the way how two fit parameters are related. The Borexino spectrum is shown in Figure 2.3. From this figure we can identify three types of correlations:

- When shape of a component is sharp, with with large value of the first derivative, it can be used to determine the detector response parameters, and the counts of this component is correlated with the corresponding parameters. For example, ^{7}Be solar neutrino interactions (see Figure 3.1) and ^{210}Po decays are sharp, and they are correlated with the resolution parameters.

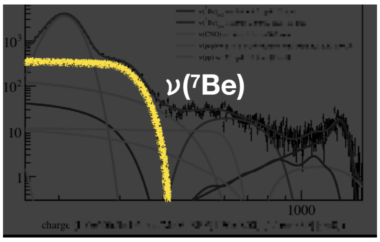


Figure 3.1: Correlation between $\nu(^7\mathrm{Be})$ and resolution parameters.

- When shape of a component is self-similar under the simultaneous transformation of some response parameters and counts, the information of the counts of such component is poor and a bias can be introduced easily by biased response parameters. The counts of these components are correlated with the corresponding resolution parameters. For example, ^{210}Bi decay energy distribution has this kinds of shape, and its count is correlated with the light yield, see Figure 3.2.

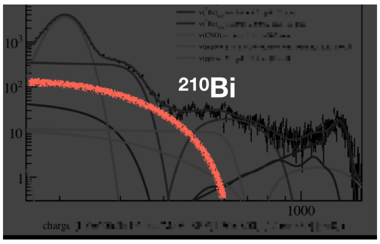


Figure 3.2: Correlation between the ^{210}Bi decay rate and the light yield.

- When the shapes of two components are similar, their counts are correlated. With huge amount of exposure the precision of single component might be improved. However, like the last type of correlation, if their shape are similar under the simultaneous transformation of some response parameters and counts, the counts will migrate between two due to bias of response parameter and the sensitivity will be spoiled. For example, the shapes of ^{210}Bi decays and CNO solar neutrino interactions are similar, see Figure 3.3. Although the experimental sensitivity to the difference of two components improves as square root of the exposure, tiny bias on the light yield will spoil it.

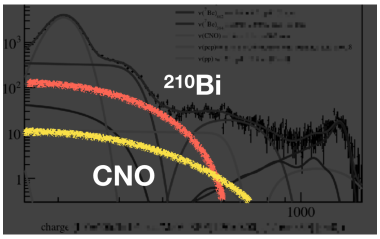


Figure 3.3: Correlation between the ^{210}Bi decay rate and the CNO solar neutrino interaction rate.

To break the correlation, the only method is to put more information and fix or constrain part of the fit parameters that are correlated. For example, the light yield can be determined with calibrations and can be fixed. The resolution parameters can be fixed to the values obtained with Monte Carlo, which is tuned against calibration data. The correlation between ^{210}Bi decays and CNO solar neutrino interactions can be broken by fixing or constraining the ^{210}Bi decay rates, and the ^{210}Bi decay rates can be determined with independent measurement using ^{210}Po decay rates and the link between ^{210}Bi decay and ^{210}Po decay rates. Sometimes we do not break the correlation and consider the degrading of the statistical uncertainty is worthy, because when fixing or constraining fit parameters, larger systematic uncertainties might be introduced. For example, in the analytical fit strategies part of resolution parameters are fixed while light yield is left free.

3.1.2 *Energy estimators*

In Borexino, three types of energy estimators are used:

- The normalized total number of triggered PMTs in a fixed window interval of 230 (400) ns ($N_p^{dt_{1(2)}}$, or `npmt_dt1` (`npmt_dt2`)). Refer total number of triggered PMTs in the window as $N_p^{dt_1(2),\mathrm{raw}}$, the number of live channels that can discriminate if the PMT is fired as N', then $N_p^{dt_{1(2)}} = N_p^{dt_1(2),\mathrm{raw}} \cdot 2000/N'$.
- The normalized total number of detected photons, including multiple photons on the same PMT (N_h, or `nhit`). For one physics event, if two photons arrives on the same PMT, both produce photo-electrons and are collected, and the separate time between two photons is long enough, they will be registered as two hits. Refer total number of detected photons as N_h^{raw} and use the same definition of N' in the previous item, then $N_h = N_h^{\mathrm{raw}} \cdot 2000/N'$.
- The normalized sum of charge of all collected hits (N_Q, or `charge`). Refer the sum of charge of all collected hits as N_Q^{raw}, the number of live channels that can collect charges as N'', then $N_Q = N_Q^{\mathrm{raw}} \cdot 2000/N''$.

More details can be found in[2].

The `charge` variable has the best energy resolution for events with energy higher than around 700 keV. It spatial non-uniformity has little energy dependence, and thus can be corrected. Two types of corrections, based on the PMT geometrical acceptance or ^{210}Po α-decay events, are developed and applied to `charge`. Compared with the other two variables, its non-linearity and resolution models are relatively simple. More details are introduced in Chapter 4.

The `nhit` variable has the best energy resolution for low energy events, and `g4bx2` is tuned based on the `nhit` variable. It is similar to `npmt_dt1` variable in both energy non-linearity and resolution. It is used as the main variable in the Monte Carlo fit.

The `npmt_dt1` variable has good energy resolution for low energy events, and it is used to extract *pp* solar neutrino interaction rate in the analytical fit strategy.

3.1.3 Fitting methods

Fitting methods here refer to the way of building the probability density functions. In Borexino, two types of fitting methods are used: the analytical strategy and the Monte Carlo strategy.

The analytical strategy builds energy distribution models of each component with analytical functions. The information of detector response parameters, such as the light yield, is extracted from the spectrum directly, rather than fixed to the values obtained from calibrations.

The Monte Carlo strategy builds energy distribution models with `g4bx2`. `g4bx2` is tuned based on calibration data and selected ^{14}C decay events. The calibration data is used to fix the light yield, non-linearity, resolution models. The ^{14}C decay events are used to determine the so-called effective quantum efficiencies, in order to track the temporal dependence of the light yield and spatial non-uniformity.

[2] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 89, no. 11, pp. 1–68, 2014, DOI: `10.1103/PhysRevD.89.112007`, arXiv: `1308.0443` Section IX

3.2 *Analytical multivariate fit method*

In this section I describe the analytical multivariate fit method, in which the energy distributions are modeled with analytical functions. In the multivariate fitting method, three observables are utilized: the energy of the event, the distance between the event vertex and the detector center (radius), and the e^+/e^- pulse shape discriminator $\mathcal{L}(\mathbf{r})$. The radius is used to differentiate the external gammas. The $\mathcal{L}(\mathbf{r})$ is used to suppress the cosmogenic ^{11}C. It is also discussed in Section 2.5.6.

It is difficult to construct an analytical model of the joint distribution of energy and radius or pulse shape discriminators. To solve this problem, Monte Carlo, or `g4bx2`, is used to predict the joint distributions. The principle of the solution is introduced first, then the detailed procedure to construct the likelihood is documented.

3.2.1 *Principle of analytical MV fit*

It is considered that the energy and radius (or pulse shape discriminators) are independent if the energy range is small enough. When they are independent, the binned likelihood built upon full two-dimensional distribution can be replaced by the sum of likelihood terms built on energy and radius, respectively. When using the naive sum of two likelihood terms, the information of the total counts is used twice and the result is thus biased. So for the likelihood of secondary observables, a self calibrating parameter is inserted to cancel such double counting effect. Although the detector energy response model may be different between the `g4bx2` and the analytical part can be different, the `g4bx2` energy response model is not important here, because only the shape of radial distribution[3] is used to differentiate between *pep* solar neutrinos and external ^{40}K decay γs, and the energy is only used to track the energy dependence of the radial distribution shape.

[3] radius here means the distance to the detector center.

3.2.2 *Likelihood construction*

Here we take radius as an example. The procedure is similar for pulse shape discriminators. To construct the analytical multivariate likelihood, first we define the concept of class. All species are classified into two classes, signal or background. Then we collect the inputs:

- The histogram of energy-radius joint distribution of collected data, see Figure 3.4.
- The histograms of predicted energy-radius joint distribution for the signal and the background class, according to the simulation with `g4bx2`, see Figure 3.5 and 3.6.
- The predicted number of events at each energy bin for each species according to analytical response functions
- The classification of all components: signals or backgrounds.
- The configurations: the energy range where multivariate analysis is applied and the number of energy slices.

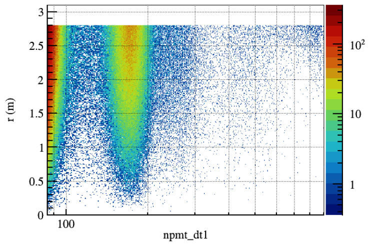


Figure 3.4: Energy-Radius joint distribution (data).

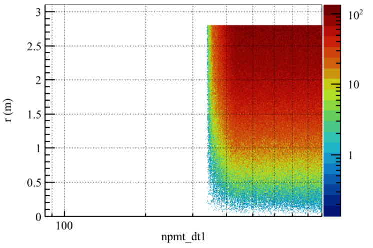


Figure 3.5: Energy-Radius joint distribution (MC, signal class).

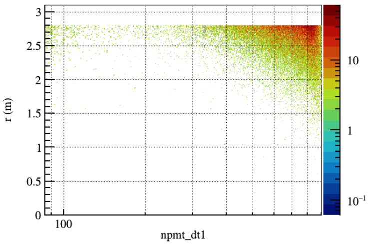


Figure 3.6: Energy-Radius joint distribution (MC, background class).

After that, project the two dimensional histograms of both the data and predictions onto histograms of radius distributions with energy range restricted according to the chosen range of the energy slices, see Figure 3.7.

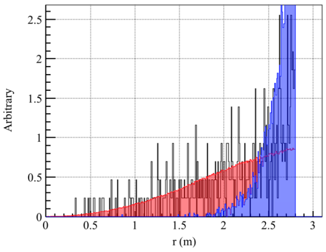


Figure 3.7: Comparison of the radius distributions of the first energy slice (`npmt_dt1` is in the range 500–550). Black: data. Red: Monte Carlo, signal class. Blue: Monte Carlo, background class.

Finally, calculate the sum of the log-likelihood of each energy slice.

The log-likelihood of one energy slice is defined as the following:

$$\ln(L) = \log\left\{\sum_{i=0}^{k}\exp\left[P(k-i,m_0,n_0)+P(i,m_1,n_1)\right]\right\} \tag{3.1}$$

$$P(k,m,n) = \ln\left[\frac{(n+k+1)!m^k}{(m+1)^{n+k+1}k!n!}\right] \tag{3.2}$$

where the parameters are defined as the following:

1. $k(j)$: number of events in j-th bin of the projected data histogram
2. $n_0(j)$: number of events in j-th bin of the projected predicted signal histogram
3. $n_1(j)$: number of events in j-th bin of the projected predicted background histogram.
4. m_0 and m_1: scaling factors.

The scaling factors are defined as

$$m_i = \frac{N_i}{N_0+N_1}\cdot\frac{\mathrm{sum}(k)}{I_i+\mathrm{Nbin}} \tag{3.3}$$

where N_i is defined as the predicted sum of species classified as i, 0: signal, 1: background; $\mathrm{sum}(k)$ is the integral of the k-th slice of the projected data histogram; I_i is the integral of the projected reference-i histogram; Nbin is the number of bins of the projected histograms.

3.3 *Multivariate analysis framework* `GooStats`

In this section we presented the development of the multivariate analysis framework `GooStats`. Without this software, it is impractical to perform the analytical multivariate analysis due to too long computation time.

Algorithms should be transformed to allow concurrent execution to be adapted to modern computing architectures. In 2005, Herb Sutter pointed out in his "The Free Lunch Is Over"[4] that while the number of transistors on CPUs is still increasing, mainstream computers are being permanently transformed into heterogeneous supercomputer clusters. The computing power is improved by increasing the number of computing units rather than by increasing the clock rate of single unit. For example, on one hand, the clock rate of Intel Pentium 4 CPU, released in 2003, is 3.8 GHz[5], while the clock rate of recent CPUs, such as Intel Xeon E5-2630 v3 released in 2014, has decreased to 2.4 GHz[6]; on the other hand, the number of computing units per CPU has increased from 2 (Pentium 4) to 8 (Xeon E5-2630 v3). What's more, General Purpose Graphics Processing Unit (GPGPU)[7], emerged at the beginning of the twenty-first century, has hundreds of computing units and can launch thousands of computation tasks simultaneously[8].

Algorithms can be transformed to allow concurrent execution if two sets of operations are identified: transformation and associative summation. Transformation is the operation of applying the same operation to independent datasets. Associative summation is the operation of applying an associative operation sequentially to a series of values. The calculation of likelihoods fulfills these two criteria: calculation of likelihood of each bin is independent with respect to each other; the arithmetic summation of likelihood of all bins is associative.

3.3.1 *Software design*

`GooStats` is a statistical analysis framework. It is initiated to facilitate the use of computation power of modern GPGPUs for fitting and statistical analysis. It serves as the middle layer between statistical analysis module written by users and the minimization engine `GooFit`. It provides a complete suit of analysis tools, supports common statistical analysis and also includes algorithms used in Borexino analysis[9,10], such as analytical detector response models and multi-observable likelihoods.

Structure of the software and responsibilities of classes The structure of `GooStats` is depicted with a simplified Unified Modeling Language (UML) graph[11] in Figure 3.8. The classes are named after their responsibilities. The workflow is managed by a singleton of `AnalysisManager`. Two manager classes are designed responsible for manipulating inputs and outputs, respectively, and the actual

---

[4] S. Herb, "The Free Lunch Is Over: A Fundamental Turn Toward Concurrency in Software", *Dr. Dobb's journal*, vol. 30, no. 3, pp. 202–210, 2005

[5] Intel, *Intel® Pentium® 4 670 processor with support for HT technology*, 2003

[6] Intel, *Intel® Xeon® Processor E5-2630 v3*, 2014

[7] Mark Harris, *GPGPU: Beyond Graphics*, Game Developer Conference, 2004

[8] nVidia Inc., *TESLA K20 GPU ACCELERATOR Board specification*, 2012

[9] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II", pp. 1–8, July 2017, arXiv: `1707.09279`

[10] The Borexino collaboration, "Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data", *Physical Review D*, vol. 96, no. 9, p. 091103, Nov. 2017, DOI: `10.1103/PhysRevD.96.091103`, arXiv: `1707.09355`

[11] Object Management Group, *Unified Modeling Language 2.5.1*, https://www.omg.org/spec/UML/2.5.1/, 2017

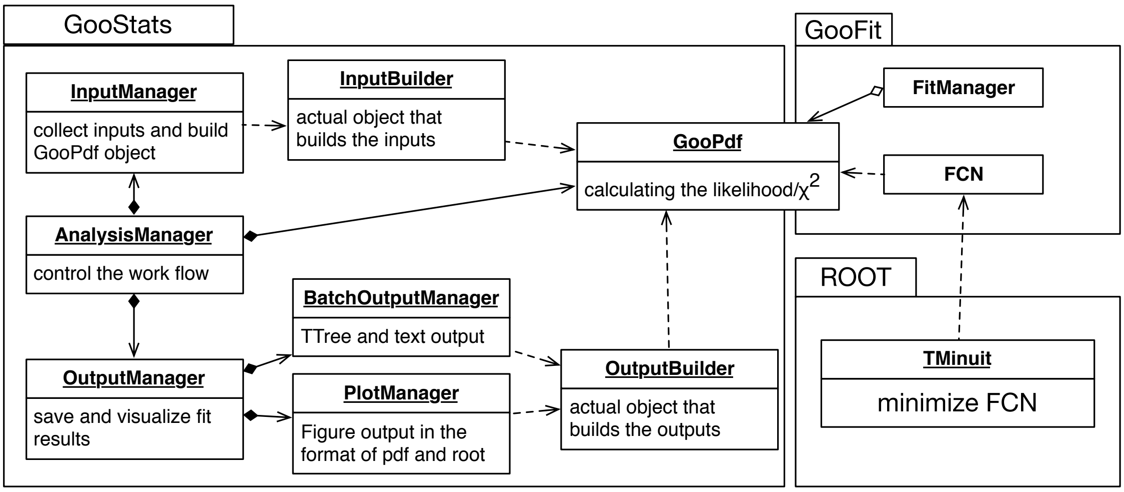


Figure 3.8: `GooStats` and its relationship with `GooFit` and `Minuit`.

work is performed by two builder objects to decouple the realization details. The input-managing object collects and parses the raw data and configurations, and builds with the help of its builder the probability density function objects, which are needed by the minimization engine. The output-managing object collects fit results and produces outputs with the help of two manager objects, responsible for graphic and non-graphic outputs, respectively, and the output builder. The default behavior can be overwritten through the polymorphism mechanism of the `C++` language.

The data model is designed to support fit of multiple datasets with independent configurations natively. The information are stored in two levels: configuration associated, which stores parsed key–value pairs, and dataset associated, which stores all essential information to construct a probability density function object including the dataset object, the configurations and the fit parameter objects. Each dataset corresponds to one configuration and one controller, and it is filled by its controller. After filling, the input builder will construct the probability density function object from the dataset object through spectrum builders. Their relationships are shown in Figure 3.9. The configuration objects are managed using the tree data structure in order to support the parameter synchronization technique, which will be explained in the next section.

The output `BatchOutputManager` and `PlotManager` get the information from `DatasetManager` objects through `OutputHelper` and `TF1Helper`, respectively. See Figure 3.10.

Parameter synchronization technique An analysis technique called parameter synchronization has been developed to improve the model precision. Consider that we can divide the collected data with respect to the volume and time and assign independent detector response parameters for each piece, while requiring different types of fit parameters to be synchronized among datasets at different levels of granularity. For example, Borexino has collected in total around five years of data in period Phase-I and period Phase-II[12,13]. With

[12] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy"

[13] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

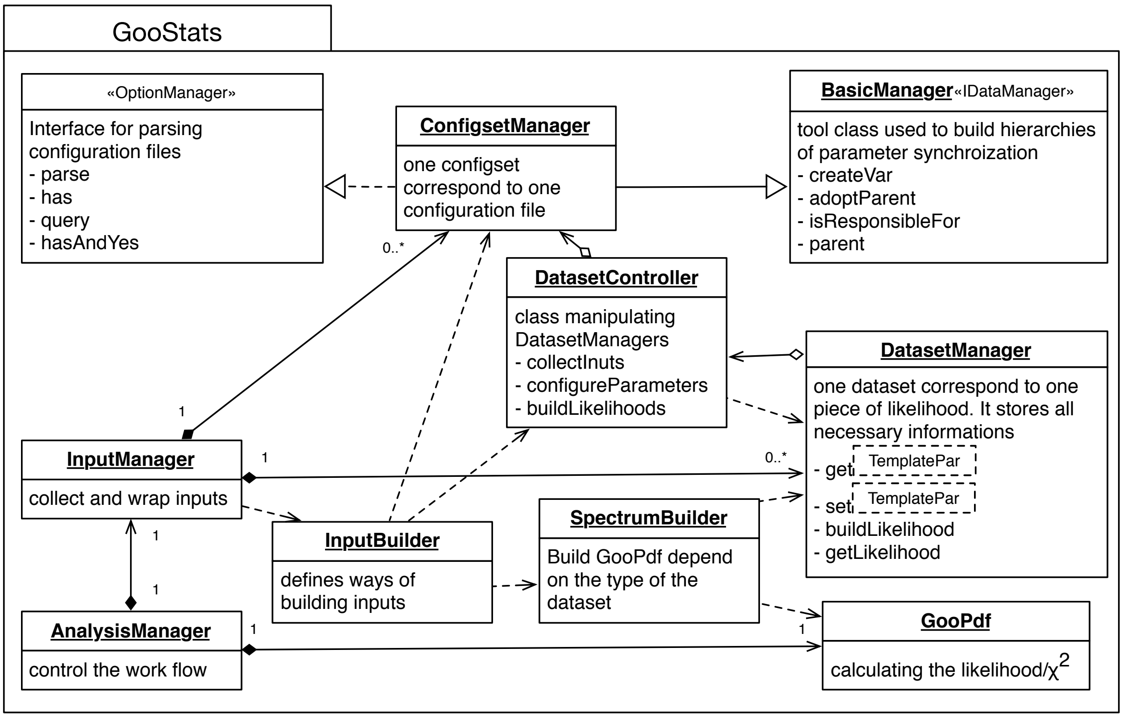


Figure 3.9: Data structure of `GooStats` and classes responsible for filling the input

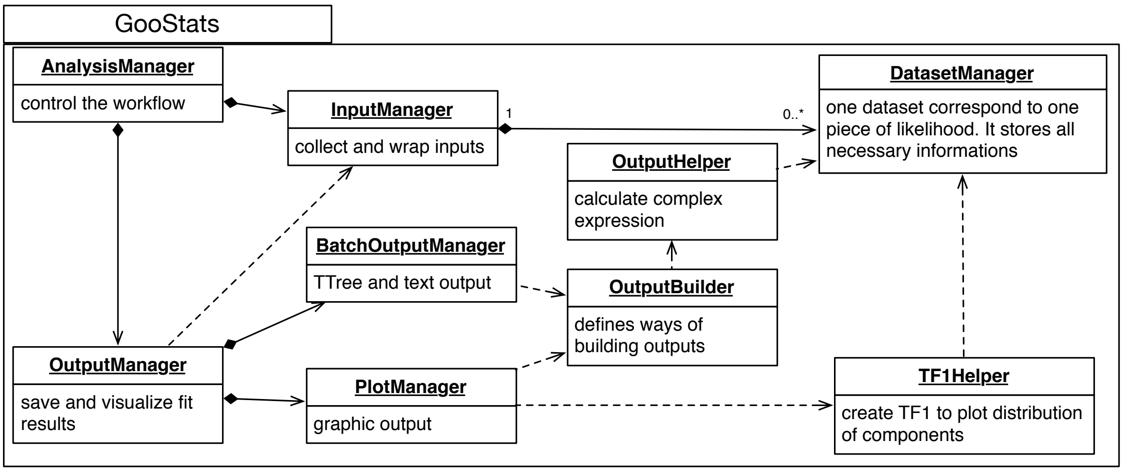


Figure 3.10: Organization of classes used by `OutputManager`.

such technique we can divide the collected data into five datasets, each corresponding to one year of data, and fit all datasets simultaneously. The solar neutrino rates should be required to be the same for all datasets because they are stable. The rates of the backgrounds, the radioactive decays, are stable within each period, while the purification campaign between two periods lowered their rates, so it is better to keep their rates the same in each period but independent between two periods. Meanwhile, in Borexino analyses, analytical functions are used to describe the detector response, and because the spectrum is very sensitive to the light yield and the PMTs are dying, it is sensible to keep detector response parameters independent among each year. Such strategy and corresponding data structure are summarized in Figure 3.11.

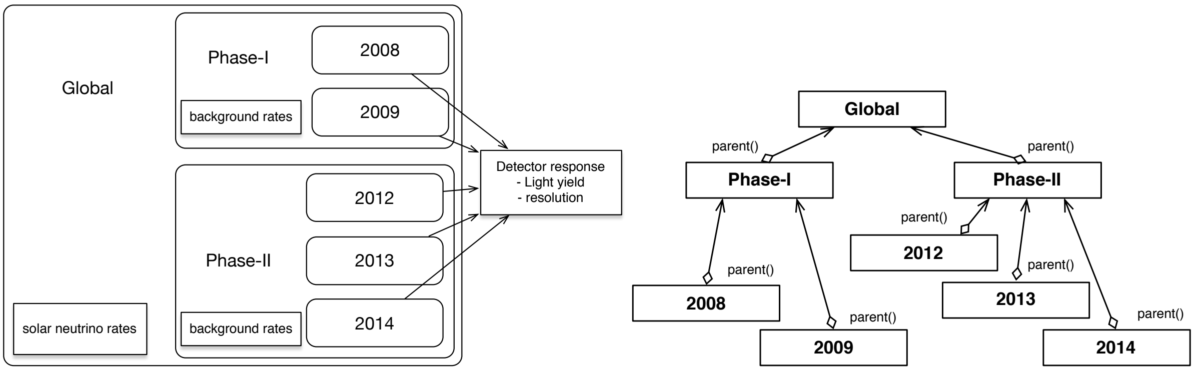


Figure 3.11: Left: An example synchronization strategy of Borexino Phase-I and Phase-II joint analysis using parameter synchronization technique. Right: The internal data structure. A parameter stored in a parent node will be synchronized among all its children nodes.

Such technique is implemented using the decorator design pattern and tree data structure. In `GooFit`, each fit parameter is represented with an independent object and if the object is shared within various probability density function objects, its value will be synchronized among them. In `GooStats`, the fit parameter objects are stored in the configuration objects. When filling the fit parameter objects into the dataset object, the synchronization manager will visit the tree until it finds the correct parent node corresponding to the required synchronization level of the parameter. The process of building the tree and that of visiting the tree are summarized in Figure 3.12.

3.3.2 Validation and benchmark

When developing the module for Borexino Phase-II analysis[14], the `GooStats` software framework was validated against the existing analysis tool. Tests were also designed to measure the overhead. In this section I will present the validation and benchmark results.

Validation processes have been designed such that the relative difference of the expected number of events between the `GooStats` Borexino module and the existing software used by the Borexino collaboration should be within 10^{-12} for all bins and all species, and the relative difference of the optimized value of the rates and the likelihood should be smaller than 10^{-4}. The comparison results are shown in Figure 3.13. As can be seen that the results are within required precision. The validation is implemented using the `GoogleTest` library

[14] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

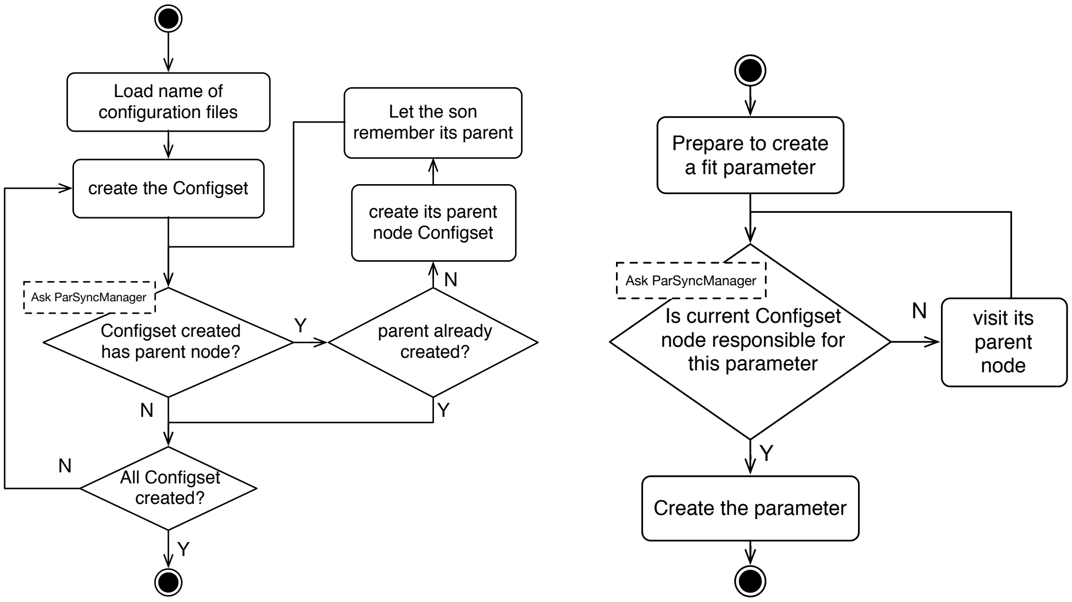


Figure 3.12: Flow chart of building the tree of configuration objects (left) and creating a fit parameter (right).

(See `https://github.com/google/googletest`).

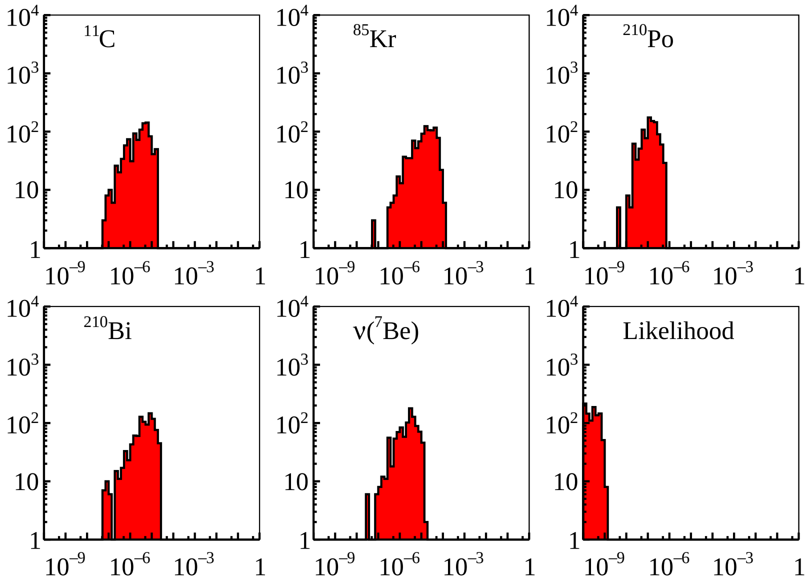


Figure 3.13: The distribution of relative difference $|\Delta_R/R|$ of the optimized value of various species rates and the likelihood between the `GooStats` Borexino module and the existing tool used by the Borexino collaboration.

The computing time needed to solve a statistical problem depends strongly on the specific application. GooStats can handle out-of-the-box three kinds of fits:

1. Fitting with the Monte Carlo detector response model, hereafter as "Monte Carlo fit";

2. Fitting with the analytical detector response function model, hereafter as "Analytical fit";

3. Fitting with multiple observables using multivariate likelihood, hereafter as "Multivariate fit".

The time spent on `GooStats` are mainly used on the parallel computing tasks executed on GPU. The fitting time almost linearly scales with the size of the problem, while the size of the problem is determined by different factors among different types of jobs:

- For the Monte Carlo fit, the computation heavy part is the evaluating of likelihood on each bin, and thus the size of the problem is the number of bins:

  $$T = T_0 + N_{E_\text{vis}} \cdot k \tag{3.4}$$

  where T is the total wall time, T_0 is the overhead used on loading and parsing input and producing figure outputs etc., N_{E_vis} is the number of bins of the visible energy spectrum, k is the speed of the program.

- For the analytical fit, each expected spectrum is the convolution between the energy spectrum and the detector response, and thus the size can be defined as the sum of product of the mesh size of the physical energy spectrum and that of the visible energy spectrum:

  $$T = T_0 + k_1 \cdot \sum_{i=1}^{N_\text{comp}} N^i_{E_\text{vis}} \cdot N^i_{E_\text{true}} \tag{3.5}$$

  where N_comp is the number of components, $N^i_{E_\text{kin}}$ is the number of bins of the physical energy spectrum.

- For the multivariate fit, much more time is spent on the calculation of the multivariate-likelihoods. The size of the problem can be defined as the following:

  $$T = T_0 + k_2 \cdot \sum_i M_i \tag{3.6}$$

  where M_i is the number of energy slices of secondary observable spectrum.

We performed tests to evaluate the overhead. Fitting tasks of different sizes are performed on an nVidia Tesla K20m GPU[15], whose double precision processing power is 1.18 TFLOPS (tera floating point operations per second). Sizes of fitting tasks are varied by changing the bin size and the fit range. The results are shown in figure 3.14. We can see that the speed of the software on the K20m GPU is satisfactory that the fit time per iteration for a typical-size fitting task in Borexino Phase-II analysis[16] ranges from sub-milli-seconds to tens of milli-seconds, depending on the type of the task. The fraction of the overhead, that is, the time spent on parsing configuration files, communicating between CPU and GPU, I/Os, etc., in the total wall time is around 50% and is also satisfactory.

[15] Inc., *TESLA K20 GPU ACCELERATOR Board specification*

[16] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

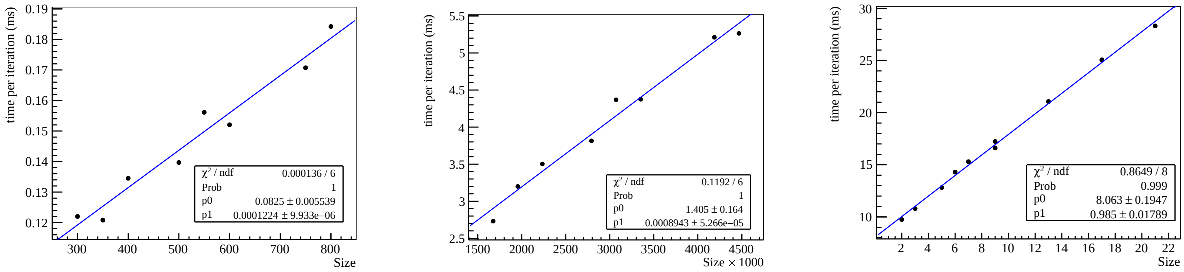


Figure 3.14: The time used per iteration during minimization for three types of jobs. Left: Monte Carlo fit. Middle: Analytical fit. Right: Multivariate fit. Least square fits with linear function are also shown.

We also compared the fitting time between `GooStats` Borexino module and the original software used by the Borexino collaboration for the "Analytical fit" type of job. The results are summarized in Table 3.1. The speed up is more than 100.

	CPU			GPU			
Type	AMD Opteron(TM) Processor 6238			nVidia Tesla K20m			
Size	T_{tot} (s)	N	T_{it} (ms)	T_{tot} (s)	N	T_{it} (ms)	speed up
400	27.6	1128	24.4	0.181	1346	0.135	181
350	29.4	1331	22.1	0.156	1294	0.121	183
300	22.5	1239	18.2	0.243	1995	0.122	149

Table 3.1: Comparison of fitting time between `GooStats` and original software used by the Borexino collaboration. T_{tot}: total execution time. N: the number of iterations taken to converge in `MINUIT`. T_{it}: average execution time per iteration. Speed up: $T_{\text{it}}(\text{CPU})/T_{\text{it}}(\text{GPU})$.

3.3.3 *Application examples*

In this section I will present the applications of `GooStats` to fitting and statistical analyses. The codes are available in the `github` repository (See `https://github.com/GooStats/GooStats`. Two pedagogical examples and one realistic problems is shown here.

Fit of Gaussian signal plus flat background Consider we would like to extract the rate of a Gaussian signal over a flat background. For this simple task, a `main` function using the default classes is enough:

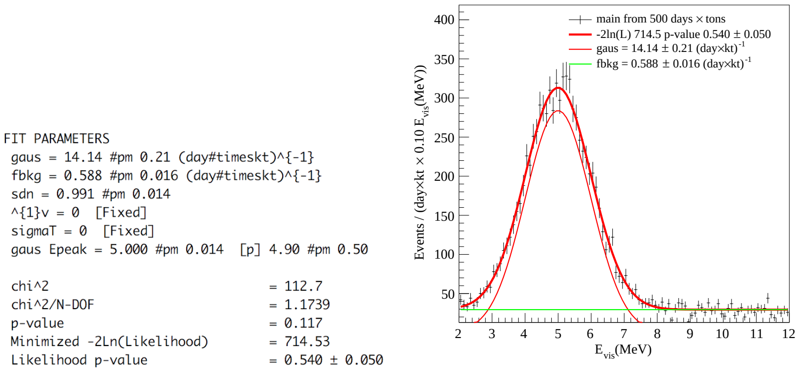
```
int main (int argc, char** argv) {
  AnalysisManager *ana = new AnalysisManager();

  InputManager *inputManager = new InputManager(argc,argv);
  inputManager->setInputBuilder(new SimpleInputBuilder());
  GSFitManager *gsFitManager = new GSFitManager();
  OutputManager *outManager = new OutputManager();
  outManager->setOutputBuilder(new SimpleOutputBuilder());
  outManager->setPlotManager(new SimplePlotManager());

  StatModule::setup(inputManager);
  StatModule::setup(gsFitManager);
  StatModule::setup(outManager);

  PrepareData *data = new PrepareData();
  SimpleFit *fit = new SimpleFit();

  ana->registerModule(inputManager);
```

```
  ana->registerModule(data);
  ana->registerModule(fit);
  ana->registerModule(outManager);

  ana->init();
  ana->run();
  ana->finish();

  return 0;
}
```

An example project is provided in the `GooStats/Modules/simpleFit` folder. A `Makefile` is provided for convenience and user can compile this example project with a `make` command. Users can run the fit with the command

```
./fit toyMC.cfg out exposure=500
```

At the end, a summary of fit results is printed on the screen. When we would like to see the impact of changing a few fit configurations quickly, this summary would be useful. See Figure 3.15 left. A figure in file format of `PDF` is also produced, which directly presents how well models describe the data. See Figure 3.15 right. `GooStats` also produce outputs in `CERN ROOT` objects. A file in the `TFile` format containing `TF1` objects of each components, etc. is produced. See Figure 3.16 left for the full list of objects saved in the file. Sometimes we need the distribution of fit results against many pseudo-experiment spectra, so it would be convenient to have output in the `TTree` form. It is included in the produced `TFile`. See Figure 3.16 right.



```
FIT PARAMETERS
 gaus = 14.14 #pm 0.21 (day#timeskt)^{-1}
 fbkg = 0.588 #pm 0.016 (day#timeskt)^{-1}
 sdn = 0.991 #pm 0.014
 ^{1}v = 0  [Fixed]
 sigmaT = 0  [Fixed]
 gaus Epeak = 5.000 #pm 0.014  [p] 4.90 #pm 0.50

 chi^2                              = 112.7
 chi^2/N-DOF                        = 1.1739
 p-value                            = 0.117
 Minimized -2Ln(Likelihood)         = 714.53
 Likelihood p-value                 = 0.540 ± 0.050
```

Figure 3.15: Left: screen shot of fit result summary. Right: produced figure in the file format of `pdf`.

The contents in the configuration file and on the command line arguments can be understood as the following. In the process of the fit, the binned log-likelihood is minimized:

$$\ln \mathcal{L} = \sum_i \ln \left(\frac{\lambda_i^k}{k_i!} e^{-\lambda_i} \right) + \sum_j \left(\frac{p_j - c_j}{\sigma_j} \right)^2 \tag{3.7}$$

where the first term is the Poisson likelihood and the second term is the sum of the pull terms; k_i is the number of events in the i-th bin of the visible energy histogram of selected events, λ_i is the expected

```
TFile**         out.root
 TFile*         out.root
  KEY: TTree    fit_results;1   Fit result of GooStats
  KEY: TStyle   Plain;1 Plain Style (no colors/fill areas)
  KEY: TCanvas  LL_p_value;1    LL_p_value
  KEY: TF1      main;1  main
  KEY: TH1D     main_data;1     main_data
  KEY: TF1      main.gaus;1     main.gaus
  KEY: TF1      main.fbkg;1     main.fbkg
```

```
======> EVENT:0
 nfit            = 1
 gaus            = 14.1404
 gaus_err        = 0.210194
 fbkg            = 0.587698
 fbkg_err        = 0.0164991
 sdn             = 0.991057
 sdn_err         = 0.0138929
 v1              = 0
 v1_err          = 0
 sigmaT          = 0
 sigmaT_err      = 0
 gaus_Epeak      = 4.99981
 gaus_Epeak_err  = 0.0144807
 main.exposure   = 500
 gaus_Epeak_pull.exposure = 500
 chi2            = 112.693
 NDF             = 96
 likelihood      = 357.266
 minim_conv      = 1
 hesse_conv      = 1
 LLp             = 0.54
 LLpErr          = 0.0498397
```

Figure 3.16: Left: content of `TFile` output. Right: content of `TTree` output.

number of events in that bin, p_j is j-th constrained fit parameters, such as the position of the peak, c_j and σ_j are the centroid and width of the corresponding pull term. The calculation of Equation (3.7) is performed in the object of `SumLikelihoodPdf`, and `GooStats` serves to construct it. In order to construct the `SumLikelihoodPdf` object, the following inputs are needed:

- The histogram of the observable of selected events.
- The list of components, their probability density function types.
- The initial guesses and ranges of rates and response function parameters
- The list of pull terms, including the name of the constrained parameter, the centroid and the width of the constraint.
- The exposure

They are loaded from the configuration file and the command line arguments in the format of key-value pairs. The pairs on the command line arguments will override the contents in the configuration file.

Statistical analysis Consider that, in the previously mentioned project, we would like to evaluate the significance of the Gaussian signal. To do so, we define the test statistic as the profile likelihood ratio[17]

$$t = -2\ln\left(\frac{\mathcal{L}(\hat{\lambda},\hat{\theta})}{\mathcal{L}(0,\hat{\hat{\theta}}(0))}\right) \tag{3.8}$$

where the numerator is the maximized likelihood with the Gaussian signal rate free, the denominator is the maximized likelihood with the Gaussian signal rate fixed to zero. Our task is to obtain its distribution under null hypothesis assumption and to obtain its value corresponding to the data histogram.

First we need to add two lines to the main function. The modified part reads like this:

[17] G. Cowan et al., “Asymptotic formulae for likelihood-based tests of new physics”, *European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011, DOI: `10.1140/epjc/s10052-011-1554-0`, arXiv: `1007.1727`

```
PrepareData *data = new PrepareData();
SimpleFit *fit = new SimpleFit();
DiscoveryTest *discovery = new DiscoveryTest();

ana->registerModule(inputManager);
ana->registerModule(data);
ana->registerModule(fit);
ana->registerModule(discovery);
ana->registerModule(outManager);
```

An example project is provided in the `GooStats/Modules/statAnalysis` folder.

After that we run the following command to obtain the test statistic value corresponding to the data histogram:

```
./stat toyMC.cfg out exposure=500 DiscoveryTest=default.gaus
```

The log-likelihood-ratio can be retrieved in the `CERN ROOT terminal` using

```
root [0] fit_results->Scan("likelihood[1]-likelihood[0]");
```

At last we run the following command to obtain the test statistic distribution under null hypothesis assumption using the Monte Carlo method:

```
./stat toyMC.cfg nullpdf exposure=500 DiscoveryTest=default.gaus
    fitFakeData=true repeat=100
```

The distribution can be retrieved in the `CERN ROOT terminal` using

```
root [0]
    fit_results->Draw("likelihood[1]-likelihood[0]","minim_conv[0]&&minim_conv[1]");
```

Medium baseline neutrino mass ordering determination The neutrino mass ordering can be determined in a medium baseline experiment with reactor anti-electron neutrinos, by considering the difference between two neutrino mass ordering on the 3-neutrino vacuum oscillation inference. By fitting the energy spectrum with normal ordering and inverted ordering sequentially then taking the difference of the χ^2 of two fits, the inverted or normal ordering hypothesis can be rejected at certain confidence level depending on the magnitude of the absolute value of the $\Delta\chi^2$.

This project is more complex, because built-in classes are not enough to describe the oscillated reactor anti-neutrino spectrum if we would like to leave the neutrino oscillation parameters free. A customized `SpectrumBuilder` class, which constructs the probability density function object, and a customized `DatasetController` class, which collects needed information, are needed. Besides, three lines need to be modified in the `main` function:

```
InputBuilder *builder = new ReactorInputBuilder();
```

```
  builder->installSpectrumBuilder(new ReactorSpectrumBuilder());
  inputManager->setInputBuilder(builder);
```

An example is given in the `GooStats/Modules/naive-Reactor` folder.

Besides, the statistical analysis also need to be customized. We need to fit assuming normal neutrino mass ordering first, then fit again assuming inverted neutrino mass ordering. To do so, we can write a new `StatModule` class called `NMOTest`:

```
bool NMOTest::run(int) {
  if(!GlobalOption()->hasAndYes("fitNMO")) return true;
  auto deltaM2s =
      getInputManager()->Datasets().front()->get<std::vector<Variable*>>("deltaM2s");
  deltaM2s[1]->value = - deltaM2s[1]->value;
  deltaM2s[1]->lowerlimit = - deltaM2s[1]->upperlimit;
  deltaM2s[1]->upperlimit = - deltaM2s[1]->lowerlimit;
  getGSFitManager()->run(0);
  getOutputManager()->subFit(0);
  deltaM2s[1]->value = - deltaM2s[1]->value;
  deltaM2s[1]->lowerlimit = - deltaM2s[1]->upperlimit;
  deltaM2s[1]->upperlimit = - deltaM2s[1]->lowerlimit;
  getGSFitManager()->run(0);
  getOutputManager()->subFit(0);
  return true;
}
```

and register it in the `main` function:

```
  NMOTest *nmo = new NMOTest();
  ...
  ana->registerModule(nmo);
```

An example fit of the normal ordering spectrum with normal ordering hypothesis is shown in Figure 3.17.

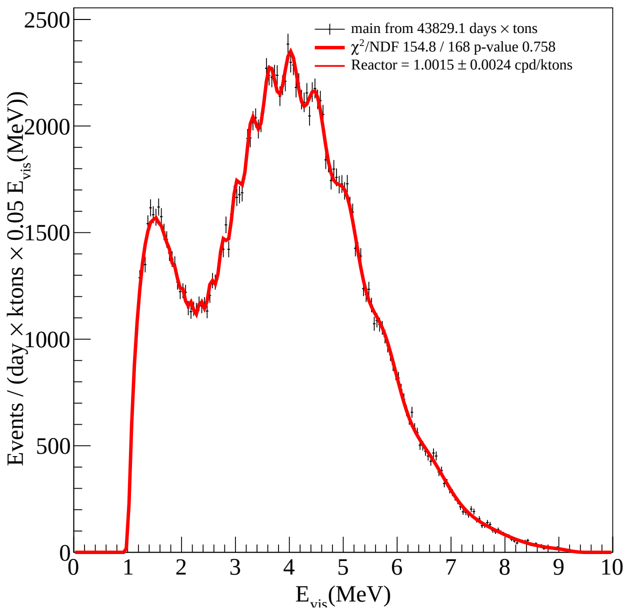


Figure 3.17: Example output of `naive-Reactor` module.

4

Characterizing and Modeling of detector energy response

In the previous chapter we discussed the method of fitting, and in this chapter we will build the model of detector response and discuss its accuracy. The models of theoretical shapes can also introduce systematic uncertainties, yet only that of the ^{210}Bi backgrounds and pile-up events need independent discussions, and for all other components either their shapes are well known, or their contribution are so small that the introduced systematic uncertainties are negligible. In Phase-I analysis analytical response functions for `npmt` and `charge` energy estimators have been developed[1]. In Phase-II analyasis, in order to achieve 3% measurement of ^{7}Be solar neutrino flux and to perform a global fit of all low energy solar neutrinos covering from 200 keV to 3 MeV energy range, new analytical response functions were developed for both `npmt_dt1` and `charge`[2]. For `npmt_dt1` higher order terms were added to enlarge the accurate energy range and proper powers of scaling factors due to PMT loss $f_{\text{eq.}}$ were added to keep parameters constant over time. For `charge` the non-linearity related to liquid scintillator and that related to electronics and clustering were separated and the resolution model was also improved. Compared with previous studies[3], in this work I emphasized the model of the resolutions and evaluated the bias of extracted rates.

This chapter is organized as in the following: Section 4.1 describes the processes that introduce non-linearity and smearing. Section 4.2 builds detector response models according to the probability theory and validates the models against Monte Carlo simulations. Section 4.3 presents the tuning of parameters against Monte Carlo and calibration data. In Section 4.4 I further discusses the reliability of the analytical models. The comparison between Monte Carlo and analytical response function strategies and the motivation of the latter are discussed in Section 3.1.3 and can be complementary materials to Section 4.4 here.

I have co-worked with Nicola Rossi, Oleg Smirnov, Simone Marcocci, Alina Vishneva and Zara Bagdasarian in this work. The model presented here has been used in [4,5,6].

---

[1] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 89, no. 11, pp. 1–68, 2014, DOI: `10.1103/PhysRevD.89.112007`, arXiv: `1308.0443` Sec. XVII

[2] Definition see Section 3.1.2

[3] ibid.

[4] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

[5] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II", pp. 1–8, July 2017, arXiv: `1707.09279`

[6] The Borexino collaboration, "Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data", *Physical Review D*, vol. 96, no. 9, p. 091103, Nov. 2017, DOI: `10.1103/PhysRevD.96.091103`, arXiv: `1707.09355`

4.1 Liquid scintillator detector principles

In Section 2.1 I have introduced the design of the whole Borexino detector. In this section I will focus on the center detector system using the main electronics, *LABEN*, and also cover the reconstruction, event selection and data histograms preparation.

Borexino detector measures the particle energy by counting the scintillation photons. When charged particles deposit energies in benzene-based liquid scintillator, mostly the π-electrons of the benzene ring of solvent molecules will be excited or ionized[7]. An interesting phenomena is that α particles induce more ionizations and thus the fraction of slow scintillation component would be higher, and this could be used to discriminate between α and β particles. See [8]. Borexino liquid scintillator consists of PC and 1.5 g/L PPO. Under such concentration, significant fraction[9,10] of energy of excited solvent molecules will be transferred to the PPO molecule through non-radiative dipole-dipole coupling. Then optical photons can be produced when these excited PPO molecules de-excites. Besides, photons will also be produced through Cherenkov radiation.

Absorption, Rayleigh scattering, Mie scattering, re-emission, Fresnel refraction and reflection might happen before the photon reach PMTs, and around half of these photons will be absorbed. Less than 30% of the remaining photons will be converted to electrons through photon-electron effect and most of them will be collected by the dynodes in the PMTs and then a pulse will be produced. Around 500 p.e. per 2000 PMTs will be produced for 1 MeV energy deposition.

Each pulse is collected by a front-end circuit first and divided into two signals, one integrated for measuring energies and the other amplified for measuring arrival time. The integrator output rises when the pulse arrives and decays exponentially to zero with a constant of 500 ns, and non-linearity from pile-up of pulses separated more than 140 ns[11] can be corrected properly[12]. The two signals are then fed to the digitizer board[13].

When trigger criteria[14] is satisfied, all hits in a gate of 6.9 μs or 16.5 μs are collected and wrapped into one event. After that the clusterization algorithm[15] is applied to separate pile-up physics events into clusters. Normalized sum of collected `charge`, or normalized number of fired PMTs in a cluster is treated as an estimate of the energy of the particle. See Sec. 3.1.2 for more descriptions for energy estimators.

Based on the reconstructed energy, interaction time and vertex of each event, part of events are removed, see Sec. 2.5. The energy of remaining events are filled into two histogram with cosmogenic ^{11}C backgrounds suppressed or enhanced by a projected likelihood based three-fold-coincidence algorithm[16]. Suppression of ^{210}Po α-decay backgrounds using a multi-layer perceptron algorithm has been considered for future studies.

---

[7] For a complete description of energy absorption see J. B. Birks, *The Theory and Practice of Scintillation Counting*, First Edit, Pergamon Press Ltd., 1964, p. 662 Chapter 2

[8] ibid. Sec. 3.7.1 pp. 56

[9] There is no direct measurement of transfer efficiency between pseudocumene and PPO in literature. Data for PPO-Xylene system can be found in I. B. Berlman, "Efficiency of Energy Transfer in a Solution of PPO in Xylene", *The Journal of Chemical Physics*, vol. 33, no. 4, pp. 1124–1127, Oct. 1960, DOI: `10.1063/1.1731345` . Some data for LAB-PPO system can be found on Mark Chen's slide on scintillation and sensitive detectors.

[10] According to Monte Carlo tuning, the product of such efficiency and PPO radiative de-excitation fraction is 82%. The Borexino collaboration, "The Monte Carlo simulation of the Borexino detector", *Astroparticle Physics*, vol. 97, pp. 136–159, Jan. 2018, DOI: `10.1016/j.astropartphys.2017.10.003`, arXiv: `1704.02291`

[11] D. D'Angelo, "Towards the detection of low energy solar neutrinos in BOREXino: data readout, data reconstruction and background identification", PhD, Munich, Tech. U., 2006, p. 334 pp. 107

[12] V. Lagomarsino and G. Testera, "A gateless charge integrator for Borexino energy measurement", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 430, no. 2-3, pp. 435–446, July 1999, DOI: `10.1016/S0168-9002(99)00170-9`

[13] F. Gatti et al., "The Borexino read out electronics and trigger system", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 461, no. 1-3, pp. 474–477, 2001, DOI: `10.1016/S0168-9002(00)01275-4`

[14] D'Angelo, "Towards the detection of low energy solar neutrinos in BOREXino: data readout, data reconstruction and background identification" Sec. 4.4.1 pp. 121

[15] ibid. Sec. 7.4.3 pp. 207

[16] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy" Sec. XV.A, pp. 32

4.2 *Analytical detector response function*

In the previous section I have introduced the processes involved with construction of the energy spectra used for analysis, and in this section I will explain how we describe these processes with analytical functions and discuss physics meaning of parameters used. The form of these functions were validated against a GEANT4 based Monte Carlo code `g4bx2`[17]. In the next section I will discuss the tuning of the parameters used against `g4bx2` and crosscheck with calibration data.

4.2.1 *Methodology of analytical response function*

A detector response function is the probability density function of an energy estimator given the kinetic energy of the particles. With such function, the distribution of an energy estimator can be predicted given the energy spectrum:

$$g(N) = \int f(E) \cdot R^{\text{det.}}(N;\ E) \cdot \mathrm{d}E \tag{4.1}$$

where $R^{\text{det.}}(N; E)$ is the detector response, $f(E)$ is the energy spectrum and $f(N)$ is the energy estimator spectrum.

In Borexino we have replaced the actual detector response with an analytical function $R^{\text{theo.}}(N;\ E)$ using the momentum-based approximation method:

$$\begin{aligned} E[R^{\text{theo.}}(N;\ E)] &= E[R^{\text{det.}}(N;\ E)] && (4.2)\\ \text{Var}[R^{\text{theo.}}(N;\ E)] &= \text{Var}[R^{\text{det.}}(N;\ E)] && (4.3)\\ \text{Skewness}[R^{\text{theo.}}(N;\ E)] &= \text{Skewness}[R^{\text{det.}}(N;\ E)] && (4.4)\\ &\ldots \end{aligned}$$

If all degrees of momentum match, we can prove $R^{\text{det.}} \equiv R^{\text{theo.}}$ rigorously. In practice, we choose functions with limited number of degrees of freedom and with fixed higher degree of momentum in prior and match only first few momentums. In Borexino three functions were used: modified gaussian(m.g., for `charge`)[18], generalized gamma(g.g., for `charge`)[19] and scaled Poisson(s.p., for `npmt_dt1`)[20,21]:

$$\begin{aligned} R(N;\ E)_{\text{m.g.}} &= \frac{1}{\sqrt{2\pi}\sqrt{a+b\cdot N}} \exp\left(-\frac{(N-\mu+b)^2}{2(a+b\cdot N)}\right) && (4.5)\\ R(N;\ E)_{\text{g.g.}} &= \frac{2\beta^{\alpha} N^{2\alpha-1}}{\Gamma(\alpha)} \exp\left(-\beta N^2\right) && (4.6)\\ R(N;\ E)_{\text{s.p.}} &= p(\frac{N}{s}; \frac{\mu}{s}) && (4.7) \end{aligned}$$

Parameters (a, b, μ), (β, α), (μ, s) are calculated according to the predicted expectation and variance, and, for m.g., skewness. As a result, in the following sections I will focus on how to predict the expectation and variance according to the particle energy.

[17] The Borexino collaboration, "The Monte Carlo simulation of the Borexino detector"

[18] R. N. Saldanha, "Precision Measurement of the 7 Be Solar Neutrino Interaction Rate in Borexino", PhD thesis, Princeton University, 2012 Sec. 5.5 pp. 201 Eq. (5.1.3)

[19] O. J. Smirnov, "An approximation of the ideal scintillation detector line shape with a generalized gamma distribution", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 595, no. 2, pp. 410–418, 2008, DOI: `10.1016/j.nima.2008.07.139`

[20] G. Bohm and G. Zech, "Statistics of weighted Poisson events and its applications", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 748, pp. 1–6, 2014, DOI: `10.1016/j.nima.2014.02.021`, arXiv: `1309.1287`

[21] Oleg Smirnov, "Analytical fit", *Borexino internal note*

4.2.2 e^- detector response

Intermediate variables In order to build the analytical detector response model, we define a few useful intermediate energy estimators that could not be observed.

1. The first one is the kinetic energy of electrons E_{kin}.
2. The second one is the number of photons produced L. Its yield follows the energy conservation law and Birk's law[22]:

$$\begin{aligned} E_{\text{kin}} &= \int_0^l \left[\left(\frac{\mathrm{d}E}{\mathrm{d}x} \right)_{\text{sci.}} + \left(\frac{\mathrm{d}E}{\mathrm{d}x} \right)_{\text{Cher.}} \right] \mathrm{d}x \\ \frac{\mathrm{d}L}{\mathrm{d}\omega} &= \int_0^l \left[k_{\text{qch.}} \cdot \left(\frac{\mathrm{d}^2E}{\mathrm{d}x\mathrm{d}\omega} \right)_{\text{sci.}} + \left(\frac{\mathrm{d}^2E}{\mathrm{d}x\mathrm{d}\omega} \right)_{\text{Cher.}} \right] \mathrm{d}x \\ k_{\text{qch.}} &= \left[1 + k_\beta \cdot \left(\frac{\mathrm{d}E}{\mathrm{d}x} \right)_{\text{sci.}} + k_\alpha \cdot \left(\frac{\mathrm{d}E}{\mathrm{d}x} \right)^2_{\text{sci.}} \right]^{-1} \end{aligned} \tag{4.8}$$

   where *sci.* denotes contribution related to scintillation and *Cher.* denotes contribution related to Cherenkov radiation; L is the number of photons produced, ω is the frequency of photons, $k_{\text{qch.}}$ is the quenching factor and k_β and k_α are parameters relevant to β and α particles, respectively.

   For its expectation, we use the following empirical formulae[23]:

$$\begin{aligned} \mu_{L_{\text{sci.}}} &= \mu_{L_{\text{sci.},0}} \cdot f_{\text{qch.}} \\ f_{\text{qch.}} &= \frac{q_1 + q_2 \ln(E) + q_3 \ln(E)^2}{1 + q_4 \ln(E) + q_5 \ln(E)^2} \end{aligned} \tag{4.9}$$

$$\begin{aligned} \mu_{L_{\text{Cher.}}} &= \Big(A_0 + A_1 \ln(E) + A_2 \ln^2(E) + A_4 \ln^3(E) \Big) \\ &\quad \cdot (1 + A_3 \cdot E) \cdot \Theta(E - E_{\text{thres.}}) \end{aligned} \tag{4.10}$$

   where E is in the unit of MeV, $\mu_{L_{\text{sci.},0}}$ is the scintillation photon yield at 1 MeV; $\text{Pol3}(x)$ is a cubic polynomial and $\Theta(x)$ is the Heaviside step function.

   For its variance, due to quenching effect, the scintillation photon yield is not fully bounded to the energy conservation law and depends on how fast the particle deposits energies along its trajectory, so additional smearing is introduced, and this can be seen with simulation results shown in Fig. 4.1. This effect is difficult to verify experimentally as it's very small

$$\frac{\sqrt{\Delta \text{var}_L}}{\mu_L} = (0.397 \pm 0.006)\%$$

   and will be more significant when the light yield is higher. Besides, the fraction of Cherenkov radiation loss will further modify the yield and the photon wavelength distribution, and the latter affects the yield of photon-electrons.

3. The third one is the number of photon-electron collected on the first dynode of PMTs for localized events $N_{\text{p.e.}}(\mathbf{r})$, or, `npe`. $\mathbf{r}$ here

---

[22] Birks, *The Theory and Practice of Scintillation Counting* Eq. (6.12) pp. 194

[23] For quenching Eq. (4.9) see J. Los Arcos and F. Ortiz, "kB: a code to determine the ionization quenching function Q(E) as a function of the kB parameter", *Computer Physics Communications*, vol. 103, no. 1, pp. 83–94, June 1997, DOI: `10.1016/S0010-4655(97)00012-X`. For Cherenkov Eq. (4.10) is chosen to describe the simulated results.

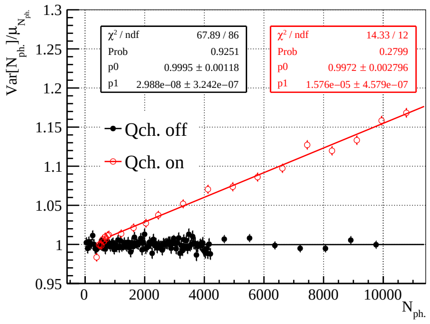


Figure 4.1: Variance/Expectation vs Expectation for the number of produced scintillation photon $L_{\text{sci.}}$ given by `g4bx2`. Black solid: quenching effect turned off ($k_\beta = 0$). The Poisson statistics is followed. Red circle: quenching effect turned on. Additional smearing due to variation of ways of losing energies can be observed. Sample below detection threshold (100 keV) are not included in the fit.

is to emphasize that it depends on the position of the event in the detector. Given the number of photons, $N_{\text{p.e.}}(\mathbf{r})$ follows binomial distribution, and because only around 5% of photons are converted to photon-electrons, it approximately follows a Poisson distribution.

First we discuss how to connect its expectation with that of photon yield. Spatial non-uniformity is introduced by the optical propagation including absorption, Rayleigh scattering, re-emission etc. and the photon-electron processes:

$$\mu_{\text{p.e.}} = \frac{\Omega(\mathbf{r})}{4\pi} \left(\mu_{L_{\text{sci.}}} \cdot \varepsilon_{\text{sci.}}(\mathbf{r}) + \mu_{L_{\text{Cher.}}} \cdot \varepsilon_{\text{Cher.}}(\mathbf{r})\right) \tag{4.11}$$

where Ω is the effective collection solid angle, $\varepsilon_{\text{sci.}}$ includes the average re-emission probability and the average quantum efficiencies, and $\mathbf{r}$ is the event vertex.

We can see that there is energy dependence of the spatial non-uniformity from Eq. (4.11), but we have ignored it in this work. Such dependence comes from the non-trivial difference between $\varepsilon_{\text{sci.}}(\mathbf{r})$ and $\varepsilon_{\text{Cher.}}(\mathbf{r})$: the scaling factor $\varepsilon(\mathbf{r})$ can be approximately written as $\varepsilon \cdot e^{-l(\mathbf{r})/L^{\text{att.}}}$, where ε includes the average re-emission probability and the average quantum efficiencies, $L^{\text{att.}}$ is the average attenuation length; because the photon wavelength distribution are different between scintillation and Cherenkov photons, $L^{\text{att.}}_{\text{sci.}}$ is different from $L^{\text{att.}}_{\text{Cher.}}$. We ignored such a difference because the contribution of Cherenkov is small and for solar neutrinos we only analyze events in the central region. With such simplification, Eq. (4.11) can be written as

$$\mu_{\text{p.e.}} = \varepsilon(\mathbf{r}) \cdot Y_{\text{p.e.}} \cdot \left(f_{\text{qch.}}(E) + \mu_{L_{\text{Cher.}}}(E) \cdot \text{fCher}\right) \tag{4.12}$$

where $\varepsilon(\mathbf{r})$ describes the non-uniformity and is normalized to 1 at the center, fCher is a relative coefficient and independent of vertex, $Y_{\text{p.e.}}$ is the photon-electron yield, $f_{\text{qch.}}$ and $\mu_{L_{\text{Cher.}}}$ are defined in Eq. (4.9) and (4.10). For uniform events, averaged $\varepsilon(\mathbf{r})$ is absorbed into $Y_{\text{p.e.}}$.

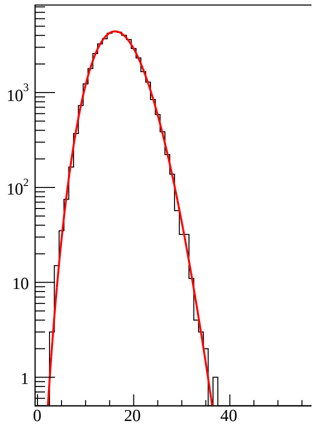


Figure 4.2: Distribution of $N_{\text{p.e.}}$ for simulated E_k 0.03 MeV electrons fitted with Poisson distribution. Here quenching effect is turned off ($k_\beta \equiv 0$) and the energy is below the threshold of the Cherenkov radiation, so the collected p.e. follows Poisson distribution.

Then let's consider the variance of $N_{\text{p.e.}}(\mathbf{r})$. For point like events, when Cherenkov contribution is small and quenching effect of scintillation is turned off, the number of photon-electron follows Poisson distribution, see Fig. 4.2;

When either of them exists, the variance deviates from expectation with observable magnitude, see Fig. 4.3 and 4.4, and it can be expressed as

$$\mathrm{Var}(N_{\mathrm{p.e.}}) = \mu_{\mathrm{p.e.}} + \left(\varepsilon^2_{\mathrm{qch.}} + \varepsilon^2_{\mathrm{Cher.}}\right) \cdot \mu^2_{\mathrm{p.e.}} \tag{4.13}$$

From Fig. 4.4 we can see without quenching effect the smearing due to Cherenkov radiation etc. is

$$\frac{\sqrt{\Delta \mathrm{var}_L}}{\mu_L} = (0.73 \pm 0.02)\%.$$

Notice that the corresponding smearing of `npe` from quenching effect is $(0.58 \pm 0.05)\%$ and is incompatible with that of photons $(0.397 \pm 0.006)\%$, which is not understood yet.

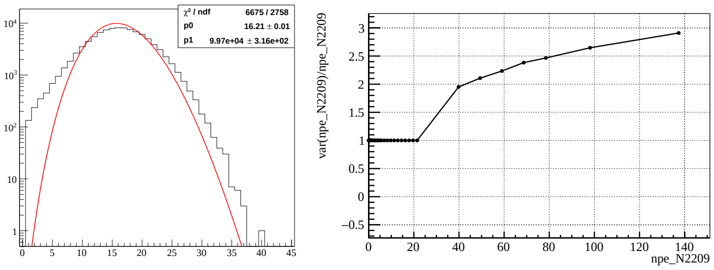


Figure 4.3: Distribution and Variance/-Expectation vs Expectation for the number of Cherenkov photons. Left: Distribution of $N_{\mathrm{p.e.}}$ fitted to Poisson distribution for simulated E_k 0.85 MeV electrons. Right: $N_{\mathrm{p.e.}}$'s variance/expectation versus its expectation. Below 20 p.e. no Cherenkov p.e. is collected and the ratio is set to 1 by hand. From both figures we can see that $N_{\mathrm{p.e.\,Cher.}}$ does not follow Poisson distribution.

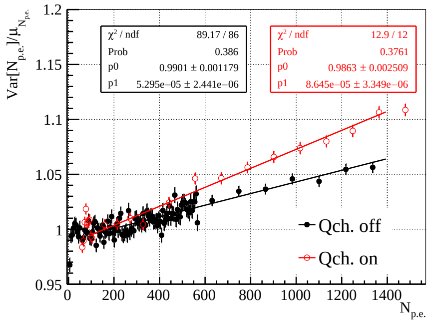


Figure 4.4: Contribution of Cherenkov radiation and quenching effect to the variance. Black solid: quenching effect turned off ($k_\beta = 0$). Red circle: quenching effect turned on. Simulated center events. Only samples in [100,1400] p.e. are used for fitting. From the figure we can see that both quenching effect and Cherenkov photons contribute around 0.6% smearing.

The `npmt` variables The `npmt` variable N_{fired}, introduced in Sec. 3.1.2, can be considered as the sum of Bernoulli distributed random variables f_i, which is 0 or 1 depending on if the i-th PMT is fired:

$$N_{\mathrm{fired}} = \sum_{i=1}^{N^t_{\mathrm{live}}} f_i \tag{4.14}$$

$$M_{N_{\mathrm{fired}}} = \prod_{i=1}^{N^t_{\mathrm{live}}} \left(1 - p_i + p_i \cdot e^t\right) \tag{4.15}$$

where N^t_{PMT} is the number of channels that can receive timing signals and registering hits, $M_{N_{\mathrm{fired}}}$ is the momentum generation function[24], p_i is the overall firing probability of i-th PMT.

---

[24] See wiki page for the definition of the "momentum generation function".

The expectation and variance of N_{fired} can be calculated using properties of the momentum generation function:

$$\mu_{\text{fired}} = \sum_{i=1}^{N^t_{\text{live}}} p_i \tag{4.16}$$

$$\text{Var}(N_{\text{fired}}) = \mu_{\text{fired}} - (1 + v_1) \cdot \mu^2_{\text{fired}} \tag{4.17}$$

where v_1 is almost the non-uniformity of firing probability p_i and is defined as

$$v_1 = \frac{\text{Var}[p_i]}{\mu^2_{\text{fired}}} - \frac{1}{(N^t_{\text{live}})^2} \tag{4.18}$$

$$\text{Var}[p_i] = \sum (p_i - \overline{p_i})^2 \tag{4.19}$$

Next let's simplify the expression of μ_{fired}. The firing probability p_i, considering loss of hits due to low amplitude, is

$$p_i = 1 - P_i(0) - P_i(1) \cdot p_{\text{ser},i} \tag{4.20}$$

where $p_{\text{ser},i}$ is the probability of a hit being rejected due to its amplitude lower than the front-end discriminator threshold, $P(0)$ and $P(1)$ are the average probability of detecting 0 and 1 photon-electron on one PMT, respectively.

$P(0)$ and $P(1)$ depend on the number of photon-electrons collected by i-th PMT $N_{\text{p.e.},i}$, and $N_{\text{p.e.},i}$ follows binomial distribution with known total number of photon-electrons $N_{\text{p.e.}}$:

$$P_i(0) = (1 - \varepsilon_i(\mathbf{r}))^{N_{\text{p.e.}}} \tag{4.21}$$

$$= 1 + \ln(1 - \varepsilon_i(\mathbf{r})) \cdot N_{\text{p.e.}} + \frac{\ln^2(1 - \varepsilon_i(\mathbf{r}))}{2} \cdot N^2_{\text{p.e.}} + \ldots \tag{4.22}$$

$$P_i(1) = N_{\text{p.e.}} \cdot \varepsilon_i(\mathbf{r}) \cdot (1 - \varepsilon_i(\mathbf{r}))^{N_{\text{p.e.}}-1} \tag{4.23}$$

$$= \frac{\varepsilon_i(\mathbf{r})}{1 - \varepsilon_i(\mathbf{r})} \cdot N_{\text{p.e.}} + \frac{\varepsilon_i(\mathbf{r}) \cdot \ln(1 - \varepsilon_i(\mathbf{r}))}{1 - \varepsilon_i(\mathbf{r})} N^2_{\text{p.e.}} + \ldots \tag{4.24}$$

where ε_i is fraction of photon-electrons that will be collected by i-th PMT and

$$\varepsilon_i = \frac{\epsilon_i}{\sum \epsilon_i} \tag{4.25}$$

where ϵ_i is the probability of a photon being converted to photon-electrons on i-th PMT.

Insert Eq. (4.22) and (4.24) into Eq. (4.20) we have

$$p_i = \left((1 - p_{\text{ser.}}) \cdot \varepsilon_i(\mathbf{r}) + (\frac{1}{2} - p_{\text{ser.}}) \cdot \varepsilon_i(\mathbf{r})^2 \right) \cdot N_{\text{p.e.}}$$
$$- \left(\frac{1}{2} - p_{\text{ser.}} \right) \cdot \varepsilon_i(\mathbf{r})^2 \cdot N^2_{\text{p.e.}} + \mathcal{O}(\varepsilon_i(\mathbf{r})^3)) \tag{4.26}$$

$$\mu_{\text{fired}} \approx (1 - p_{\text{ser.}} + v_2) \cdot N_{\text{p.e.}} - v_2 \cdot N^2_{\text{p.e.}} \tag{4.27}$$

where in the second equation we have utilized the fact that $\sum \varepsilon = 1$ and v_2 is defined as

$$v_2 = \frac{1}{2} \left(\frac{1}{(N^t_{\text{live}})^2} + v'_2 \right) \cdot (1 - 2p_{\text{ser.}}) \tag{4.28}$$

$$v'_2 = \sum (\varepsilon_i - \overline{\varepsilon_i})^2 \tag{4.29}$$

Eq. (4.27) reflects that the `npmt−npe` dependence for localized events has two degrees of freedom, one related to the $p_{\text{ser.}}$ and the other related to v_2 defined in Eq. (4.28).

In analysis, we use the normalized version of `npmt_dtl` with the number of live channels normalized to a fixed number to remove smearing due to shift of energy scale induced by PMT loss:

$$\mu_{\text{n.fired}} = \frac{N_{\text{n.live}}}{N^t_{\text{live}}} \cdot \mu_{\text{fired}} \tag{4.30}$$

where $N_{\text{n.live}}$ is a chosen number and 2000 in Borexino, N^t_{live} is the number of channels that can register hits.

One expression of $\mu_{\text{n.fired}}$ useful for later discussion on $\mu^{\text{FV}}_{\text{n.fired}}$ (pp. 71) is the following:

$$\begin{aligned} r &= \mu_{\text{n.p.e.}} / N_{\text{n.live}} \\ \mu_{\text{n.fired}} &= \varepsilon_{\text{sc.}} \cdot N_{\text{n.live}} \cdot \left[1 - (1 + p_{\text{ser.}} \cdot r) \cdot e^{-r}\right] \qquad (4.31) \\ \frac{\mu_{N_{\text{n.fired}}}}{\mu_{N_{\text{p.e.}}}} &= \varepsilon_{\text{sc.}} \cdot (1 - p_{\text{ser.}}) - \varepsilon_{\text{sc.}} \left(\frac{1}{2} - p_{\text{ser.}}\right) \cdot r + \mathcal{O}(r^2) \qquad (4.32) \end{aligned}$$

where $\varepsilon_{\text{sc.}}$ is a scaling factor, r is defined as $\mu_{N_{\text{p.e.}}} / N_{\text{PMT}}$. Fitting with such formula is shown in Fig. 4.5.

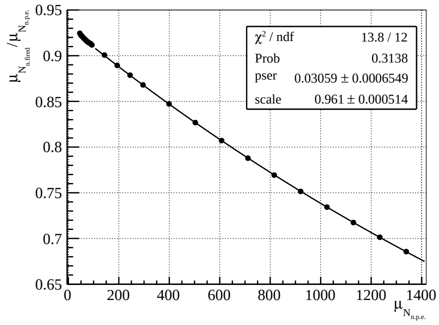


Figure 4.5: `npmt−npe` dependence based on simulation of center events according to run 17407. `npmt_dtl` is used here.

The determined value in Fig. 4.5 are listed as the following:

$$\begin{aligned} \varepsilon_{\text{sc.}} &= 0.9610 \pm 0.0006 \qquad (4.33) \\ p^{\text{Phase−I}}_{\text{ser.}} &= 0.0306 \pm 0.0007 \qquad (4.34) \end{aligned}$$

and we have

$$\lim_{\mu_{\text{n.p.e.}} \to 0} \mu_{\text{n.fired}} / \mu_{\text{n.p.e.}} = 0.9316 \pm 0.0008.$$

With this we can get two conclusions:

First, if we fix $\varepsilon_{\text{sc.}}$ to 1, fCher and Y will absorb it and become effective parameters. We can see that if we fix it to 1, we would have

$$\lim_{\mu_{\text{n.p.e.}} \to 0} \mu_{\text{n.fired}} / \mu_{\text{n.p.e.}} = 1.$$

and 7% in $\mu_{\text{n.fired}} / \mu_{\text{n.p.e.}}$ is missing. The light yield Y alone cannot make up such discrepancy because of the non-linear dependence of $\mu_{\text{n.fired}}$ on $\mu_{\text{n.p.e.}}$.

Second, $p_{\text{ser.}}^{\text{Phase-I}}$ is an effective parameter and not really the probability of a hit being rejected due to its low amplitude. We can see this by comparing Eq. (4.27) and (4.32). We have

$$p_{\text{ser.}} = (6.86 \pm 0.08) \times 10^{-2} \tag{4.35}$$
$$v_2 = (2.256 \pm 0.004) \times 10^{-4} \tag{4.36}$$

As you can see $p_{\text{ser.}}^{\text{Phase-I}}$ is only half of $p_{\text{ser.}}$. Actually $p_{\text{ser.}}$ is also too small compared with the measured probability 12%[25].

If we average it over the fiducial volume, the situation becomes much more complex, as the energy dependence of non-uniformity is so strong that it could not be ignored. We can use the following empirical formulae tuned on Monte Carlo to express its expectation:

$$r = \mu_{\text{n.p.e.}} / N_{\text{n.live}}$$
$$\frac{\mu_{\text{n.fired}}^{\text{FV}}}{N_{\text{PMT}}} = \left[1 - (1 + p_{\text{ser}} \cdot r) \cdot e^{-r}\right] \cdot \left(1 - g_{\text{LC}} \cdot r + g'_{\text{LC}} \cdot r^2\right) \tag{4.37}$$
$$\begin{aligned}\frac{\mu_{\text{n.fired}}^{\text{FV}}}{\mu_{N_{\text{p.e.}}}} &= (1 - p_{\text{ser.}}) - \left(g_{\text{LC}}(1 - p_{\text{ser.}}) + (\frac{1}{2} - p_{\text{ser.}})\right) \cdot r \\ &+ \left(g_{\text{LC}}(\frac{1}{2} - p_{\text{ser.}}) + g'_{\text{LC}}(1 - p_{\text{ser.}}) + (\frac{1}{6} - \frac{p_{\text{ser.}}}{2})\right) \cdot r^2 \\ &+ \mathcal{O}(r^3)\end{aligned} \tag{4.38}$$

where $\mu_{\text{p.e.}}$ is given by Eq. (4.12), $p_{\text{ser.}}$ is the probability of being rejected due to low charge for single p.e. hits, g_{LC} and g'_{LC} counts for less fired PMTs due to saturation in the periphery region. The fit of such formula against Monte Carlo is shown in Fig. 4.7. Notice that this time

$$\lim_{\mu_{\text{n.p.e.}} \to 0} \mu_{\text{n.fired}} / \mu_{\text{n.p.e.}} = 0.9336 \pm 0.0002.$$

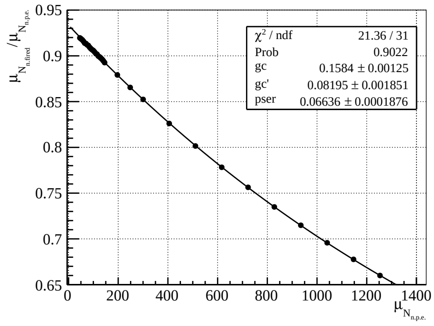


Figure 4.6: $\mu_{\text{n.fired}}^{\text{FV}}$–`npe` dependence based on simulation of uniform events in ^{7}Be fiducial volume according to run 17407. `npmt_dt1` is used here.

In[26], in order to remove the discrepancy on fCher between value from fitting spectrum and that from fitting simulated mono-energetic electrons, the following equation was been proposed:

$$\frac{\mu_{\text{n.fired}}^{\text{FV}}}{\mu_{N_{\text{p.e.}}}} = \varepsilon_{\text{sc.}} \cdot \left[1 - (1 + p_{\text{ser}} \cdot r) \cdot e^{-r}\right] \cdot (1 - g_{\text{LC}} \cdot r) \tag{4.39}$$

where $\varepsilon_{\text{sc.}}$ was inspired by Eq. (4.31) and fixed to the value obtained by fitting point-like events. However, such model cannot describe the data well and the goodness of fitting simulated mono-energetic

[25] A. Ianni et al., "The measurements of 2200 ETL9351 type photomultipliers for the Borexino experiment with the photomultiplier testing facility at LNGS", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 537, no. 3, pp. 683–697, 2005, DOI: `10.1016/j.nima.2004.07.249`, arXiv: `0406138 [physics]`

[26] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

electrons using such model was very bad. This was also confirmed by the results obtained with the method described by Sec. 4.4 that after introducing $\varepsilon_{\text{sc.}}$ the fitted fCher did not change and the bias of ^{210}Bi and ^{85}Kr became even larger. Finally the working group has decided to use Eq. (4.39) without $\varepsilon_{\text{sc.}}$.

For the variance, in[27] we use the following formulae tuned on `g4bx2`:

$$
\begin{aligned}
r_v &= \mu^{\text{FV}}_{\text{n.fired}} / N_{\text{n.live}} \\
f^t_{\text{eq.}} &= \text{average}\{2000/N^t_{\text{PMT}}\} \\
\text{Var}(N^{\text{FV}}_{\text{n.fired}}) &= f^t_{\text{eq.}} \cdot \mu^{\text{FV}}_{\text{n.fired}} \cdot [1 - r_v \cdot (1 + v_1)] \\
&\quad + \beta_0 \cdot \left(\mu^{\text{FV}}_{\text{n.fired}}\right)^3 \cdot \left(f^t_{\text{eq.}}\right)^{-1} \\
&\quad + \beta_1 \cdot \mu^{\text{FV}}_{\text{n.fired}} \cdot f^t_{\text{eq.}} \\
&\quad + \beta_2 \cdot (N_{\text{n.live}} \cdot (1 - r_v) \cdot \ln(1 - r_v))^2 \qquad (4.40)
\end{aligned}
$$

where N^t_{PMT} is the number of PMTs that can receive timing signals and register hits, v_1 accounts for variance of $p_{\text{ser.}}$ among channels, β_0 and β_2 accounts for the non-uniformity induced smearing, β_1 accounts for additional smearing due to quenching effect and Cherenkov contribution of electrons and is set to zero for ^{210}Po events.

The fit against `g4bx2` is shown in Fig. 4.7. We can see that the quenching contributes additional smearing at high energies, similar to the case of `npe`.

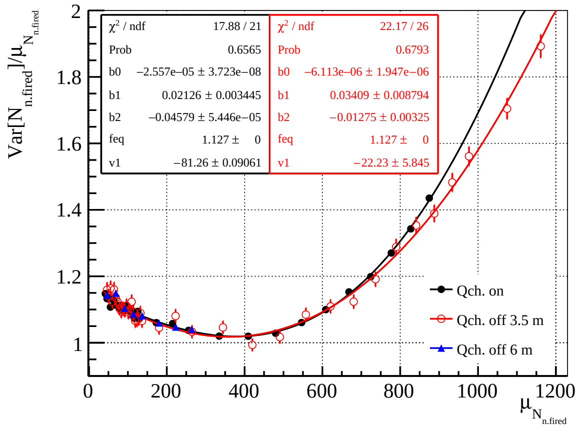


Figure 4.7: $\text{Var}(N^{\text{FV}}_{\text{n.fired}})$–$\mu^{\text{FV}}_{\text{n.fired}}$ dependence based on simulation of uniform events in ^{7}Be fiducial volume according to run 17407. `npmt_dt1` is used here. Fit range is [80, 900] `npmt_dt1`. The last point ($\mu_{\text{n.fired}} \sim 900$) correspond to 2.6 MeV electron events. Black solid: quenching off. Red circle: quenching on. Events are generated in a sphere of radius 3.5 meters. In order to address the effect of bad position reconstruction for low energy events, samples generated within 6 meters are also shown in blue solid triangle. As can be seen the position reconstruction does not modify the variance dependence significantly.

To understand the meaning of each parameter, we rewrite Eq. 4.40 as the following:

$$
\begin{aligned}
\frac{\text{Var}(N^{\text{FV}}_{\text{n.fired}})}{\mu_{\text{n.fired}}} &= f_{\text{eq.}}(1 + \beta_1) - \left(f_{\text{eq.}}(1 + v_1) - \beta'_2\right) r \\
&\quad + (f^{-1}_{\text{eq.}}\beta'_0 - \beta'_2) r^2 - \frac{1}{12}\beta'_2 r^3 + \mathcal{O}(r^5) \qquad (4.41) \\
&\equiv a_0 + a_1 r + a_2 r^2 + a_3 r^3 + \mathcal{O}(r^5) \qquad (4.42) \\
\beta'_0 &= \left(N^t_{\text{PMT}}\right)^2 \cdot \beta_0 \\
\beta'_2 &= N^t_{\text{PMT}} \cdot \beta_2
\end{aligned}
$$

[27] The Borexino collaboration, “First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II”

We can see that a_1 constrains β_2' when v_1 is fixed while a_3 constrains β_2', too. Unfortunately a_3 prefers different β_2' compared with that preferred by $(a_1 + \text{physical}\, v_1)$, and v_1 becomes far from its physics meaning. In[28] the working group decided to keep v_1 fixed to its physics value 0.16 and accept the introduced systematic uncertainties.

In order to inspire future work, we tried to fit Eq. (4.42) directly and listed the result in Tab. 4.1.

	Qch. on	Qch. off
a_0	1.161 ± 0.008	1.17 ± 0.02
a_1	-1.2 ± 0.2	-1.7 ± 0.3
a_2	1.3 ± 0.9	4.5 ± 1.3
a_3	6.5 ± 1.3	1.1 ± 1.5

Table 4.1: Fitted parameters of `npmt_dt1` variance model according to `g4bx2` using a third degree polynomial.

The `charge` variables Here we discuss the `charge` variables Q, also introduced in Sec. 3.1.2.

For localized events, `charge` can be approximately considered as the sum of $N_{\text{p.e.}}$ Gaussian distributed random variables, whose average is calibrated to be 1 and variance is the single p.e. variance v_1 and has been measured experimentally. For the variance of `charge`, see also Eq. (4.13), regardless of additional linear term introduced by single p.e. variance, the quadratic term of variance of charge is reduced from 0.7% in `npe` to around 0.4%, see Fig. 4.20. The mechanism of decreasing is not understood yet.

If we averaged it over the fiducial volume, it can be considered as the compound distribution of event vertex and the localized `charge`. Compared with localized `charge`, smearing due to non-uniformity is introduced and light yield is also modified.

To determine the expectation of `charge` from that of `npe`, we absorb all effects from the clustering and residual non-linearity from the `charge` integrator, the contribution of dark noise etc. into a second degree polynomial:

$$\mu^{\text{FV}}_{\text{n.Q}} = p_{\text{dn}} + (1 + p_{\text{mis.}}) \cdot \mu_{\text{n.p.e.}} + p_{\text{quadr.}} \left(\mu_{\text{n.p.e.}}\right)^2 \tag{4.43}$$

where $\mu_{\text{n.p.e.}}$ is given by Eq. (4.12). Here we do not include f^{Q}_{eq} because for both $N_{\text{n.p.e.}}$ and $N^{\text{FV}}_{\text{n.Q}}$ we have normalized the number of live channels to 2000. Fit of Eq. (4.43) against `g4bx2` is shown in Fig. 4.9. Two curve of the non-linearity of `charge` and non-uniformity corrected `charge` are compatible.

Now let's predict the magnitude of smearing introduced by non-uniformity. Due to the fact that less PMTs survived at bottom, the spatial non-uniformity of detector energy response is approximately a linear function of z in the fiducial volume, see Fig. 4.8:

$$\varepsilon(\mathbf{r}) = 1 + k(t) \cdot z \tag{4.44}$$

$$k = (2 \sim 4) \times 10^{-2}\,\text{m}^{-1} \tag{4.45}$$

where k increased over time and is measured with ^{210}Po α-decay background events selected by a multi-layer perceptron algorithm.

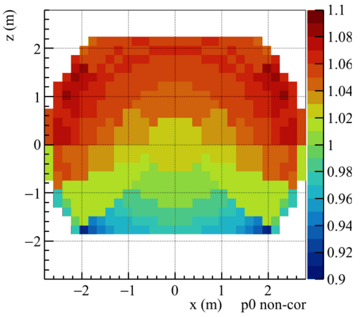


Figure 4.8: Energy non-uniformity in x-z plane measured with peak positions of ^{210}Po α-decay events from 2011 December 14th. to 2012 July 1st.

[28] ibid.

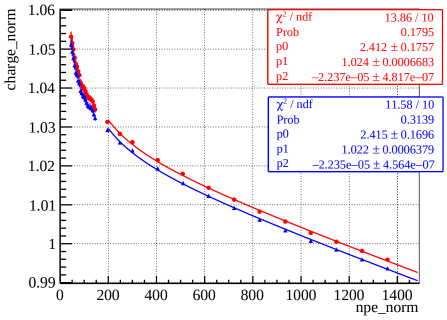


Figure 4.9: Comparison of non-linearity of `charge` (red, solid) and geo-corrected `charge` (blue, triangle) variable against photon-electrons for simulated uniform distributed events. Run 17407 is simulated.

Consider the definition of *pep* and ^{7}Be fiducial volume and using the sum rule of compound distribution we have

$$\begin{aligned}
\mathrm{Var}(N^{\mathrm{FV}}_{\mathrm{n.Q}}) &= E_{\mathbf{r}}[\mathrm{Var}[N_{\mathrm{n.Q}}](\mathbf{r}) + \mu_{\mathrm{n.Q}}(\mathbf{r})^2] - E_{\mathbf{r}}[\mu_{\mathrm{n.Q}}(\mathbf{r})]^2 \\
&= f^{\mathrm{Q}}_{\mathrm{eq.}} \cdot (1+v_1) \cdot \mu^{\mathrm{FV}}_{\mathrm{n.Q}} \\
&\qquad + \left(\sigma_v^2 + k^2 \cdot c'^2\right) \cdot \left(\mu^{\mathrm{FV}}_{\mathrm{n.Q}}\right)^2 \qquad (4.46) \\
&= f^{\mathrm{Q}}_{\mathrm{eq.}} \cdot (1+v_1) \cdot \mu^{\mathrm{FV}}_{\mathrm{n.Q}} + v_T \cdot \left(\mu^{\mathrm{FV}}_{\mathrm{n.Q}}\right)^2 \qquad (4.47) \\
c'(pep) &\approx 1.11\,\mathrm{m} \qquad (4.48) \\
c'(^7\mathrm{Be}) &\approx 0.845\,\mathrm{m} \qquad (4.49)
\end{aligned}$$

Notice that the definition of v_1 here is not the variance of the single p.e. charge distribution.

The simulated dependence between variance and expectation is shown in Fig. 4.10. σ_T for run 17407 is $(2.1 \pm 0.1)\%$, which is consistent with the predicted $kc'(^7\mathrm{Be}) = 2\%$. We can see that the smearing introduced by non-uniformity will contribute $\sim$ 30% of total variance when $\mu_{\mathrm{n.Q}}$ reaches 1000. For non-uniformity corrected variable $N^{\mathrm{FV}}_{\mathrm{n.Qgeo}}$, k is significantly reduced and $\sigma_T(N_{\mathrm{n.Qgeo}})$ is $(0.94 \pm 0.02)\%$.

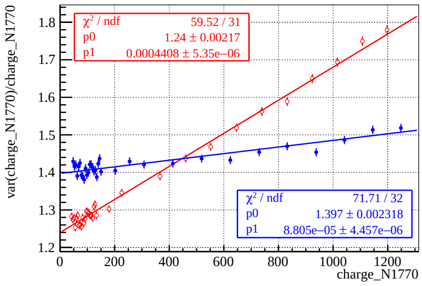


Figure 4.10: Comparison of variance of `charge` (red, circle) and geo-corrected `charge` (blue, solid dot) variable for simulated uniform distributed events. Run 17407 is simulated where $k \sim 2 \times 10^{-2}\ \mathrm{m}^{-1}$. p_0 is $f_{\mathrm{eq.}} \cdot (1 + v_1)$. $f_{\mathrm{eq.}}$ is 1 for `charge` and around 1.13 for geo-corrected `charge`.

4.2.3 γ *and* e^+ *detector response*

First let's discuss the non-linearities. A gamma transfers its energy to electrons through Compton scattering and photon-electron effect or to electrons and positrons through pair-production and then the energy is deposited through electrons and positrons, so the quenching effect and Cherenkov contribution is different from that of electrons. For positrons, their kinetic energy is deposited in a way similar to electrons while additional energy is deposited from two 0.511 MeV γs produced by annihilation. In solar neutrino analysis only ^{11}C is relevant and we have assigned one additional parameter describing the overall non-linearity of annihilation γs:

$$\mu_{N^{e^+}_{\text{p.e.}}} = \mu_{N^{e^-}_{\text{p.e.}}}(E^{e^+}_{\text{kin}}) + 2 \cdot \varepsilon^{\gamma}_{\text{NL.}}(0.511\,\text{MeV}) \cdot Y_{\text{p.e.}} \cdot E^{\gamma}_{\text{kin.}} \tag{4.50}$$

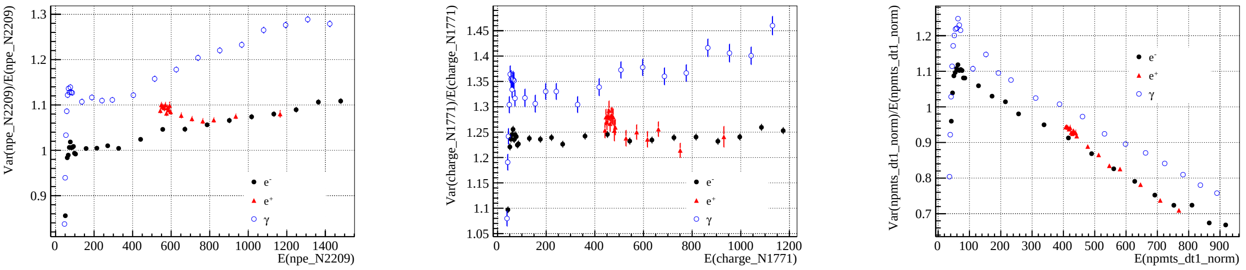
There are discussions about predicting γ's non-linearity using that of electrons and the primary electron energy distribution according to Monte Carlo simulation[29,30]. This can be used to fit the calibration γ peak positions as pull terms and connect their peak positions with the β analytical response function.

Second let's discuss the variance. Similar to the quenching effect, due to different ways of transferring energies to electrons, gammas have additional smearing. For positrons, the extra energy resolution coming from γs has no energy dependence:

$$\text{Var}[N^{\gamma}_{X}] = \text{Var}[N^{e^-}_{X}](\mu^{\gamma}_{X}) + \delta v^{X}_{1} \cdot \mu^{\gamma}_{X} + \delta v^{X}_{T} \cdot (\mu^{\gamma}_{X})^2 \tag{4.51}$$

$$\text{Var}[N^{e^+}_{X}] = \text{Var}[N^{e^-}_{X}](\mu^{e^+}_{X}) + 4 \cdot \delta\text{Var}[N^{\gamma-e^-}_{X}](0.511\,\text{MeV}) \tag{4.52}$$

where δ-marked term is related to the additional resolution, X can be `npe`, `charge` or `npmt`. See also Figure 4.11.



(a) (b) (c)

Figure 4.11: Comparison of variance dependence of different particles. Based on simulations of events at detector center. Black dot: electrons. Solid circle: gammas. Red triangle: positrons. (a): photon-electrons. (b): `charge`. (c): npmts. The decrease below 100 keV is not understood yet.

4.2.4 *External γ detector response*

Compared with uniform electron events, external γs are concentrated in the periphery regions and without non-uniformity correction their average light yield will be larger[31], and non-uniformity induced resolution, *i.e.* c' in Eq. (4.46), will be smaller. Non-uniformity correction can remove these two effects. At the same time, γ has additional resolution contributions. In Figure 4.12 it is shown the fit of simulated external γ spectrum using such models.

4.2.5 *$f_{\text{eq.}}$ and ^{210}Po α-decay detector response*

Like electrons, α particles deposit energy also through ionization and excitation etc., but with much larger dE/dx and thus have different non-linearity, and we have assigned a float parameter for the only relevant component ^{210}Po α-decay:

$$\mu_{N^{\alpha}_{\text{p.e.}}}(^{210}\text{Po}) = \varepsilon^{\alpha}_{\text{NL.}}(^{210}\text{Po}) \cdot Y_{\text{p.e.}} \cdot E^{\alpha}_{\text{kin.}} \tag{4.53}$$

---

[29] M. Yang and Y. Cheng, "Energy Non-linearity study at Daya Bay", *International Journal of Modern Physics: Conference Series*, vol. 31, p. 1460312, Mar. 2014, DOI: `10.1142/S2010194514603123`, arXiv: `1403.1550`

[30] In Chinese. Y. Zeyuan, "Study of the calibration, reconstruction and neutrino disappearance at Daya Bay neutrino experiment", PhD thesis, Chinese Acamedy of Science University, 2013 Chapter 3.5.2 pp. 67

[31] Scaling factor 1.0228 is introduced in Phase-I analysis, see A. E. Chavarria, "Study on pep and CNO solar neutrino interaction rates in Borexino", PhD thesis, Princeton University, 2012 pp. 91

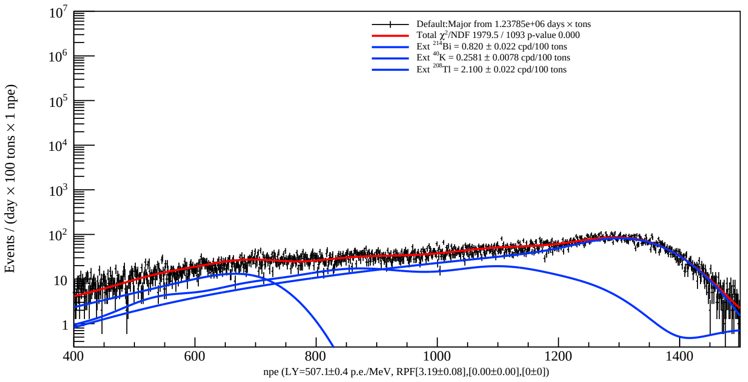


Figure 4.12: Fit on Monte Carlo external backgrounds. The light yield is scaled by 1.0228. The best fit and HESSA uncertainty of the response function is shown as the x axis title. As shown g_3, corresponding to v_T, rails to zero and g_1, corresponding to $1 + v_1$, is larger than expected.

For the variance, according to the `g4bx2` simulation, there is no significant additional contribution ($|\delta v_1| < 0.006, 90\%\text{C.L.}$):

$$\mathrm{Var}(N^{\alpha}_{\text{p.e.}}) = (1 + \delta v_1)\mu_{\text{p.e.}}$$

As a result, α has better energy resolution due to absence of smearing introduced by quenching and Cherenkov-radiation, i.e. $\left(\varepsilon^2_{\text{qch.}} + \varepsilon^2_{\text{Cher.}}\right)$ in Eq. (4.13).

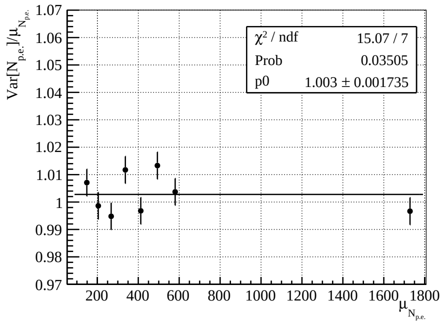


Figure 4.13: Variance/expectation vs expectation of `npe` for alpha particles events according to `g4bx2` simulation at detector center. From the figure we can see that unlike e^-, γs, α has no additional variance.

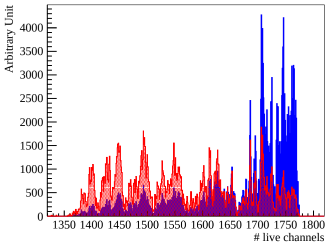


Figure 4.14: Distribution of number of live PMTs for stable events (red dotted) and decaying ^{210}Po events (blue solid). The data is collected during so called "PeriodAll" (see Sec. 6.1).

Besides, due to around 50 PMTs loss per year, the $f_{\text{eq.}}$ keeps growing from ~ 1.2 at the beginning of 2012 to ~ 1.8 now. For dataset spanning long time, we need to use the average of $f_{\text{eq.}}$ weighted by the number of events. Be careful that it's not $2000/\mathrm{avg}(N_{\text{live}})$. For ^{210}Po, due to its decay, corresponding averaged $f_{\text{eq.}}$ is smaller because the weight of early and small f_{eq} is higher.

4.2.6 *Pile-up and Dark Noise Convolution*

Pile-up is the effect that another physics event falls into the DAQ window of the main event and is not separated by the clustering algorithm. Pile-up events consisting of two ^{14}C events or one ^{14}C with another external γ are the crucial background of *pp* solar neutrinos[32]. In the analytical fit, the pile-up effect is either described by adding an independent component called synthetic pile-up[33], or convolving the probability density function with the random trigger spectrum, and the latter method is called the Dark Noise Convolution.

[32] S. Marcocci, "Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with Xe -loaded Liquid Scintillators", PhD thesis, Gran Sasso Science Insitute, 2016 pp. 192

[33] The Borexino collaboration, "Neutrinos from the primary proton–proton fusion process in the Sun", *Nature*, vol. 512, no. 7515, pp. 383–386, Aug. 2014, DOI: `10.1038/nature13702`, arXiv: `1508.05379`

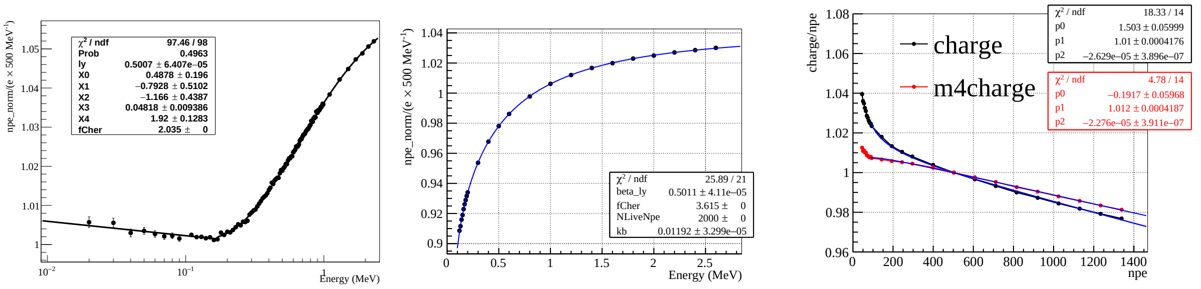


Figure 4.15: Left: tuning of Cherenkov parameters. Middle: tuning of quenching parameters. k_b is fitted instead of $q_1 ... q_5$. A program has been used to calculate qs from k_b based on stopping power. Right: tuning of electronics non-linearity parameters. Simulated electron events at detector center are used for tuning. The fCher in the left is not fixed to 3.615 because it's defined in a different way.

4.3 *Tuning against Monte Carlo and calibration data*

In the previous section we have built the analytical functions for the dependence of the expectation and variance of energy estimators and we have validated the trends against Monte Carlo simulation samples using `g4bx2`. In this section we will focus on the tuning of parameters in the models. We need to tune and fix part of parameters, because we have no sensitivity or they are correlated with the neutrino rates. The tuning against calibration data is also presented here.

4.3.1 *non-linearity*

The full non-linearity model for `charge` is given by Eq. (4.12), (4.9), (4.10) and (4.43). For npmt it's given by also Eq. (4.37). The strategy for parameters involved are summarized in Tab. 4.2.

Type	Symbol	Strategy
Light yield	$Y_{\mathrm{p.e.}}$	Free
Quenching par.	q_1 ... q_5	Fixed
Cherenkov par.	A_0 ... A_4	Fixed
relative eff.	fCher	Fixed
Electronics *etc.*	$p_{\mathrm{dn.}}$, $p_{\mathrm{mis.}}$, $p_{\mathrm{quadr.}}$	Fixed
Threshold eff.	$p_{\mathrm{ser.}}$	Fixed
Light collection	g_{LC}, g'_{LC}	Fixed
^{11}C γ qch.	$\varepsilon^{\gamma}_{\mathrm{NL.}}(0.511\,\mathrm{MeV})$	Free
Ext. γ qch.	-	Free / M.C. p.d.f.
^{210}Po α qch.	$\varepsilon^{\alpha}_{\mathrm{NL.}}(^{210}\mathrm{Po})$	Free

Table 4.2: Strategy for parameters in non-linearity models.

In[34], for `npmt` analysis, the quenching, Cherenkov parameters and $p_{\mathrm{ser.}}$ used in Phase-I analysis are used. g'_{LC} is not used and g_{LC} is tuned using simulation samples with no threshold and with $p_{\mathrm{ser.}}$ set to zero. fCher is fixed to the best fit value on neutrino spectrum with it free. For `charge` analysis, firstly Cherenkov parameters are tuned against samples with quenching turned off, see Fig. 4.15 left. Then quenching parameters and relative efficiency are fitted with Cherenkov shape fixed with quenching turned on, see Fig. 4.15 middle. The electronics related parameters are fitted independently, see Fig. 4.15 right. The values are summarized in Tab. 4.5 in Sec. 4.3.7.

[34] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

4.3.2 *Variance*

The full variance model for `charge` is given by (4.47). For npmt it's given by Eq. (4.40). The strategy for parameters involved are summarized in Tab. 4.3.

Type	Symbol	Strategy
PMT loss	$f_{\text{eq.}}$	Fixed
α var. model	-	see text below
extra var. for ext. γ	-	Free
$p_{\text{ser.}}$ variance	`npmt` v_1	Fixed
Empirical par.	`npmt` β_0 ... β_2	β_2 Fixed, others free
single p.e. var.	`charge` v_1	Free
Non-uniformity *etc.*	`charge` v_T	Fixed

Table 4.3: Strategy for parameters in variance models.

PMT loss correction factor is calculated according to the weight histogram shown in Fig. 4.14. In `npmt` analysis, independent β_0 and β_2 are used for α particles, and `g4bx2` based probability density function is used for external species. v_1 is given by the working group fixed according to a toy model and β_2 is tuned against `g4bx2`. In `charge` analysis same v_1 and $v_T(f_{\text{eq.}})$ (see Eq. (4.54)) are used for β and αs considering the conclusion given by `g4bx2` that the variance of ^{210}Po peak is almost the same as that of e^- at the ^{210}Po peak ($\sim$ 220 p.e.). Single p.e. variance v_1 is free and constrained by the ^{210}Po peak width. It is not fixed against `g4bx2` because its value in `g4bx2` is smaller than that in data. Non-uniformity variance v_T is fixed according to `g4bx2`, see Fig. 4.10. The values are also summarized in Tab. 4.5.

4.3.3 *Stability and trend of parameters*

The temporal trends of model parameters are needed when studying the trends of background rates in order to cross-check our understanding of the background. The non-linearity parameters, except the light yield, are treated as constants. According to `g4bx2`, the loss of PMTs and increasing non-uniformity does not induce observable change of non-linearity, see Fig. 4.16 left. Due to its importance, the light yield trend is studied dedicatedly based on data in Sec. 8.2. Resolution parameters trends based on `g4bx2` are shown in Fig. 4.16 middle and right. We have

$$2000 \cdot v_T(N_{\text{n.Q}}) = (6.01 \pm 0.06) \cdot (f_{\text{eq.}} - 1) \tag{4.54}$$

$$2000 \cdot v_T(N_{\text{n.Q-geo}}) = 0.184 \pm 0.011 \tag{4.55}$$

Consider some backgrounds, such as ^{85}Kr, can be non-uniform in the fiducial volume, and it's necessary to understand the uniformity of ^{210}Pb contamination concentration, it's also useful to get the temporal trend of model parameters for top-half and bottom-half of the fiducial volume. The trends of `charge` v_1 and v_T for top and bottom

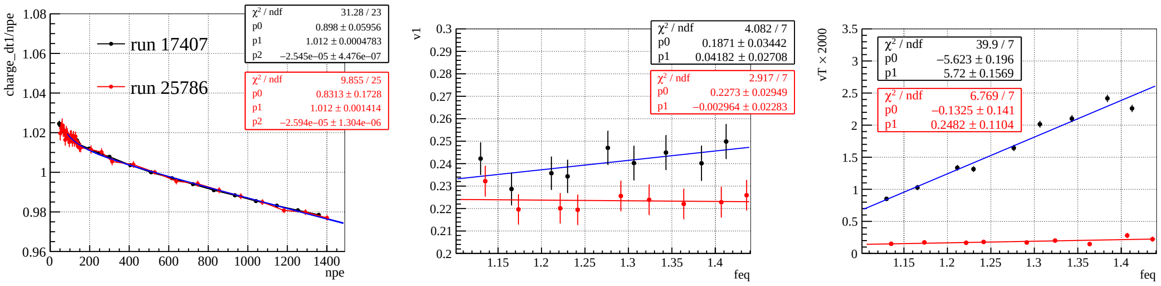


Figure 4.16: Left: comparison between run 17407 and run 25786 of electronics non-linearity. Middle: `charge` v_1 stability. Right: v_T stability. Black and red are without and with non-uniformity correction `charge`, respectively.

of ^{7}Be fiducial volume can be found in Fig. 4.17. We have

$$
\begin{aligned}
\text{Top} \quad 2000 \cdot v_T(N_{\text{n.Q}}) &= (2.5 \pm 0.2) \cdot (\text{feq} - 1) - (0.04 \pm 0.05) \\
\text{Bottom} \quad 2000 \cdot v_T(N_{\text{n.Q}}) &= (2.2 \pm 0.2) \cdot (\text{feq} - 1) + (0.25 \pm 0.06) \\
\text{Top} \quad 2000 \cdot v_T(N_{\text{n.Q-geo}}) &= (0.6 \pm 0.2) \cdot (\text{feq} - 1) + (0.08 \pm 0.05) \\
\text{Bottom} \quad 2000 \cdot v_T(N_{\text{n.Q-geo}}) &= 0.15 \pm 0.02
\end{aligned}
$$

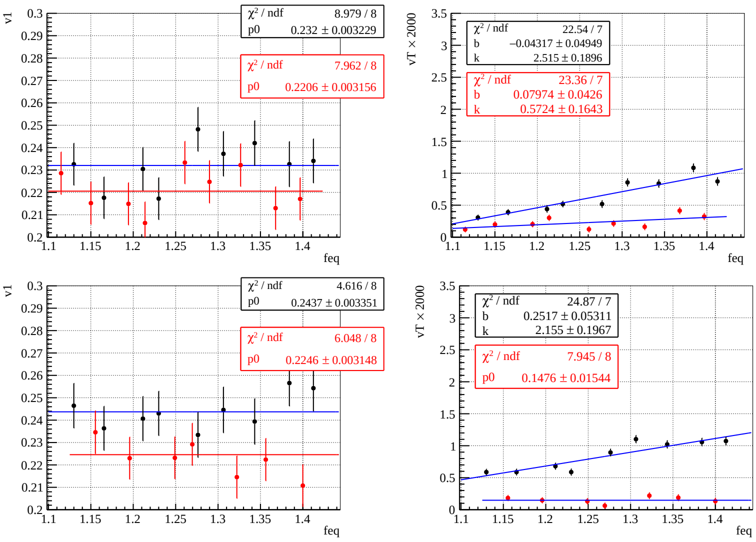


Figure 4.17: Trend of variance of events uniformly distributed in ^{7}Be Top / Bottom fiducial volume. Top and bottom rows of figures correspond to top and bottom part of ^{7}Be fiducial volume, respectively. Left: `charge` v_1. Right: v_T. Black and red are without and with non-uniformity correction `charge`, respectively. As we can see the v_1 are consistent with of the one of the whole FV, while the v_T is only around half of the one of the whole FV, which is expected.

4.3.4 Validation against calibration data

Although `g4bx2` is calibrated against calibration data such that the non-linearity accuracy is better than 1%, we would like to cross check the tuned analytical non-linearity model against calibration data directly. The `npe` is not observable so it's impossible to validate the $E_{\text{kin.}}$–`npe` relationship Eq. (4.12), `npe`–`npmt_dt1` relationship Eq. (4.37) and `npe`–`charge` relationship Eq. (4.43) directly.

If we trust that Monte Carlo can reproduce the `npe`–`nhit` dependence, which is reliable because the effects from PMT instability and electronics are suppressed, we can use `nhit` to reconstruct `npe` and validate `npe`–`charge` dependence. We used the following formula to reconstruct `npe`:

$$\frac{\mu_{\text{p.e.}}}{\mu_{\text{n.hit}}} = \frac{a + b \cdot \log\left(\frac{\mu_{\text{n.hit}}}{500}\right) + c \cdot \log^2\left(\frac{\mu_{\text{n.hit}}}{500}\right)}{1 + d \cdot \log\left(\frac{\mu_{\text{n.hit}}}{500}\right)} \tag{4.56}$$

where $\mu_{\text{n.hit}}$ is the average of the nhits. The best fit is shown in Fig. 4.18 left. Using npe reconstructed with this formula, we can fit `npe`–`charge` dependence, and the result is shown in Fig. 4.18 right. $p_{\text{dn.}}$, $p_{\text{mis.}}$ and $p_{\text{quadr.}}$ are consistent with that derived from Monte Carlo.

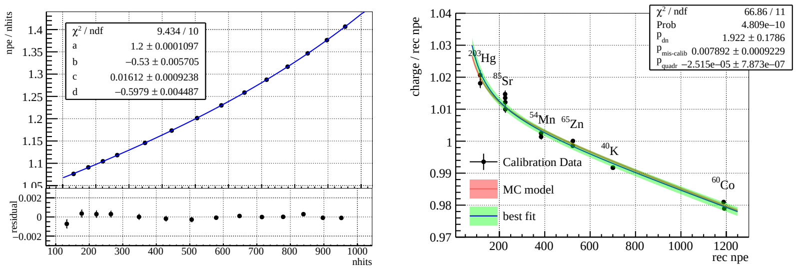


Figure 4.18: Left: the `nhit` to `npe` non-linearity and its residual. Simulated according to run 17407. Right: relationship between reconstructed npe and Echidna `charge` from calibration data.

4.3.5 *`charge` v_1 parameter*

Although `charge` v_1[35] is not fixed, because it's so important for determining ^{7}Be solar neutrino, it's worth a dedicated section for robust ^{7}Be solar neutrino measurement.

We decided to let v_1 free in the fit, because:

1. It can be pinned down by the ^{210}Po peak in the analytical fit when v_T is fixed properly according to `g4bx2`.
2. It's correlated with ^{7}Be solar neutrino.
3. `g4bx2` did not reproduce its value in data.

`charge` v_1 in `g4bx2` To convince you with point 3, let's see how large is v_1 in `g4bx2`. The single p.e. `charge` is simulated according to a Gaussian distribution, whose variance v_0 is set to the value measured with laser-calibration events collected at the beginning of each run. Because the `charge` is measured as the difference of peak and baseline voltage, the jitter of the baseline contributes additional variance $v_{\text{ped.}}$. The width of the baseline jitter is hard-coded to be 7 ADC unit and is normalized to 7/gain (p.e.). See Fig. 4.19 for the trends of v_0 and $v_{\text{ped.}}$. Sum of these two term should be consistent with v_1 derived by fitting simulated mono-energetic electrons, and they are consistent:

$$v_1^{\text{MC}}(Q, \text{fit}) = 0.230 \pm 0.035 \tag{4.57}$$

$$v_0 + v_{\text{ped.}} = 0.258 \pm 0.003 \,(\text{stability}) \tag{4.58}$$

However this is inconsistent with the value of v_1 in data, which is determined to be 0.298±0.005(stat.)±0.012(sys.). Retuning of Monte Carlo for `charge` has been planned to reduce such discrepancy.

---

[35] Notice here v_1 is defined in Equation (4.47) and is not the variance of single p.e. charge distribution

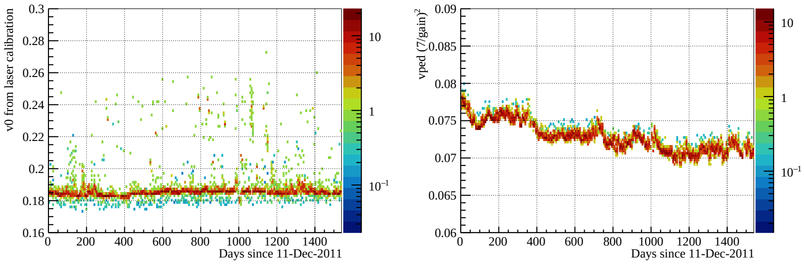


Figure 4.19: The stability of the measured variance of the peak (left) and baseline (right) of `charge` for single photon signal from the laser calibration events. As we can see the peak variance is quite stable and the baseline variance is increasing because of the increase of the gain. The average value for v_0 and v_{ped} from this graph are 0.186±0.002 and 0.072±0.003, respectively.

`charge`* v_1 *in Data We have investigated the following methods:

1. Constrain v_1 through γ calibrations
2. Consrain v_1 through α peaks from calibrations or backgrounds.
3. Fit v_1 with v_T constraint to tuned value on data.
4. The same as point 3, but only on the 2012 data.

For point 1, first we need to quantify δv_1 and δv_T in Eq. (4.51). They can be tuned against `g4bx2`, see Fig. 4.20.

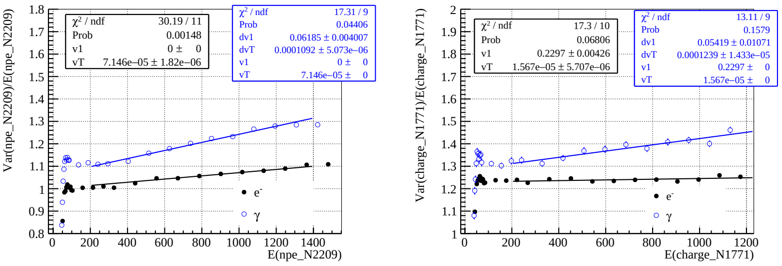


Figure 4.20: Additional smearing of γ compared with electrons. Left: `npe` variable. Right: `charge` variable.

We can see that

$$\delta v_1(\text{from p.e.}) = 0.062 \pm 0.004 \tag{4.59}$$

$$\delta v_T(\text{from p.e.}) = (1.09 \pm 0.06) \times 10^{-4} \tag{4.60}$$

$$\delta v_1(\text{from } \texttt{charge}) = 0.054 \pm 0.011 \tag{4.61}$$

$$\delta v_T(\text{from } \texttt{charge}) = (1.24 \pm 0.15) \times 10^{-4} \tag{4.62}$$

which are consistent between them.

We re-analyzed the γ calibration conducted at 2009 and they are summarized in Tab. 4.4.

With such correction we can fit v_1, see Fig. 4.21, and we got

$$v_1(\gamma) = 0.37 \pm 0.01(\text{stat.}) \tag{4.63}$$

The value here is inconsistent with (23% larger than) the value got with point 3 and 4. However the goodness of fit is very bad and ^{65}Zn and one point of ^{85}Sr are off the trend, indicating the calibration is not fully understood under such precision.

Table 4.4: Expectation, variance and their uncertainty of `charge` for γ calibration runs.

Type	Run	$f_{\text{eq.}}$	$\mu_{\text{n.Q}}$	$\delta\mu$	var	δvar
^{57}Co	10550	1.09027	49.91	0.13	75.0	2.2
^{57}Co	10551	1.09051	49.98	0.13	78.9	2.3
^{57}Co	10552	1.09051	49.88	0.13	72.7	2.1
^{139}Ce	10564	1.09202	71.80	0.19	176.9	5.2
^{139}Ce	10565	1.0923	71.71	0.19	180.0	5.5
^{139}Ce	10566	1.0917	71.73	0.20	175.8	5.7
^{203}Hg	10404	1.08976	117.09	0.12	193.8	3.5
^{203}Hg	10405	1.09037	117.17	0.13	190.6	3.5
^{85}Sr	10396	1.08924	228.82	0.30	326	10
^{85}Sr	10397	1.08929	228.25	0.30	362	11
^{85}Sr	10577	1.09195	228.44	0.21	368.7	7.8
^{85}Sr	10579	1.09199	228.71	0.21	361.5	7.7
^{54}Mn	10416	1.09135	386.30	0.22	632	11
^{65}Zn	10396	1.08924	524.72	0.67	749	33
^{65}Zn	10397	1.08929	524.27	0.61	727	32
^{40}K	10416	1.09135	698.19	0.34	1157	22
^{60}Co	10396	1.08924	1171.45	0.70	1813	56
^{60}Co	10397	1.08929	1170.92	0.66	1767	55

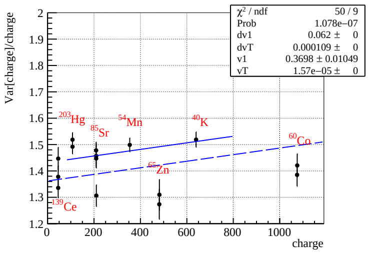


Figure 4.21: Variance/Expectation vs expectation of electrons, positrons and gammas for the calibration runs in 2009.

For point 2, we only considered the ^{214}Po α peak from ^{14}C-^{222}Rn calibrations. 182 ^{14}C-^{222}Rn calibration runs were taken with the source deployed at different positions inside the inner vessel. The ratio of width over peak position is shown in the Fig. 4.22 left and we got

$$v_1(\alpha) = 0.55 \pm 0.05(\text{stat.}) \tag{4.64}$$

The value here is also too large. Actually if we plot the ratio based on Tab. 5.5 of Saldanha[36] we can see that

$$(1 + v_1) \cdot f_{\text{eq.}} \sim 1.6 \tag{4.65}$$

while $f_{\text{eq.}} \sim 1.1$ for calibration runs in 2009, and the ^{214}Po α peak is off the trend. Considering this, the v_1 determined from ^{214}Po in calibration is not trusted.

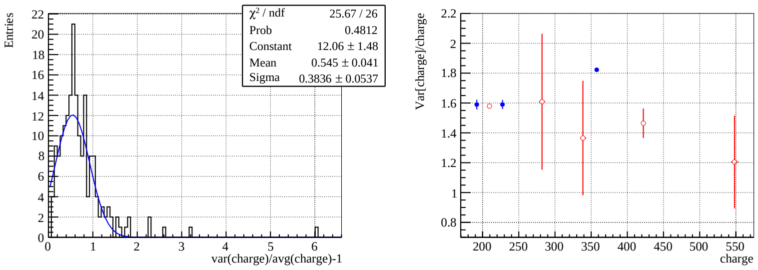


Figure 4.22: Variance of α particles from calibration runs in 2009. Left: distribution of Variance/Expectation of ^{214}Po peak in 182 ^{14}C-^{222}Rn calibration runs. Right: Variance/Expectation vs expectation according to Tab. 5.5 of Dr. Saldanha's thesis.

For point 3 and 4, we extract the v_1 from data directly.

For point 3, we use the whole Phase-II data (from 2011 December 14th. till 2016 May 15th.). v_T is determined according to `g4bx2` to be

$$2000 v_T(N_{\text{n.Q}}) = 1.00 \pm 0.13 \tag{4.66}$$

[36] R. N. Saldanha, "Precision Measurement of the 7 Be Solar Neutrino Interaction Rate in Borexino", PhD thesis, Princeton University, 2012.

where the uncertainty is evaluated assuming 1% precision of `g4bx2` on non-uniformity.

For geo-corrected variable, we have

$$2000v_T(N_{\text{n.Q-geo}}) = 0.18 \pm 0.18 \tag{4.67}$$

where we simply assumed 100% uncertainty.

By performing fits on data with v_T fixed to two extremes, we got

$$v_1(N_{\text{n.Q}}) = 0.295 \pm 0.006(\text{stat.}) \pm 0.014(\text{sys.}) \tag{4.68}$$

$$v_1(N_{\text{n.Q-geo}}) = 0.293 \pm 0.006(\text{stat.}) \pm 0.017(\text{sys.}) \tag{4.69}$$

For point 4, we use 2012 data (from 2012 January 1st. till 2012 December 23th.). v_T is determined according to `g4bx2` to be

$$2000v_T(N_{\text{n.Q}}; e^-) = 1.72 \pm 0.14 \tag{4.70}$$

$$2000v_T(N_{\text{n.Q}}; {}^{210}\text{Po}) = 1.21 \pm 0.16 \tag{4.71}$$

where the uncertainty is also evaluated assuming 1% precision of `g4bx2` on non-uniformity, and the ^{210}Po spatial distribution non-uniformity is ignored here and v_T is different between ^{210}Po and e^- only because ^{210}Po decays. Such simplification seems reasonable as determined v_1 is consistent with other methods in Point 3 and 4.

For geo-corrected variable v_T is the same as the case of whole Phase-II.

By performing fits on data with v_T fixed to two extremes, we got

$$v_1(N_{\text{n.Q}}) = 0.298 \pm 0.005(\text{stat.}) \pm 0.012(\text{sys.}) \tag{4.72}$$

$$v_1(N_{\text{n.Q-geo}}) = 0.295 \pm 0.005(\text{stat.}) \pm 0.016(\text{sys.}) \tag{4.73}$$

4.3.6 Skewness

Additional parameters describing the energy dependence of skewness are introduced in the Modified Gaussian response function. It is parametrized as

$$\kappa(N_{\text{n.Q}}) = g_2' \cdot f_{\text{eq.}}^2 \cdot \mu + 3(1 + v_1) f_{\text{eq.}} v_T \cdot \mu^2 \tag{4.74}$$

Compared with its expression used in Phase-I

$$\kappa(N_{\text{n.Q}}) = g_2 \cdot \mu + 3 \cdot g_1 g_3 \cdot \mu^2 + g_4 \cdot \mu^3 \tag{4.75}$$

We found g_4 is not needed ($g_4 \equiv 0$) and g_2' is constant after introducing $f_{\text{eq.}}^2$, see Fig. 4.23. Besides, after geometrical correction the g_2' is more stable. In[37], g_4 is not used and g_2' is free.

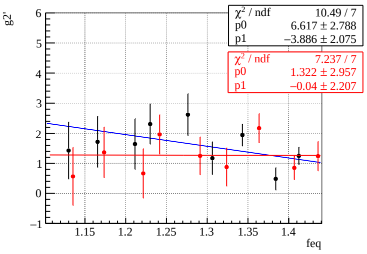


Figure 4.23: The stability of g_2'. Blue is the original charge, and red is the geometrical corrected charge.

For the Generalized Gamma response function there is no free parameters for the skewness, it's automatically fixed by the expectation and variance.

4.3.7 Tuned Parameters Summary

The model and parameters are summarized in Tab. 4.5.

[37] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

Type	
Type	`npmt_dt1` (Scaled Poisson)
N.L.	q_1 ... q_5: $(0.972338, 0.200838, 0.010463, 0.195038, 0.014228)$
	A_0 ... A_4: $(1.415307, -3.396670, 1.106921, 0.072419, 1.336687)$
	fCher, $p_{\mathrm{ser.}}, g_{\mathrm{LC}}$: (1.0, 0.12, 0.101 (*pep* FV))
Var.	use α var. model
	$f_{\mathrm{eq.}}(\mathrm{stable}), f_{\mathrm{eq.}}(^{210}\mathrm{Po})$: $(1.27996, 1.19409)$
	$v_1, 10^4\beta_2$: $(0.16, 7)$
Type	Echidna or Mach4 `charge` (Modified Gaussian)
N.L.	q_1 ... q_5: calculated according to $k_B = 0.0119386$
	A_0 ... A_4: $(0.487789, -0.792818, -1.16555, 0.0481817, 1.92020)$
	fCher: 3.615
Var.	use α var. model
	$f_{\mathrm{eq.}}(\mathrm{stable}), f_{\mathrm{eq.}}(^{210}\mathrm{Po})$: same as `npmt` (normal)
	$(1.31021, 1.22481)$ (geo.)
	$(1.32443, 1.23457)$ (zcorr.)
	$2000 \cdot v_T$: $(6.01 \pm 0.06) \cdot (\mathrm{feq} - 1)$(normal)
	0.184 ± 0.011(geo./zcorr.)
Type	Echidna `charge`
	$p_{\mathrm{dn.}}, p_{\mathrm{mis.}}, p_{\mathrm{quadr.}}$: $(1.47, 0.011, -2.4 \times 10^{-5})$
Type	Mach4 `charge`
	$p_{\mathrm{dn.}}, p_{\mathrm{mis.}}, p_{\mathrm{quadr.}}$: $(-0.19, 0.012, -2.0 \times 10^{-5})$

Table 4.5: Summary of response function models and parameters. *N.L.* = non-linearity. *Var.* = variance. *geo.* = geometrical correction. Comparison between Echidna `charge` and Mach4 `charge` and that between normal and geometrical corrected `charge` is described in Sec. 3.1.2.

4.4 *Validation against Monte Carlo based detector model*

In order to account for various simplification we made, we fit thousands of pseudo-experiment spectra generated according to Monte Carlo based probability density function and take the bias of the median value of the best fit rates compared with the injected value as the corresponding systematic uncertainty of the analytical fit approach. The bias given by such method is also the final criteria for tuning the detector response models.

First, we used the `g4bx2` based probability density function to generate thousands of pseudo-experiment spectra. Then we fitted each of them using analytical response function and get the distribution of the best fit. At last we compared the median values of the best fit rates with the injected rates.

The validation of `npmt_dt1` and `charge` using the configuration specified in Tab. 4.5, except few points, are shown in Fig. 4.24 and Fig. 4.25 and summarized in Tab. 4.6. For `npmt_dt1`, it is not possible to fit the pseudo-experiment spectra generated with `g4bx2` including pile-up events using the Dark Noise Convolution method. The reason is still not understood. Besides, we do not have enough computing resources to simulate enough ^{14}C events to produce smooth ^{14}C probability density function, and such effect might introduce bias. Due to this reason, when validating the `npmt_dt1`, the ^{14}C and pile-up are not injected, Dark Noise Convolution is not applied, and the fit range start from 100 `npmt_dt1`. As a result, the *pp* solar neutrino rate precision is artificially higher in such tests.

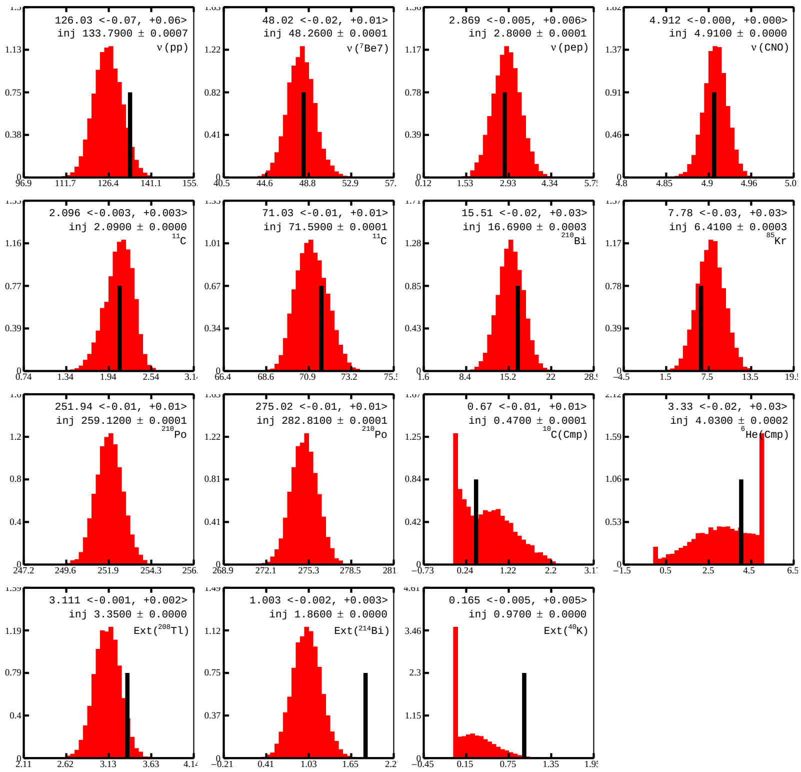


Figure 4.24: Results on biases of analytical response function for `npmt_dt1`. Red: Distribution of best fit using analytical response function against pseudo-experiment spectra generated with `g4bx2` probability density function. Black: injected rate. The median value is shown on the first row at the top region of each pad. The injected value with statistical fluctuation of the expectation is shown at the second row, and the name of the component is shown at the third row.

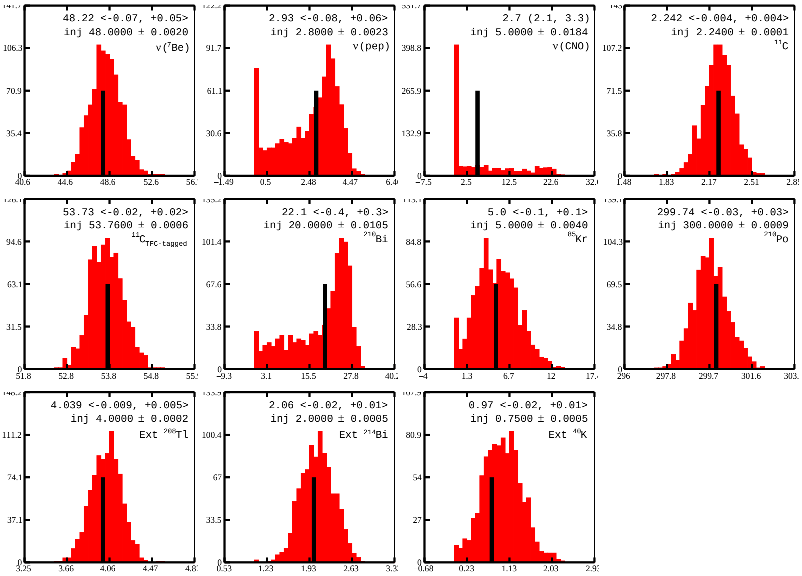


Figure 4.25: Results on biases of analytical response function for `charge`. Red: Distribution of best fit using analytical response function against pseudo-experiment spectra generated with `g4bx2` probability density function. Black: injected rate. The median value is shown on the first row at the top region of each pad. The injected value with statistical fluctuation of the expectation is shown at the second row, and the name of the component is shown at the third row.

	`npmt_dt1`		`charge`	
Exposure	582.326/323.622		377.5/377.5	
Type	Inj.	best fit	Inj.	best fit
$\nu(pp)$	133.79	126.05 ± 0.09	133	Fix
$\nu(^7\mathrm{Be})$	48.26	48.00 ± 0.02	48.00	48.22 ± 0.07
$\nu(pep)$	2.800	2.866 ± 0.007	2.80	2.93 ± 0.08
$\nu(\mathrm{CNO})$	4.91	penalty	5.0	$2.7^{+3.3}_{-2.1}$
$^{11}\mathrm{C}$	2.090	2.098 ± 0.003	2.240	2.242 ± 0.004
$^{11}\mathrm{C}_2*$	71.59	71.01 ± 0.01	53.76	53.73 ± 0.02
$^{210}\mathrm{Bi}$	16.69	15.53 ± 0.04	20.0	22.1 ± 0.4
$^{85}\mathrm{Kr}$	6.41	7.79 ± 0.03	5.0	5.0 ± 0.1
$^{210}\mathrm{Po}$	259.12	251.96 ± 0.01	300.00	299.74 ± 0.03
$^{210}\mathrm{Po}_2*$	282.81	275.04 ± 0.02	300.00	299.74 ± 0.03
$^{10}\mathrm{C}_2*$	0.47	0.65 ± 0.01	-	
$^{6}\mathrm{He}_2*$	4.03	3.36 ± 0.03	-	
Ex. $^{208}\mathrm{Tl}$	3.350	3.114 ± 0.003	4.000	4.039 ± 0.009
Ex. $^{214}\mathrm{Bi}$	1.860	1.000 ± 0.003	2.00	2.06 ± 0.02
Ex. $^{40}\mathrm{K}$	0.970	0.164 ± 0.006	0.75	0.97 ± 0.02

Table 4.6: Summary of biases in Fig. 4.24 and 4.25. The unit of exposure is day × 100 ton. *Inj.* = injected rate. The unit of best fit of rates and injected rates are counts per day per 100 tons, except for ^{14}C, which is Bq per 100 tons (1 Bq = 86400 counts per day per 100 tons). For `npmt_dt1` ^{14}C and pile-up are not injected and Dark Noise Convolution is not applied, and fit range start from 100, and thus *pp* solar neutrino rate bias is not reliable. CNO solar neutrino is constrained to the injected value with a pull term and thus the bias is also not reliable. For `charge` analysis, *pp* solar neutrino rate is fixed to the injected rate.

4.5 *Summary and Discussions*

In this chapter we described the analytical response functions used in[38], discussed the physics meaning of each parameter and possible ways to improve the models, and evaluated the biases using a Monte Carlo method. The models and parameters used in[39] are summarized in Tab. 4.5 and the biases are summarized in Tab. 4.6.

In this work we have studied carefully the resolution model. In `charge` analysis the ^{7}Be solar neutrino rate is correlated with the resolution parameters. Due to the degraded resolution, in Phase-II analysis, the spectrum does not contain enough information in `charge` about the resolution, and it is necessary to constrain the v_T and use the same variance model parameters for ^{210}Po and e^-. In `npmt` analysis, the resolution parameters are constrained by the high statistics of ^{14}C shape and thus there is no need to constrain the variance model parameters of e^- to be the same as that of ^{210}Po.

In[40] the `npmt` group has introduced a free parameter β_1 to account for the variance introduced by quenching effect and Cherenkov contribution. As discussed in Sec. 4.2 pp. 72, β_1 is an effective parameter, and its value won't decrease to 0 when quenching effect and Cherenkov radiation are turned off.

---

[38] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

[39] ibid.

[40] ibid.

5

Statistical and systematic uncertainties

In this chapter the evaluation of the statistical and systematic uncertainties on solar neutrino elastic scattering interaction rates in Borexino Phase-II analysis is presented. This chapter provides also the benchmark of the accuracy of the detector response model described in the last chapter. In the Borexino solar neutrino analysis, since the signal-to-background ratio in the region of interest is low and the separation between signals and backgrounds heavily relies on the energy distribution differences, the sensitivity and the systematic uncertainties have been throughly evaluated before performing fitting of the data. There are challenges in applying standard methods in evaluating the systematic uncertainties, and a toy Monte Carlo method was developed to overcome these issues and will be described in this chapter. The optimizations of the analysis parameters, including range and that of choice of the so-called Three-Fold-Coincidence method[1], are performed by minimizing the total uncertainty based on the results of this work, and such blinded optimization avoids biases.

This chapter is organized as in the following: Section 5.1 introduces the concepts and challenges. Section 5.2 describes the toy Monte Carlo method developed for Borexino analysis. Section 5.3 describes the results of the exposure related systematic uncertainties. Section 5.4 presents the optimization and crosschecks of the fit methods. Section 5.5 presents the evaluation of statistical and systematic uncertainties in the Monte Carlo approach. Section 5.6 reports the evaluated uncertainties for the analytical approach. At last conclusions and discussions are presented in Section 5.7.

I have co-worked with Matteo Agostini, Daniele Guffanti, Nicola Rossi, Simone Marcocci and Anna Janny in this work. The results of this work has been used in [2,3].

[1] See Section 2.5.4 page 41

[2] The Borexino collaboration, “Comprehensive measurement of pp-chain solar neutrinos”, *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

[3] The Borexino collaboration, “First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II”, pp. 1–8, July 2017, arXiv: `1707.09279`

5.1 Introduction

Statistical and systematic uncertainties are quantities specifying the precision of measurements. In this work, the width of the distribution of results of identical measurements is used as the estimator of the statistical uncertainty. The difference between the average expected measurement and the true value is the bias, and the range of possible biases is used as the systematic uncertainty. The statistical uncertainty is due to random processes, and the systematic uncertainty is due to the limited accuracy of the calibration of the tool used.

The median statistical uncertainties provide a reference on the potential measurement precision to the analysis on data. The evaluated systematic uncertainty can be taken as the systematic uncertainty of the measurement. It is independent of the statistical realization of the data. The systematic uncertainty measures the stability of the results against model inaccuracies under various assumptions of models and guarantee that the result is robust regardless of variation of models. Besides, the parameters used in the analysis procedure, such as fitting range, can be optimized in order to reduce the total uncertainty.

There are challenges applying standard methods for evaluating systematic uncertainties. The systematic uncertainty is usually estimated by adding nuisance parameters[4]. The cut efficiencies and detector response, such as the non-linearity, are allowed to vary in the fit and be optimized versus the data. Usually subsidiary datasets are included to constrain the nuisance parameters and the product of the likelihood of the main dataset and subsidiary datasets are maximized. Ideally, all the parameters in the Monte Carlo simulation, such as the attenuation length of each wavelength and quantum efficiencies, are treated as nuisance parameters and the Monte Carlo is tuned during the fit, and calibration data is embedded in the pull terms, however it requires huge amount of computing power. We can make a rough estimation here. To produce Borexino Monte Carlo probability density functions for all components it takes around 2 weeks of 2000 AMD Opteron(tm) Processor 6320 CPUs. One fit needs around $50N$ iterations, where N is the number of free parameters and is around 2×10^2. Considering the Borexino collaboration currently can at most have 2000 CPUs on CNAF for a short period, to perform one fit it would take 380 years, so this method is impractical for Borexino[5].

The probability density function obtained by Monte Carlo simulation can be distorted according to some function, usually a polynomial, during the fit, and the coefficients of the function become the nuisance parameters. Although this method takes much less CPU time than the example before, it still takes too much time to obtain the distorted probability density function during the fit. In order to overcome this challenge, a Monte Carlo based method is developed and the distortion is applied when generating pseudo-experiment

[4] Particle Data Group, "Review of Particle Physics", *Physical Review D*, vol. 98, no. 3, p. 030001, Aug. 2018, DOI: `10.1103/PhysRevD.98.030001` Section 39.2.2.2, page 529 and Section 39.3.2.1, page 533

[5] Supercomputer might work. The computing power of AMD Opteron(tm) Processor 6320 CPUs is around 1 GFlops/s, then one fit consumes a computation power of 5×10^{10} TFlops. The computation power of current world's top super computer `Summit` is 1.2×10^5 TFlop/s, so it takes 5 days of `Summit` to perform such a fit for Borexino.

spectra. First of all, the distortion is injected once for all. Second, the number of events is much less. So the time is much shorter and reasonable.

The analytical fit approach[6] can be considered as the nuisance parameter method. The systematic uncertainty due to the light yield etc. in the Monte Carlo fit are embedded into the statistical uncertainty.

Another method is called the covariance matrix method. In this method, the covariance matrix of the standard χ^2 method is modified. The covariance matrix is approximated with the simulated distribution of number of events in each bin, and the distribution is distorted with predefined magnitude. Borexino cannot use this method directly because it requires all uncertainties to be Gaussian, which is not true for components like CNO solar neutrinos. The CNO solar neutrino interaction rate is constrained to be positive, and this makes the profiled likelihood of CNO solar neutrino interaction rate asymmetrical.

The toy Monte Carlo method used in this work is intrinsically different from the nuisance parameter method in the way of treating the information of the nuisance parameter in the main dataset. In this method, the information in the main dataset is abandoned and the distribution of the nuisance parameter is set to the one obtained with subsidiary datasets by force, while in the nuisance parameter method information in both datasets are used. For example, in Borexino spectrum there is high statistics of ^{14}C events, and thus the light yield in the ^{14}C region can be extracted with high precision and even better than that from calibrations. Besides, ^{14}C events are uniformly distributed, while calibrations using radioactive sources usually provides detector responses scanned over a finite set of positions. The systematic uncertainty of the light yield evaluated with the covariance matrix method will be much higher than that using the nuisance parameter method.

5.2 Methods to evaluate systematic uncertainties

In the last section the challenges in applying usual methods for Borexino analysis have been explained, and in this section, I will describe a toy Monte Carlo method that overcomes these challenges. First of all, the procedure of the toy Monte Carlo method is described. Secondly, the list of systematic uncertainties considered, the corresponding forms of deformations and precisions are given.

5.2.1 Procedure of toy Monte Carlo method

Datasets with and without various deformation are prepared and they are fitted with un-deformed probability density function. The width of the distribution of the reconstructed parameters of interests without deformations is taken as the expected statistical uncertainties, and the width with deformations is considered as the quadratic

[6] See Section 3.1.3

sum of the expected statistical uncertainties and the systematic uncertainties.

The deformations are applied by modifying the energy and e^+/e^- pulse shape variable based on the models accounting for the possible imprecision of detector response.

5.2.2 *Treatment in Monte Carlo fit*

Four deformation models have been chosen to represent the Monte Carlo inaccuracies and are described below. The normalized `nhit` variable and `g4bx2` based Monte Carlo probability density function (except ^{14}C, for which the analytical model is used) are used.

Non-linearity The non-linearity accounts for the imprecision on the tuning of quenching parameters, the re-emission efficiency of ultraviolate photons etc. To decouple this non-linearity from the trivial non-linearity due to the multinomial distribution[7], `nhit` is converted to `npe` [8] first, then the deformation is applied, then the deformed `npe` is converted back to `nhit`.

The deformation model is described in Eq. (5.2). Two data points are prepared with x chosen between 0 and 2.5 MeV, y sampled from a Gaussian distribution of average 1 and $\sigma = 0.2\%$. Then p_0 and p_1 are determined requiring crossing the two points.

$$\texttt{nhit} \mapsto \texttt{npe} = \mathrm{NL}_{\texttt{nhit}\to\texttt{npe}}(\texttt{nhit}) \tag{5.1}$$

$$\texttt{npe} \mapsto \texttt{npe}' = \texttt{npe} \cdot (p_0 + E \cdot p_1) \tag{5.2}$$

$$\texttt{npe}' \mapsto \texttt{nhit}' = \mathrm{NL}_{\texttt{npe}\to\texttt{nhit}}(\texttt{npe}') \tag{5.3}$$

where NL is the conversion function, E is the MC truth energy.

The density of the distorted non-linearity model is illustrated in Figure 5.1. We also performed a fit of the distorted non-linearity model to see the corresponding variation of the kb and light yield. We can see from the figure that the distortion is very small and we can still describe the model with the analytical model quite well, and the change of kb is less than 10% at 95% C.L. .

z dependence Here the z dependence refers to the z dependence of the light yield. The local light yield depends on the total acceptance solid angle of PMT photocathodes, and the z dependence is due to the fact that there are more live PMTs on the top hemisphere than on the bottom.

The deformation model is described in Eq. (5.5). The center is fixed to 1 to decouple this effect from the non-linearity. Another point at z = 2.2 m is sampled following the Gaussian distribution with average 1 and $\sigma = 0.28\%$. We take 0.28% because the discrepancy between data and Monte Carlo is less than 0.4% in the fiducial volume as shown in Fig. 5.2, and this is taken as 2σ. Because we fix the radial dependence at the center, it's as if we are varying the ratio between the top of FV and center, and thus the σ is chosen as

[7] Strictly speaking only `npmt` follows the multinomial distribution. The non-linearity of `nhit` and `npmt` are similar as can be seen on the spectrum

[8] `npe`: the number of photo-electrons produced on the photocathode

$\sqrt{2} \times 0.2\% = 0.28\%$.

$$\texttt{nhit} \mapsto \texttt{npe} = \mathrm{NL}_{\texttt{nhit}\to\texttt{npe}}(\texttt{nhit}) \tag{5.4}$$

$$\texttt{npe} \mapsto \texttt{npe}' = \texttt{npe} \cdot (1 + z \cdot p_1) \tag{5.5}$$

$$\texttt{npe}' \mapsto \texttt{nhit}' = \mathrm{NL}_{\texttt{npe}\to\texttt{nhit}}(\texttt{npe}') \tag{5.6}$$

where NL is the conversion function, z is the MC truth vertex coordinate along top-bottom direction.

The $\mathcal{L}(\mathbf{r})$ variable To avoid the fake sensitivity due to limited Monte Carlo statistics and catching of fluctuations, the third dimension, the $\mathcal{L}(\mathbf{r})$ variable, or the position reconstruction likelihood, is replaced by three heuristic models for electrons, positrons and external γs, respectively. The models are determined by fitting the Monte Carlo sample and then the main Gaussian peak is shifted to the value found in the fit on data. They are described in Eq. (5.7)–(5.9)[9]

$$h_{\text{electrons}}(\mathcal{L}(\mathbf{r})) \sim \mathrm{Gaus}(\mathcal{L}(\mathbf{r}); \mu, \sigma) \tag{5.7}$$

$$h_{\text{fl}}(\mathcal{L}(\mathbf{r})) \sim 0.5 \cdot \exp\left[0.5l \cdot (2\mu + l\sigma^2 - 2x)\right] \mathrm{erfc}(\frac{\mu + l\sigma^2 - x}{\sqrt{2}\sigma}) \tag{5.8}$$

$$\begin{aligned} h_{\text{positron}}(\mathcal{L}(\mathbf{r})) \sim & (1-\epsilon) \cdot 0.5 \cdot \exp\left[0.5l \cdot (2\mu + l\sigma^2 - 2x)\right] \mathrm{erfc}(\frac{\mu + l\sigma^2 - x}{\sqrt{2}\sigma}) \\ & + \epsilon \cdot \begin{cases} 0 & \mathcal{L}(\mathbf{r}) < \mu_2 - \sigma_2 \\ \mathrm{Gaus}(\mathcal{L}(\mathbf{r}); \mu_2, \sigma_2) & \mathcal{L}(\mathbf{r}) >= \mu_2 - \sigma_2 \end{cases} \end{aligned} \tag{5.9}$$

To account for the imprecision of the MC models or the heuristic model itself and the statistical fluctuations, the datasets are prepared with $\mathcal{L}(\mathbf{r})$ following the same models but with different parameters sampled according to Gaussian distributions. The main peak position is varied with the fit error on data. For positrons the other parameters are also varied within the fit error on MC.

The fit is shown in Fig. 5.3 and the errors on $\mathcal{L}(\mathbf{r})$ variable are summarized in Tab 5.1.

Par	Electrons	Positrons	External γ
μ	4.7010 ± 0.003 [10]	4.6945 ± 0.003	4.6580 ± 0.003
σ	0.063261 (fixed)	0.05719± 0.0001	0.05552 (fixed)
l	NA	19.13 ± 0.1	21.14 (fixed)
ϵ	NA	0.7279%±0.01%	NA
μ_2	NA	4.9128±0.003	NA
σ_2	NA	0.1818±0.001	NA

Table 5.1: The parameters and errors of the heuristic model used in the systematic study.

[9] D. Guffanti, "Measurement of solar neutrino with BOREXINO", PhD thesis, Gran Sasso Science Institute, 2019

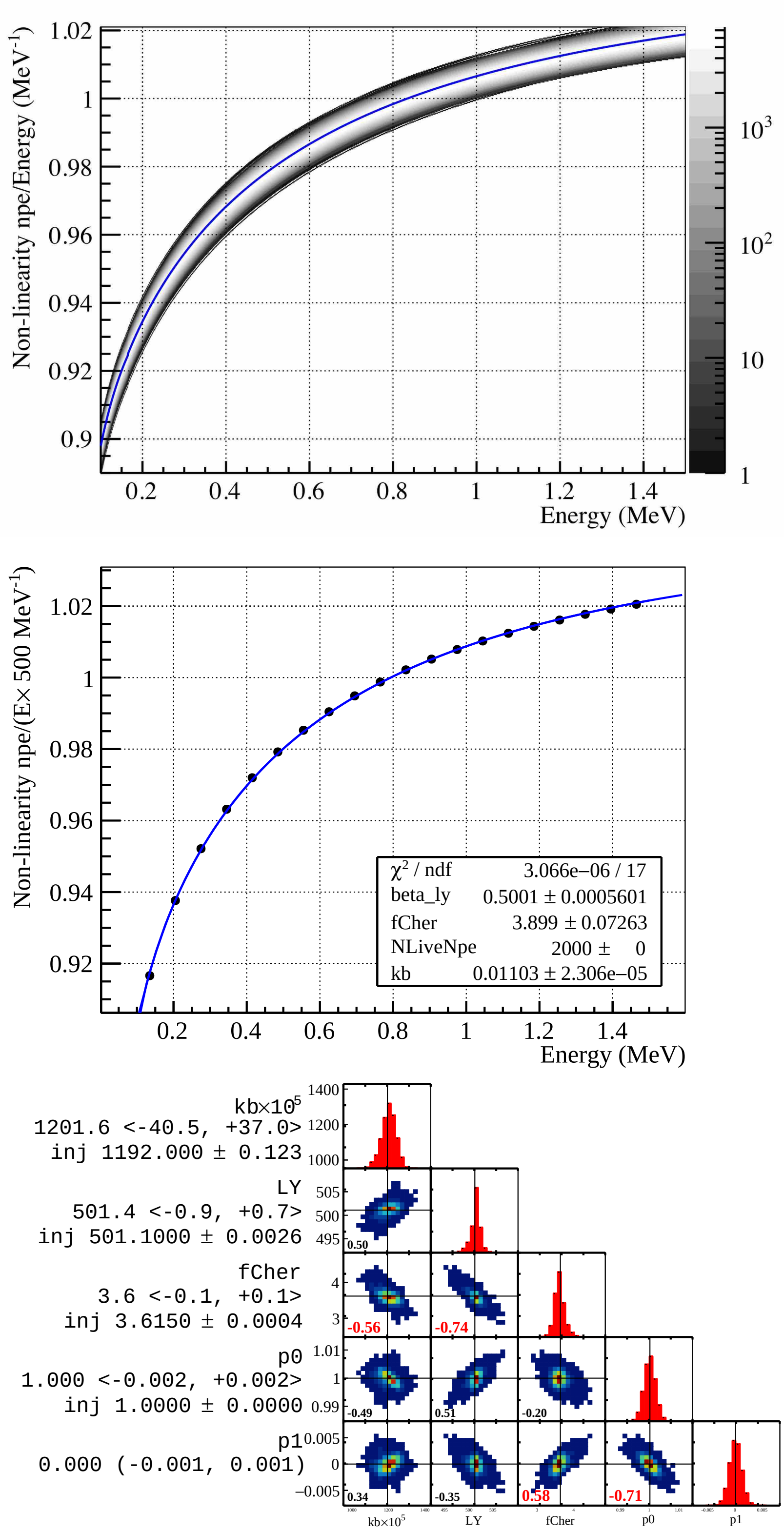


Figure 5.1: Top: the density of the non-linearity model. 10^5 realization is plotted. The kink at 165 keV is due to the Cherenkov threshold. Left bottom: An example least-square fit of the distorted non-linearity model to extract the effective kb and light yield. Right bottom: The extracted kb, light yield, Cherenkov contribution and distortion parameters. The 68% center interval is shown in brackets. Here 0.2% non-linearity distortion is injected.

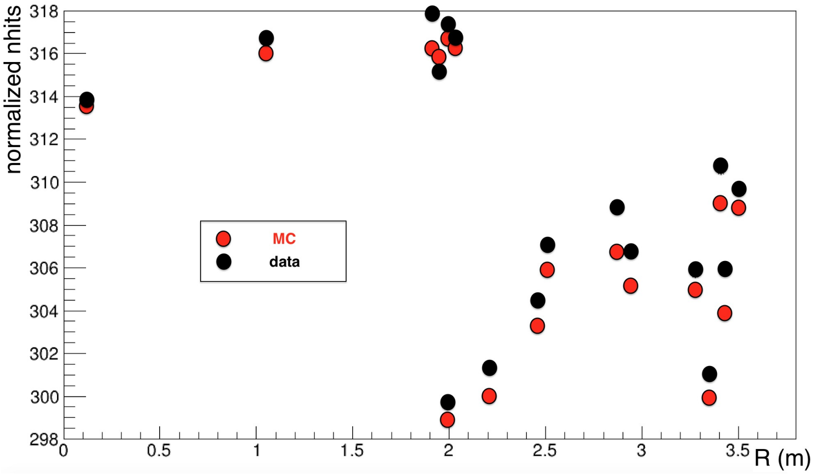


Figure 5.2: The comparison of MC and calibration datas with ^{214}Po events. Reproduced from S. Marcocci, "Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with Xe -loaded Liquid Scintillators", PhD thesis, Gran Sasso Science Insitute, 2016, Figure 5.17, page 166.

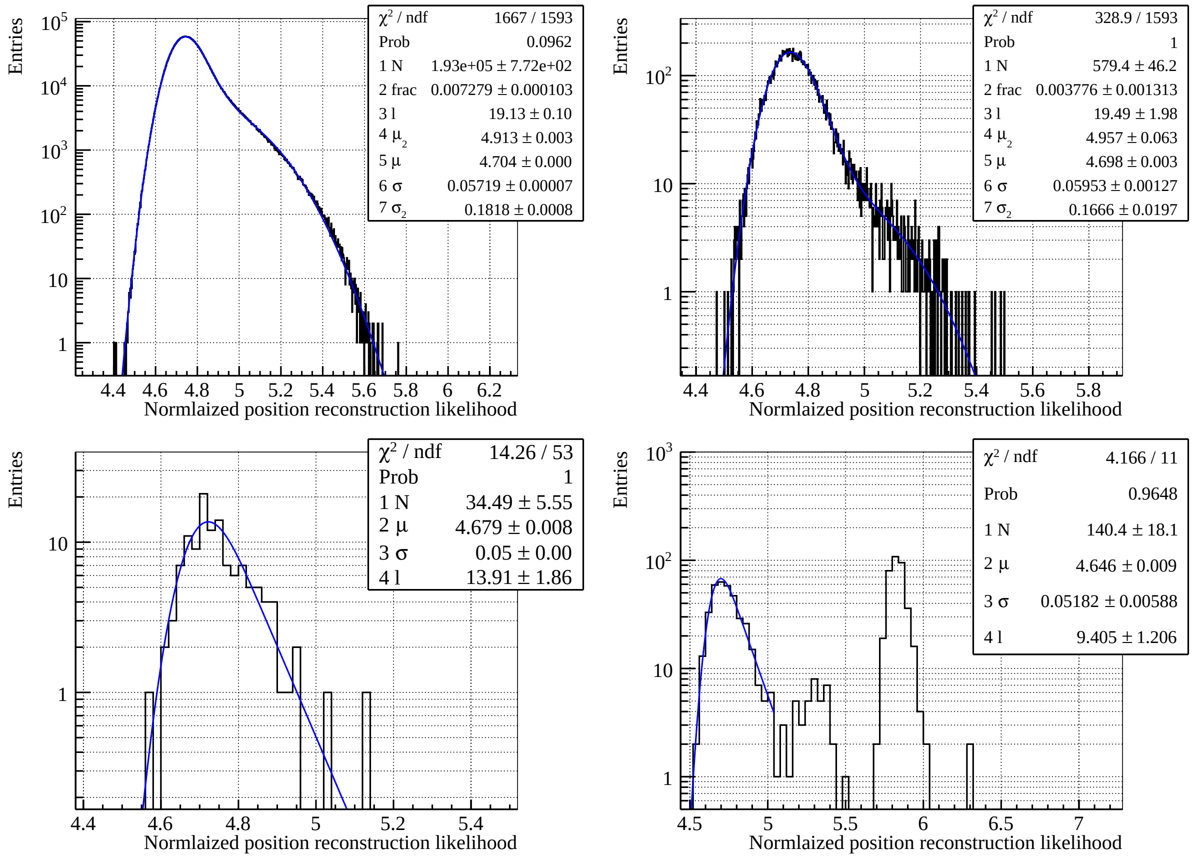


Figure 5.3: The fit of the $\mathcal{L}(\mathbf{r})$ distribution. Left top: the MC ^{11}C sample. Right top: the ^{11}C sample from data selected with the TFC algorithm. Left bottom: Simulated external ^{208}Tl calibration. Right bottom: External ^{208}Tl calibration. The events under the high-value peak in data come from the pileup events and are excluded from fitting.

The ^{210}Bi shape To account for the systematic uncertainties induced by the uncertainty of the ^{210}Bi shape, a distortion is applied. For each sampled ^{210}Bi event, a survival probability is applied to deform the energy spectrum:

$$p_{\text{survival}} = p_0 + (p_1 - p_0)\frac{E}{E_{\text{End}}} \tag{5.10}$$

where p_0 and p_1 is the deformation parameter and are sampled uniformly in the range 0.99–1, E_{End} is the end point of the ^{210}Bi spectrum and it is 1.162 MeV. The total number of sampled ^{210}Bi events after fiducial volume cut and above 90 `nhit` threshold are kept constant for this deformation.

The ^{14}C spectrum ^{14}C is one of the major backgrounds for ν(pp). There is systematic uncertainties coming from both the fluctuations in the Monte Carlo sample due to limited statistics and the precision of the Monte Carlo.

The rate of ^{14}C events is 40 Bq/100t. Because the position reconstruction is not reliable in the low energy region, ^{14}C from the whole inner vessel might be mis-reconstructed in the fiducial volume, so events in the whole inner vessel must be simulated, and the total amount of events to be simulated is

$$40\ (\text{Bq}/100\text{t}) \times 300\ \text{ton} \times 1330\ \text{days} = 1.4 \times 10^{10}\ \text{events} \tag{5.11}$$

Consider the simulation speed is around 2.55 (second × MeV^{-1}), the CPU time needed would be

$$\begin{aligned} t &= N \cdot \overline{E} \cdot \Delta T \\ &= 1.4 \times 10^{10} \times 100\ \text{keV} \times 2.55\ (\text{second} \times \text{MeV}^{-1}) \\ &= 400\ \text{days} \times 100\ \text{CPU} \end{aligned} \tag{5.12}$$

Considering the whole collaboration has only around 2000 CPUs during the peak usage, it takes 20 days to simulate the same amount of events we collect in data, and it is not feasible to simulate 100 times more to get a smooth shape for fitting.

We could simulate around 10^9 event, only 1/15 statistics of data, and after fiducial volume cut above 90 normalized `nhit` there are only 6277 events. Their energy spectrum is shown in Fig. 5.4.

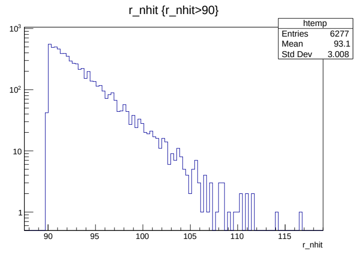


Figure 5.4: The energy distribution of the full Monte Carlo statistics for ^{14}C. Due to limited computing resources we only simulate 1/15 of the amount collected on data. The number of events with energy higher than 90 normalized nhit is only 6277, and you can see the fluctuations in the region of interest for ν(pp) is very strong. In Simone's thesis a scaled Poisson method is applied to the bin below 100 to suppress the systematic uncertainties from this fluctuations.

An analytical ^{14}C probability density function is got by fitting the `npmt_dt1` variable of the second cluster. The fit result and the comparison with the MC shape is shown in Fig. 5.5.

With this analytical shape we don't have the energy and position information, and we use the converted energy from the nhit, and a randomized z:

$$E' = \text{NL}_{\text{nhit}\to E}(\text{nhit}) \tag{5.13}$$

$$z \quad \text{randomly sampled in the fiducial volume} \tag{5.14}$$

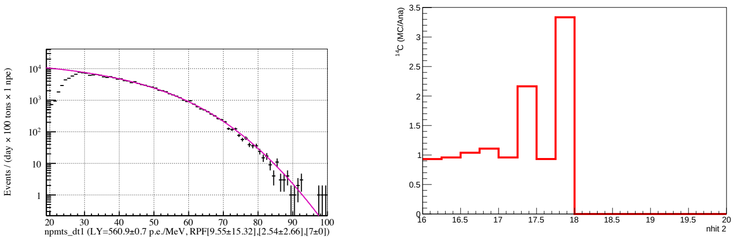


Figure 5.5: Left: the fit on the second cluster. Right: the comparison bewteen MC and Analytical shape. As we can see the fluctuation in the region of interests is very strong.

5.2.3 *Treatment in analytical fit*

The principle is the same as the one for studying the Monte Carlo systematic uncertainties. Only the fitting procedure is changed. The evaluation of the systematic uncertainties of the analytical fit is based on the normalized `npmt_dt2` variable and analytical probability density functions. The fit start from 85 `npmt_dt2` and ends at 900.

The same set of deformations of the Monte Carlo fit are considered exception of the pulse shape variable $\mathcal{L}(\mathbf{r})$. The systematic uncertainties are evaluated using the fit of energy distributions rather than the multivariate fit. The result obtained with fit of energy distributions is reliable. First of all, when fitting data, the multivariate likelihood only changes the statistical uncertainty and does not modify the best estimate of the component rates. Secondly, according to the experience in the Monte Carlo fit, the multivariate fit only affects the *pep* and CNO neutrinos, while their systematic uncertainties are dominated by the non-linearity.

The analytical probability density functions only provide the predicted energy distribution. The inaccuracy in reproducing the energy response spatial non-uniformity can induce a mismatch between the energy resolution of data and Monte Carlo, and thus it is one of the systematic uncertainties. To study this, true and reconstructed vertexes need to be sampled for each event when making the pseudo-experiment spectra. They are sampled according to uniform distribution for uniform components or the distribution according to Monte Carlo simulations for components such as external γ events.

5.3 *Exposure related systematic uncertainties*

The exposure related systematic uncertainties are listed in tab. 5.2. For the density change

$$\frac{\Delta\rho}{\rho} = \frac{\Delta V}{V} = \beta \Delta T \tag{5.15}$$

β = 1.25×10^{-3} and $\Delta T < 2.6$ °C.

Item	σ_L	σ_R
Fiducial Volume	-1.1%	+0.6%
LS density	<0.5%	
Live time	0.05%	

Table 5.2: The exposure related systematic unceratinties. Here and after σ_L and σ_R represent the uncertainty in the positive and negative direction.

5.4 *Optimization and crosschecks*

In this section the optimization of the fit range based on total uncertainties and comparison between three Three-Fold-Coincidence algorithms are also presented.

5.4.1 *Optimization of the fitting range*

The optimization of the fit range is performed based on the total uncertainty on ν(pp). The result is shown in Fig. 5.6. From the figure we can see that below 100 the total uncertainty decrease, however when fitting data they are excluded because the extracted ^{14}C decay rate is not stable. Thus we pick the minimum at `nhit2` = 17[11], which correspond to `nhit` 106.6. Here the analytical ^{14}C shape is used.

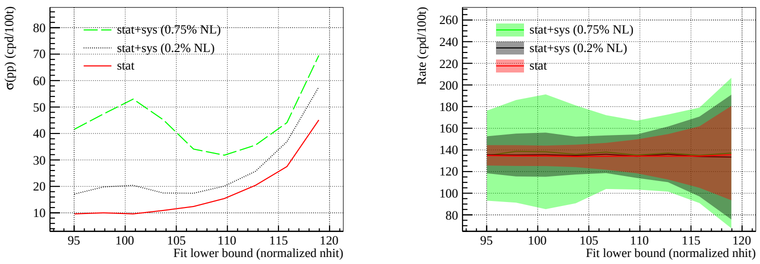


Figure 5.6: A series of test was performed to study how the total uncertainty depends on the fitting start range. For each test 200 pseudo experiment were fit.

5.4.2 *Comparison of different TFCs*

Three TFC[12] algorithms exist and they differ in the efficiency and rejection power against cosmogenic ^{11}C. The statistical and systematic uncertainties of ν(pp), $\nu(^7$Be) and ν(pep) are estimated and listed in Table 5.3. In the comparison, the same average ^{11}C rate and measured enhanced ^{11}C rate from the fit on data value with corresponding algorithm are assumed. Because the Mainz (denoted as MZ) algorithm, which is based on projected likelihood, has the best efficiency and rejection power, the statistical and systematic uncertainty of ν(pep) using it is slightly better than the other two.

For both the sensitivity and the systematic uncertainties, Mainz TFC has the best performance, but the gain is very small. For ν(pp) and $\nu(^7$Be) no gain, for ν(pep) we gain 1% sensitivity, gain 0.5% systematic uncertainties compared with Milano TFC and gain 2.5% systematic uncertainties compared with LNGS TFC.

5.4.3 *Impact of wrong ν(CNO) constraint*

Because the measurement of ν(pp), $\nu(^7$Be) and ν(pep) are given with ν(CNO) constrained, the impact of wrong constraint is reported in Table 5.4. The change of sensitivity and systematic uncertainty is less than 0.1% for $\nu(^7$Be) and less than 1% for ν(pp). However the reconstructed pep rate can be biases up to 30% if the injected rate differs from the centroid of the pull.

[11] Definition see Appendix A.

[12] TFC: Three Fold Coincidence. It is an algorithm to suppress the neutron-associated cosmogenic isotopes. See Section 2.5.4.

Item	ν(pp) %			ν(^{7}Be) %			ν(pep) %		
	$\Delta\mu$	σ_L	σ_R	$\Delta\mu$	σ_L	σ_R	$\Delta\mu$	σ_L	σ_R
Injected	133.26			48.26			2.8		
MZ	0.3	-6.0	6.8	-0.0	-2.4	2.3	-2.2	-14.4	10.4
MZ sys	0.0	18.5	14.7	-0.2	1.6	1.1	-3.0	10.3	6.1
LNGS	0.2	-6.3	6.7	-0.1	-2.4	2.2	-2.1	-14.9	11.1
LNGS sys	-0.2	18.3	14.3	-0.1	1.7	1.3	-2.2	10.1	8.9
MI	0.4	-6.4	7.1	-0.1	-2.4	2.3	-2.1	-14.8	11.1
MI sys	-0.9	18.8	14.1	-0.2	1.9	0.9	-2.7	9.9	7.1

Table 5.3: The comparison of three TFCs. Here ^{210}Bi shape sys. is not included. ν(CNO) is injected and constrained to be HZ. The unit of injected rate is cpd/100t. The uncertainties are in percent. Here and after $\Delta\mu$ represents the bias of the median value.

Item	ν(pp) %			ν(^{7}Be) %			ν(pep) %		
	μ	σ_L	σ_R	μ	σ_L	σ_R	μ	σ_L	σ_R
Injected	133.26			48.26			2.8		
inj CNO fit CNO	0.3	-6.0	6.8	-0.0	-2.4	2.3	-2.2	-14.4	10.4
inj CNO fit CNO sys	0.0	18.5	14.7	-0.2	1.6	1.1	-3.0	10.3	6.1
inj noCNO fit CNO	0.6	-5.3	7.2	0.1	-2.2	2.3	-30.1	-42.4	-18.7
inj noCNO fit CNO sys	0.2	17.8	15.0	-0.1	1.5	1.2	-30.7	16.3	0.0
inj CNO fit no CNO	-0.3	-6.7	6.2	-0.2	-2.6	2.2	26.1	13.8	38.8
inj CNO fit no CNO sys	-0.7	19.1	14.2	-0.3	1.5	0.9	25.6	0.0	12.5
inj noCNO fit no CNO	0.1	-6.0	6.6	-0.1	-2.4	2.1	-2.0	-14.1	9.6
inj noCNO fit no CNO sys	-0.3	18.2	14.6	-0.2	1.6	1.2	-2.5	9.6	6.9

Table 5.4: Comparison of fit w/ w/o ν(CNO) injected and fit w/ w/o ν(CNO). Here the starting point is around 95 nhit (nhit2 = 16) and ^{210}Bi shape sys. is not included. NL is injected with 8 anchor points (old method). When fit CNO, the CNO is constrained to HZ. when fit no CNO, CNO is fixed to 0.

5.5 *Evaluation of systematic uncertainties for MC fit*

In this section the evaluation of the statistical and systematic uncertainties of the Monte Carlo fit are reported.

The *pp*, ^{7}Be and *pep* solar neutrino interaction rates are extracted with the CNO solar neutrino interaction rate constrained according to the SSM predictions. The breakdown of the systematic uncertainties is listed in table 5.5.

Table 5.5: The breakdown of the systematic uncertainties. ν(CNO) is injected as 4.91 cpd/100t and constrained to be 4.91±0.56 cpd/100t. The unit is cpd/100 tons.

Item	ν(pp)			$\nu(^7$Be)			ν(pep)		
	μ	σ_L	σ_R	μ	σ_L	σ_R	μ	σ_L	σ_R
Injected	133.26			48.26			2.8		
stat	0.75	12.15	13.23	-0.00	1.14	1.14	-0.06	0.41	0.31
z	1.02	2.80	6.52	-0.01	0.14	0.00	-0.06	0.04	0.00
$\mathcal{L}(\mathbf{r})$	0.57	0.00	0.00	-0.03	0.38	0.00	-0.07	0.13	0.00
Bi210	0.62	0.00	0.92	-0.02	0.00	0.00	-0.05	0.00	0.00
NL	1.12	7.12	10.65	-0.01	0.31	0.13	-0.06	0.16	0.08
all	1.36	9.68	13.14	-0.04	0.47	0.12	-0.06	0.17	0.07

The systematic uncertainty of ν(CNO) limit with ν(pp)/ν(pep) constraint is estimated with ν(CNO) injected according to high metallicity model, and it is listed in Tab. 5.6.

Item	Injected	95% C.L. Upper limit	σ(sys)
statistical	4.91	8.60	
all		10.1	3.6

Table 5.6: The systematic uncertainties for ν(CNO) 95% confidence level upper limit. The unit is cpd/100t. ν(pp)–ν(pep) ratio is constrained.

We study also the systematic uncertainties with ν(CNO) injected according to low metallicity or no CNO case, and the systematic uncertainties does not change and follow Eq. 5.16:

$$(\mu_{\text{up}}^{\text{stat.+sys.}} - \hat{\mu})^2 = (\mu_{\text{up}}^{\text{stat.}} - \hat{\mu})^2 + \sigma^2_{\text{sys.}} \tag{5.16}$$

$$1 - \alpha_{\text{up}} = 90\% \tag{5.17}$$

where μ_{up} is the upper limit, $\hat{\mu}$ is the injected CNO rate, $1 - \alpha_{\text{up}}$ is the confidence level, and σ_{sys} is the systematic uncertainty of the CNO rate limit.

Summary The break down of the systematic uncertainties is summarized in Table 5.7 and the total systematic uncertainties considering the non-linearity, detector response uniformity, user model and the ^{210}Bi shape are given as [13]

- ν(pp) 133.79$^{+13.3}_{-12.2}$(stat.) $^{+13.2}_{-9.7}$(mod. sys.) $^{+1.1}_{-1.6}$(expo. sys.)
- $\nu(^7$Be) 48.26 $^{+1.14}_{-1.14}$(stat.) $^{+0.12}_{-0.47}$(mod. sys.) $^{+0.39}_{-0.58}$(expo. sys.)
- ν(pep) 2.8$^{+0.31}_{-0.41}$(stat.) $^{+0.07}_{-0.17}$(mod. sys.) $^{+0.03}_{-0.04}$(expo. sys.)

The systematic uncertainties of the ν(CNO) 95% C.L. limit is studied. We find the sensitivity and systematic uncertainty of ν(CNO) is independent of the injected ν(CNO) rate, and the systematic uncertainties is 3.6 cpd/100t.

[13] Here the CNO is injected with high metallicity and constrained to be of high metallicity. The stat. denotes statistical uncertainty, mod. sys. denotes fit model related systematic uncertainty, and expo. sys. denotes exposure related systematic uncertainty. unit: cpd/100t.

5.6 *Results of analytical fit*

The evaluated systematic uncertainties are listed in table 5.7. As we can see the span of the light yield is almost proportional to the magnitude of the injected non-linearity distortion.

Item	ν(pp)			ν(^{7}Be)			ν(pep)			LY (p.e./MeV)		
	μ	σ_L	σ_R	μ	σ_L	σ_R	μ	σ_L	σ_R	μ	σ_L	σ_R
Injected	134.2			47.9			2.64			557.4		
no deform	-3.86	14.80	6.87	-0.23	1.60	1.09	-0.06	0.50	0.37	0.45	1.56	2.59
all	-4.55	4.67	0.00	-0.27	0.00	0.00	-0.05	0.00	0.09	0.37	1.15	1.01

Table 5.7: The evaluated systematic uncertainties. ν(CNO) is injected as 4.91 cpd/100t and constrained to be 4.91±0.56 cpd/100t. The unit is cpd/100 tons.

The systematic uncertainties are studied and the recommended values for the total systematic uncertainties considering the non-linearity, detector response uniformity, user model and the ^{210}Bi shape are given (in cpd/100t):

- ν(pp) 134.2±10.0(stat.) $^{+0.0}_{-4.7}$(det. sys.) $^{+1.1}_{-1.6}$(expo. sys.)
- ν(^{7}Be) 47.9$^{+1.09}_{-1.60}$(stat.) $^{+0.00}_{-0.00}$(det. sys.) $^{+0.39}_{-0.58}$(expo. sys.)
- ν(pep) 2.64$^{+0.37}_{-0.50}$(stat.) $^{+0.09}_{-0.00}$(det. sys.) $^{+0.03}_{-0.04}$(expo. sys.)

5.7 *Conclusions and discussions*

With a toy Monte Carlo method, we evaluate the expected statistical and systematic uncertainties of Borexino Phase-II solar neutrino analysis. The systematic uncertainties induced by the inaccuracy in the modeling of the detector response include the energy scale, uniformity of the energy response, pulse-shape discrimination shape, and uncertainties in the theoretical energy spectra used in the fit are included. Three TFC algorithms are compared and the Mainz TFC is found to be slightly better compared with the other two. The wrong ν(CNO) penalty only affects ν(pep) significantly and the bias can be up to 30%.

For the Monte Carlo fit, the fit range is optimized to start from `nhit2` = 17, or `nhit` = 106.6, and the fit model related systematic uncertainty of ν(pp), ν(^{7}Be), ν(pep) and ν(CNO) limit are estimated to be around 10%, 1%, 6% and 3.6 cpd/100ton, respectively. For the analytical fit, the fit model systematic uncertainties of ν(^{7}Be) and ν(pep) are negligible and is 3.5% for ν(pp).

For the Monte Carlo fit, the measurement of ν(pp) is mainly driven by two pull terms on ^{14}C and pile-up. When the non-linearity deformation is injected, the number of evens in the fit range increases or decreases because of the migration of events above/below the lower edge of the fitting range, and all these changes are absorbed by the only free component ν(pp). When we extend the fit range to lower energies, the pull term is not strong enough and the ^{14}C can accommodate the event migration, and thus the systematic uncertainties on ν(pp) decreases.

It might be useful to use thinner bins to enhance the information from the shape and weaken the pull or even remove the pull so that ν(pp) systematic uncertainties can decrease even more if we know precisely the ^{14}C and the pile-up energy distributions. However this is not possible since the ^{14}C and the pile-up energy distributions are not modeled with enough accuracy.

The ν(CNO) is narrowly peaked around the center value of the pull. This is as expected. Considering that the spectra does not provide information, the minimum of the likelihood corresponds to the centroid of the pull term.

For the analytical fit, we performed the study with ^{14}C constrained to the injected value. However we see large biases on the *pep* solar neutrino interaction rate, see table 5.8.

Item	ν(pp)			$\nu(^7$Be)			ν(pep)			LY (p.e./MeV)		
	μ	σ_L	σ_R	μ	σ_L	σ_R	μ	σ_L	σ_R	μ	σ_L	σ_R
Injected	134.2			47.9			2.64			557.4		
no deform	-2.20	12.40	7.60	-0.17	1.31	1.09	-0.04	0.42	0.37	-0.13	0.92	0.75
all	-3.45	6.06	0.00	-0.39	0.82	0.00	-0.41	0.70	0.00	-0.45	1.88	0.83

Table 5.8: The evaluated systematic uncertainties with ^{14}C constrained to the injected value. On data we can constrain ^{14}C with the fit on the second cluster. The unit is cpd/100t.

When ^{14}C is free the bias on pep disappear. The systematic uncertainties on $\nu(^7$Be) also disappear. The systematic uncertainties on ν(pp) is smaller and the statistical uncertainty increase. This is as expected: when we let ^{14}C free, part of the systematic uncertainties goes to the statistical uncertainty for ν(pp). The systematic uncertainties on ν(pp) is still one-sided. It could be possible that they comes from the deformation of the z-dependence, which is a one-side-resolution-increase-deformation.

The uncertainty of ν(pp) is much smaller than the Monte Carlo method, because the fit can accommodate the deformed non-linearity. As we can see the span of the light yield increases almost linearly with the injected non-linearity deformation.

6

High precision measurement of solar neutrino fluxes

In this chapter, the measurement of solar neutrino interaction rates is presented. It is the first time that we measure all low energy solar neutrino branches using one single detector and one unique dataset. The solar neutrino interaction rates are extracted simultaneously using analytical `charge`, analytical `npmt_dt1` and Monte Carlo `nhit` fitting methods. The results of the analytical fit are obtained with the software `bx-GooStats`[1] and the fitting procedure is accelerated by using the Graphics Processing Units (GPUs) and parallel computing techniques. Results of Monte Carlo `nhit` fit are obtained with the software `bx-stats`[2]. All results have been extracted using a multivariate analysis based on energy, distance from the detector center and a pulse shape estimator. Two analytical fit results use the new multivariate fitting method described in Chapter 3. The Monte Carlo fit result uses three-dimensional probability density functions and data distribution histograms. The evaluation of the systematic uncertainties are described in Chapter 5.

Using the measured solar neutrino interaction rates, we test its compatibility with two Standard Solar Models (SSMs) assuming the MSW-LMA model of neutrino oscillations, and measure the solar neutrino survival probabilities assuming the SSM model of neutrino fluxes.

This chapter is organized as in the following: Section 6.1 presents the event selection. Section 6.2 presents the measured solar neutrino rates using `charge` and analytical probability density functions. Section 6.3 presents the results using `npmt_dt1` and analytical probability density functions. Section 6.4 presents the results using `nhit` and Monte Carlo probability density functions. Section 6.6 presents the test on the Standard Solar Models (SSMs). Section 6.7 presents the tests on the model of the solar neutrino survival probabilities. Section 6.8 is devoted to conclusions and discussions.

I have co-worked with Nicola Rossi in the work of analytical `charge` fit, with Alina Vishneva, Oleg Smirnov, Ömer Penek, Zara Bagdasarian and Mariia Redchuk in the work of analytical `npmt_dt1` fit, and have co-worked with Daniele Guffanti and Matteo Agostini in the work of Monte Carlo `nhit` fit. The results of this work are used in[3].

[1] X. Ding, "GooStats: A GPU-based framework for multi-variate analysis in particle physics", *Journal of Instrumentation*, vol. 13, no. 12, P12018–P12018, Dec. 2018, DOI: `10.1088/1748-0221/13/12/P12018`

[2] M. Agostini, *M-STATS: framework for frequentist statistical analysis*, https://github.com/mmatteo/m-stats, 2015

[3] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

6.1 *Event selection*

In this section we present the list of criteria to select the event sample.

The dataset used in the analysis expand from 14th December 2011 (Run 17328) to 21st May 2016 (Run 26568), corresponding to 1291.2 days of DAQ time and 1270.3 days of live time after all cuts. The duty cycle is shown in Figure 6.1.

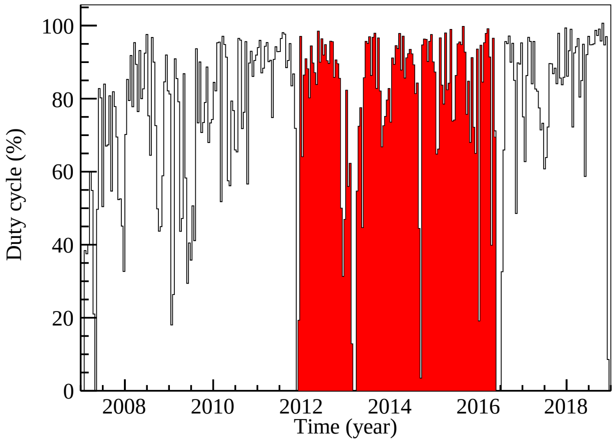


Figure 6.1: The duty cycle of Borexino data. The data used in this thesis is marked as red. The loss is due to operational reasons or degraded data quality.

The analysis is based on clusters, i.e. group of hits selected by an algorithm[4] that are believed to belong to one physics event. The following cuts are applied sequentially:

1. muon cut. Clusters tagged as muons are rejected.
2. muon daughter cut. All clusters within 300 ms of a muon are rejected.
3. zero cluster cut. Events without clusters are rejected. This cut is essential because an event without clusters is wrapped as a special cluster.
4. fast coincidence cut. If two clusters are within 2 micro-seconds and 1.5 meters, they are both rejected.
5. ^{214}Bi–^{214}Po cut. If two clusters are within 20–1180 micro-seconds and 1 meter, have $90 < Q(\text{prompt}) < 1800$ and $200 < Q(\text{delayed}) < 500$, then they are both rejected.
6. trigger cut. Clusters of events of service triggers, such as laser calibration triggers, random triggers, are rejected.
7. multi-cluster cut. A cluster is kept only if the associated event contains only one cluster or this cluster is the first cluster of the two clusters contained in the event.
8. start time cut. If the start time of the cluster is out of the expected time window, the event is rejected.
9. fiducial volume cut. Clusters with reconstructed position within the region $r < 2.8$ m and $z > -1.8$ m and $z < 2.2$ m are accepted. This region is called the *pep* fiducial volume[5].

[4] Section 7.4.3, page 207 D. D'Angelo, "Towards the detection of low energy solar neutrinos in BOREXino: data readout, data reconstruction and background identification", PhD, Munich, Tech. U., 2006, p. 334

[5] A. E. Chavarria, "Study on pep and CNO solar neutrino interaction rates in Borexino", PhD thesis, Princeton University, 2012

10. noise cuts. explained below.

The noise cuts used in this analysis are almost the same as the one used in the pp analysis[6] and the one described in [7]:

1. Charge consistency cut. The energy estimated by `charge` and `npmt` should be statistically consistent. The energy estimated by `charge` and `charge_noavg` should be consistent.

2. Mach4 charge consistency cut. The energy estimated by Mach4 `charge` and Mach4 `charge_noavg` should be consistent. Here Mach4[8] is name of an alternative data analysis package, and the clusterization algorithm in it is different.

3. crate fraction. If more than 75% of fired PMTs are physically attached to the same crate (out of $\mathcal{O}(10)$ crates), the cluster is considered to be most likely crate noise and rejected.

Compared with previous studies, the cut requiring the consistency between `charge` and `charge_noavg` is new. It is needed in order to remove unphysical events in 2015 in the sensitive region of pp solar neutrinos. See the comparison of spectra of `charge` and `npmt_dt1` between 2014 and 2015 in Figure 6.2. For `npmt_dt1` the difference between two spectra comes from the loss of PMT and thus degrades the energy resolution. For `charge` there are extra events in the critical region that cannot be explain with this reason.

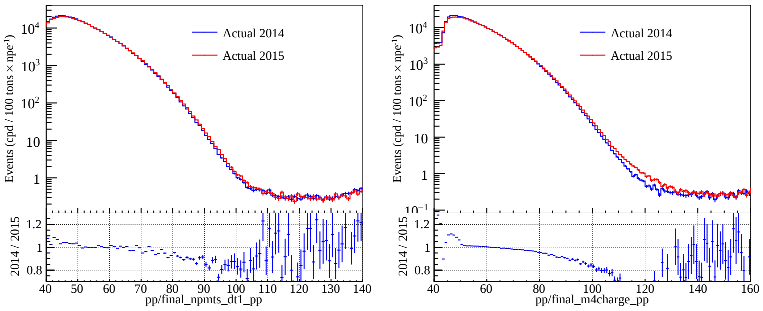


Figure 6.2: Comparison of spectra between 2014 and 2015. Left: `npmt_dt1`. Right: `charge`. We can see that there are extra events in the region around 120 p.e. for `charge` and they are not present in the `npmt_dt1` spectra.

These events have been attributed to biased ^{14}C events. Sometimes one or more channel receives extra-ordinary large charge, then the `npmt_dt1` variable is biased only by 1 or 2, while the `charge` variable is usually biased by a few p.e.'s. This effect is corrected in the `charge_noavg` variable. If the collected `charge` is much larger than expected, then the charge collected on this channel is set to zero. More specifically, the cut criteria is

$$Q_{\mathrm{ch.}} > 3 + \frac{Q_{\mathrm{n.cl.}}}{500} \tag{6.1}$$

where $Q_{\mathrm{ch.}}$ is the charge collected on one channel, and $Q_{\mathrm{n.cl.}}$ is the normalized charge of the cluster. As a result, these events can be selected using the ratio between `charge` and `charge_noavg`. See the distribution of the ratio in Figure 6.3.

[6] P. Mosteiro, "First measurement of pp neutrinos in real time in the Borexino detector", PhD thesis, Princeton University, 2014 Section 4.2, page 65

[7] S. Marcocci, "Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with Xe -loaded Liquid Scintillators", PhD thesis, Gran Sasso Science Insitute, 2016 Section 6.3, page 195

[8] R. N. Saldanha, "Precision Measurement of the 7 Be Solar Neutrino Interaction Rate in Borexino", PhD thesis, Princeton University, 2012 Section 4.0.12, page 99

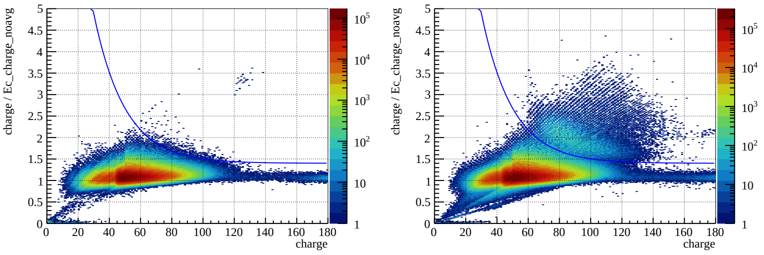


Figure 6.3: `charge/charge_noavg` distribution of data collected in time other than 2015 (left) and in 2015 (right).

The new charge consistency cut is defined based on Figure 6.3. It is shown in the figure as a blue solid line and is defined as the following:

$$\frac{\texttt{charge}}{\texttt{charge\_noavg}} > 1.3 + e^{-0.043\cdot(\texttt{charge}-65)} \tag{6.2}$$

The stability of event rejection rate of the charge consistency cut is shown in Figure 6.4.

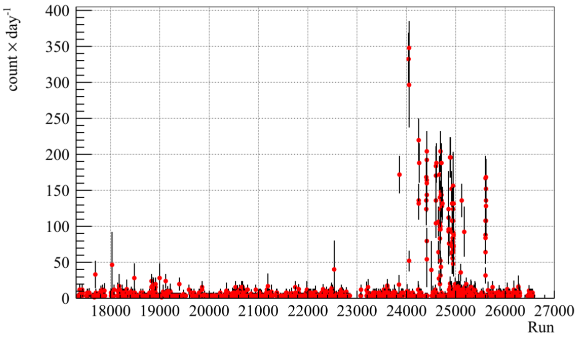


Figure 6.4: Stability of event rejection rate of the charge consistency cut

The noise cut removes few physics events[9] and after the selection the unphysical events contribution is negligible.

6.2 *Results with* `charge`

In this section we present the results of using analytical probability density functions of the `charge` variable. Compared with the analytical `npmt_dt1` fit and the Monte Carlo `nhit` fit, in the `charge` fit the low energy region is not included, and thus there is not sensitivity of the *pp* solar neutrinos.

Compared with the Phase-I analysis, in order to achieve 3% precision of the ^{7}Be solar neutrino interaction rate, one of two resolution parameters is fixed and the other is fixed through the width of the ^{210}Po peak. If both are left free, because the ^{7}Be solar neutrino interaction rate is correlated with the two resolution parameters, its statistical uncertainty will be around 4%. The parameter which represents the energy response non-uniformity, is fixed to the value determined according to `g4bx2`. This approach is motivated, because the non-uniformity of `g4bx2` is tuned to match that of the calibration data. In the analytical `charge` fit, same model of the resolution has been used and thus the width of the ^{210}Po peak can be used to fix one degree

[9] Less than 0.2% for all energies, see Table XIV The Borexino collaboration, “Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy”, *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 89, no. 11, pp. 1–68, 2014, DOI: `10.1103/PhysRevD.89.112007`, arXiv: `1308.0443`

of freedom in the resolution model. It is important to put the correct f_{eq} parameter in the model, which is smaller for ^{210}Po events compared with β events. As discussed in Section 4.2.5 , the resolution of α events is slightly better than that of β events even when the average and f_{eq} is the same, due to quenching effect and the Cherenkov radiation contribution, but the difference is small for low energy events, such as the ^{210}Po decays, and the residual difference is included in the systematic uncertainties.

Because ^{210}Bi, the *pep* and CNO solar neutrinos are correlated with each other, two strategies are applied. The ^{7}Be solar neutrino and *pep* solar neutrino interaction rates are measured with the CNO solar neutrino interaction rates constrained to the SSM-HM model or SSM-LM model. To evaluate the CNO solar neutrino interaction rate, the *pp* solar neutrino interaction rate is constrained to the value obtained from Phase-I and the *pp*–*pep* interaction rate ratio is constrained according to SSM prediction.

Because the low energy part of the spectrum is not included, the ^{85}Kr decay rate is constrained according to the result obtained with a fast coincidence analysis. This analysis only gives an upper limit of 4.9 cpd/100t (95% C.L.)[10]. This subsidiary measurement is included as a pull term of the ^{85}Kr decay rate, and a half-Gaussian penalty function is used:

$$
\begin{aligned}
-\log \mathcal{L}(\mathcal{R}) &= \left(\frac{R}{\sigma}\right)^2 && \text{if } R \geqslant 0 \\
&= +\infty && \text{if } R < 0
\end{aligned}
\qquad (6.3)
$$

where R is the ^{85}Kr decay rate, and σ is set to 2.5 cpd/100t.

6.2.1 Statistical and systematic uncertainties

The statistical uncertainties are obtained from the profiled likelihoods. We have considered three types of systematic uncertainties:

- Exposure related. This term includes the uncertainties of the live time, the scintillation density and the fiducial volume and is dominated by the fiducial volume uncertainty. It is evaluated to be the same as that of phase-I $^{+0.6\%}_{-1.1\%}$[11].
- Fit model related. Pseudo-experiment spectra are produced based on the probability density functions produced with `g4bx2`, and the biases are taken as the systematic uncertainties.
- ^{85}Kr penalties. Fits with and without the pull term on ^{85}Kr decay rate were performed, and the difference is considered as the systematic uncertainties.

6.2.2 The ^{7}Be and *pep* solar neutrino interaction rate measurements

CNO HM Here the CNO solar neutrino interaction rate is constrained to the prediction of the High Metallicity model (4.92 ± 0.56

---

[10] Section 5.2.2 page 119 I. Drachnev, "New Spectral Analysis of Solar B Neutrino with the Borexino Detector", PhD thesis, Gran Sasso Science Institute, 2016

[11] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy"

count/day/100t). The results of charge fit is shown in Figure 6.5 and the centroid values and systematic uncertainties of each species are listed in Table 6.1. The profiled likelihoods of ^{7}Be and *pep* solar neutrino interaction rates are shown in Figure 6.6.

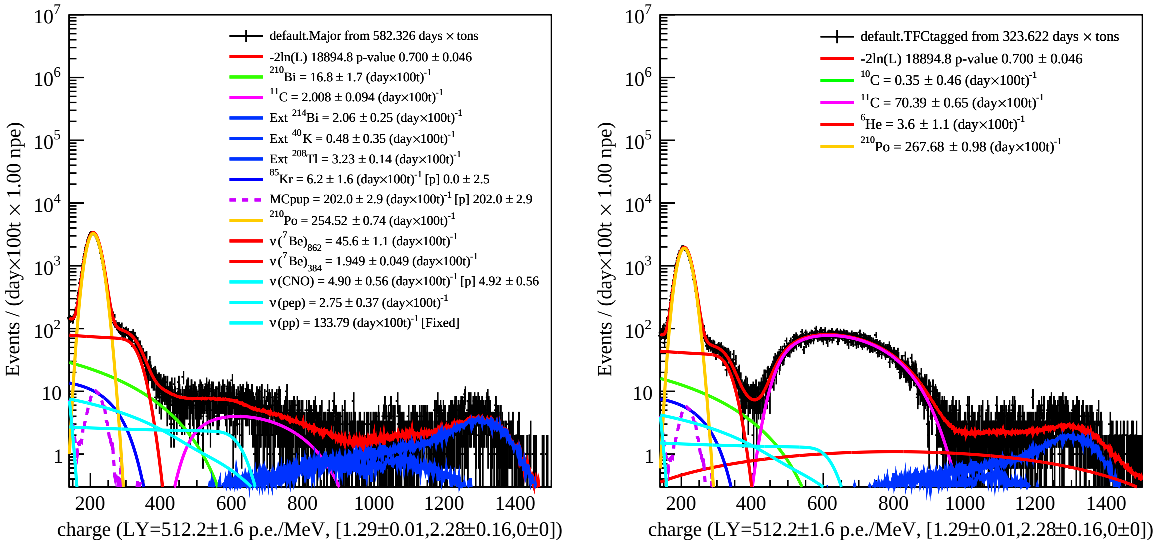


Figure 6.5: Fit results of analytical charge fit. The CNO solar neutrino interaction rate is constrained to the High Metallicity model.

Table 6.1: Fit results of ^{7}Be and *pep* solar neutrino interaction rates using `charge` with CNO solar neutrino interaction rate constrained to the High Metallicity model. Unit: count per day per 100 tons.

Species	best fit	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$	$\sigma^{\text{exp.}}_{\text{sys.}}$	$\sigma^{\text{RPF}}_{\text{sys.}}$	$\sigma^{\text{Kr}}_{\text{sys.}}$
$\nu(^{7}\text{Be})$	47.5	±1.2	$^{+0.3}_{-1.7}$	$^{+0.3}_{-0.5}$	$^{+0.2}_{-0}$	$^{+0.0}_{-1.6}$
$\nu(pep)$	2.75	±0.37	$^{+0.36}_{-0.03}$	$^{+0.02}_{-0.03}$	$^{+0.31}_{-0}$	$^{+0.18}_{-0}$
^{210}Bi	16.8	±1.7	$^{+2.6}_{-1.7}$	$^{+0.1}_{-0.2}$	$^{+2.6}_{-0}$	$^{+0}_{-1.7}$
^{85}Kr	6.2	±1.6	$^{+4.4}_{-0.1}$	$^{+0.0}_{-0.1}$	$^{+0}_{-0}$	$^{+4.4}_{-0}$
^{11}C sub	2.008	±0.094	$^{+0.012}_{-0.040}$	$^{+0.012}_{-0.022}$	$^{+0}_{-0}$	$^{+0}_{-0.033}$
^{11}C comp	70.39	0.65	$^{+0.43}_{-0.77}$	$^{+0.42}_{-0.77}$	$^{+0}_{-0}$	$^{+0.09}_{-0}$
^{210}Po sub	254.52	0.74	$^{+1.53}_{-2.83}$	$^{+1.53}_{-2.80}$	$^{+0}_{-0.21}$	$^{+0}_{-0.38}$
^{210}Po comp	267.68	0.98	$^{+1.61}_{-2.98}$	$^{+1.61}_{-2.94}$	$^{+0}_{-0.22}$	$^{+0}_{-0.37}$
Ext ^{208}Tl	3.23	0.14	$^{+0.02}_{-0.04}$	$^{+0.02}_{-0.04}$	$^{+0}_{-0}$	$^{+0}_{-0}$
Ext ^{214}Bi	2.06	0.25	$^{+0.01}_{-0.02}$	$^{+0.01}_{-0.02}$	$^{+0}_{-0}$	$^{+0}_{-0}$
Ext ^{40}K	0.48	0.35	$^{+0.12}_{-0.02}$	$^{+0}_{-0.01}$	$^{+0.12}_{-0}$	$^{+0}_{-0.02}$

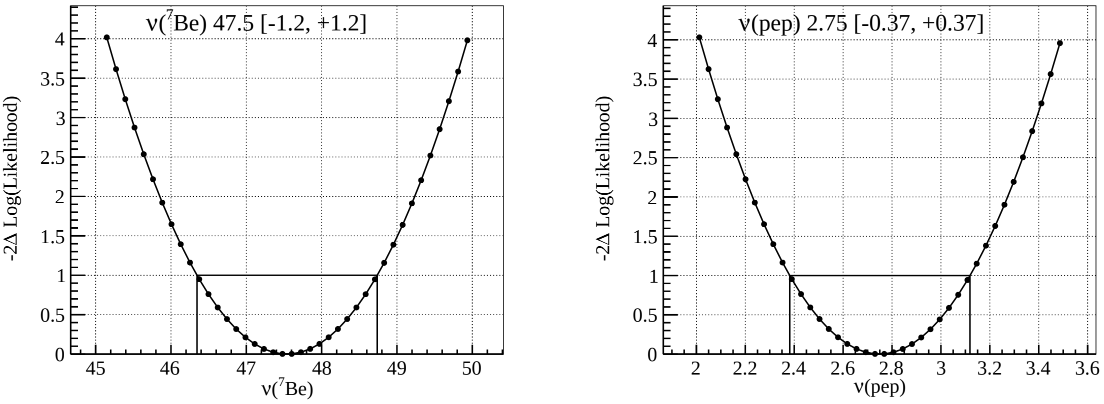


Figure 6.6: Profiled likelihoods of the ^{7}Be and *pep* solar neutrino interaction rates in charge fit (CNO HM). The 68% confidence interval is shown on the figures. Units: count per day per 100 tons.

CNO LM Here the CNO solar neutrino interaction rate is constrained to the prediction of the Low Metallicity model (3.52 ± 0.37 count/day/100t). The centroid values and systematic uncertainties of neutrino species are listed in Table 6.2. From the table we can see that compared to the results obtained with CNO solar neutrino interaction rate constrained to the High Metallicity model, the best fit value and statistical and systematic uncertainties of the ^{7}Be solar neutrino interaction rate do not change, while the best fit value of the *pep* solar neutrino interaction rate increases by 8%, and its statistical and systematic uncertainties do not change.

Table 6.2: Fit results of ^{7}Be and *pep* solar neutrino interaction rates using `charge` with CNO solar neutrino interaction rate constrained to the Low Metallicity model. Unit: count per day per 100 tons.

Species	best fit	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$	$\sigma^{\text{exp.}}_{\text{sys.}}$	$\sigma^{\text{RPF}}_{\text{sys.}}$	$\sigma^{\text{Kr}}_{\text{sys.}}$
$\nu(^7\text{Be})$	47.5	± 1.2	$^{+0.3}_{-1.7}$	$^{+0.3}_{-0.5}$	$^{+0.2}_{-0}$	$^{+0.0}_{-1.6}$
$\nu(pep)$	2.97	± 0.36	$^{+0.39}_{-0.03}$	$^{+0.02}_{-0.03}$	$^{+0.34}_{-0}$	$^{+0.19}_{-0}$

6.2.3 The CNO solar neutrino interaction rate limit

The CNO solar neutrino interaction rate limit is obtained with *pp* solar neutrino interaction rate constrained to the value reported in Borexino Phase-I analysis, and the ratio of the *pp* and *pep* interaction rate constrained to the value predicted by the SSMs, which is almost the same between the High and Low Metallicity models.

The upper limit obtained here is 5.1 cpd/100t (95% C.L.) without systematic uncertainties. The profiled likelihoods of CNO solar neutrino interaction rate is shown in Figure 6.7.

6.3 Results with `npmt_dt1`

In this section we present the results of using analytical probability density functions of the `npmt_dt1` variable. Compared with the other two fit strategies, `npmt_dt1` fit has the best sensitivity to *pp* solar neutrino interaction rate.

In the analytical `npmt_dt1` fit approach, two free parameters are assigned to the detector resolution model, accounting for the intrinsic resolution of β-like particles caused by δ-electrons and the resolution caused by spatial dependence of the number of triggered PMTs; one

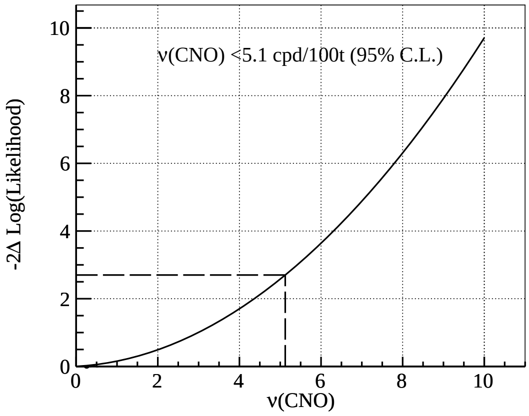


Figure 6.7: Profiled likelihoods of the CNO solar neutrino interaction rates in `charge` fit with pp/pep constrained. The pp solar neutrino interaction rate is constrained to the value reported in Borexino Phase-I analysis. The 95% confidence interval is shown on the figures. Units: count per day per 100 tons.

free parameter, the light yield, is assigned to the detector energy response non-linearity, and other parameters are fixed to `g4bx2`. Compared with that of Phase-I, the fitting range is increased to include the whole ^{11}C decay and the high energy external γ backgrounds, and thus an additional fixed parameter is added in the resolution model, and the Monte Carlo probability density functions are used for external γs.

Same as the `charge` analysis, the pp, ^{7}Be and pep solar neutrino interaction rates are measured with the CNO solar neutrino interaction rates constrained to the SSMs; To obtain the CNO solar neutrino limit, the pp–pep ratio is constrained according to SSM prediction.

Because the low energy part is included in the fit, the ^{85}Kr decay rate can be left free and the statistical uncertainty of neutrino interaction rate is still acceptable, the final results were obtained without the ^{85}Kr decay rate constraint.

6.3.1 *The systematic uncertainties*

We have considered five types of systematic uncertainties:

- Exposure related. Same as `charge`.
- Fit model related. This term includes the uncertainties due to possible inaccuracies in the modeling of the detector response (energy scale, uniformity of the energy response, shape of pulse-shape discriminators) and uncertainties in the theoretical energy spectra (^{210}Bi). Pseudo-experiment spectra with deformations are produced and the width of the best fit distribution is used to evaluate this systematic uncertainties. The magnitude of the injected deformation is chosen to be within the range allowed by the calibration data.
- ^{85}Kr penalties. Same as `charge`.
- Treatment of pile-up effect. The pile-up effect can be either included by convolving the spectrum with a solicited-trigger spectrum[12], or included as an independent component obtained from data or Monte Carlo. The difference of the results obtained with

[12] The Borexino collaboration, "Neutrinos from the primary proton–proton fusion process in the Sun", *Nature*, vol. 512, no. 7515, pp. 383–386, Aug. 2014, DOI: `10.1038/nature13702`, arXiv: `1508.05379`

the second or third methods with respect to that obtained with the first method is taken as the systematic uncertainty.

6.3.2 The ^{7}Be and *pep* solar neutrino interaction rate measurements

CNO HM Here the CNO solar neutrino interaction rate is constrained to the prediction of the High Metallicity model (4.92 ± 0.56 count/day/100t). The results are shown in Figure 6.8 and the centroid values and systematic uncertainties of each species are listed in Table 6.3. The profiled likelihoods of *pp*, ^{7}Be and *pep* solar neutrino interaction rates are shown in Figure 6.9.

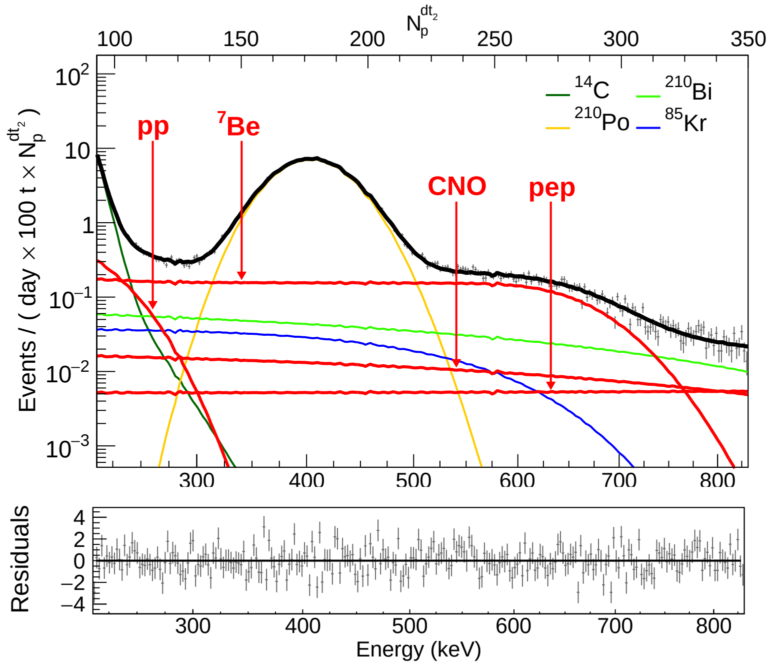


Figure 6.8: Results of the fit for TFC-subtracted energy spectrum zoomed in to the lowest energy region (an example obtained with the analytical method) and residuals.

Species	μ	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$	$\sigma_{\text{sys.}}^{\text{exp.}}$	$\sigma_{\text{sys.}}^{\text{RPF}}$	$\sigma_{\text{sys.}}^{\text{Kr}}$	$\sigma_{\text{sys.}}^{\text{pup}}$
$\nu(pp)$	133.8	±10.1	$^{+6.7}_{-5.9}$	$^{+0.8}_{-1.5}$	$^{+0}_{-4.7}$	$^{+6.6}_{-0}$	$^{+0.7}_{-3.3}$
$\nu(^7\text{Be})$	47.8	±1.3	$^{+0.9}_{-0.5}$	$^{+0.3}_{-0.5}$	$^{+0}_{-0}$	$^{+0.9}_{-0}$	$^{+0}_{-0}$
$\nu(pep)$	2.61	±0.40	$^{+0.09}_{-0.12}$	$^{+0.02}_{-0.03}$	$^{+0.09}_{-0}$	$^{+0}_{-0.12}$	$^{+0}_{-0}$

Table 6.3: Fit results of *pp*, ^{7}Be and *pep* solar neutrino interaction rates using `npmt_dt1` with CNO solar neutrino interaction rate constrained to the High Metallicity model. Unit: count per day per 100 tons.

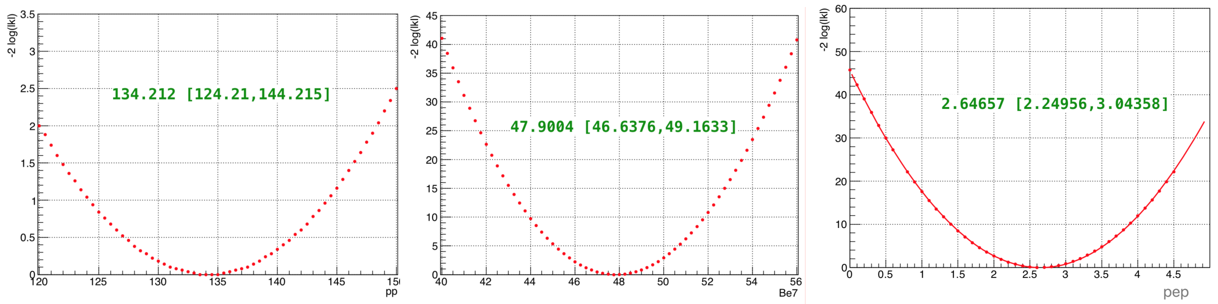


Figure 6.9: Profiled likelihoods of the *pp* (left), ^{7}Be (middle) and *pep* (right) solar neutrino interaction rates in `npmt_dt1` fit (CNO HM). The 68% confidence interval is shown on the figures. Units: count per day per 100 tons.

CNO LM Here the CNO solar neutrino interaction rate is constrained to the prediction of the Low Metallicity model (3.52 ± 0.37 count/-day/100t). The centroid values and systematic uncertainties of neutrino species are listed in Table 6.4. Similar to `charge`, compared with results with CNO solar neutrino interaction rate constrained to the High Metallicity model, only the best fit value of the *pep* solar neutrino interaction rate increased by 8%.

Species	μ	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$	$\sigma^{\text{exp.}}_{\text{sys.}}$	$\sigma^{\text{RPF}}_{\text{sys.}}$	$\sigma^{\text{Kr}}_{\text{sys.}}$	$\sigma^{\text{pup}}_{\text{sys.}}$
$\nu(pp)$	133.8	± 10.1	$^{+6.5}_{-8.3}$	$^{+0.8}_{-1.5}$	$^{+0}_{-4.7}$	$^{+6.4}_{-0}$	$^{+0.7}_{-6.7}$
$\nu(^7\text{Be})$	47.8	± 1.2	$^{+0.9}_{-0.5}$	$^{+0.3}_{-0.5}$	$^{+0}_{-0}$	$^{+0.9}_{-0}$	$^{+0}_{-0}$
$\nu(pep)$	2.84	± 0.40	$^{+0.09}_{-0.11}$	$^{+0.02}_{-0.03}$	$^{+0.09}_{-0}$	$^{+0}_{-0.11}$	$^{+0}_{-0}$

Table 6.4: Fit results of *pp*, ^{7}Be and *pep* solar neutrino interaction rates using `npmt_dt1` with CNO solar neutrino interaction rate constrained to the Low Metallicity model. Unit: count per day per 100 tons.

6.3.3 *The CNO solar neutrino interaction rate limit*

The CNO solar neutrino interaction rate limit is obtained with the ratio of the *pp* and *pep* interaction rate constrained to the value predicted by the High Metallicity model 47.8 ± 0.8. The upper limit obtained here is 7.8 cpd/100t (95% C.L.). The profiled likelihoods of CNO solar neutrino interaction rate is shown in Figure 6.10. The systematic uncertainties are not included here.

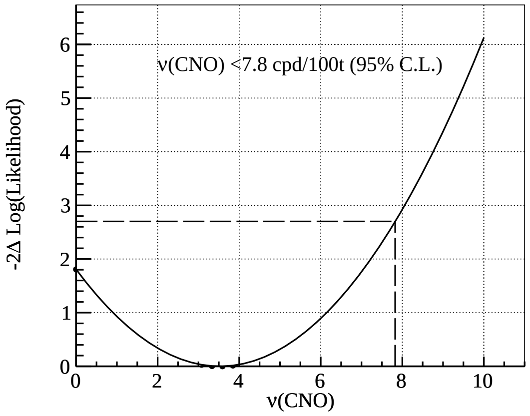


Figure 6.10: Profiled likelihoods of the CNO solar neutrino interaction rates in `npmt_dt1` fit with *pp*/*pep* constrained. The 95% confidence interval is shown on the figures. Units: count per day per 100 tons.

6.4 *Results with* `nhit`

In this section we present the results of using Monte Carlo probability density functions of the `nhit` variable. The results are obtained with the software `bx-stats`[13]. Compared with analytical multivariate analysis or the multivariate fit in[14] which assumes the radial distribution of events, such as cosmogenic ^{11}C decays, are uniform, this analysis uses the multi-dimensional probability density functions and data histograms, and takes into account the dependence of radial distribution on the energies. This describes better the ^{11}C component. The ^{11}C is almost uniformly distributed in the fiducial volume, but because the events of the same energy at large radii receive less p.e.'s, the distribution of r^3 (where r is the radius) with

[13] Agostini, *M-STATS: framework for frequentist statistical analysis*

[14] S. Davini, "Measurement of the pep and CNO solar neutrino interaction rates in Borexino-I", *The European Physical Journal Plus*, Springer Theses, vol. 128, no. 8, p. 89, Aug. 2013, DOI: `10.1140/epjp/i2013-13089-9`

`nhit` restricted to a certain range is no more uniform, but less at large radii, see Figure 6.11.

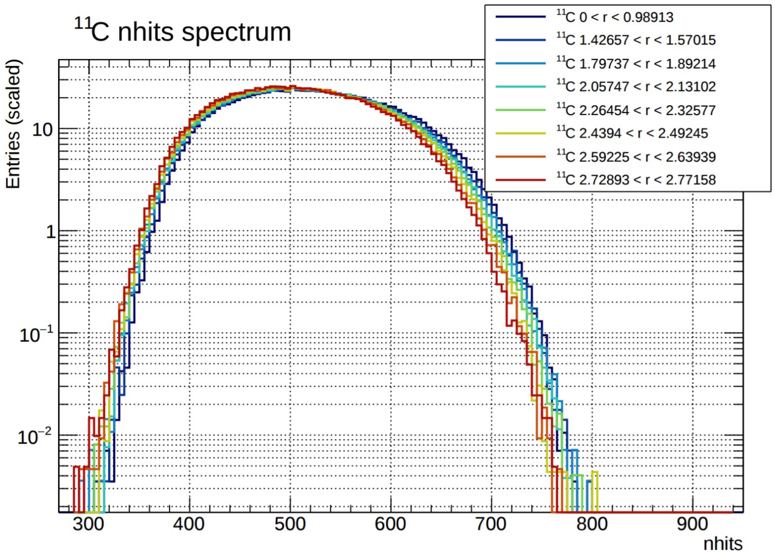


Figure 6.11: Distribution of `nhit` for ^{11}C at different radii. From the figure we can see that at the same `nhit` there is less ^{11}C events at larger radii when `nhit` is larger than 550

In the Monte Carlo `nhit` fit approach, all probability density functions are produced with `g4bx2`. The pile-up effect is modeled with an independent component[15]. The Monte Carlo probability density function shape is fixed during the fit. `nhit` is transformed into `nhit2` to bin data according to the detector energy resolution.

Same as the `charge` analysis, the *pp*, ^{7}Be and *pep* solar neutrino interaction rates are measured with the CNO solar neutrino interaction rates constrained to the SSMs. To obtain the CNO solar neutrino limit, the *pp*–*pep* ratio is constrained according to SSM prediction.

6.4.1 *The systematic uncertainties*

We have considered three types of systematic uncertainties:

- Exposure related.
- Fit model related.
- ^{85}Kr penalties.

More details about these them can be found in Section 6.3.1 on page 108.

6.4.2 *The ^{7}Be and pep solar neutrino interaction rate measurements*

CNO HM Here the CNO solar neutrino interaction rate is constrained to the prediction of the High Metallicity model (4.92 ± 0.56 count/day/100t). The results are shown in Figure 6.12 and the centroid values and systematic uncertainties of each species are listed in Table 6.5. The profiled likelihoods of *pp*, ^{7}Be and *pep* solar neutrino interaction rates are shown in Figure 6.13.

---

[15] Section 5.11 Marcocci, "Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with Xe -loaded Liquid Scintillators"

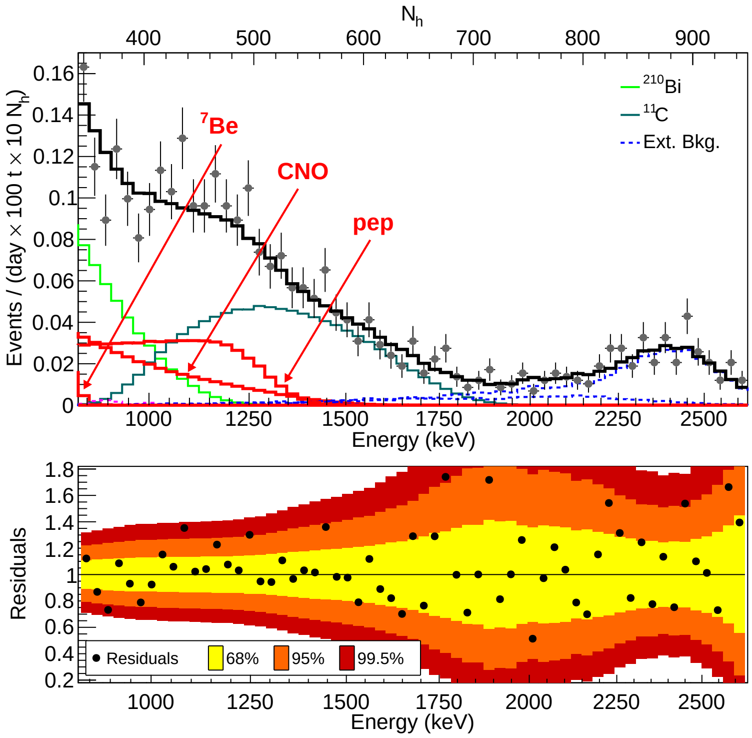


Figure 6.12: Fit results of `bx-stats` Monte Carlo `nhit` fit. The CNO solar neutrino interaction rate is constrained to the High Metallicity model.

Species	best fit	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$	$\sigma^{\text{exp.}}_{\text{sys.}}$	$\sigma^{\text{RPF}}_{\text{sys.}}$	$\sigma^{\text{Kr}}_{\text{sys.}}$
$\nu(pp)$	133.41	± 12.73	$^{+13.92}_{-9.79}$	$^{+0.80}_{-1.47}$	$^{+13.14}_{-9.68}$	$^{+4.51}_{-0}$
$\nu(^{7}\text{Be})$	48.13	± 1.13	$^{+0.49}_{-0.71}$	$^{+0.29}_{-0.53}$	$^{+0.12}_{-0.47}$	$^{+0.38}_{-0}$
$\nu(pep)$	2.38	± 0.36	$^{+0.07}_{-0.19}$	$^{+0.01}_{-0.03}$	$^{+0.07}_{-0.17}$	$^{+0}_{-0.07}$

Table 6.5: Fit results of *pp*, ^{7}Be and *pep* solar neutrino interaction rates using `npmt_dt1` with CNO solar neutrino interaction rate constrained to the High Metallicity model. Unit: count per day per 100 tons.

CNO LM Here the CNO solar neutrino interaction rate is constrained to the prediction of the Low Metallicity model (3.52 ± 0.37 count/day/100t). The centroid values and systematic uncertainties of neutrino species are listed in Table 6.6. Similar to `charge` and `npmt_dt1` fit, compared with results with CNO solar neutrino interaction rate constrained to the High Metallicity model, only the *pep* solar neutrino interaction rate increased by 10%.

Species	best fit	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$	$\sigma^{\text{exp.}}_{\text{sys.}}$	$\sigma^{\text{RPF}}_{\text{sys.}}$	$\sigma^{\text{Kr}}_{\text{sys.}}$
$\nu(pp)$	133.12	± 12.73	$^{+13.92}_{-9.79}$	$^{+0.80}_{-1.47}$	$^{+13.14}_{-9.68}$	$^{+4.51}_{-0}$
$\nu(^{7}\text{Be})$	48.11	± 1.13	$^{+0.49}_{-0.71}$	$^{+0.29}_{-0.53}$	$^{+0.12}_{-0.47}$	$^{+0.38}_{-0}$
$\nu(pep)$	2.61	± 0.35	$^{+0.07}_{-0.19}$	$^{+0.01}_{-0.03}$	$^{+0.07}_{-0.17}$	$^{+0}_{-0.07}$

Table 6.6: Fit results of *pp*, ^{7}Be and *pep* solar neutrino interaction rates using `nhit` with CNO solar neutrino interaction rate constrained to the Low Metallicity model. Unit: count per day per 100 tons.

6.4.3 *The CNO solar neutrino interaction rate limit*

The CNO solar neutrino interaction rate limit is obtained with the ratio of the *pp* and *pep* interaction rate constrained to prediction of the High Metallicity models 47.8 ± 0.8.

The upper limit obtained here is 6.6 cpd/100t (95% C.L.). The profiled likelihoods of CNO solar neutrino interaction rate is shown in Figure 6.14. The systematic uncertainties are not included here.

The total systematic uncertainty on CNO solar neutrino inter-

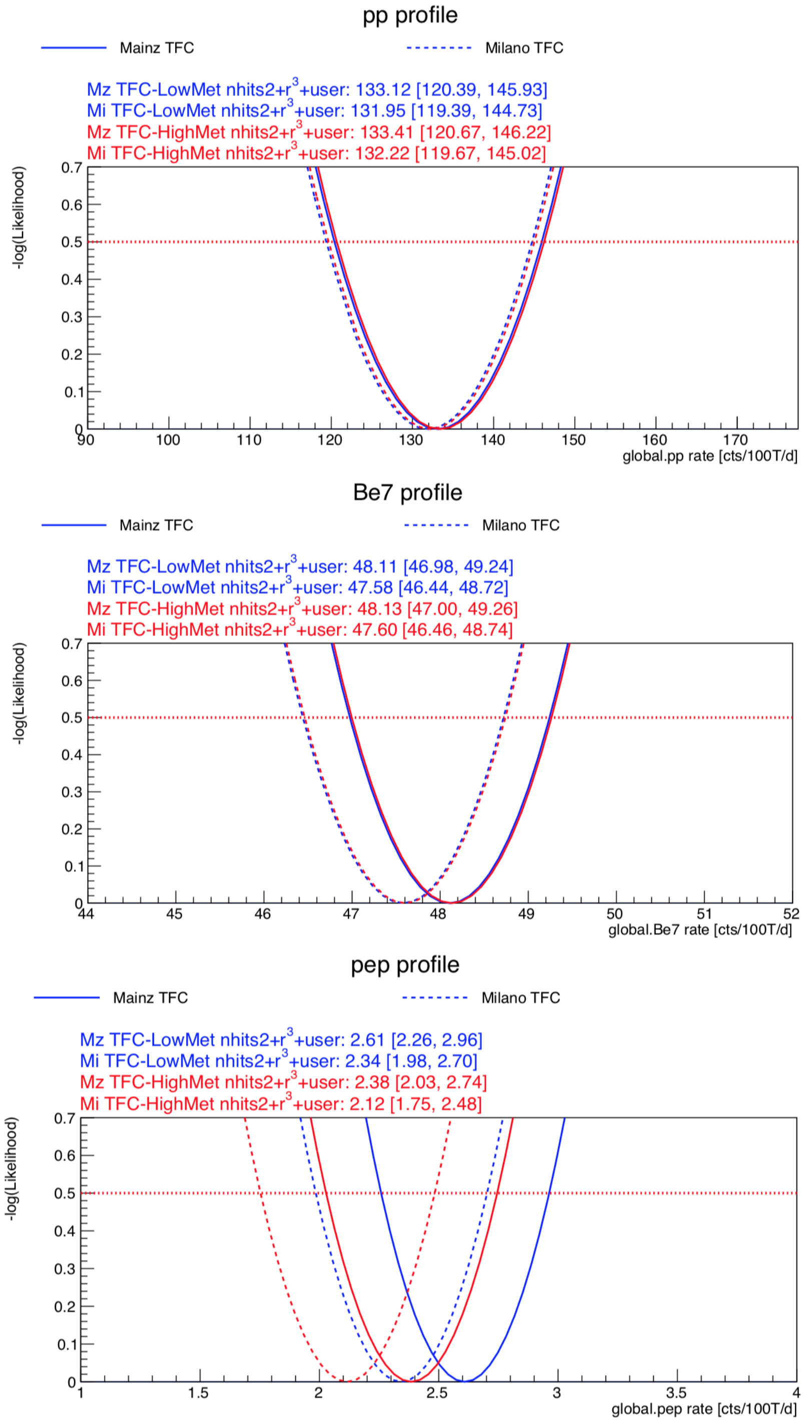


Figure 6.13: Profiled likelihoods of the *pp*, ^{7}Be and *pep* solar neutrino interaction rates in `nhit` fit (CNO HM+LM). The 68% confidence interval is shown on the figures. Units: count per day per 100 tons.

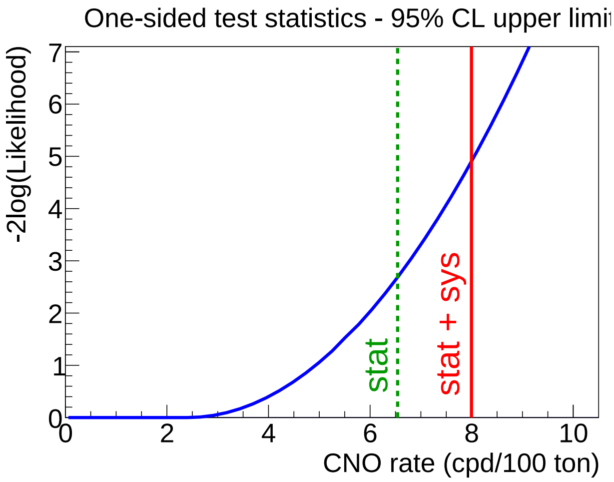


Figure 6.14: Profiled likelihoods of the CNO solar neutrino interaction rates in charge fit with *pp*/*pep* constrained. The 95% confidence interval is shown on the figures. Units: count per day per 100 tons. Reproduced from D. Guffanti, "Measurement of solar neutrino with BOREXINO", PhD thesis, Gran Sasso Science Institute, 2019

action rate, considering the three types of systematic uncertainties mentioned in this section, is evaluated to be 2.2 cpd/100t. Considering the statistical uncertainty is 2.5 cpd/100t, the total uncertainty is

$$\sigma_{\text{tot}}^2 = \sigma_{\text{stat}}^2 + \sigma_{\text{sys}}^2 \tag{6.4}$$

so it is 3.4 cpd/100t. The best fit of CNO solar neutrino interaction rate is 2.4 cpd/100t. The upper limit of CNO solar neutrino interaction rate including the systematic uncertainty is[16]

$$\mu_{\text{up}} = \mu + 1.645\sigma \tag{6.5}$$

so it is 8.0 cpd/100t. According to Monte Carlo simulation, the median 95% C.L. upper limit for CNO solar neutrino interaction rate to be ~9 cpd/100 t and 10 cpd/100 t, for low and high metallicity, respectively. The upper limit we obtained is slightly stronger than the median limit expected from the toy Monte Carlo study. This result, using a weaker hypothesis on *pep* solar neutrino interaction rate, confirms the current best limit on CNO solar neutrino interaction rate previously obtained with Borexino Phase-I data[17].

6.5 *Combination of results: final results*

In this section we present the final results. As described in Sections 6.2–6.4, the neutrino interaction rates can be measured with different fitting methods and different energy estimators. The analytical `charge` results do not include the *pp* solar neutrino interaction rate measurement, so it is not included and only used to cross-check the validity of fit results. The final results are obtained by combining the analytical `npmt_dt1` fit results and the Monte Carlo `bx-stats` results. The difference between two results is taken as a source of systematic uncertainties. The combined list of systematic uncertainties is listed in Table 6.7. The final results of neutrino interaction rate and statistical/systematic uncertainties are listed in Table 6.8.

[16] Equation (60) G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011, DOI: `10.1140/epjc/s10052-011-1554-0`, arXiv: `1007.1727`

[17] The Borexino collaboration, "First Evidence of *pep* Solar Neutrinos by Direct Detection in Borexino", *Physical Review Letters*, vol. 108, no. 5, p. 051302, Feb. 2012, DOI: `10.1103/PhysRevLett.108.051302`, arXiv: `1110.3230`

	pp		^{7}Be		*pep*	
Source of uncertainty	−%	+%	−%	+%	−%	+%
Fit method (analytical/MC)	-1.2	1.2	-0.2	0.2	-4.0	4.0
Choice of energy estimator	-2.5	2.5	-0.1	0.1	-2.4	2.4
Pile-up modeling	-2.5	0.5	0	0	0	0
Fit range and binning	-3.0	3.0	-0.1	0.1	1.0	1.0
Fit models (see text)	-4.5	0.5	-1.0	0.2	-6.8	2.8
Inclusion of ^{85}Kr constraint	-2.2	2.2	0	0.4	-3.2	0
Exposure related	-1.1	0.6	-1.1	0.6	-1.1	0.6
Total sys. uncertainties (%)	-7.1	4.7	-1.5	0.8	-9.0	5.6

Table 6.7: Relevant sources of systematic uncertainty and their contribution to the measured neutrino interaction rates.

Table 6.8: Borexino Phase-II results on *pp*, ^{7}Be (862 +384 keV), *pep* and CNO solar ν's: interaction rates inferred assuming the MSW-LMA oscillation parameters. Unit: cpd/100t. The first error is the statistical derived by profiling the likelihood under Wilk's approximation. The interval extracted is coherent with the expectation from the toy-MC. The second error is the systematic uncertainty. Different contributions to the systematic error are detailed in Table 6.7. The result on *pep* ν's depends on whether we assume HZ or LZ metallicity for CNO ν's. The remaining columns show the theoretical interaction rates predicted by the Standard Solar Model under the high (HZ) and low (LZ) metallicity assumptions N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: `10.3847/1538-4357/835/2/202`, arXiv: `1611.09867`.

Solar ν	Borexino results	B16(GS98)-HZ	B16(AGSS09)-LZ
pp	$134 \pm 10 \, ^{+6}_{-10}$	131.0 ± 2.4	132.1 ± 2.3
^{7}Be	$48.3 \pm 1.1 \, ^{+0.4}_{-0.7}$	47.8 ± 2.9	43.7 ± 2.6
pep (HZ)	$2.43 \pm 0.36 \, ^{+0.15}_{-0.22}$	2.74 ± 0.05	2.78 ± 0.05
pep (LZ)	$2.65 \pm 0.36 \, ^{+0.15}_{-0.24}$	2.74 ± 0.05	2.78 ± 0.05
CNO	< 8.1 (95% C.L.)	4.91 ± 0.56	3.52 ± 0.37

6.6 *Neutrino fluxes*

With the neutrino elastic scattering interaction rates measured, the neutrino fluxes can be calculated assuming the MSW-LMA neutrino oscillation model of the survival probabilities. More details about calculation can be found in Section 10.1 and Section 10.2. The obtained neutrino fluxes are listed in Table 6.9.

Solar ν	Borexino results	B16(GS98)-HZ	B16(AGSS09)-LZ
pp	$(6.1 \pm 0.5\ ^{+0.3}_{-0.5}) \times 10^{10}$	$5.98\,(1 \pm 0.006) \times 10^{10}$	$6.03\,(1 \pm 0.005) \times 10^{10}$
^{7}Be	$(4.99 \pm 0.11\ ^{+0.06}_{-0.08}) \times 10^{9}$	$4.93\,(1 \pm 0.06) \times 10^{9}$	$4.50\,(1 \pm 0.06) \times 10^{9}$
pep (HZ)	$(1.27 \pm 0.19\ ^{+0.08}_{-0.12}) \times 10^{8}$	$1.44\,(1 \pm 0.009) \times 10^{8}$	$1.46\,(1 \pm 0.009) \times 10^{8}$
pep (LZ)	$(1.39 \pm 0.19\ ^{+0.08}_{-0.13}) \times 10^{8}$	$1.44\,(1 \pm 0.009) \times 10^{8}$	$1.46\,(1 \pm 0.009) \times 10^{8}$
CNO	$< 7.9 \times 10^{8}$ (95% C.L.)	$4.88\,(1 \pm 0.11) \times 10^{8}$	$3.51\,(1 \pm 0.10) \times 10^{8}$
^{8}B	$(5.68^{+0.39}_{-0.41}\ ^{+0.03}_{-0.03}) \times 10^{6}$	$5.46\,(1 \pm 0.12) \times 10^{6}$	$4.50\,(1 \pm 0.12) \times 10^{6}$

Table 6.9: Borexino Phase-II results on pp, ^{7}Be (862 +384 keV), pep and CNO solar ν's fluxes inferred assuming the MSW-LMA oscillation parameters. Unit: [cm^{-2}s^{-1}]. The result on ^{8}B from another analysis is also included. Reference: The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure", vol. 016, pp. 1–13, Sept. 2017, arXiv: `1709.00756`

Here the result of the ^{8}B solar neutrino flux measurement[18] is included in Table 6.9. The ratios of the measured ^{7}Be and ^{8}B neutrino flux by Borexino to that predicted by B16-HZ and B16-LZ are shown in Figure 6.15. The ratios corresponding to the global fit results are also shown in the figure. From the figure we can see that Borexino alone prefers the B16-HZ model.

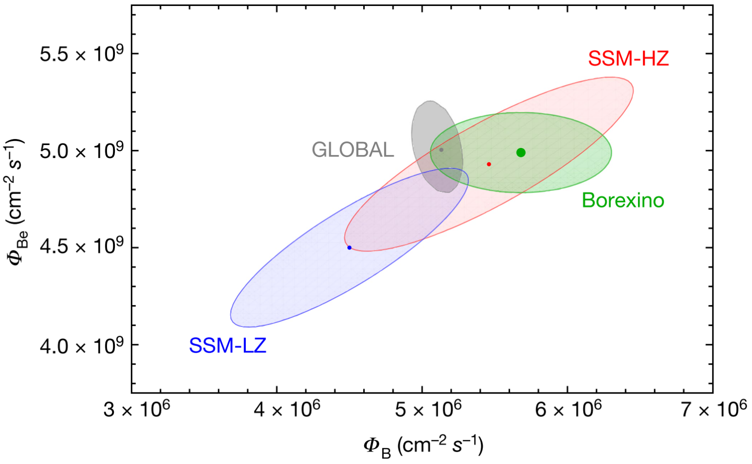


Figure 6.15: Borexino results for ^{7}Be and ^{8}B neutrino fluxes (green point and shaded area). The combined results of this work and all solar and KamLAND data is shown as gray dot and shaded area and marked as GLOBAL. In both results, the oscillation parameters θ_{12} and Δm_{12^2} are left free. The theoretical prediction for the low-metallicity (LZ) (blue) and the high-metallicity (HZ) (red) Standard Solar Models (SSM) are also shown. All contours correspond to 68.27% C.L.

In order to quantify the preference towards one of two SSMs, a hypothesis test is performed. The test statistic is defined as

$$t_{\mathrm{SSM}} = -2\log\left(\frac{\mathcal{L}(\mathrm{HZ})}{\mathcal{L}(\mathrm{LZ})}\right) \tag{6.6}$$

where $\mathcal{L}$(HZ) ($\mathcal{L}$(LZ)) is the optimized likelihood with neutrino rates constrained according to the B16-HZ (B16-LZ) prediction. The distribution of t assuming B16-HZ and B16-LZ, and its value on data are shown in Figure 6.16. Its value corresponding to Borexino results is $t^{\mathrm{BX}}_{\mathrm{SSM}} = -3.49$, and the p-value is 0.034, corresponding to a rejection of B16-LZ model with a C.L. of 96.6%. A Bayesian analysis is also performed, and the Bayes factor is 4.9, also showing a mild preference for HZ[19].

[18] The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure", vol. 016, pp. 1–13, Sept. 2017, arXiv: `1709.00756`

[19] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of pp, ^{7}Be, and pep Solar Neutrinos with Borexino Phase-II", pp. 1–8, July 2017, arXiv: `1707.09279`

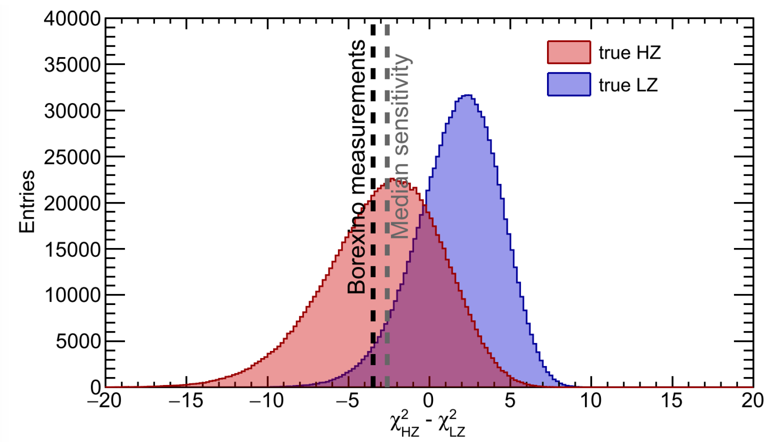


Figure 6.16: The probability distribution of the test statistics t_{SSM} obtained by Monte Carlo method. The dotted black line corresponds to the Borexino results.

6.7 *Neutrino survival probability*

With the neutrino elastic scattering interaction rates measured, the neutrino survival probability can be calculated assuming the SSM model of the neutrino flux[20]. The obtained neutrino survival probability are listed in Table 6.10 and shown in Figure 6.17.

Table 6.10: Borexino Phase-II results on *pp*, ^{7}Be (862 +384 keV) and *pep* solar neutrino survival probability assuming B16-HZ model of the neutrino fluxes. The results on ^{8}B from another analysis are also included. Here the errors include uncertainties of the SSM prediction of neutrino fluxes. Reference: The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure", vol. 016, pp. 1–13, Sept. 2017, arXiv: `1709.00756`

Solar ν	average E_ν (MeV)	P_{ee}
pp	0.267	0.57 ± 0.09
^{7}Be	0.862	0.53 ± 0.05
pep	1.44	0.41 ± 0.11
^{8}B HER-I	7.9	0.39 ± 0.09
^{8}B HER-II	9.9	0.35 ± 0.09
^{8}B HER	8.7	0.37 ± 0.08

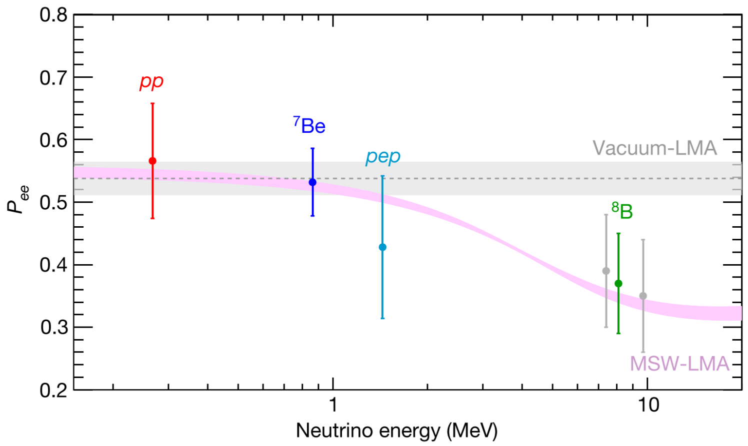


Figure 6.17: Electron neutrino survival probability Pee as a function of neutrino energy. The pink band is the $\pm 1\sigma$ prediction of MSW-LMA with oscillation parameters. The grey band is the vacuum-LMA case. Data points represent the Borexino results for pp (red), ^{7}Be (blue), pep (cyan) and ^{8}B (green for the HER range, and grey for the separate HER-I and HER-II sub-ranges), assuming HZ-SSM. ^{8}B and pp data points are set at the mean energy of neutrinos that produce scattered electrons above the detection threshold. The error bars include experimental and theoretical uncertainties.

Similarly, a hypothesis test is performed on rejecting the vacuum oscillation model. The test statistic is defined as

$$t_{\mathrm{Pee}} = -2\log\left(\frac{\mathcal{L}(\mathrm{MSW})}{\mathcal{L}(\mathrm{vacuum})}\right) \tag{6.7}$$

where $\mathcal{L}(\mathrm{MSW})$ ($\mathcal{L}(\mathrm{vacuum})$) is the optimized likelihood with neutrino rates constrained according to the B16-HZ (B16-LZ) prediction

[20] D. Guffanti, "Measurement of solar neutrino with BOREXINO", PhD thesis, Gran Sasso Science Institute, 2019

and the survival probability given by the MSW (vacuum oscillation) model. The p-value obtained is 0.018, corresponding to a rejection of the vacuum oscillation model with a C.L. of 98.2%[21].

6.8 *Conclusions*

In this section we summarize the conclusions of the solar neutrino measurement:

- The interaction rate of the 862 keV + 384 keV ^{7}Be solar neutrinos in Borexino is found to be $48.3 \pm 1.1(\text{stat.})^{+0.4}_{-0.7}(\text{sys.})$ cpd/100t. In the assumption of the MSW-LMA model of solar neutrino oscillations this corresponds to a flux of $(4.99 \pm 0.11^{+0.06}_{-0.08}) \times 10^9$ cm^{-2}s^{-1} . The observed survival probability is $P_{ee} = 0.53 \pm 0.05$ for the High Metallicity SSM prediction of the flux.

- The interaction rate of the *pp* solar neutrinos in Borexino is found to be $134 \pm 10(\text{stat.})^{+6}_{-10}(\text{sys.})$ cpd/100t. In the assumption of the MSW-LMA model of solar neutrino oscillations this corresponds to a flux of $(6.1 \pm 0.5^{+0.3}_{-0.5}) \times 10^10$ cm^{-2}s^{-1} . The observed survival probability is $P_{ee} = 0.57 \pm 0.09$ for the High Metallicity SSM prediction of the flux.

- With the CNO solar neutrino flux constrained to the High (Low) Metallicity model prediction, the interaction rate of the *pep* solar neutrinos in Borexino is found to be $2.43 \pm 0.36(\text{stat.})^{+0.15}_{-0.22}(\text{sys.})$ cpd/100t ($2.65 \pm 0.36(\text{stat.})^{+0.15}_{-0.24}(\text{sys.})$ cpd/100t). With this constraint, in the assumption of the MSW-LMA model of solar neutrino oscillations, this corresponds to a flux of $(1.27 \pm 0.19^{+0.08}_{-0.12}) \times 10^9$ cm^{-2}s^{-1} ($(1.39 \pm 0.19^{+0.08}_{-0.13}) \times 10^9$ cm^{-2}s^{-1}). The observed survival probability is $P_{ee} = 0.41 \pm 0.11$ for the High Metallicity SSM prediction of the flux.

- With the ratio of the *pp* and *pep* solar neutrino interaction rate constrained according to the SSM prediction, the interaction rate of the CNO solar neutrinos is not detected and an upper limit of 8.1 cpd/100t (95% C.L.) is obtained. In the assumption of the MSW-LMA model of solar neutrino oscillations this corresponds to an upper limit of the flux of 7.9×10^8 cm^{-2}s^{-1}.

- Combined with the results of ^{8}B neutrino interaction rates and fluxes, a hypothesis test using only Borexino results is performed on discriminating the B16-HZ and B16-LZ models. The B16-LZ model is rejected with a C.L. of 96.6%. The Bayes factor of B16-HZ against B16-LZ model is 4.9.

- Combined with the results of ^{8}B neutrino interaction rates and fluxes, a hypothesis test using only Borexino results is performed on discriminating the MSW and vacuum oscillation models. With neutrino fluxes constrained according to B16-HZ prediction, the vacuum oscillation model is rejected with a C.L. of 98.2%.

[21] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

7

Sensitivity of Borexino to CNO solar neutrino

In this chapter, the evaluation of the sensitivity to CNO solar neutrinos with Borexino Phase-II data is presented. In the Sun, approximately 1% of total energy is produced by the CNO-cycle process, in which hydrogen is converted to helium with C, N and O as catalysts. It is predicted to be the major source of energy in more massive ($>$1.2 $M_{\odot}$) or more evolved stars, and neutrinos produced in this process has never been observed. Neutrinos produced during this process in the Sun, namely CNO solar neutrino, can be observed through their elastic scattering on electrons with large liquid scintillator detectors, such as Borexino. The main challenge of the CNO solar neutrino detection in Borexino is to constrain its major background, the ^{210}Bi decays. The ongoing effort to measure the ^{210}Bi decay rate is discussed dedicatedly in the next chapter. In this chapter, using the standard hypothesis test method, the sensitivity to the existence of CNO solar neutrinos is estimated as a function of the precision of the hypothetical constraint on the ^{210}Bi decay rate. The systematic uncertainty from the light yield and the detector resolution model are included by using the analytical response functions. Furthermore, an alternative method using the covariance matrix is presented. The usage of the covariance matrix has the advantage of making evident the correlation among CNO solar neutrinos, ^{210}Bi decays and *pep* solar neutrinos intuitively.

This chapter is organized as in the following: Section 7.1 introduces the motivations and reviews the sensitivity of current and future experiments to CNO solar neutrinos. Section 7.2 presents the correlation analysis using the covariance matrix. Section 7.3 presents the method and results of the standard hypothesis test. Section 7.4 is devoted to the conclusions and discussions.

I have co-worked with Matteo Agostini, Daniele Guffanti and Nicola Rossi in this work.

This work is in synergy with the work of Daniele Guffanti[1]. In this work, the analytical fit approach is used, so the non-linearity and resolution systematic uncertainties are included. In Daniele's work, the Monte Carlo approach is used and different strengths of ^{210}Bi and *pep* solar neutrino constraints are considered. Besides, when performing the hypothesis test, I have used the `charge` estimator, while in the study of Daniele the `nhit` estimator is used.

[1] D. Guffanti, "Measurement of solar neutrino with BOREXINO", PhD thesis, Gran Sasso Science Institute, 2019

7.1 Introduction

An experimental evidence of the existence of the `CNO`-cycle would be a groundbreaking achievement in astrophysics. The `CNO`-cycle process is predicted to be the major source of energy in stars more massive ($>1.2\ M_\odot$) or older than the Sun. It is one of two processes in which hydrogen is converted to helium, and it uses ^{12}C as catalysts[2]. The detection of CNO solar neutrinos is currently the only accessible way to directly prove the existence of this process. So far, Borexino[3] has set the most stringent upper limits[4,5]. Several future experiments have sensitivities to it, and LArTPC experiments can achieve 10% precision measurement in the next decades[6]. By then, the solar metallicity problem might be resolved[7].

In Borexino[8], CNO solar neutrinos are detected through their elastic scattering on electrons in the liquid scintillator. The interaction rates are extracted by spectral fitting. A novel three-fold-coincidence method[9] is used to suppress cosmogenic ^{11}C background by about 90%. In Phase-I analysis[10], the *pep* solar neutrino interaction rate was constrained according to the standard solar model and MSW-LMA prediction. In Phase-II analysis[11,12], only the ratio of *pp* and *pep* interaction rate was constrained. In both results, the ^{210}Bi decay rate was left free, and only upper limits of the CNO solar neutrino interaction rate were reported.

Borexino is pursuing a method to improve the CNO solar neutrino sensitivity. Due to the similarity in the spectral shapes, only the sum of the CNO solar neutrino interaction and ^{210}Bi decay rates are measured, and there is little information about the fraction of the CNO solar neutrino interaction rate in the sum. An independent measurement of ^{210}Bi decay rate would significantly improve the sensitivity to the CNO solar neutrinos[13].The feasibility of a measurement of the ^{210}Bi background is discussed in the following chapter.

In this chapter, the correlation between the CNO solar neutrino interaction and ^{210}Bi decay rates is illustrated using the covariance matrix of fit results. After that, the median sensitivity to the CNO solar neutrino interaction rate is reported.

[2] H. A. Bethe, "Energy Production in Stars", *Physical Review*, vol. 55, no. 5, pp. 434–456, Mar. 1939, DOI: `10.1103/PhysRev.55.434`

[3] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 89, no. 11, pp. 1–68, 2014, DOI: `10.1103/PhysRevD.89.112007`, arXiv: `1308.0443`

[4] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II", pp. 1–8, July 2017, arXiv: `1707.09279`

[5] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

[6] D. G. Cerdeño et al., "CNO neutrino Grand Prix: The race to solve the solar metallicity problem", *Journal of Cosmology and Astroparticle Physics*, vol. 2018, no. 4 2018, DOI: `10.1088/1475-7516/2018/04/037`

[7] J. Bergström et al., "Updated determination of the solar neutrino fluxes from solar neutrino data", *Journal of High Energy Physics*, vol. 2016, no. 3, p. 132, Mar. 2016, DOI: `10.1007/JHEP03(2016)132`, arXiv: `1601.00972`

[8] See Chapter 6

[9] See Section 2.5.4 page 41

[10] The Borexino collaboration, "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy"

[11] The Borexino collaboration, "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II"

[12] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos"

[13] F. L. Villante et al., "A step toward CNO solar neutrino detection in liquid scintillators", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 701, no. 3, pp. 336–341, 2011, DOI: `10.1016/j.physletb.2011.05.068`, arXiv: `1104.1335`

7.2 *Correlation analysis using Covariance matrix*

The intuitive argument that Borexino only measures the sum of the CNO solar neutrino interaction rate and the ^{210}Bi decay rate can be quantitatively derived using the covariance matrix of the fit results. With this method, it is found that the spectrum holds information on the weighted sum of the ^{210}Bi decay rate, the *pep* solar neutrino interaction rate and the CNO solar neutrino interaction rate. Besides, the precision of the difference between the CNO solar neutrino interaction rate and the ^{210}Bi decay rate, or the strength of the shape information, is also found with this method.

The formula of the weighted sum is consistent with the results obtained by simply counting the number of events in an energy range and that obtained by analyzing the distribution of best fits of many pseudo-experiment spectra. These two results are reported in the PhD thesis of Daniele Guffanti[14].

7.2.1 *Spectral analysis with covariance matrix*

The result of the spectral fit is a set of model parameters that minimize the discrepancy between the data and the model. The optimized values depend on the statistical realization of the data and thus are random variables, so their covariance matrix can be defined:

$$V_{ij} = E[(\theta_i - \mu_i)(\theta_j - \mu_j)]\,, \tag{7.1}$$

where V_{ij} is the element of the covariance matrix, E denotes the expectation and μ is the expectation of the parameter θ.

When the fit parameters follow Gaussian distributions, the inverse of their covariance matrix can be estimated as the Hessian matrix of the logarithm of the likelihood function[15]:

$$\left(\widehat{V}^{-1}\right)_{ij} = -\frac{\partial^2 \ln \mathcal{L}}{\partial\theta_i \cdot \partial\theta_j}\bigg|_{\widehat{\theta}}\,, \tag{7.2}$$

where V is the covariance matrix, $\mathcal{L}$ is the likelihood and θ are the fit parameters. A set of quantities that can be determined independently can be found by diagonalizing the covariance matrix

$$P^{-1} \cdot V \cdot P = \mathrm{diag}\left\{\sigma_1, \sigma_2, ...\sigma_n\right\}\,, \tag{7.3}$$

where the set of quantities are measured to be

$$\widehat{\tau} \sim \left(P^{-1}\widehat{\theta}\right) \pm \widehat{\sigma}\,, \tag{7.4}$$

then the statistical uncertainties of the interested parameters can be written as

$$\sigma(\theta_i) = \sqrt{\sum_j \left[P_{ij} \cdot \sigma(\tau_j)\right]^2}\,. \tag{7.5}$$

When diagonalizing the covariance matrix, we can use its submatrix, then the expression (7.5) only involve a few quantities.

This analysis has been implemented in the `bx-GooStats` software[16].

[14] Guffanti, "Measurement of solar neutrino with BOREXINO"

[15] G. Cowan, *Statistics* Particle Data Group, "Review of Particle Physics", *Physical Review D*, vol. 98, no. 3, p. 030001, Aug. 2018, DOI: `10.1103/PhysRevD.98.030001` Section 39.2.2, Equation (39.12), page 528

[16] X. Ding, "GooStats: A GPU-based framework for multi-variate analysis in particle physics", *Journal of Instrumentation*, vol. 13, no. 12, P12018–P12018, Dec. 2018, DOI: `10.1088/1748-0221/13/12/P12018`

7.2.2 *Rate and shape information*

Results on the rate and shape information using the method described in the previous section are presented in this section. The definition of `charge`, `npmt_dt1` and `nhit` are described in Section 3.1.2 on page 47. The Asimov dataset is used. The detector response parameters configurations are described in Table 4.5.

The correlation matrix, defined as

$$\rho_{ij} = \frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}}\,, \tag{7.6}$$

is visualized in Figure 7.1.

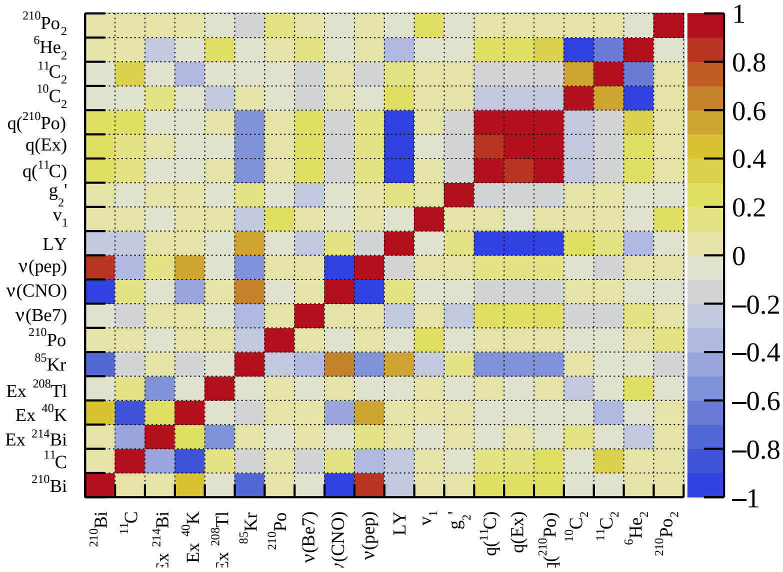


Figure 7.1: The correlation coefficient of fit parameters. Echidna charge and analytical probability density functions are used. Light yield and resolution parameters are left free. Subscript 2 denotes the rate of components in the TFC-selected dataset. The Asimov dataset (explained on Section 7.3) is used.

From the figure we can see that only the ^{210}Bi decay and *pep* solar neutrino interaction rates are strongly anti-correlated with the CNO solar neutrino interaction rate. The rate of external ^{40}K decay backgrounds is weakly anti-correlated with the CNO solar neutrino interaction rate. The ^{85}Kr decay rate has positive correlation with it, and this positive correlation is brought by the anti-correlation between the ^{85}Kr and ^{210}Bi decay rates.

Choosing the submatrix consisting of the CNO solar neutrino interaction rate, the *pep* solar neutrino interaction rate and the ^{210}Bi decay rate, by diagonalizing the covariance matrix, we obtained a set of independent quantities:

$$q_{\text{sum}} = R_{\nu(\text{CNO})} + 0.57R_{^{210}\text{Bi}} + 2.55R_{\nu(\text{pep})} = 21.18 \pm 0.79 \tag{7.7}$$

$$q_{\text{pep}} = R_{\nu(\text{CNO})} + 1.13R_{^{210}\text{Bi}} - 0.66R_{\nu(\text{pep})} = 21.96 \pm 2.55 \tag{7.8}$$

$$q_{\text{Bi}} = R_{\nu(\text{CNO})} - 0.98R_{^{210}\text{Bi}} - 0.16R_{\nu(\text{pep})} = -11.83 \pm 34.74\,, \tag{7.9}$$

where the unit on the right side of the equation is count per day per 100 tons. Here the CNO solar neutrino interaction rate is allowed to be negative so that its distribution follows approximately the Gaussian distribution.

q_{sum} can be considered as the number of events in a small energy range where its three components dominate. q_{pep} can be considered as the shape difference between the CNO solar neutrino interactions/^{210}Bi decays and the *pep* solar neutrino interactions. q_{Bi} can be considered as the shape difference between the CNO solar neutrino interactions and the ^{210}Bi decays. Such an interpretation is corroborated by how the precision of three quantities change with respect to the energy resolution of detectors. This is discussed in the Section 7.4.

Comparison of coefficients and precisions of q_x in Equation (7.7)–(7.9) between `npmt_dt1` and `charge` are shown in Table 7.1 and Table 7.2, respectively. From the table we can see that `charge` and `npmt_dt1` have similar precisions and coefficients.

Table 7.1: Coefficients of expressions of q_x. Define $q_x = R_{\nu(\text{CNO})} + \epsilon_{x2} R_{^{210}\text{Bi}} + \epsilon_{x3} R_{\nu(\text{pep})}$. In each cell the numbers are $(\epsilon_{x2}, \epsilon_{x3})$. Unit: per day per 100 tons.

	`charge`	`npmt_dt1`
q_{sum}	(0.61, 2.55)	(0.57, 2.81)
q_{pep}	(1.13, −0.66)	(1.16, −0.59)
q_{Bi}	(−0.98, −0.16)	(−0.95, −0.16)

Table 7.2: Values and precisions of q_{sum}, q_{pep} and q_{Bi}. Unit: per day per 100 tons.

	`charge`	`npmt_dt1`
q_{sum}	22.18 ± 0.79	22.26 ± 0.82
q_{pep}	21.96 ± 2.55	22.56 ± 2.51
q_{Bi}	-11.83 ± 34.74	-11.38 ± 39.15

Because q_{sum}, q_{pep} and q_{Bi} are independent, the precision of the CNO solar neutrino interaction rate can be simply written as[17]

$$R_{\nu(\text{CNO})} = 0.13 q_{\text{sum}} + 0.37 q_{\text{pep}} + 0.51 q_{\text{Bi}} \tag{7.10}$$

$$\sigma_{\nu(\text{CNO})} = 0.13 \sigma_{\text{sum}} \oplus 0.37 \sigma_{\text{pep}} \oplus 0.51 \sigma_{\text{Bi}} \tag{7.11}$$

$$\approx 0.10 \oplus 0.94 \oplus 17.56 \text{ cpd/100t} \tag{7.12}$$

where $a \oplus b = \sqrt{a^2 + b^2}$. Obviously the uncertainty on q_{Bi} dominated.

Impact of ν(pep) constraint The *pep* solar neutrino interaction rate can be constrained, if we constrain the ratio between $R_{\nu(\text{pp})}$ and $R_{\nu(\text{pep})}$. If the solar luminosity is considered, the *pp* solar neutrino interaction rate can be constrained within 2%, and so does the *pep* solar neutrino interaction rate. After that, only the ^{210}Bi decay rate is correlated with the CNO solar neutrino interaction rate:

$$q_{\text{sum}} = R_{\nu(\text{CNO})} + 0.63 R_{^{210}\text{Bi}} = 15.39 \pm 0.81 \tag{7.13}$$

$$q_{\text{Bi}} = R_{\nu(\text{CNO})} - 1.59 R_{^{210}\text{Bi}} = -21.66 \pm 12.17 \tag{7.14}$$

where the unit is cpd/100t, and the CNO solar neutrino interaction rate can be expressed as

$$R_{\nu(\text{CNO})} = 0.72 q_{\text{sum}} + 0.28 q_{\text{Bi}} \tag{7.15}$$

$$\sigma_{\nu(\text{CNO})} = 0.72 \sigma_{\text{sum}} \oplus 0.28 \sigma_{\text{Bi}} \tag{7.16}$$

$$\approx 0.6 \oplus 5.5 \text{ cpd/100t} \tag{7.17}$$

[17] `charge` is used here.

Impact of ^{210}Bi constraint A constraint on the ^{210}Bi decay rate can replace the information of q_{Bi}. Consider a 10% constraint of the ^{210}Bi decay rate, then the expression of $\sigma_{\nu(\text{CNO})}$ becomes

$$R_{\nu(\text{CNO})} = q_{\text{sum}} - 0.63 q_{\text{Bi}}^{\text{ext}} \tag{7.18}$$

$$\sigma_{\nu(\text{CNO})} = \sigma_{\text{sum}} \oplus 0.63 \sigma_{\text{Bi}}^{\text{ext}} \tag{7.19}$$

$$\approx 0.8 \oplus 1.1 \text{ cpd}/100\text{t}, \tag{7.20}$$

where $\sigma_{\text{Bi}}^{\text{ext}}$ is the precision of the ^{210}Bi constraint and it is 1.75 cpd/100t. The precision of CNO solar neutrino interaction rate is improved to 1.3 cpd/100t.

The critical precision of the ^{210}Bi decay rate constraint is when the contribution of σ_{sum} and $0.63\sigma_{\text{Bi}}^{\text{ext}}$ are equal:

$$\sigma_{\text{sum}} \sim 0.63 \sigma_{\text{Bi}}^{\text{ext}} \tag{7.21}$$

or

$$\sigma_{\text{Bi}}^{\text{ext}} \sim 1.3 \text{ cpd}/100\text{t} \tag{7.22}$$

When the ^{210}Bi constraint is much stronger than this value, the precision of the CNO solar neutrino interaction rate is dominated by the precision of σ_{sum} and would not be improved much with a stronger ^{210}Bi constraint. Vice versa, when the ^{210}Bi constraint is much weaker than this value, the precision of CNO solar neutrino interaction rate is dominated by the precision of the ^{210}Bi constraint, and accumulating more data would not improve the precision of the CNO solar neutrino interaction rate much.

7.3 Statistical methods to define sensitivity to signal

In this section, the results of the sensitivity to the CNO solar neutrinos using standard hypothesis test methods[18] are presented. The ^{210}Bi decay background is in this section assumed to be known with a 10% precision and the *pep* solar neutrino interaction rate with a 2% precision.

The experimental sensitivity is characterized with the expected significance, which is defined in Equation 7.26. The essential concepts are introduced here.

Definition of hypothesis Define the null hypothesis H_0 as $R_{\text{CNO}} = 0$. Define the alternative hypothesis H_1 as $R_{\text{CNO}} = R_{\text{CNO}}^{\text{HM}}$, where $R_{\text{CNO}}^{\text{HM}}$ is the predicted CNO solar neutrino interaction rate according to the high metallicity model (B16-GS98)[19]. The signal strength is defined as $\mu = R_{\text{CNO}}/R_{\text{CNO}}^{\text{HM}}$. For the low metallicity model (B16-AGSS09met)[20], $\mu = R_{\text{CNO}}^{\text{LM}}/R_{\text{CNO}}^{\text{HM}} = 0.717$ and the corresponding hypothesis is denoted as $H_{0.717}$.

Profile Likelihood Ratio The profile likelihood ratio is defined as

$$\lambda(\mu) = \frac{\mathcal{L}\left(\mu, \hat{\hat{\theta}}(\mu)\right)}{\mathcal{L}\left(\hat{\mu}, \hat{\theta}\right)} \tag{7.23}$$

[18] G. Cowan, *Statistics* Particle Data Group, "Review of Particle Physics" Section 39.3, page 532 and Section 39.5, page 540

[19] N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: `10.3847/1538-4357/835/2/202`, arXiv: `1611.09867`

[20] ibid.

where the nominator is the profile likelihood function, $\hat{\hat{\theta}}(\mu)$ is the value of θ that maximizes the likelihood when μ is fixed to a certain value, and the denominator is the maximized likelihood with μ free.

Test statistic In order to establish the discovery of CNO solar neutrinos, the test statistic is chosen as

$$t_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{7.24}$$

p-value For the CNO solar neutrino discovery, the goodness-of-fit, or the p-value, is defined as

$$p_0 = \int_{t_{0,\,\mathrm{obs}}}^{\infty} f(t_0|H_0)\mathrm{d}t\,, \tag{7.25}$$

where f is the probability density function of t_0 assuming H_0, $t_{0,\,\mathrm{obs}}$ is the value of the test statistic obtained in the actual experiment.

Significance The significance is defined as

$$Z_\mu = \Phi^{-1}(1 - p_\mu)\,, \tag{7.26}$$

where Φ is the standard Gaussian cumulative distribution and its inverse Φ^{-1} is the standard Gaussian quantile; μ is the signal strength ($\mu = 1$ for the HM model) and p_μ is the corresponding p-value.

To claim a 3 σ evidence of CNO solar neutrino, Z_0 need to be larger than 3.

Experimental sensitivity The experimental sensitivity is characterized with the expected significance, which is the median of the distribution of Z_μ under a hypothesis $H_{\mu'}$: $\mathrm{med} f(Z_\mu|H_{\mu'})$.

The sensitivity to the discovery of CNO solar neutrinos assuming the high metallicity model is $\mathrm{med} f(Z_0|H_1)$.

$f(t)$ from Monte Carlo method The distribution $f(t_0|H_0)$, $f(t_0|H_1)$ and $f(t_0|H_{0.717})$ can be obtained using the Monte Carlo method. In each case, many pseudo-experiment spectra assuming the corresponding hypothesis are generated and fitted, the corresponding t_0 is calculated, and then the distribution of t_0 is obtained. They are shown in Figure 7.2

The Asimov dataset The Asimov dataset is the spectrum where the number of events in each bin equals the expected number of events, and thus they are not necessarily integers. Assuming the hypothesis is $H_{\mu'}$, the likelihood, the profile likelihood ratio and the test statistic calculated using the Asimov dataset are denoted as $\mathcal{L}^{A;\mu'}$, $\lambda^{A;\mu'}(\mu)$ and $t_\mu^{A;\mu'}$, respectively.

$f(t)$ using asymptotic formulae The asymptotic formulae of the distribution $f(t_0|H'_\mu)$ is the following[21]:

$$f(t_0|H'_\mu) = \left[1 - \Phi\left(\frac{\mu'}{\sigma}\right)\right]\delta(t_0) + \frac{1}{2\sqrt{2\pi}\sqrt{t_0}}\exp\left[-\frac{1}{2}\left(\sqrt{t_0} - \frac{\mu'}{\sigma}\right)^2\right], \tag{7.27}$$

where $\mu' = 1$ or 0.717.

Using the Asimov dataset, it is found that $t_0^{A;1} = 15.02$ and $t_0^{A;0.717} = 7.87$, so according to Equation (32) of Cowan et al.,[22] $\sigma_0^{A;1} = 0.258$ and $\sigma_0^{A;0.717} = 0.256$. The predicted distribution $f(t_0|H_0)$, $f(t_0|H_1)$ and $f(t_0|H_{0.717})$ are shown in Figure 7.2. It can be seen that the asymptotic formulae can describe the distributions well.

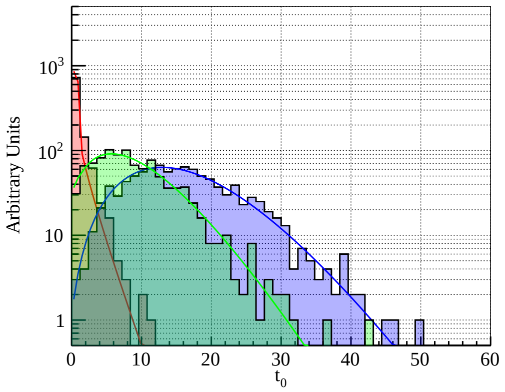


Figure 7.2: Distributions of the test statistic defined in Equation (7.24) for CNO solar neutrino discovery assuming High Metallicity (HM, blue), Low Metallicity (LM, green) and no CNO (red). The distribution predicted by the asymptotic formula is overlaid as solid lines of corresponding colors.

Result The median value of a test statistic is just its Asimov value, so we have a simple formula for the experimental sensitivity:

$$\text{med}f(Z_0|H_{\mu'}) = \sqrt{t_0^{A;\mu'}}, \tag{7.28}$$

where $t_0^{A;\mu'}$ is the test statistic value of the Asimov dataset assuming $H_{\mu'}$. The median sensitivities to CNO solar neutrinos assuming the High Metallicity and Low Metallicity models are 3.9 σ and 2.8 σ, respectively.

[21] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011, DOI: `10.1140/epjc/s10052-011-1554-0`, arXiv: `1007.1727`

[22] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011, DOI: `10.1140/epjc/s10052-011-1554-0`, arXiv: `1007.1727`.

7.4 *Conclusions and discussions*

From the correlation analysis we can see that there is little sensitivity in the Borexino spectrum to the difference between the CNO solar neutrino interactions and the ^{210}Bi decay backgrounds, so it is necessary to measure the ^{210}Bi decay rate independently and constrain it in the spectral fit. From the hypothesis test we can see that if the ^{210}Bi decay rate is constrained within 10% and the *pep* solar neutrino interaction rate is constrained within 2%, the median sensitivity assuming the High Metallicity model is 3.9 σ, so under this assumption there are more than 50% chances that we can claim CNO solar neutrino evidence[23]. It will be a groundbreaking result in astro-physics.

On Section 7.2.2 we give an intuitive interpretation of q_{sum} , q_{pep} and q_{Bi}, and this can be seen from their trends with respect to the increasing exposure and improving energy resolutions, see Figure 7.3. The comparison between spectra with the Borexino Phase-II exposure and 10^4 times more exposure is shown in Figure 7.4. The comparison of the spectra under different resolutions is shown in Figure 7.5.

From Figure 7.3 we can see that all three quantities scale with $\left(\sqrt{MT}\right)^{-1}$ given by the active mass M and data taking time T as expected. The precision of q_{sum} (q_1) improves slightly while that of q_{pep} (q_2) and q_{Bi} (q_3) improve significantly when the resolution is improved. As interpreted, q_{sum} represents the sum of rates of three components and it should not depend on the energy resolution. It is slightly improved because the energy range contributing to its precision, or the valley between the ^{7}Be edge and the ^{11}C peak, in the TFC selected (^{11}C enhanced) spectrum increased. q_{pep} is the energy distribution difference between the CNO solar neutrinos/^{210}Bi decays and the *pep* solar neutrinos, so its precision is significantly improved when the energy resolution is improved and the difference between two component becomes more clear. The same applies for q_{Bi}.

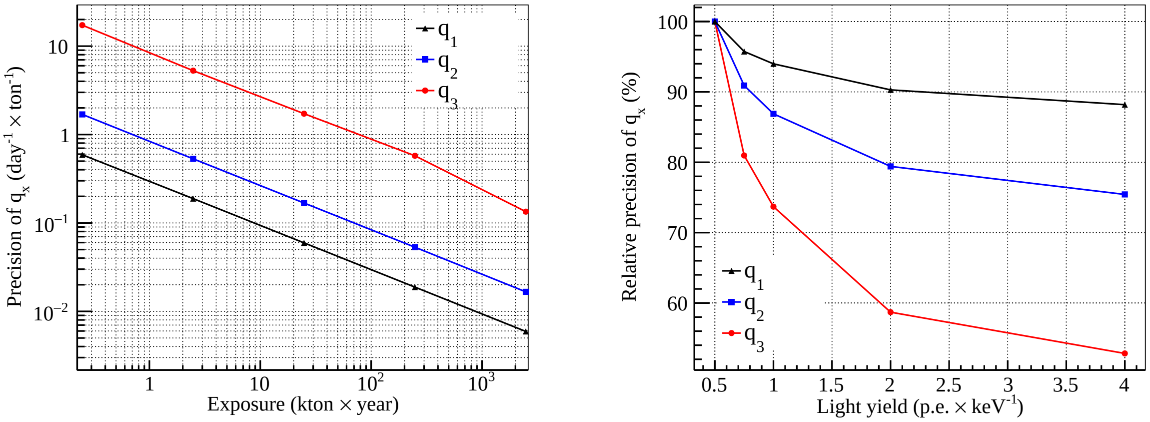


Figure 7.3: Left: precision of q_x versus exposure. Right: relative precision of q_x versus light yield. To get stable results, the light yield on the left is set to 6 p.e. per keV and the exposure on the right is set to 10 times of Borexino Phase-II dataset.

---

[23] 3 σ: evidence. 5 σ: discovery.

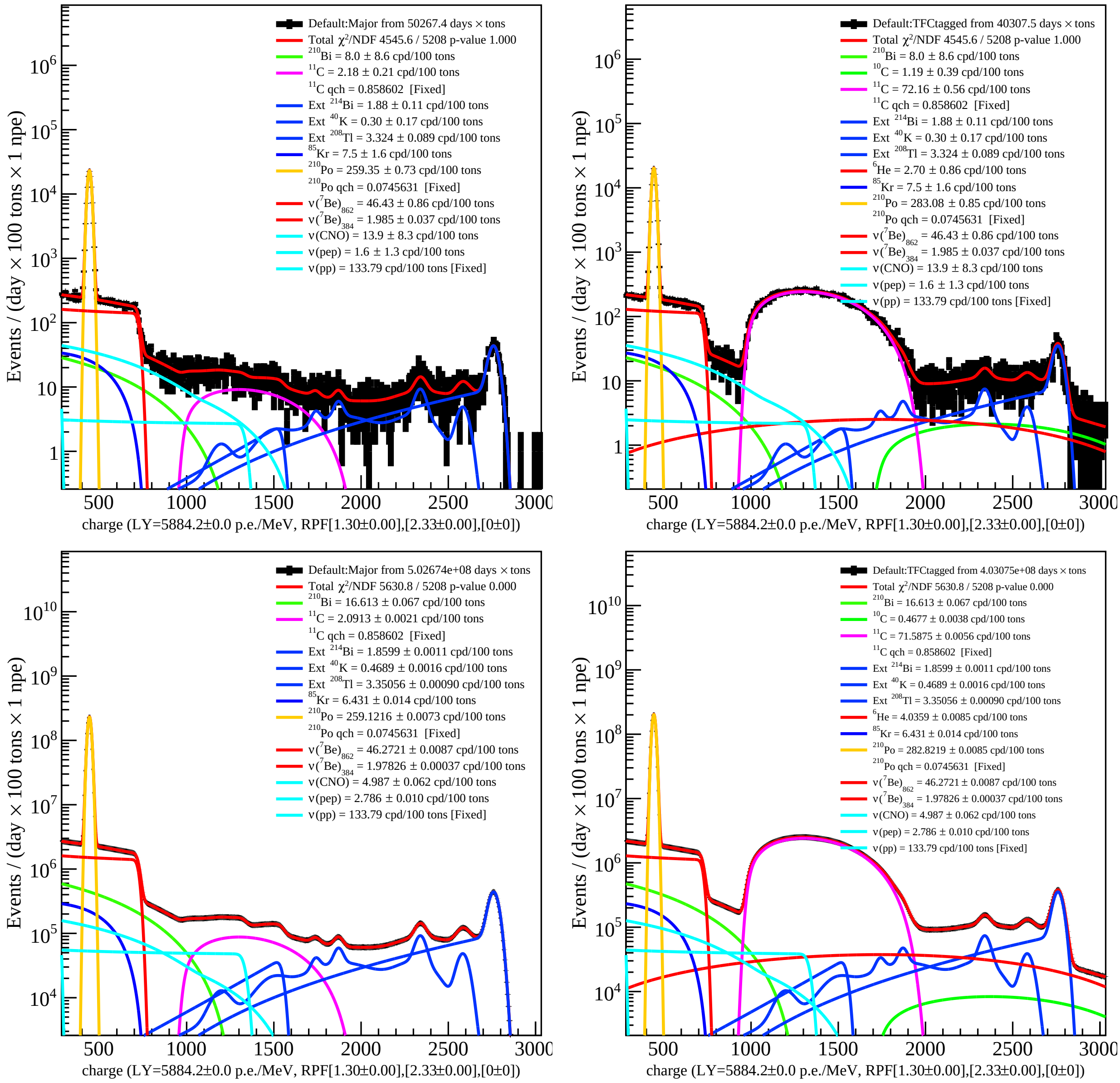


Figure 7.4: Comparison of spectra between low and high exposure. Top: Borexino Phase-II exposure. Bottom: 10^4 times top. To get stable results, the light yield is set to 6 p.e. per keV.

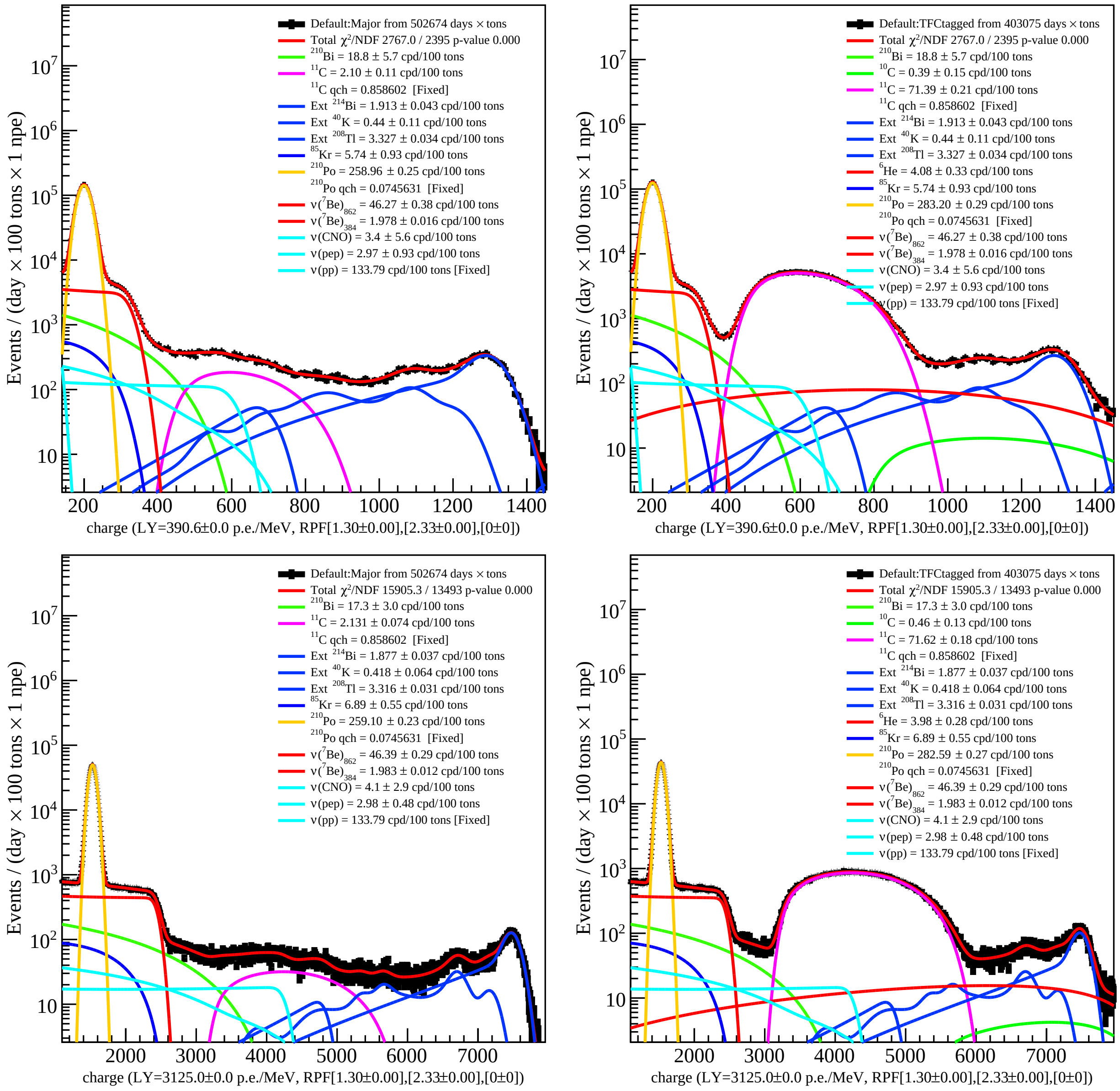


Figure 7.5: Comparison of spectra between poor and good energy resolutions. Top: Borexino Phase-II resolution. Bottom: 8 times light yield of top. To get stable results, the exposure is set to 10 times of Borexino Phase-II dataset.

8

^{210}Bi decay rate measurement and `A1000` dataset

In this chapter, the ^{210}Bi decay rate measurement and the application of the `A1000` dataset, a special dataset enabling only 1000 selected survived PMTs, are presented. The `A1000` dataset is produced to reduce systematic uncertainties from the change of detector response induced by the PMT loss. This dataset is applied to not only the ^{210}Bi decay rate measurement, but also to improving the Monte Carlo simulation. In the previous chapter, it is demonstrated that the ^{210}Bi measurement is strongly motivated in order to measure the CNO solar neutrinos, and this is further discussed in this chapter. To constrain the ^{210}Bi decay rate, we can measure the decay rate of its daughter ^{210}Po[1], and the major background is the ^{210}Po that comes from the inner vessel and is brought into the center region by the convective motion. In this chapter, we present the methods developed to get rid of the convection induced ^{210}Po and the methods used to test the assumptions made in order to measure the ^{210}Bi decay rate with the ^{210}Po decay rate.

This chapter is organized as in the following: Section 8.1 introduces motivation and principle of the ^{210}Bi decay rate measurement. Section 8.2 reviews the motivation of `A1000` dataset. Section 8.3 presents the production, validation and application of the `A1000` dataset. Section 8.4 describes methods used to extract the supported ^{210}Po decay rate. Section 8.5 presents the results of ^{210}Bi decay rate from spectral fits. Section 8.6 is devoted to conclusions and discussions.

I have co-worked with Nicola Rossi, Marcin Misiaszek, Frank Calaprice and Alessandra Re in the work of `A1000`. It is in synergy with the work of Mariia Redchuk. Mariia carried out the work of comparing different ways of implementing new effective quantum efficiencies using the `A1000` dataset. I have co-worked with Nicola Rossi in the work of ^{210}Bi decay rate measurement. It is in synergy with the work of Daniele Guffanti, Nicola Rossi and Davide Basilico.

[1] F. L. Villante et al., "A step toward CNO solar neutrino detection in liquid scintillators", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 701, no. 3, pp. 336–341, 2011, DOI: `10.1016/j.physletb.2011.05.068`, arXiv: `1104.1335`

8.1 *Introduction*

If we would like to claim strong results on CNO solar neutrino interaction rate, it is necessary to measure the ^{210}Bi decay rate independently. Because the energy distribution of ^{210}Bi decay events is similar to that of CNO solar neutrino interactions, the ^{210}Bi decay rate and the CNO solar neutrino interaction rate are highly correlated and the statistical uncertainty of the CNO solar neutrino interaction rate is large. It is so high that only an upper limit can be claimed. When ^{210}Bi decay rate is constrained from an independent measurement, the correlation is broken. From the last chapter we know that assuming the HZ model[2] of CNO neutrino flux, if the ^{210}Bi decay rate is constrained within 10% precision, the median sensitivity of CNO solar neutrino discovery is 3.9 σ.

The ^{210}Bi decay rate can be measured with the ^{210}Po decay rate[3], if a few assumptions are valid. This is possible, because they are in the same decay chain:

$$^{210}\text{Pb} \rightarrow {}^{210}\text{Bi} \rightarrow {}^{210}\text{Po}$$

Fundamental assumptions Because the time and volume used to measure the ^{210}Bi decay rate is different from that used to perform the spectral-fit to extract the CNO solar neutrino interaction rate, it is necessary to assume that the ^{210}Bi decay rate is uniform in the analyzed region and there is no source of ^{210}Pb or ^{210}Bi.

The following hypotheses are assumed:

- **^{210}Pb is out of equilibrium** The ^{210}Pb is not supported by its parent ^{214}Po, or there is no ^{214}Po in the liquid scintillator.
- **No source term of ^{210}Pb** There is no ^{210}Pb entering the fiducial volume, which is different from ^{210}Po. There is a lot of ^{210}Pb on the surface of the Inner Vessel. It is assumed that they are attached to the surface and not washed off into the scintillator. It is also assumed that they do not diffuse into the scintillation bulk.

These two points imply that the amount of ^{210}Pb in the scintillator decays with time with half-life of 22 years. The first assumption is validated, because decay of ^{210}Pb's parents ^{214}Bi and ^{214}Po can be easily tagged through fast coincidence and such kinds of events are not found. Tests are performed to prove the second assumptions and presented in Section 8.3.2 and Section 8.5.

By counting the number of events in a chosen energy range, we have studied the stability of ^{210}Bi events, and it is observed that the event rate decreased at the beginning of Phase-II period and then stayed constant. This is explained as the following. At the beginning of Phase-II, the ^{210}Bi (^{210}Pb) activity were not uniformly distributed in the liquid scintillator. The top of the fiducial volume had higher rates than the bottom. Then convective motions caused mixing, and it is assumed that after certain time the ^{210}Bi (^{210}Pb) decay rate in the fiducial volume has become uniform. Further tests are being performed to validate this assumption.

---

[2] N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: `10.3847/1538-4357/835/2/202`, arXiv: `1611.09867`

[3] Villante et al., "A step toward CNO solar neutrino detection in liquid scintillators"

It is also assumed that

- **No source of ^{210}Bi in the fiducial volume** Considering that from the evolution of the ^{210}Po event distribution, the migration time to reach from the inner vessel surface to the fiducial volume is order of months, even if there is ^{210}Bi detached or diffused from the inner vessel, they decays and will not enter the fiducial volume.

- **The ^{210}Po events consist of three parts** The ^{210}Po events come from the decay of ^{210}Bi dissolved in the liquid scintillator, or are left in the liquid scintillator after the water extraction, or are brought into the fiducial volume by convective motion. This is discussed in detail in the next paragraph.

^{210}Po event types The ^{210}Po decay events are α decays and thus can be selected and counted precisely. The ^{210}Po counts, however, is the sum of three source:

- **the un-supported ^{210}Po** left after the water extraction which decay exponentially;

- **the supported ^{210}Po** in secular equilibrium with ^{210}Bi;

- **the migration ^{210}Po** that originally sits on the inner surface of the inner balloon, are detached from the surface and brought into the center region by the convective motion of the liquid scintillator.

This can be seen from the distribution of ^{210}Po decay rate along z-direction, which is shown in Figure 8.1. Before 2015 (and after around 2013), the convective motion is strong and the third component distribute across the whole inner volume. The relatively clean region, where ^{210}Po decay rate is low, move up and down periodically due to the seasonal temperature variation and thus the modulation of the convection strength. In 2015 summer, the collaboration took a series of operations to reduce the convection current[4]. The idea is to thermally insulate the detector and install heating system surrounding the detector and in the experimental hall, in order to

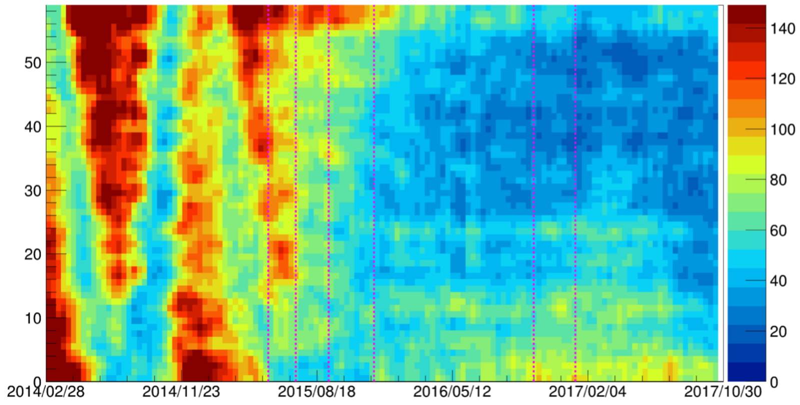


Figure 8.1: The temporal trend of the ^{210}Po decay rate along z direction. The y-axis is the index of the cubes. There are 59 cubes of 1.5 m edge fitting approximately a sphere of radius of 3 m, and index from 0 to 59 represent approximately the z distribution of the ^{210}Po activity from the bottom to the top. Vertical magenta dashed lines represent the most important operation in the thermal stabilization program. From the left to the right: 1. insulation starts, 2. Water Loop off, 3. 5th ring completed, 4. Organ Pipes and 6th ring completed, 5. CR4 floor and Organ Pipes tops and 6. TACS start-up.

[4] See Section 4.9 of S. Marcocci, "Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with Xe -loaded Liquid Scintillators", PhD thesis, Gran Sasso Science Insitute, 2016

maximize the temperature stratification in a stable way[5]. After that it can be seen that the clean region stays in the center region above the equator, and is expanding. In this period, the convection current is weak so that the third component hardly reaches the center region and the concentration almost decrease exponentially along the radial direction, giving chances to separate it from the supported ^{210}Po.

The main challenge of the ^{210}Bi decay rate measurement becomes to select a quiet region where the migration ^{210}Po is negligible while keeping enough exposure, and to quantify systematic uncertainties due to the residual migration ^{210}Po.

Tests of fundamental assumptions Two tests are performed to validate these assumptions and discussed in the following sections:

- obtain ^{210}Bi decay rate with spectral fit of dataset divided by years and divided by top and bottom region
- count the event rate of dataset divided by every 10 days and divided by top and bottom region.

8.2 *Light yield stability and proposal of A1000 dataset*

In this section we presented the study of the detector light yield stability and the proposal of A1000 dataset. It is important to study the detector light yield stability, because:

- In Borexino frequent calibration is avoided to reduce the risk of contaminating the liquid scintillator with radioactive sources inserted in the inner vessel volume.
- In the Monte Carlo fitting method, the detector light yield is assumed to be stable apart from the loss of PMTs. Possible change would lead to bias in the spectral fit results, especially of ^{210}Bi decay rate. Spectral fit results of ^{210}Bi in each year is needed to test the fundamental assumption of the ^{210}Bi measurement.
- The assumptions are also tested by counting the number of events in fixed energy ranges. The change of light yield is a source of systematic uncertainty of this method.

^{210}Po decay events are α-decays and can be tagged with pulse shape discriminators with high efficiency. The peak position of energy distribution of ^{210}Po decay events can be used to measure the detector light yield precisely, and this is investigated in this section.

The A1000 dataset is the dataset found to have stable light yield and can be used as a standard candle. It is the dataset reconstructed with the best and surviving 1000 PMTs. The number 1000 is chosen because it is almost the maximum number of PMTs that survive the whole Phase-II, see Figure 8.2. Because the number of PMTs are stable in this dataset, changes related to the number of live PMTs, the energy resolutions, the fraction of good PMTs are minimized, and thus using this dataset is an excellent way to reduce the systematic

[5] D. Bravo-Berguño et al., "Fluid-dynamics in the Borexino Neutrino Detector: behavior of a pseudo-stably-stratified, near-equilibrium closed system under asymmetrical, changing boundary conditions", *Detectors* 2017, arXiv: `1705.09658`

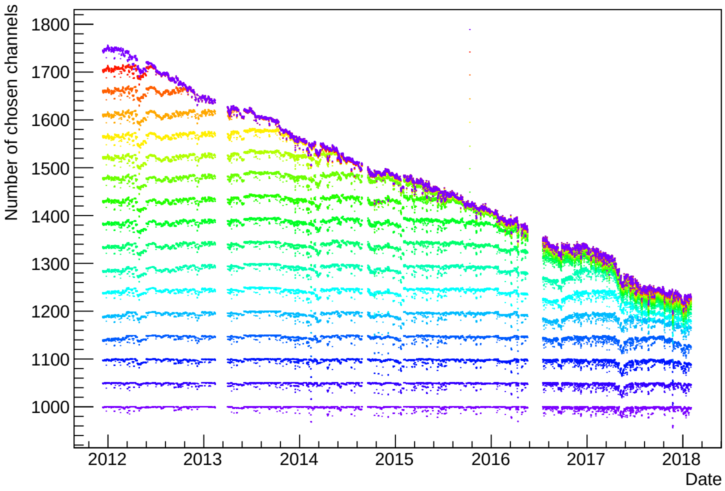


Figure 8.2: The number of survived PMTs versus time under different cut. From the figure we can see that when 1050 or more PMTs are selected, the number of survived PMTs decrease at the last period.

uncertainties. The disadvantage of this dataset is that, because live PMTs are manually disabled in this dataset, the average energy resolution is worse compared with the normal dataset. As a result, the precision of pp and ^{7}Be solar neutrino interaction rates are worse using this dataset.

To obtain the light yield, ^{210}Po decay events were selected and the corrected energy of each event was filled in a histogram. A gaussian plus flat backgrounds model was used to fit the histogram, and the average of the gaussian distribution was used as the value of the light yield.

The correction is necessary, because ^{210}Po decay events are not uniform, and because the detector energy response is not uniform. This is obvious in Figure 8.3. The light yield in the top region is higher than that of the bottom region. When the ^{210}Po decay events are concentrated in the top region, even if the detector light yield doesn't change, the light yield obtained will be higher compared with that obtained when ^{210}Po decay events are concentrated in the bottom. In order to remove this effect, either we remove the ^{210}Po decay events non-uniformity, which can be done by assigning a weight to each event according to the event density, or we remove the detector energy response non-uniformity, which can be done by correcting the event energy. Here the second strategy is taken.

The detector energy response non-uniformity can be corrected with the geometrical correction or the z-correction. The results obtained with two methods were consistent. The detector energy response is not uniform mainly due to the variation of the total geometrical acceptance of live PMTs. The geometrical correction calculate the total acceptance solid angle of all PMTs for each event and normalize the event energy to that in center:

$$E' = E \cdot \frac{\sum_i \Omega_i(\mathbf{0})}{\sum_i \Omega_i(\mathbf{r})} \tag{8.1}$$

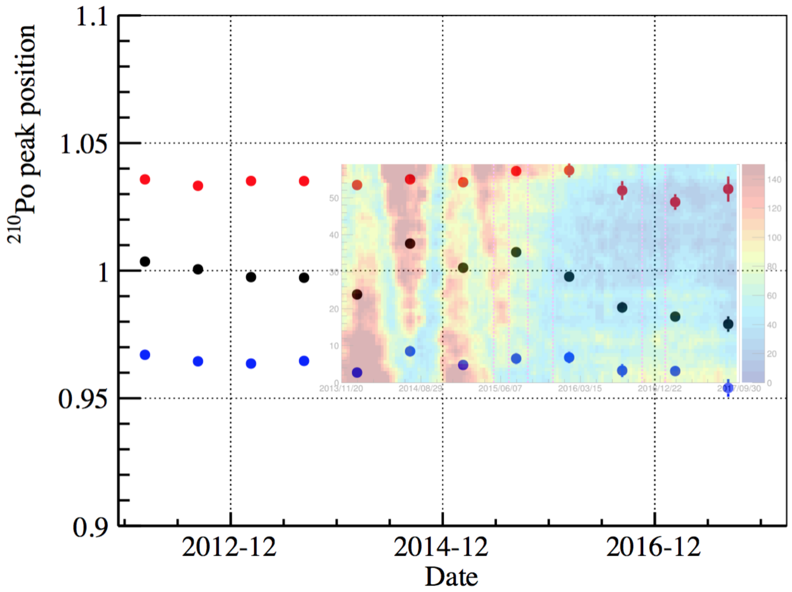


Figure 8.3: The light yield obtained with ^{210}Po events in the whole (black), top (red) and bottom (blue) region. It is obvious that the variation of the obtained light yield is correlated to the ^{210}Po events spatial distribution and the liquid scintillator convective motion.

where E is the energy estimator without correction, E' is the corrected energy estimator, $\Omega_i(\mathbf{r})$ is the acceptance solid angle of i-th fired PMT with respect to the coordinate $\mathbf{r}$, and $\mathbf{r}$ is the reconstructed vertex of the event. The z-correction considers only the z-dependence of the energy non-uniformity, which is the major dependence. This correction uses the peak position of energy distribution of ^{210}Po events, and thus include effects beyond geometrical effects. In this analysis, the z-correction was used.

An example of z correction curve is shown in Figure 8.4. The dependence is well described with a second order polynomial. The obtained value and uncertainty at $z = 0$ is taken as the average light yield at the detector center.

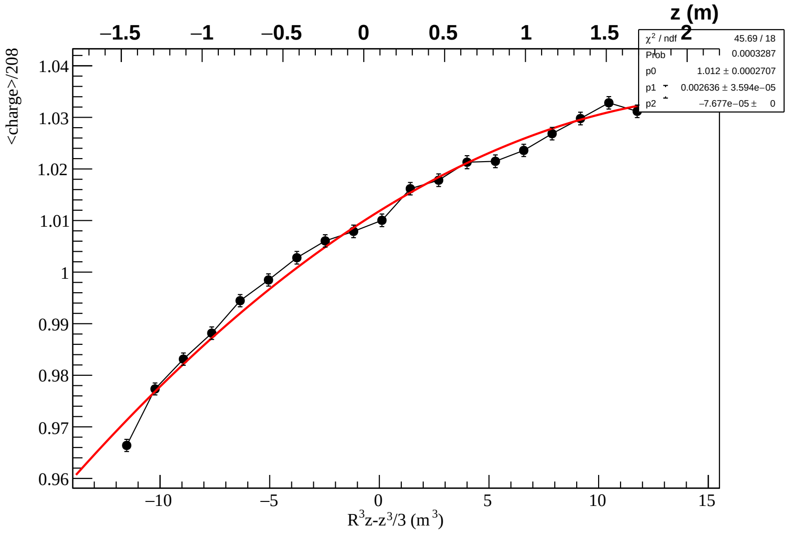


Figure 8.4: The z correction curve corresponding to data collected between 2011 December and 2012 May. ^{210}Po events reconstructed in the sphere of radius R were selected with a multilayer perceptron algorithm. The sphere is divided into 20 iso-volumetric layers. Each point corresponds to one layer in this figure.

The obtained light yield trend is shown in Figure 8.6 as the red band. It can be seen that the light yield has decreased by around 1%. In order to cross-check the correction, an energy estimator with strong correction that even removes that variation of the correction factor at the center is produced, and its stability is shown in Figure 8.5. As expected, it is constant. In order to decouple the factors related to the loss of PMTs, a special dataset reconstructed with only 1000 PMTs was prepared and the procedure mentioned above was repeated to obtain the light yield trend, and the result is shown in Figure 8.6 as the black band. It can be seen that the light yield with this dataset is stable, so that the temperature variation and thermal insulation operations did not cause the change of optical properties of the liquid scintillator.

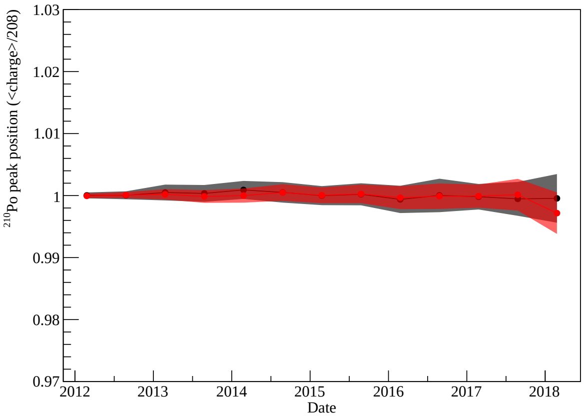


Figure 8.5: Energy estimator with strong z correction that removes also the temporal dependence of the light yield at the detector center. As expected, it is constant over time with all or surviving PMTs. Red: dataset reconstructed with all live PMTs. Black: dataset reconstructed with 1000 selected best PMTs.

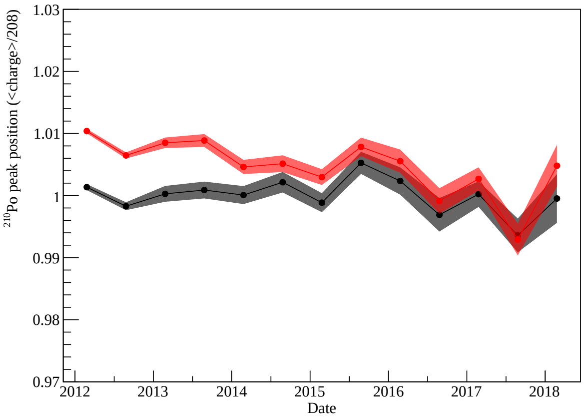


Figure 8.6: The light yield at the detector center obtained with ^{210}Po events corrected with z-correction. Red: dataset reconstructed with all live PMTs. Black: dataset reconstructed with 1000 selected best PMTs.

This result is further confirmed with ^{14}C events. In Borexino, ^{14}C events are selected to calculate the so-called effective quantum efficiencies[6]. The effective quantum efficiencies (EQE) is a measure of the variation of the light yield. In Figure 8.7 the ratio of average of EQEs of all PMTs and that of 1000 selected best PMTs is shown and compared with the ratio of the light yield between two datasets. It can be seen that two trends are consistent.

[6] Chapter 4.4 I. Drachnev, "New Spectral Analysis of Solar B Neutrino with the Borexino Detector", PhD thesis, Gran Sasso Science Institute, 2016

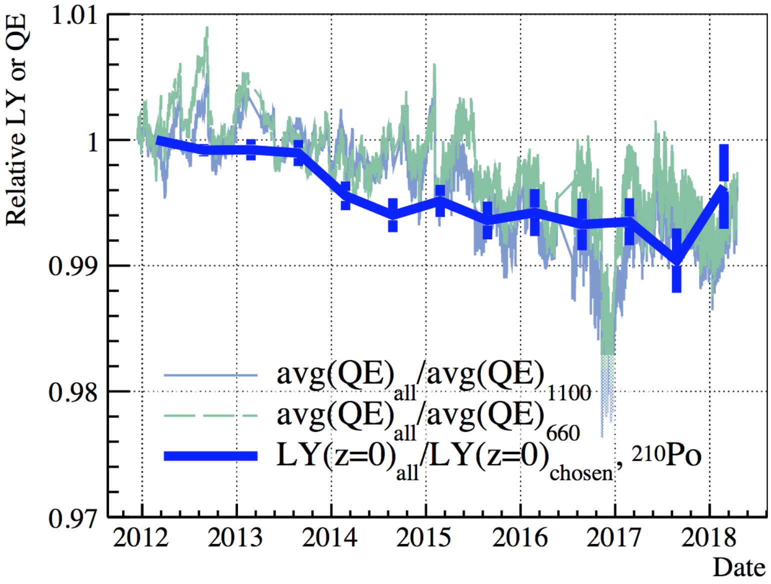


Figure 8.7: The ratio of the sum of EQEs of all PMTs and that of 1000 selected PMTs. The ratio of the light yield of all PMTs and that of 1000 selected PMTs is also shown. From the figure we can see that two ratios are consistent.

This is result is also consistent with the results obtained from the analytical spectral-fits. The light yield from the analytical `npmt_dt1` fit with all PMTs and with 1000 selected PMTs are shown in Figure 8.8. From the figure we can see that the light yield of all PMTs decreased while that of 1000 selected PMTs is stable.

In conclusion, with three different methods, it is found that the light yield of the normal dataset with all PMTs has decreased by 1% throughout the Phase-II period, and the light yield of the `A1000` dataset is constant. From this we can further conclude that the optical properties of the liquid scintillator have not changed in the Phase-II and the light yield decrease is due to the loss of PMTs.

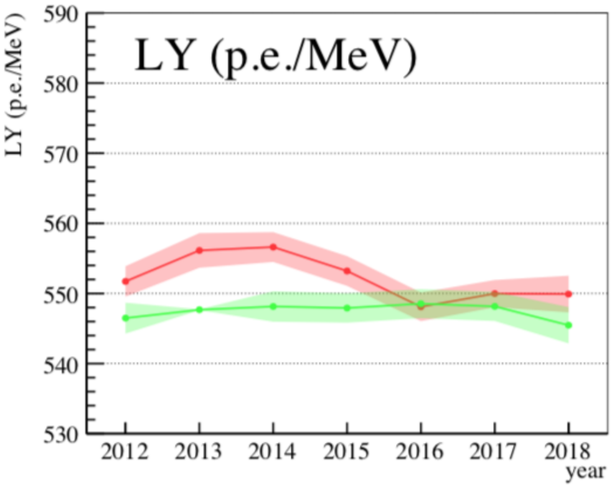


Figure 8.8: The light yield obtained from the analytical `npmt_dt1` fit of dataset reconstructed with all PMTs and with 1000 selected PMTs.

8.3 The `A1000` dataset

In this section, the production, validation and the application of `A1000` is presented.

Generally, `A1000` is a set of concepts proposed to reduce the temporal variation induced systematic uncertainty due to PMT dying. Specifically, `A1000` is the set of 1000 chosen PMTs that ranked best regarding their live time, average Effective Quantum Efficiencies (`EQE`) and dark noise rate in Phase-II. `A1000` dataset refers to the dataset reprocessed enabling only `A1000` PMTs. In this dataset, the light yield and the energy resolution is much more stable, and thus can improve the stability of various studies, including the two methods of study the ^{210}Bi decay rate stability mentioned in the last section. Besides, using this dataset, it is possible to track the temporal dependence of light yield with ^{14}C events when calculating the effective quantum efficiencies, and thus the treatment of light yield in the Monte Carlo simulation can be improved.

8.3.1 Procedure of selecting `A1000`

The set of `A1000` is selected by ranking good PMTs according to their live time. A PMT is marked as good if its Calibrated Effective Quantum Efficiencies (`CEQE`) is high and its dark noise rate is low. 7768 validated runs, from run 17328 till run 30368, are analyzed.

The Effective Quantum Efficiencies (`EQE`) The `EQE` is a set of values proportional to the number of collected hits on each PMT in the selected ^{14}C events. It is used as the input of the Monte Carlo simulation in order to reproduce the light yield and energy response non-uniformity in data. In this work, the values from the database (`Qe` of table `QuantumEfficiency` in the database `bx_calib`) is used.

The live time of PMTs A PMT is marked as being used in a run if and only if all these criteria are fulfilled:

1. The PMT is associated with a non-zero channel number in the `HolesMapping` table of `bx_geometry` database. This means a PMT is installed and it is connected with the electronics.

2. The corresponding channel is not in the `BadChannelsList` or `OffChannelsList` of table `LabenPrecalibDecodingQuality` in the `bx_precalib` database. This means the channel collects enough hits in the first 1000 pulser events.

3. The `ChannelDescription` of the channel is not marked as a type different from `Ordinary` in the `LabenChannelMapping` table of `daq_config` database. This means the channel is not for service purpose, but used to collect photoelectrons.

4. The channel is not marked as disabled in the timing channel in the first 2000 events based on `DisabledChannels` table of `bx_calib`

database. This means the channel is not disabled by the reconstruction algorithms.

The number of live PMTs should be almost consistent with the value reported in the processing log files of the reconstruction software `Echidna`.

Once a PMT is marked as live in a run, the whole run duration is counted as the live time of the PMT. Such algorithm can be further improved.

The `A1100` PMT set The `A1100` PMT set is the top 1100 PMTs that have longest live time in the analyzed runs.

The `A1100p` PMT set The `A1100p` PMT set is the PMTs that are in `A1100` and fullfill `EQE`>0.1 in the corresponding run. `A1100p` is defined run by run.

The Calibrated Quantum Efficiencies (`CEQE`) `CEQE` is defined as the ratio of `EQE` and the average `EQE` of `A1100p` PMT set.

The Dark Noise rate The dark noise rate from the database (`DarkNoise` of table `InnerPmtsDarkRate` in the database `bx_calib`) is used.

The `A1000` PMT set The `A1000` PMT set is selected in the following steps:

1. The PMT should have a positive live time in the analyzed runs.
2. The live time weighted average `CEQE` of the PMT is larger than 0.6
3. The live time weighted average dark noise rate of the PMT is less than 1
4. Rank the rest PMTs according to their live time. Pick the top 1000 PMTs.

The threshold 0.6 and 1 is chosen such that the PMT qualities are improved while the loss of PMTs is acceptable. Around 2% PMTs are removed for each of two cut. To keep the stability of number of live PMTs, 10% PMTs are removed compared with `A1100` PMT set in which no cut is applied.

`CEQE` rather than `EQE` is used as cut criteria because it is believed to be more stable, and because `EQE` is believed to be bugged such that the `EQE` average is supposed to be constant while found to be increasing, see Figure 8.9.

Production of `A1000` dataset The `A1000` dataset is produced by executing `Echidna` on raw data disabling all logic channels (Channel 32 is disabled, too) that are

- not associated with an `A1000` PMT

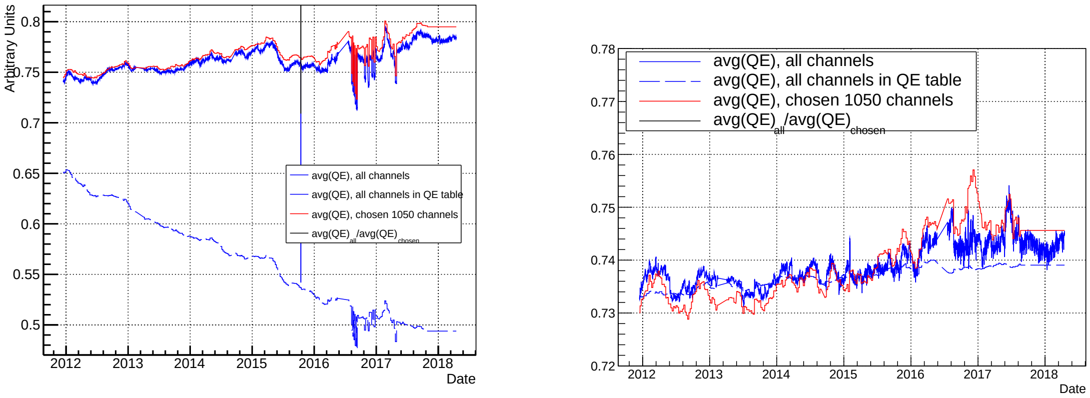


Figure 8.9: Average `EQE` from database (left) and ASCII files (right, `g4bx2` input). Two source of EQE treat the backgrounds in different way. On the left, the backgrounds are suppressed by imposing cuts. On the right, the backgrounds are subtracted. Each point correspond to one run. The most recent `EQE` value set is used for the corresponding run. Y axis is the arithmetic average `EQE` of all live PMTs in that run. From the figure it is obvious that in both methods the average is not constant. In `g4bx2` the average `EQE` increased by 1% in 2016.

- (and) `ChannelDescription` not marked as a type other than `Ordinary` in the `LabenChannelMapping` table of `daq_config` database.

The whole processing chain is executed. Clusters are redefined. Energy and vertex are re-calculated.

Validation The number of live PMTs are validated to be as stable as expected. The decrease only appears after second half of 2017 and the magnitude is around 30 PMTs. See Figure 8.10.

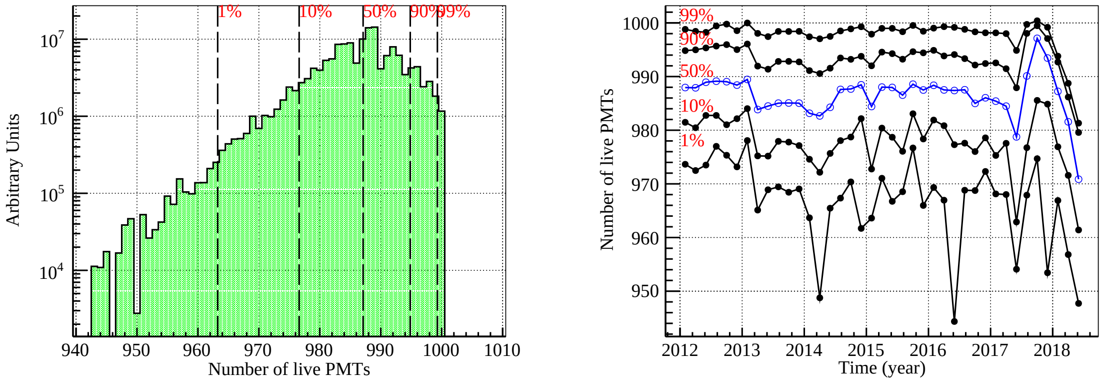


Figure 8.10: Left: The distribution of number of live PMTs in the output of `bx_nusol` module. Different levels of quantiles are indicated with dashed lines. Right: quantile of distribution of number of live PMTs versus the event time stamp. Five quantile curves with level 1%, 10%, 50%(median), 90% and 99% are shown. From the figure we can see there is 2.5% decrease of number of PMTs starting from around 2017 September. Before that the number of live PMTs is stable within 0.5%.

8.3.2 *Applications*

The `A1000` has been applied in three studies: the ^{210}Bi decay rate stability measurement with the spectral fit method and the fixed window counting method, and the new effective quantum efficiencies. The first result has been described in Section 8.5, and here we will introduce the other two.

^{210}Bi uniformity with fixed window counting method The two fundamental assumptions mentioned in Section 8.1 can be tested by counting the number of events in a fixed energy range. The results obtained with the normal dataset and `A1000` dataset is shown in Figure 8.11. From the figure we can see that the event rate of top and bottom region is closer when `A1000` is used. Discrepancy found using the normal dataset is probably due to variation of the light yield, and discrepancy found using the `A1000` dataset is being investigated and believed to be related to the radial dependence of the efficiency of the pulse shape variables used to select ^{210}Po events.

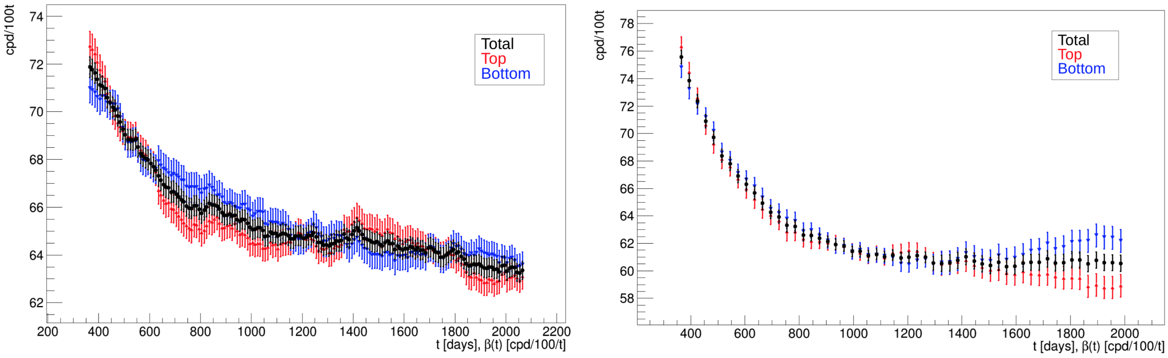


Figure 8.11: ^{210}Bi uniformity with fixed window counting method. Left: using normal dataset reconstructed with all PMTs. Right: using `A1000` dataset reconstructed with 1000 selected PMTs.

New effective quantum efficiencies The effective quantum efficiencies[7] (EQE) serve as a input in MC to describe the isotropy of light collections as well as current only way to change the temporal trend of light yield. The EQE is calculated using the number of hits on each PMT in the selected number of events. It is expressed as

$$Q_j^i = \epsilon_j \cdot \frac{N_j^i}{M_j} \tag{8.2}$$

where M_j is the number of selected events, N_j^i is the number of hits collected on i-th PMT, ϵ_j is a calibration factor, j represents the index of the dataset and correspond to around one week of data, and Q_j^i is the effective quantum efficiency.

In the past, ϵ_j is chosen such that the average of Q_j^i is constant over time, assuming the light yield does not change. The new effective quantum efficiencies include the actual temporal trend of the light yield. Naively, if ϵ_j is set to constant over time, Q_j^i will track the temporal trend of the light yield through N_j^i/M_j. However N_j^i is also affected by the average energy of the selected events. If normal dataset is used, because we select events based on the visible energy, the average energy of selected events is decreasing due to worsening of the energy resolution and increase of contribution of low energy events.

Using `A1000` dataset can solve this problem, because in this dataset the number of live PMTs is almost constant, and as a result the energy resolution almost does not change. Calculate the average EQE of all `A1000` PMTs using the `A1000` dataset, then calculate the EQE of all PMTs using the normal dataset requiring the average EQE of all PMTs to be constant. The final EQE is given by

$$Q_j^{i\prime} = \frac{\bar{Q}_j(A)}{\bar{Q}_c(A)} \cdot Q_j^i \tag{8.3}$$

[7] See Section 4.4 of Drachnev, "New Spectral Analysis of Solar B Neutrino with the Borexino Detector"

where Q_j^i is the EQE of i-th PMT in the j-th normal dataset, $\bar{Q}_j(A)$ is the average EQE of `A1000` PMTs in the j-th `A1000` dataset, $\bar{Q}_c(A)$ corresponds to one calibration run and is a constant.

8.4 *Methods to get rid of the migration ^{210}Po*

In this section we introduced the analysis methods used to get rid of the migration ^{210}Po in order to measure the supported ^{210}. I have worked with Nicola Rossi and Francesco Villante on this work. It is worth emphasizing again that the thermal insulation and installation of the heating system are surely the basis of all contents of this section. Four methods have been developed: the vortex fit method, the minimum path method, the plateau finder and the radial fit.

8.4.1 *The vortex fit method*

This method relies on the fact that, for the recent data, the spatial distribution of ^{210}Po events around the minimum rate position has the shape of a vortex. Assuming the rotational symmetry along the z-axis (the vertical direction), then the coordinate system is reduced to ρ, z where $\rho^2 = x^2 + y^2$ and x, y, z are the Cartesian coordinates. The distribution of ^{210}Po events[8] is shown in Figure 8.12. From the figure we can see that there is a deep vortex shape around the minimum. The distribution is approximated with the paraboloid function, and when the included volume is less than 10 tons, the paraboloid function can describe the distribution well:

$$R(\rho^2, z) = MT \cdot \left(N_0 + \frac{\rho^2}{a^2} + \frac{(z - z_0)^2}{b^2} \right) \tag{8.4}$$

where MT is the exposure.

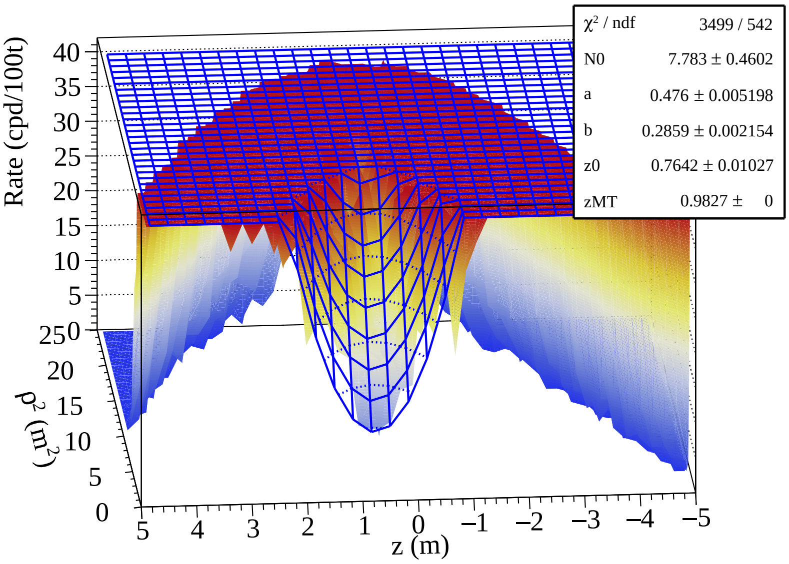


Figure 8.12: The fit of the distribution of the vertexes of ^{210}Po events using a paraboloid function.

The ^{210}Po decay rate and the vortex position extracted from the paraboloid fit is shown in Figure 8.13. From the figure we can see

[8] selected with 150<geo-normalized `charge`<270 and MLP<0.5

that the vortex minimum position is relatively stable in time, so it is unlikely due to fluctuations and minimally contaminated by (or even free from) the migration ^{210}Po.

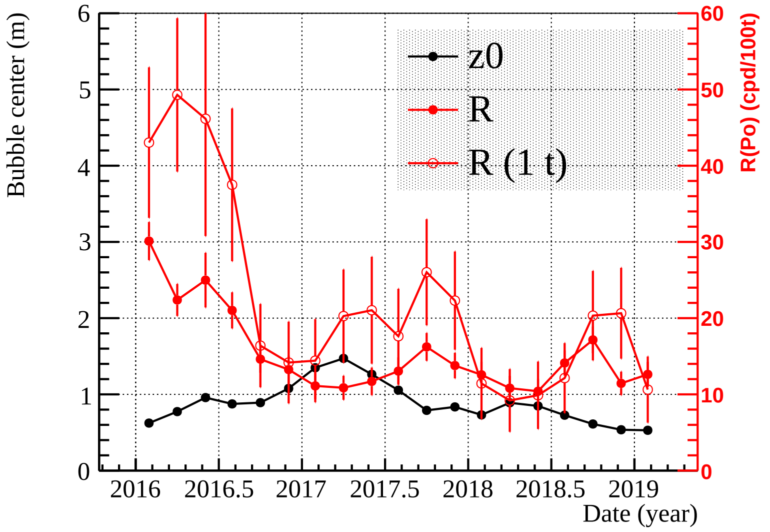


Figure 8.13: The vortex bottom position (black dot, z_0 in the legend) and the corresponding rate of the ^{210}Po distribution (red dot, R in the legend). The red open dot (R (1t) in the legend) shows the average event rate in the inner most 1 ton volume. If we use the data after 2018 and fit the ^{210}Po event rate spatial distribution with a paraboloid function, the minimum of the function is 9.7 ± 1.6 cpd/100t. With ^{210}Bi constrained to this and the *pep* solar neutrino interaction rate constrained to the HZ model with a 2% precision, the CNO solar neutrino interaction rate from the spectral fit is 4.7 ± 1.1 cpd/100t without any systematic uncertainties.

8.4.2 *The minimum path method*

This method relies on the fact that due to incompressibility of liquids, the migration ^{210}Po vanishes at the minimum or maximum. Denote the density of liquid scintillator as $\rho(\mathbf{r}, t)$ which can be approximated as a constant ρ_0. Denote its velocity as $\mathbf{v}(\mathbf{r}, t)$, the concentration of ^{210}Pb, ^{210}Bi and ^{210}Po as $X_{^{210}\mathrm{Pb}}(\mathbf{r}, t)$, $X_{^{210}\mathrm{Bi}}(\mathbf{r}, t)$ and $X_{^{210}\mathrm{Po}}(\mathbf{r}, t)$, respectively, then they are related by:

$$\frac{\partial X_{\mathrm{Bi}}}{\partial t} = X_{\mathrm{Pb}} \cdot \lambda_{\mathrm{Pb}} - X_{\mathrm{Bi}} \cdot \lambda_{\mathrm{Bi}} + \nabla \cdot (D_{\mathrm{Bi}} \cdot \nabla X_{\mathrm{Bi}} - \mathbf{v} X_{\mathrm{Bi}}) \tag{8.5}$$

$$\frac{\partial X_{\mathrm{Po}}}{\partial t} = X_{\mathrm{Bi}} \cdot \lambda_{\mathrm{Bi}} - X_{\mathrm{Po}} \cdot \lambda_{\mathrm{Po}} + \nabla \cdot (D_{\mathrm{Po}} \cdot \nabla X_{\mathrm{Po}} - \mathbf{v} X_{\mathrm{Po}}) \tag{8.6}$$

$$0 = \nabla \cdot \mathbf{v} \tag{8.7}$$

where D is the diffusion coefficient and is around 10^{-5} atoms/s/cm^2[9], λ is the decay rate of the corresponding isotopes. Eq. (8.7) is due to that liquid cannot be compressed.

If we focus on Eq. (8.6), and track the temporal dependence of the concentration of ^{210}Po at the minimum at each moment, we see that

$$\begin{aligned}
\frac{\mathrm{d}X_{\mathrm{Po}}(\mathbf{r}_{\mathrm{min}}(t), t)}{\mathrm{d}t} &= \left[\frac{\partial X(\mathbf{r}, t)}{\partial \mathbf{r}} \frac{\mathrm{d}\mathbf{r}_{\mathrm{min}}}{\mathrm{d}t} + \frac{\partial X}{\partial t}\right]\Big|_{\mathbf{r}_{\mathrm{min}}(t)} = \frac{\partial X}{\partial t}\Big|_{\mathbf{r}_{\mathrm{min}}(t)} = \\
&= X_{\mathrm{Bi}} \cdot \lambda_{\mathrm{Bi}} - X_{\mathrm{Po}} \cdot \lambda_{\mathrm{Po}} + \nabla \cdot (D_{\mathrm{Po}} \cdot \nabla X_{\mathrm{Po}} - \mathbf{v} X_{\mathrm{Po}}) = \\
&= X_{\mathrm{Bi}} \cdot \lambda_{\mathrm{Bi}} - X_{\mathrm{Po}} \cdot \lambda_{\mathrm{Po}} + D_{\mathrm{Po}} \nabla^2 X_{\mathrm{Po}} - (\nabla \cdot \mathbf{v}) X_{\mathrm{Po}} - \mathbf{v} \cdot \nabla X_{\mathrm{Po}} = \\
&= X_{\mathrm{Bi}} \cdot \lambda_{\mathrm{Bi}} - X_{\mathrm{Po}} \cdot \lambda_{\mathrm{Po}} + D_{\mathrm{Po}} \nabla^2 X_{\mathrm{Po}}
\end{aligned} \tag{8.8}$$

So the contribution from the convection vanishes at the minimum path.

---

[9] J. Winkelmann, "Diffusion coefficient of acetic acid in benzene", in: *Diffusion in Gases, Liquids and Electrolytes: Non-electrolyte Liquids and Liquid Mixtures - Part 1: Pure Liquids and Solute in Solvent Systems*, ed. by M. D. Lechner, Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 623–623, DOI: `10.1007/978-3-540-73735-3_404`

Now we consider the term from the diffusion $D_{\mathrm{Po}}\nabla^2 X_{\mathrm{Po}}$. First, it is positive. Second, it is not big. Use a quantity with the unit of length L_Δ to describe the $\nabla^2 X_{\mathrm{Po}}$:

$$X_{\mathrm{Po}} \approx X_{\mathrm{Po}}(\mathbf{r}_0)\left[1+\left(\frac{|\mathbf{r}-\mathbf{r}_0|}{L_\Delta}\right)^2\right] \tag{8.9}$$

where $\nabla^2 X_{\mathrm{Po}} = X_{\mathrm{Po}}/(2L_\Delta^2)$. Consider the distribution of the ^{210}Po will be smeared out by the diffusion, the L_Δ cannot be too small. Consider initially ^{210}Po is only at one point, then after certain time it becomes a gaussian like:

$$X_{\mathrm{Po}}(\mathbf{r}, t=0) = \delta(\mathbf{0}) \tag{8.10}$$

$$X_{\mathrm{Po}}(\mathbf{r}, t \to +\infty) = \exp(-\lambda_{\mathrm{Po}} t) \cdot G(|\mathbf{r}|; 0, \sigma) \tag{8.11}$$

where G is the gaussian distribution with average 0 and standard deviation of σ. Obviously,

$$\sigma = \sqrt{D\tau} \approx 10^1 \text{ cm} \tag{8.12}$$

So $L_\Delta \gtrsim \sigma$ and

$$\frac{D\nabla^2 X_{\mathrm{Po}}}{X_{\mathrm{Po}} \cdot \lambda_{\mathrm{Po}}} \approx \frac{D}{2L_\Delta^2 \lambda_{\mathrm{Po}}} \tag{8.13}$$

$$\lesssim \frac{10^{-5}\ \mathrm{s}^{-1}\mathrm{cm}^2}{2 \times (10^1\ \mathrm{cm})^2 \cdot 0.005\ \mathrm{day}^{-1}} \tag{8.14}$$

$$= 1 \tag{8.15}$$

In reality the variation of ^{210}Po spatial distribution is much weaker, and L_Δ is typically around 1 meter after thermal insulation, then the contribution of the diffusion term is negligible.

In practice, the system is simplified to be a one dimensional system. ^{210}Po events in spheres of chosen radii are selected, and the spheres are divided into iso-volumetric layers along vertical direction. This is reasonable, because in the vertical direction we observe the maximal effect of the convective motions.

The minimum decay rate is estimated with a weighted binned likelihood based on the event rates of all layers instead of taking naively the lowest rate among layers, which is biased by statistical fluctuations. The estimator is also validated to be statistically unbiased with toy Monte Carlo tests assuming various types of distributions. The minimum path obtained is shown in Figure 8.14.

The estimated minimum decay rate is shown in Figure 8.15. From the figure we can see that the rate reaches a plateau of around 20 cpd/100t.

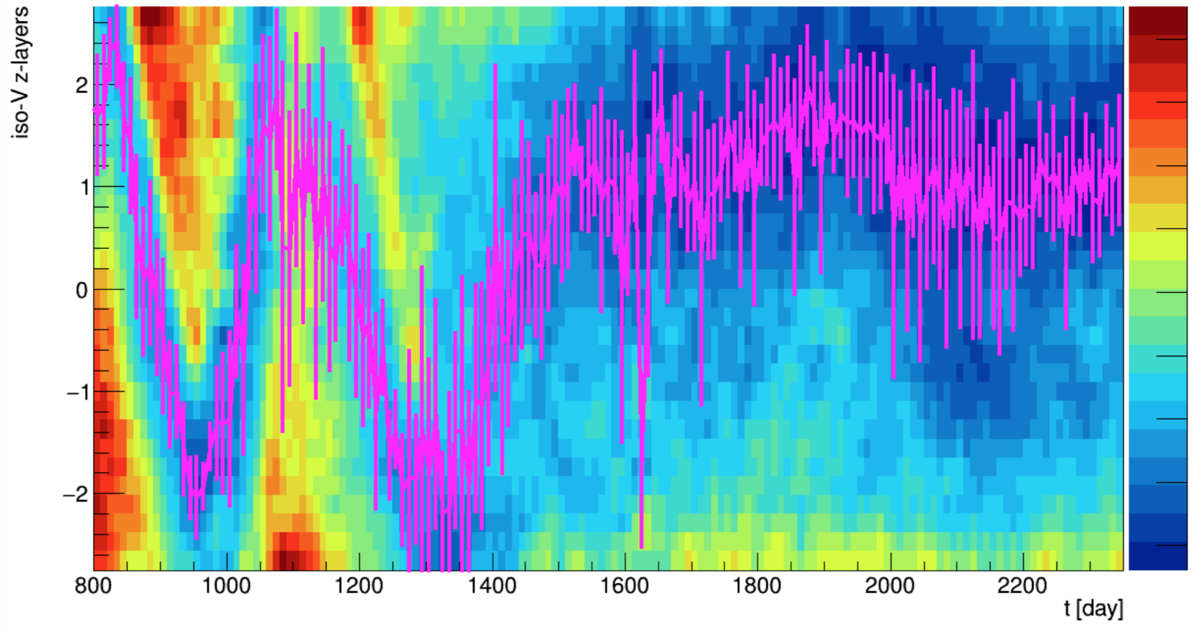


Figure 8.14: The location of minimum of ^{210}Po decay rate along vertical direction found by the weighted likelihood method versus time.

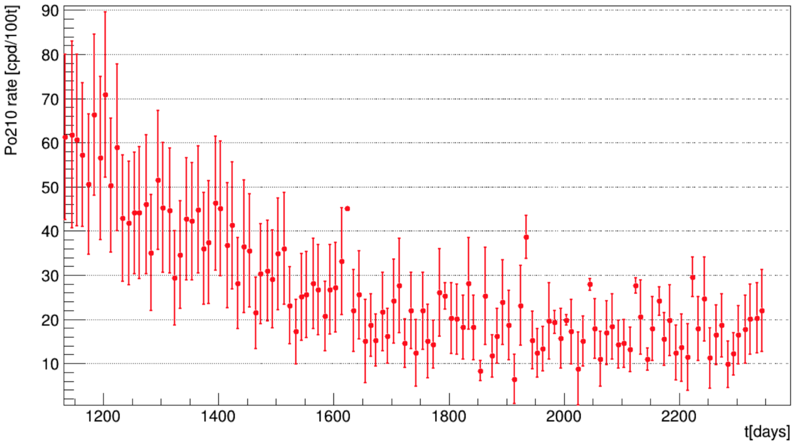


Figure 8.15: The ^{210}Po decay rate along the found minimum path.

8.4.3 The plateau finder

The plateau finder relies on the assumption that the ^{210}Po decay rate should be uniform in the migration ^{210}Po free region, and thus it is a "plateau". Due to limited statistics, the problem is simplified to be one dimensional. Events in sphere of chosen radii were selected and the sphere were divided in to iso-volumetric layers along vertical direction. More details can be found in the PhD thesis of Daniele Guffanti[10].

In this method, first the distribution of ^{210}Po decay rate along vertical direction is obtained by the kernel density estimator[11], then the plateau region is found by expanding the minimum until the first derivative of two edges reach certain threshold. The threshold is tuned with toy Monte Carlo tests such that the bias is under control. Other ideas have also been attempted, such as fixing the size of the plateau or constraining the plateau size to be within certain range, and minimize the bias by looking for the plateau size where the decay rate in the plateau becomes stable with respect to the plateau size.

8.4.4 The radial fit

The decay rate distribution of ^{210}Po events in the top part along radial direction is modeled with three components:

- A gaussian component, representing the ^{210}Po sitting in the inner vessel, or the inner balloon.

- A exponential component, or diffusion-like component. They are modeled by $\frac{\sinh(r/\lambda)}{r/\lambda}$[12], representing the ^{210}Po diffused from the inner vessel and the migration ^{210}Po.

- A flat component, representing the supported ^{210}Po.

This can be seen from figure 8.16.

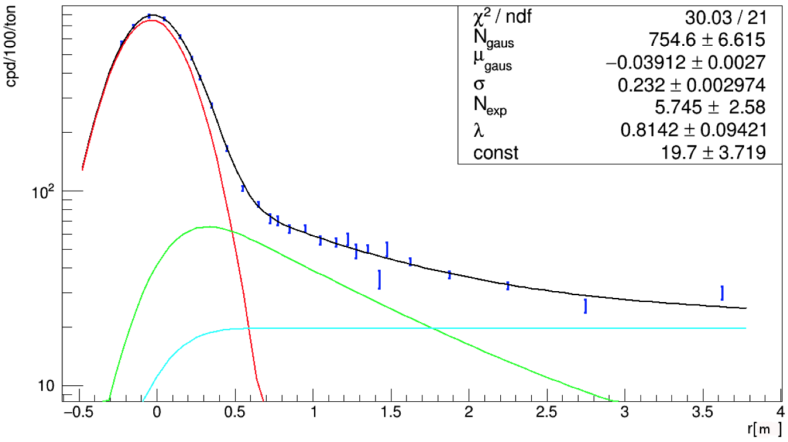


Figure 8.16: The ^{210}Po decay rate along the radial direction.

---

[10] D. Guffanti, "Measurement of solar neutrino with BOREXINO", PhD thesis, Gran Sasso Science Institute, 2019

[11] E. Parzen, "On Estimation of a Probability Density Function and Mode", *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, Sept. 1962, DOI: `10.1214/aoms/1177704472`, arXiv: `arXiv:1011.1669v3`

[12] This is the steady solution of the diffusion equation with decays:

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} = D\nabla^2 \rho$$

When the migration ^{210}Po is not present, the constant of the diffusion-like component distribution is given by

$$\lambda = \sqrt{D\lambda_{\text{Po}}} \approx 10^1 \text{ cm} \tag{8.16}$$

where D is the diffusion coefficient, λ_{Po} is the decay rate of ^{210}Po. When the migration ^{210}Po is present, λ is larger, and increases with stronger convection current.

8.5 ^{210}*Bi decay rate temporal trend*

The ^{210}Bi decay rate can be obtained with two methods: the spectral fit and the fixed window counting method.

The spectral fit results are shown in Figure 8.17.

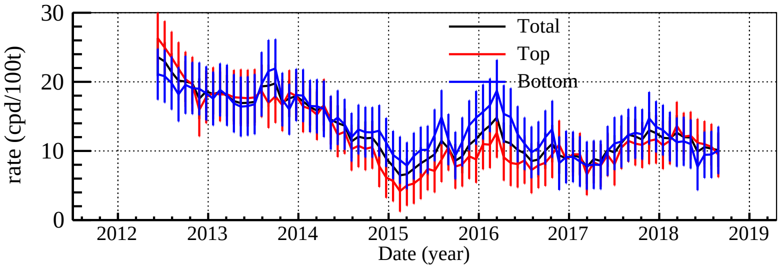


Figure 8.17: ^{210}Bi rate from spectral fit. Geo-normalized `charge` variable is used here. The neutrino rates are constrained to the HZ model. Each point is the fit of one year data, and the x-coordinate is the center of the year. From the figure we can see that the ^{210}Bi decay rates become almost stable after around 2015, and they are consistent between the top and the bottom.

For the counting analysis, the energy region is optimized to be 284–471 in geo-normalized `charge`, where the contribution of ^{210}Po events is negligible and the ratio of the ^{210}Bi event count to the total statistical uncertainty is the maximum. In this energy region, the contribution of ^{210}Bi is around 20%. The event rate versus time is shown in Figure 8.18.

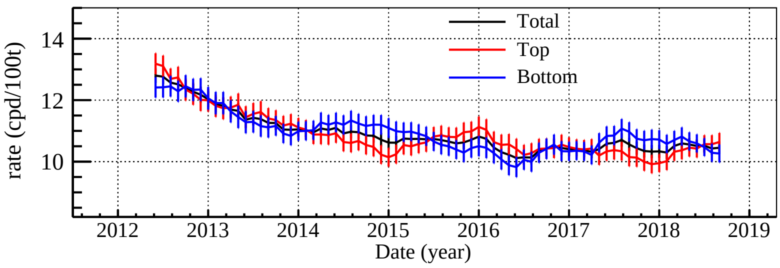


Figure 8.18: Event rate in 284–471 p.e. region. Geo-normalized `charge` variable is used here. Each point corresponds to one year data, and the x-coordinate is the center of the year. From the figure we can see that the ^{210}Bi decay rates become almost stable after around 2015, and they are consistent between the top and the bottom.

Both results show that the ^{210}Bi decay rates become almost stable after around 2015, and they are consistent between the top and the bottom.

8.6 *Conclusions and discussions*

We have developed a method to measure the ^{210}Bi decay rate with its daughter ^{210}Po. This study is of essential importance for discovering the CNO solar neutrinos in Borexino. In this method, the ^{210}Bi decay rate can be measured if a few assumptions are made: there is no source of ^{210}Pb or ^{210}Bi, and both of them are uniform distributed in the fiducial volume in a certain period. When the assumptions are fulfilled, the next challenge is to get rid of the detached ^{210}Po from the inner surface of the nylon vessel brought into the fiducial volume by the convective motion of the liquid scintillator. When this type of ^{210}Po (convection ^{210}) is absent and early data is abandoned so that ^{210}Po left from the water extraction decays to negligible level, the ^{210}Po fully comes from the decay of ^{210}Bi, the ^{210}Bi and ^{210}Po are in secular equilibrium and their decay rates are linked. Tests are performed to validate the assumptions made in this method and are presented. Four methods are developed to get rid of the migration ^{210}Po: the vortex fit method, the minimum path method, the plateau finder and the radial fit.

In order to reduce the systematic uncertainty, we have proposed a special dataset enabling only surviving PMTs is proposed, and this dataset is found to be useful in various studies, including testing fundamental assumptions made in the method to measure the ^{210}Bi decay rate and improving the light yield monitoring in the Monte Carlo simulation.

It is promising that we can make a 10% ^{210}Bi decay rate measurement with this method. The next challenge is to evaluate the systematic uncertainties due to the residual migration ^{210}Po.

9

The JUNO experiment

In this chapter we introduce the Jiangmen Underground Neutrino Observatory (JUNO) experiment. JUNO is an under-construction reactor anti-neutrino experiment. Due to its kilo-ton scale target mass and excellent energy resolution, it provides opportunities to vast number of physics topics. The experiment site is chosen to be on the middle line of Yangjiang and Taishan Nuclear Power Plants and 53 kilo-meters away from them. It is optimized for the main goal of JUNO: determining the neutrino mass ordering. In this chapter, an introduction to the JUNO detectors and its physics programs are given.

This chapter is organized as in the following: Section 9.1 describes the JUNO center and veto detectors. Section 9.2 describes the physics programs.

9.1 The JUNO Detectors

In this section we introduced the JUNO detectors, which consist of three sub-systems[1]: the central detector, the veto detector and the calibration system.

9.1.1 The central detector

The central detector is the detector that acquires physics events. It is depicted in Figure 9.1. Twenty thousand tons of liquid scintillator are contained in an acrylic ball whose inner diameter is 35.4 meters. Outside the acrylic ball it is filled with water. Around eighteen thousand twenty inches PMTs and twenty five thousand three inches PMTs are installed outside the acrylic ball in the water, supported by stainless steel structure. The photocathode coverage is as high as 78%.

The central detector is a liquid scintillator based calorimeter. It is required to reach 3% energy resolution and keep the precision of non-linearity model better than 1%.

JUNO uses the Linear Alkyl Benzene (LAB) as solvent with 2.5 g/L 2,5-diphenyloxazole (PPO) and 1∼3 mg/L p-bis-(o-methylstyryl)-benzene (bis-MSB) dissolved. To achieve good energy resolution, it is required to have hight light yield, and long attenuation length of

---

[1] T. Adam et al., "JUNO Conceptual Design Report", p. 328, 2015, arXiv: `1508.07166`

LAB is the key. Currently the result is promising that the attenuation length is longer than 25 meters for 430 nm photons. Good radiopurities are necessary to improve the sensitivity to various physics targets. Liquid scintillator are processed by alumna extraction, distillation, water extraction and N_2 stripping. A dedicated detector called OSIRIS is designed to monitor the liquid scintillator purity onsite quickly.

Among all twenty inches PMTs, fifteen thousand are MCP-PMT and five thousand are dynode PMTs. It is chosen like this because dynode PMTs have better time resolution, which affects the vertex reconstruction precision. Five thousand are enough for the JUNO requirement.

Three inches PMTs system always work in the photon counting regime, and thus can be used to calibrate the non-linearity of charge reconstruction, and also to reduce the non-stochastic resolution term. Because the dynamic range is increased, they help with large signals, such as shower μ events, and thus vetoing of cosmogenic ^{12}B, ^{9}Li and ^{8}He, etc. Besides, they can reach similar precision on $\sin^2\theta_{12}$ and Δm^2_{12} with respect to that of 20 inches PMTs, so they can provide a rather independent crosscheck of the results. At last, they can also be used as a complementary readout system for supernova signals and to ensure unbiased energy and rate measurement.

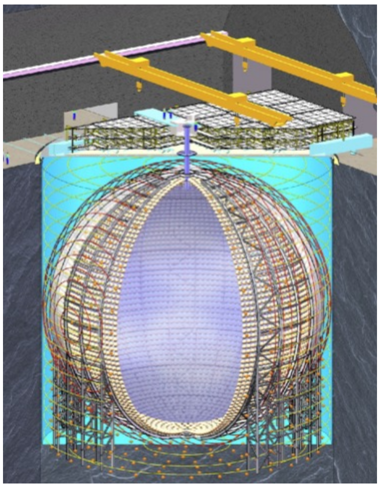


Figure 9.1: The center detector of JUNO.

9.1.2 The veto detector

The veto detector is used to tag muon events and to reconstruct the track of μs. JUNO has two veto detectors: the water Cherenkov detector and the top tracker.

The top tracker is a plastic scintillator detector used for precise muon tracking. It covers half of the top area. The water pool Cherenkov detector consists of 35 kilo-tons of ultra-pure water and around two thousand 20 inches MCP PMTs. It is used both to reconstruct the μ track and to shield the γs from rocks and fast neutrons.

9.1.3 The calibration system

The calibration system consists of three complementary systems

- The one dimensional system: the Automated Calibration Unit (ACU)
- The two dimensional system: the cable loop system and the guide tube
- The three dimensional system: the Remotely Operated under-liquid-scintillator Vehicles (ROV).

They are used in different ranges and frequencies, and are depicted in Figure 9.2.

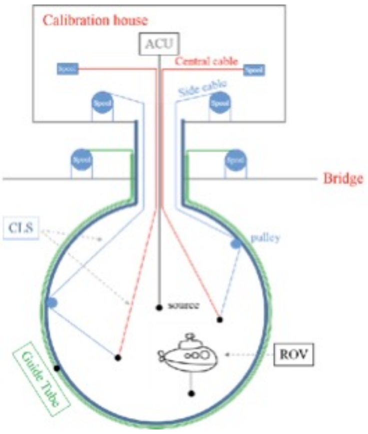


Figure 9.2: The calibration system of JUNO.

9.2 *Physics program*

JUNO provides opportunities to vast number of physics programs. In this section, we introduced the physics potential of JUNO based on the contents of[2].

9.2.1 *Identifying the neutrino mass ordering*

The main goal of JUNO is to identify the neutrino mass ordering through the 3-neutrino vacuum oscillation interference.

The determination of the neutrino mass ordering is important. It is an important step towards the generalized Standard Model. It is also needed to reduce the uncertainty on δ_{CP}, to understand the requirement for $0\nu\beta\beta$ experiment and to help to understand core-collapse supernova.

JUNO identifies the neutrino mass ordering by comparing the shift of little wiggles with respect to the large oscillation patterns. For the normal ordering, the shift is positive, and for the inverted ordering, the shift is negative. This is obvious when the survival probability is written as

$$P_{ee}(\xi) = a_0(\xi) + a_1(\xi) \cdot \cos\left[1.27 \cdot (2\Delta m_{32} \pm \Delta m^2_\phi) \cdot \xi\right] \tag{9.1}$$

where $\xi = L(\mathrm{m})/E(\mathrm{MeV})$, L is the distance between the detector and the reactor, E is the neutrino energy, a_0 and a_1 are two irrelevant function independent of the neutrino mass ordering and Δ_{32} and ϕ are defined as

$$\Delta_{ij} = 1.27 \cdot |\Delta m^2_{ij}| \cdot \xi \tag{9.2}$$

$$\sin\phi = \frac{c^2_{12}\sin 2\Delta_{21}}{1 - 4s^2_{12}c^2_{12}\sin^2\Delta_{21}} \tag{9.3}$$

$$\cos\phi = \frac{c^2_{12}\cos 2\Delta_{21} + s^2_{12}}{1 - 4s^2_{12}c^2_{12}\sin^2\Delta_{21}} \tag{9.4}$$

$$\phi \equiv 1.27 \cdot \Delta m^2_\phi \cdot \xi \tag{9.5}$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$. So the sensitivity to the neutrino mass ordering is low when ϕ is small. The sensitivity also degenerate when the little wiggles oscillate too fast and is smeared out by the energy resolution of the detector, in which case the detector cannot see the oscillation patterns. The critical region is defined as

$$2\Delta_{32} \cdot \frac{\delta E}{E} = 0.68 \cdot 2\pi \tag{9.6}$$

where 0.68 is determined by simulation[3]. This is illustrated in Figure 9.3. In this figure, the magnitude of Δm^2_ϕ in eV^2 is denoted as the color. The most sensitive region is denoted as red. The black lines correspond to equation 9.6. The region on the left and above the black them cannot contribute to determining the neutrino mass ordering due to detector energy resolution. The solid, dashed and

[2] F. An et al., "Neutrino physics with JUNO", *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, no. 3, p. 030401, 2016, DOI: `10.1088/0954-3899/43/3/030401`, arXiv: `1507.05613`

[3] X. Qian et al., "Mass hierarchy resolution in reactor anti-neutrino experiments: Parameter degeneracies and detector energy response", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 87, no. 3, pp. 1–7, 2013, DOI: `10.1103/PhysRevD.87.033005`, arXiv: `1208.1551`

dotted lines correspond to experiments with energy resolution similar to JUNO, Borexino and KamLAND, respectively. The purple solid line represents the approximate boundary of degenerate mass-squared difference. The Δm^2_ϕ is too small in the region below it.

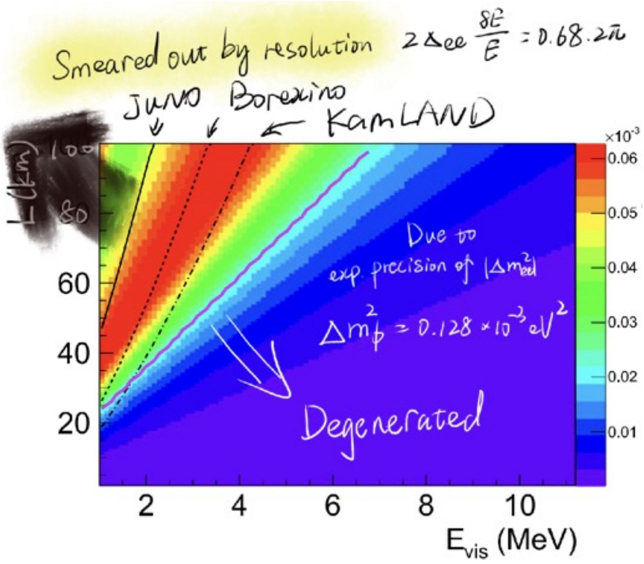


Figure 9.3: Reproduced from Qian et al., 2013. The phase shifts versus energy and distance. Region on the right bottom of the purple solid line provides little sensitivity due to small Δm^2_ϕ. Three black lines (solid, dotted and dashed) correspond to energy resolution of JUNO, Borexino and KamLAND experiments. Region on the left top of the black lines provides little sensitivity due to smearing of energy resolution.

Several statistical methods have been proposed, yet the main idea is to use the chi-square difference to do the hypothesis test:

$$T = \chi^2(\hat{\theta}_{\mathrm{NH}}) - \chi^2(\hat{\theta}_{\mathrm{IH}}) \tag{9.7}$$

where T is the test statistic, $\chi^2(\hat{\theta}_{\mathrm{NH}})$ is the chisquare obtained with fit parameters, such as Δm^2_{ij}, that minimize the chisquare assuming the normal mass ordering, and $\chi^2(\hat{\theta}_{\mathrm{IH}})$ is the chisquare with fit parameters that minimize the chisquare assuming the inverted mass ordering. When data has been collected, calculate p-value of the test statistic value of the corresponding data. It is the confidence level to reject the hypothesis assumed.

$$p = \int_{T(\mathrm{data})}^{+\infty} f(T;\ \mathrm{IH})\mathrm{d}T \tag{9.8}$$

where $f(T;\ \mathrm{IH})$ is the distribution of T assuming Inverted Neutrino Ordering.

When data is not available, we usually use the median sensitivity to express how competitive an experiment is. The median sensitivity is the p-value against one assumption (such as NO) under the median case under another assumption (such as IO). The median value is usually evaluated using the Monte Carlo method. There is a quick way to evaluate the median sensitivity: the median significance, or the number of sigmas, is just the square root of the value of the test statistic of the Asimov dataset[4].

9.2.2 *Precision measurement of neutrinos*

Another important goal of JUNO is to measure neutrino properties precisely. JUNO will be able to measure $\sin^2\theta_{12}$, Δm^2_{21} and $|\Delta m^2_{ee}|$

---

[4] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011, DOI: `10.1140/epjc/s10052-011-1554-0`, arXiv: `1007.1727`

with unprecedented precision to 0.7%, 0.6% and 0.5%, respectively. Besides, JUNO can help in testing the unitarity relation of Equation 9.9 and the mass sum rule of Equation 9.10 to the levels of around 1.2% and 1.8%, respectively.

$$|U_{e1}|^2 + |U_{e2}|^2 + |U_{e3}|^2 = 1 \tag{9.9}$$

$$\Delta m^2_{13} + \Delta m^2_{21} + \Delta m^2_{32} = 0 \tag{9.10}$$

where U_{ei} is the neutrino mixing matrix element.

9.2.3 *Supernova burst neutrinos*

JUNO can collect high statistics detection of neutrinos from a galactic supernova. For a galactic supernova at a distance of 10 kpc, there are around 5000 events in the IBD channel, 2000 events for elastic neutrino–proton scattering, and 300 events for elastic neutrino–electron scattering in the JUNO detector[5]. This provides opportunities to study the explosion mechanism and intrinsic properties of neutrinos themselves.

[5] An et al., "Neutrino physics with JUNO"

9.2.4 *Diffuse supernova neutrino background (DSNB)*

The Diffuse Supernova Neutrino Background is the integrated neutrino flux from all past core-collapse events in the visible universe, and it holds information on the cosmic star-formation rate, the average core-collapse neutrino spectrum and the rate of failed supernovae. It is never detected. It is estimated that JUNO may be able to provide a DSNB detection at the 3σ level. It is detected through the inverse beta decay (IBD) channel. The critical backgrounds include the atmospheric neutral-current background, and it is crucial to develop efficient pulse-shape-discrimination algorithms to suppress them.

9.2.5 *Solar neutrinos*

Like Borexino experiment, JUNO can detect solar neutrinos through elastic scattering on electrons. The main challenge of detecting solar neutrinos is to achieve good radiopurity of liquid scintillator, especially of ^{210}Pb. It is also important to develop the three fold coincidence techniques to suppress the cosmogenic isotope decay events.

JUNO has the potential to measure ^{7}Be solar neutrinos with percent level precision, and to detect ^{8}B solar neutrinos with 2 MeV detection threshold and 2:1 signal to background ratio. The expected precision of the ^{7}Be solar neutrino interaction rate is 7% assuming 0.5% light yield precision and will be dominated by the systematic uncertainty. For CNO solar neutrinos, if the light yield precision can reach 0.1% and the cosmogenic ^{11}C can be suppressed by 80% with the three-fold-coincidence algorithm, the expected precision can reach 20%. These will be discussed in detail in Chapter 10 and Chapter 11.

9.2.6 *Atmospheric neutrinos*

By studying the atmospheric neutrino signal and in particular the different kind of corrections due to the interaction with matter inside the Earth in the two cases of normal and inverted neutrino mass ordering one can exploit an alternative way to discriminate the two possible mass hierarchies. Considering only the atmospheric ν_μ and $\bar{\nu}_\mu$ with track length larger than 5 meters, the sensitivity of JUNO to the mass ordering is 0.9 σ for a 200 kilo-ton-years exposure; the sensitivity to wrong θ_{23} octant is 1.8 σ (0.9 σ) for normal (inverted) ordering; the sensitivity to the Charge-Parity violation is very small. More refined studies of the JUNO potentiality, for what concerns the atmospheric neutrinos are ongoing and they are based on the attempt to reconstruct the charged lepton direction, by looking at the timing and shape pattern of the first PMT hit[6].

9.2.7 *Geo-neutrinos*

Due to its unprecedented size, JUNO can record 300 to 500 geo-neutrino interaction per year. In six months JUNO would match the present world sample of recorded geo-neutrino interaction, which is less than 150 events. Using a well constrained estimate of the reactor signal and reasonable estimates of the non-antineutrino sources, the conclusion of the presented analysis is that geo-neutrinos are indeed observable at JUNO. There is unprecedented opportunities to explore the origin and thermal evolution of the Earth by recording geo-neutrino interactions with JUNO.

9.2.8 *Sterile neutrinos*

JUNO has sensitivity to sterile neutrinos through the collected IBD events from reactor anti-neutrinos. The most sensitive region is $10^{-5} < |\Delta m^2_{41}| < 10^{-2}$ eV^2. Other projects have been planned to extend the covered region of $|\Delta m^2_{41}|$. Placing a 50–100 kCi source of antineutrinos extracted from spent reactor fuel, inside or outside the detector for a 1.5 year run, JUNO would be sensitive to sterile neutrinos with $|\Delta m^2_{41}|$ in the region of 0.1–10 eV^2. Another project that using a ^{8}Li antineutrino source produced from a 60 MeV/amu cyclotron accelerator (IsoDAR@JUNO) located 5 meters away from the detector would provide sensitivity to 1 eV^2-scale light sterile neutrinos.

9.2.9 *Nucleon decays*

JUNO is in an excellent position to search for nucleon decays as a large liquid scintillator detector deep underground. In particular, in the $p \to K^+ + \bar{\nu}$ channel, JUNO will be competitive or complementary with those experiments using water Cherenkov or liquid argon detectors.

$p \to K^+ + \bar{\nu}$ is a two-body decay. If the decaying proton is from hydrogen, then it decays at rest. The kinetic energy of K^+ is fixed to

---

[6] G. Settanta et al., "e-μ Discrimination at High Energy in the JUNO Detector", no. Cc, pp. 3–6, Jan. 2019, arXiv: `1901.10340`

be 105 MeV, and it is the prompt signal. The K^+ meson decays within 12.4 nanoseconds, and around 84% of kaons will decay through

$$K^+ \to \mu^+ \nu_\mu \qquad (63.43\%) \tag{9.11}$$

$$K^+ \to \pi^+ \pi^0 \qquad (21.13\%) \tag{9.12}$$

In both cases, two temporal separated signals will be produced, and they together with the prompt signal forms three fold coincidence. Due to the high efficiency in measuring this two channels, JUNO will surpass Super-Kamiokande in a short time.

9.2.10 Neutrinos from Dark Matter

Dark Matter eventually present in the Sun could annihilate into leptons, quarks and gauge bosons. Neutrinos in the decay products of these particles can be detected by JUNO. The sensitivity of JUNO to spin-dependent DM-nucleon scattering cross section $\sigma_{\chi p}^{\mathrm{SD}}$ is much better than the current direct detection constraints, and for the spin-independent cross section $\sigma_{\chi p}^{\mathrm{SI}}$, JUNO is competitive with direct detection experiments for $m_\chi < 7$ GeV.

10

Two MeV detection threshold ^{8}B solar neutrino

In this chapter, the evaluation of the expected event rate for ^{8}B solar neutrino elastic scattering and the major backgrounds in 2–3 MeV region in the JUNO experiment is presented. At present, the neutrino survival probability in the MSW transition region is not measured yet and one of the unresolved puzzles of neutrino research is that the value of the neutrino mass splitting parameter Δm^2_{21} measured with solar neutrinos differs from the value measured with reactor neutrinos by the KamLAND experiment, at 2 σ level[1]. High precision measurement of active ^{8}B solar neutrino flux may provide an answer. Up to now, Super-Kamiokande[2] and Borexino[3] have independently provided measurement results at a detection threshold of 3.5 MeV and 3 MeV, respectively. However, the sensitivity of the results is limited, as the corresponding neutrino energy is at the high energy boundary of the transition region. In a recent GSSI PhD thesis[4] a measurement with 2.2 MeV detection threshold using Borexino data was reported, but no significant indication was given due to small detector size. In this work I will demonstrate that in the JUNO experiment we can reach 2 MeV detection threshold for ^{8}B solar neutrinos with signal to background ratio of $\mathcal{O}(1)$ in the 2–3 MeV electron recoil energy range. The 2 MeV threshold is set by the high rate of cosmogenic ^{11}C background. The leading reducible backgrounds include cosmogenic ^{10}C and internal contamination by ^{214}Bi and ^{212}Bi.

This chapter is organized as follows: Section 10.1 introduces the physics background and reviews current experiment results on the ^{8}B solar neutrino flux. Section 10.2 presents the estimation of the event rate of signal and backgrounds in the region of interest. The analysis methods to reduce backgrounds are presented in Section 10.3. Section 10.4 presents the results on the sensitivity to the upturn. Section 10.5 is devoted to the conclusions and discussions.

I have co-worked with Vito Antonelli, Emanuela Meroni, Gioacchino Ranucci, Alessandra Re, Barbara Ricci, Giuseppe Salamanna in this work. This work is in synergy with the work of Jie Zhao and Yufeng Li.

[1] P. F. de Salas et al., "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 782, pp. 633–640, 2018, DOI: `10.1016/j.physletb.2018.06.019`, arXiv: `1708.01186` Figure 6

[2] The Super-Kamiokande collaboration, "Solar neutrino measurements in Super-Kamiokande-IV", *Physical Review D*, vol. 94, no. 5, pp. 1–33, 2016, DOI: `10.1103/PhysRevD.94.052010`, arXiv: `1606.07538`

[3] The Borexino collaboration, "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector", *Physical Review D*, vol. 82, no. 3, p. 033006, Aug. 2010, DOI: `10.1103/PhysRevD.82.033006`, arXiv: `0808.2868`; The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure", vol. 016, pp. 1–13, Sept. 2017, arXiv: `1709.00756`

[4] I. Drachnev, "New Spectral Analysis of Solar B Neutrino with the Borexino Detector", PhD thesis, Gran Sasso Science Institute, 2016

10.1 Introduction

Observation of deficit of measured solar neutrino flux in the Homestake experiment[5,6] has triggered vast experimental and theoretical development in neutrino physics which lasted for 50 years, until nowadays. Interested readers can refer to review articles such as M. Maltoni and A. Yu. Smirnov, "Solar neutrinos and neutrino physics", *The European Physical Journal A*, vol. 52, no. 4, p. 87, Apr. 2016, DOI: `10.1140/epja/i2016-16087-0`, arXiv: `1507.05287`, W. Haxton, R. Hamish Robertson, and A. M. Serenelli, "Solar Neutrinos: Status and Prospects", *Annual Review of Astronomy and Astrophysics*, vol. 51, no. 1, pp. 21–61, Aug. 2013, DOI: `10.1146/annurev-astro-081811-125539`, arXiv: `1208.5723`, A. B. McDonald, "Nobel Lecture: The Sudbury Neutrino Observatory: Observation of flavor change for solar neutrinos", *Reviews of Modern Physics*, vol. 88, no. 3, p. 030502, July 2016, DOI: `10.1103/RevModPhys.88.030502` and V. Antonelli et al., "Solar Neutrinos", *Advances in High Energy Physics*, vol. 2013, no. supp01a, pp. 1–34, 2013, DOI: `10.1155/2013/351926`. With reactor neutrinos, the KamLAND experiment[7] rejected other solutions except the large mixing angle solution[8] to the solar neutrino problem. Nowadays, we know that flavor of neutrinos varies during propagation due to the mixing of neutrino flavor eigenstates and mass eigenstates[9,10,11,12].

At present, the neutrino survival probability in the MSW transition region is not measured yet and one of the unresolved puzzles of neutrino research is that the value of the neutrino mass splitting parameter Δm^2_{21} measured with solar neutrinos differs from the value measured with reactor neutrinos by the KamLAND experiment, at 2 σ level[13].The KamLAND experiment measured Δm^2_{21} according to the period of neutrino vacuum oscillation in L/E, where L is the distance from the detector to the neutrino source and E is the neutrino energy. In solar neutrino experiments, instead, it is measured in presence of matter effect. When neutrinos are produced or detected in medium, the neutrino mass eigenstates are modified due to coherent forward scattering of electron neutrinos on matters, and the survival probability depend on both the mixing angle at the production point and the detection point, and this is called the Mikheyev–Smirnov–Wolfenstein effect[14]. The solar electron neutrino survival probability is given by

$$P^{\text{day}}_{ee} = c^2_{13} c^{m2}_{13} P^{\text{ad}}_2 + s^2_{13} s^{m2}_{13} \tag{10.1}$$

$$P^{\text{ad}}_2 = \frac{1}{2}\left(1 + \cos 2\theta_{12} \cos 2\theta^m_{12}\right) \tag{10.2}$$

where c_{13}, s_{13} are the cosine and sine of neutrino mixing angle at the detection point (vacuum), respectively, and superscript m denotes the mixing angle at the region in the Sun where neutrinos are produced.

---

[5] R. Davis, D. S. Harmer, and K. C. Hoffman, "Search for neutrinos from the sun", *Physical Review Letters*, vol. 20, no. 21, pp. 1205–1209, 1968, DOI: `10.1103/PhysRevLett.20.1205`, arXiv: `arXiv:1011.1669v3`

[6] B. T. Cleveland et al., "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector", *The Astrophysical Journal*, vol. 496, no. 1, pp. 505–526, 1998, DOI: `10.1086/305343`

[7] The KamLAND collaboration, "First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance", *Physical Review Letters*, vol. 90, no. 2, p. 6, 2003, DOI: `10.1103/PhysRevLett.90.021802`, arXiv: `0212021 [hep-ex]`

[8] G. L. Fogli et al., "Getting the most from the statistical analysis of solar neutrino oscillations", *Physical Review D*, vol. 66, no. 5, pp. 1–30, 2002, DOI: `10.1103/PhysRevD.66.053010`, arXiv: `0206162 [hep-ph]`

[9] The Super-Kamiokande collaboration, "Solar neutrino measurements in Super-Kamiokande-IV"

[10] The KamLAND collaboration, "Reactor on-off antineutrino measurement with KamLAND", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 88, no. 3, pp. 1–10, 2013, DOI: `10.1103/PhysRevD.88.033001`, arXiv: `1303.4667 [hep-ex]`

[11] The T2K collaboration, "Measurement of neutrino and antineutrino oscillations by the T2K experiment including a new additional sample of veinteractions at the far detector", *Physical Review D*, vol. 96, no. 9, pp. 1–49, 2017, DOI: `10.1103/PhysRevD.96.092006`, arXiv: `1707.01048`

[12] The Daya Bay collaboration, "Measurement of electron antineutrino oscillation based on 1230 days of operation of the Daya Bay experiment", *Physical Review D*, vol. 95, no. 7, p. 072006, Apr. 2017, DOI: `10.1103/PhysRevD.95.072006`

[13] Salas et al., "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity" Figure 6

[14] L. Wolfenstein, "Neutrino oscillations in matter", *Physical Review D*, vol. 17, no. 9, pp. 2369–2374, May 1978, DOI: `10.1103/PhysRevD.17.2369`; S. P. Mikheyev and A. Y. Smirnov, "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos", *Sov. J. Nucl. Phys.* Vol. 42, pp. 913–917, 1985

The comparison of allowed band of P_{ee} with Δm^2_{21} corresponding to the value from KamLAND[15] and that from global fit of solar neutrino experiments[16] is shown in Figure 10.1. The measured P_{ee} from solar neutrinos in Borexino[17] are also shown as black dots in the figure. As can be seen solar neutrino experiments shows the upturn, though less than that obtained from KamLAND, and by lowering the current detection threshold of ^{8}B solar neutrino flux measurements, the sensitivity to such upturn will be significantly improved.

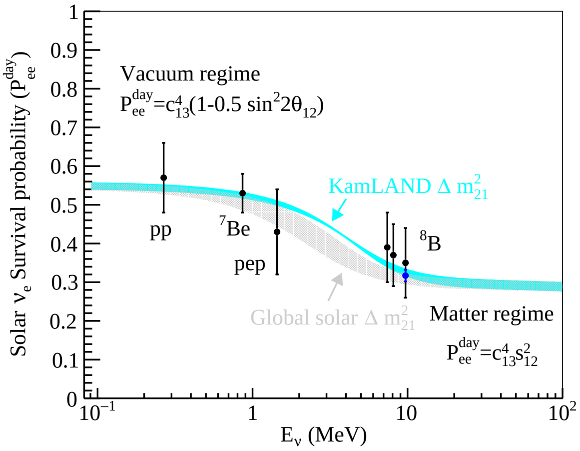


Figure 10.1: The solar neutrino survival probability. 1 σ allowed band based on Δm^2_{21} measured by KamLAND using reactor neutrinos (Cyan) and that measured by solar neutrino experiments (Gray) are shown. The measured survival probability are also shown on the graph. Black dots are from the Borexino experiment. Blue dot is the joint analysis between the Super-Kamiokande and SNO experiment.

Up to now, Borexino[18,19] have provided measurement results of ^{8}B solar neutrino neutrino flux at the lowest detection threshold, 3.2 MeV, and the corresponding average neutrino energy is 8.7 MeV. In a recent GSSI PhD thesis[20] a measurement with 2.2 MeV detection threshold using Borexino data has been reported, however, due to the limited exposure, no significant indication can be given and it is only demonstrating the methodology needed for future experiments. In this work I will demonstrate that in the JUNO experiment we can reach 2 MeV detection threshold for ^{8}B solar neutrino with signal to background ratio of $\mathcal{O}(1)$ in the 2–3 MeV electron recoil energy range.

[15] The KamLAND collaboration, "Reactor on-off antineutrino measurement with KamLAND"

[16] The Super-Kamiokande collaboration, "Solar neutrino measurements in Super-Kamiokande-IV"

[17] The Borexino collaboration, "Comprehensive measurement of pp-chain solar neutrinos", *Nature*, vol. 562, no. 7728, pp. 505–510, Oct. 2018, DOI: `10.1038/s41586-018-0624-y`

[18] The Borexino collaboration, "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector"

[19] The Borexino collaboration, "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure"

[20] Drachnev, "New Spectral Analysis of Solar B Neutrino with the Borexino Detector"

10.2 *Signal and backgrounds rates*

In this section, an estimation of the event rate of the signal process and various backgrounds, including the external γs from radioactive decays and neutron captures, internal natural radioactive decays, cosmogenic isotope decays and other backgrounds, in an region of interest of 2–3 MeV without background reduction is presented. The reduction method and its efficiency will be presented in the next section. In this analysis, the elastic scattering between solar neutrinos and electrons in the target is considered as the signal process.

10.2.1 *The ^{8}B solar neutrino elastic scattering signal*

In this section, first the solar neutrino energy spectrum and rate at the Sun is reported, then the solar electron neutrino survival probability is presented, then the cross-section of the elastic scattering process is reviewed, and at last the visible energy spectrum, which includes the effect of the detector energy response non-linearity and the smearing due to detector energy resolution, is reported.

The neutrino energy spectrum The ^{8}B solar neutrino is produced during the *pp* fusion through the *pp* chain in the core region of the sun. Its energy spectrum[21] is shown in Figure 10.2. Earlier prediction of this spectrum[22] is also shown in the figure for comparison.

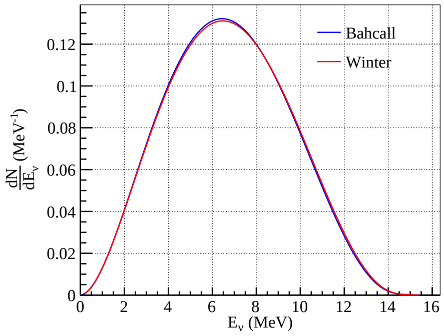


Figure 10.2: Comparison between two ^{8}B solar neutrino energy spectra. Details see text.

According to the Standard Solar Model[23], the ^{8}B solar neutrino flux is predicted to be $5.46(1 \pm 0.12) \times 10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ according to the high metallicity model (B16-GS98) and $4.50(1 \pm 0.12) \times 10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ according to the low metallicity model (B16-AGSS09met). In this work the high metallicity is used as a reference, and it is expected that the results are similar using the low metallicity model.

The neutrino survival probability Flavor of neutrinos varies during propagation due to mixing of neutrino flavor eigenstates and mass eigenstates. For solar neutrinos, because they propagate through ultra high density of electrons before leaving the Sun, the mass eigenstate is modified due to forward scattering of electron neutrino on matters[24]. The survival probability of solar neutrinos can be described with Equation (10.1)–(10.2)[25]. The neutrino vacuum oscilla-

[21] W. T. Winter et al., "The B8 neutrino spectrum", *Physical Review C - Nuclear Physics*, vol. 73, no. 2, pp. 1–15, 2006, DOI: `10.1103/PhysRevC.73.025503`, arXiv: `0406019 [nucl-ex]`, Table IV

[22] J. N. Bahcall, M. Kamionkowski, and A. Sirlin, "Solar neutrinos: Radiative corrections in neutrino-electron scattering experiments", *Physical Review D*, vol. 51, no. 11, pp. 6146–6158, June 1995, DOI: `10.1103/PhysRevD.51.6146`, arXiv: `arXiv:1011.1669v3`, Table IV

[23] N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: `10.3847/1538-4357/835/2/202`, arXiv: `1611.09867` Table 6

[24] Wolfenstein, "Neutrino oscillations in matter"; Mikheyev and Smirnov, "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos"

[25] Maltoni and Yu. Smirnov, "Solar neutrinos and neutrino physics" Equation (24)

tion parameters, including the mass splitting and mixing angles, are taken from result of global fit[26]. The average value for $n_e^{\odot}(^8\mathrm{B}\,\nu)$, in the solar region of ^{7}Be solar neutrino production, is taken from Vissani[27] as 89.6 mol/cm^3. During the night when solar neutrinos propagate through the Earth, the electron neutrinos are regenerated. The regeneration effect caused the day-night asymmetry of detected solar electron interaction rates and is described in the following equation:

$$P_{ee}^{\text{night}} = P_{ee}^{\text{day}} + c_{13}^2 \cos 2\theta_{12}^m F_{\text{reg}} \tag{10.3}$$

The expression of F_{reg} in [28] is used. The survival probability is shown in Figure 10.3.

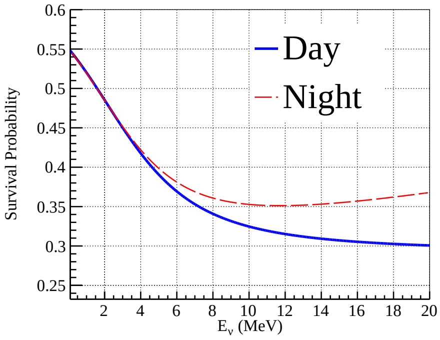


Figure 10.3: The survival probability of ^{7}Be solar neutrinos. The day-night effect is also shown. Black solid line: P_{ee} in the day. Red dashed line: P_{ee} at night.

Elastic scattering cross-section The cross-section of elastic scattering of neutrinos on electrons with radiative correction[29] is

$$\frac{\mathrm{d}\sigma^{\text{ES}}}{\mathrm{d}T_e} = \frac{\sigma_0}{m_e} \cdot \left\{ g_L^2(T)\left[1+\frac{\alpha}{\pi}f_-(z)\right] + g_R^2(T)\,(1-z)^2 \cdot \left[1+\frac{\alpha}{\pi}f_-(z)\right] \right. \\ \left. - g_L(T)g_R(T)\frac{m_e}{E_\nu}z\left[1+\frac{\alpha}{\pi}f_\pm(z)\right] \right\} \tag{10.4}$$

where $z = \dfrac{T}{E_\nu}$, T is the kinetic energy of the recoil electron, and the correction functions are described in [30]. The differential cross-section of 6.3 MeV neutrinos is shown in Figure 10.4.

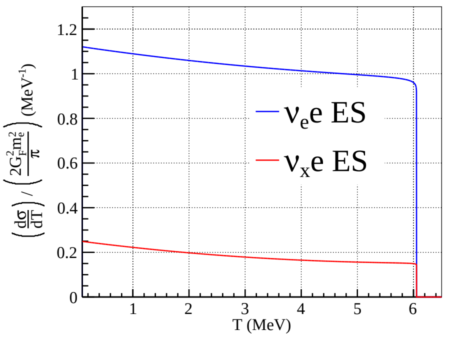


Figure 10.4: The cross-sections of elastic scattering of electron neutrinos (denoted as ν_e) and μ or τ neutrinos (denoted as ν_x) on electrons. T is the kinetic energy of the recoil electron.

[26] Salas et al., "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity" Table 1

[27] F. Vissani, "Joint analysis of Borexino and SNO solar neutrino data and reconstruction of the survival probability", pp. 1–12, Sept. 2017, arXiv: `1709.05813`, Equation (5).

[28] F. Vissani, "Solar neutrino physics on the beginning of 2017", *Nuclear Physics and Atomic Energy*, vol. 18, no. 1, pp. 5–12, Mar. 2017, DOI: `10.15407/jnpae2017.01.005`

[29] Bahcall, Kamionkowski, and Sirlin, "Solar neutrinos: Radiative corrections in neutrino-electron scattering experiments"

[30] ibid. Appendix A

Recoiled electron spectrum The recoiled electron energy spectrum is given by

$$\frac{\mathrm{d}^2N^{\mathrm{ES}}}{\mathrm{d}T\mathrm{d}t} = \int_{E_\nu^{\mathrm{min}}}^{E_\nu^{\mathrm{max}}} \phi_\nu(E_\nu)\cdot \mathrm{d}E_\nu \cdot (P_{ee}\cdot \sigma^{\nu_e} + (1-P_{ee})\cdot \sigma^{\nu_{\mu,\tau}})\cdot \frac{\mathrm{d}\sigma^{\mathrm{ES}}}{\mathrm{d}T}\cdot N_e\cdot \varepsilon_{\mathrm{det.}}$$

where E_ν is the neutrino energy, T is the kinetic energy of the recoil electron, t is the live time, $\varepsilon_{\mathrm{det.}}$ is the detection efficiency, E_ν^{min} is the lowest neutrino energy to induce an recoiled electron with kinetic energy T and it is given by $E_\nu^{\mathrm{min}} = T\cdot\left(\frac{1}{2}+\frac{1}{2}\sqrt{1+\frac{2m_e}{T}}\right)$, E_ν^{max} is the end point of ^{8}B solar neutrino, which is taken as 16 MeV, N_e is the number of electrons. For N_e we have

$$N_e = m\cdot\frac{(1-\epsilon_{1\mathrm{H}})\cdot 0.5 + \epsilon_{1\mathrm{H}}\cdot 1}{1\ \mathrm{g/mol}}\cdot N_A \tag{10.5}$$

where N_A is the Avogadro's constant, $\epsilon_{1\mathrm{H}}$ is the mass fraction of hydrogen atoms, and it is 0.1201 for JUNO LAB[31], m is the target mass and for the JUNO detector it is 20 kilo-tons, so

$$\frac{N_e}{m} = 3.373\times 10^{32}\ (\text{kilo-tons})^{-1} \tag{10.6}$$

Final visible energy spectrum Finally the detector resolution is considered by convolving the recoiled electron energy spectrum with a Gaussian function:

$$\frac{\mathrm{d}N^{\mathrm{ES}}}{\mathrm{d}E^e_{\mathrm{vis}}} = \int_0^{T^{\mathrm{max}}} \frac{\mathrm{d}N^{\mathrm{ES}}}{\mathrm{d}T}\cdot f_{\mathrm{det.}}\left(E_{\mathrm{vis.}};\mu_{\mathrm{vis.}},\sigma_{\mathrm{vis.}}\right)\cdot \mathrm{d}T \tag{10.7}$$

$$\sigma_{\mathrm{vis.}} = \sqrt{p_0^2\cdot\mu^2_{\mathrm{vis.}} + p_1^2\cdot\mu_{\mathrm{vis.}} + p_2^2} \tag{10.8}$$

where $f_{\mathrm{det.}}$ is the detector response taken as the Gaussian distribution with expectation $\mu_{\mathrm{vis.}}$ and width $\sigma_{\mathrm{vis.}}$. $(p_0, p_1, p_2) = (1, 2.8, 0.4)\%$. The final visible energy spectrum of ^{8}B solar neutrino elastic scattering interaction on electrons is shown in Figure 10.5.

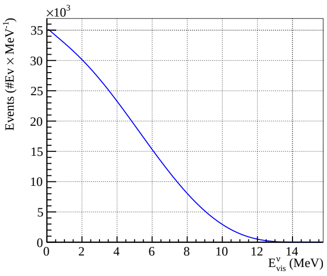


Figure 10.5: Visible energy spectrum of ^{8}B solar neutrino elastic scattering interaction on electrons.

Region of interests The region of interest is defined as 2–3 MeV in this work. The number of events in the region of interests is given by

$$\frac{N^{\mathrm{ROI}}_{\mathrm{sig}}}{t} = \int_{E^{\mathrm{ROI}}_{\mathrm{min}}}^{E^{\mathrm{ROI}}_{\mathrm{max}}} \frac{\mathrm{d}^2N^{\mathrm{ES}}}{\mathrm{d}E^e_{\mathrm{vis}}\mathrm{d}t}\mathrm{d}E^e_{\mathrm{vis}} = 90.55\,\text{per day} \tag{10.9}$$

The total event rate and ROI event rate of ^{8}B solar neutrino elastic scattering signal are summarized in Table 10.1.

Table 10.1: ^{8}B solar neutrino ES signals

X	$R^{\mathrm{tot}}_{\mathrm{X}}$ (day^{-1})	$R^{\mathrm{ROI}}_{\mathrm{X}}$ (day^{-1})
^{8}B ν ES	90.55	13.23

10.2.2 *External γs*

The γ produced during the natural radioactive decay in external components, especially ^{208}Tl decay γs in the acrylic vessel, can go

[31] F. An et al., "Neutrino physics with JUNO", *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, no. 3, p. 030401, 2016, DOI: `10.1088/0954-3899/43/3/030401`, arXiv: `1507.05613` Table A3

into the liquid scintillator and produce lights, and this becomes a strong backgrounds. Such background can be reduced to negligible level with a tight fiducial volume cut and will be discussed in Section 10.3.

As estimated in Li et al.,[32] the total event rate of radioactive decays from external components above 0.7 MeV is 8.36×10^6 per day. The energy spectra without fiducial volume cut and with a 17 m cut are shown in Figure 10.6. From the figure we can see that in the region of interest the nH capture 2.2 MeV γ peak and ^{208}Tl 2.6 MeV γ peak are critical backgrounds. Their contributions to the ROI are summarized in Table 10.2.

Table 10.2: External backgrounds

X	R_X^{tot} (day^{-1})	R_X^{ROI} (day^{-1})
^{8}B ν ES	90.55	13.23
External γs	3.333×10^7	9.105×10^5

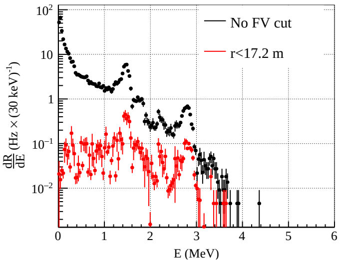


Figure 10.6: Energy spectrum of external γs.

10.2.3 *External neutrons*

Radiogenic neutrons produced in external components through (α, n) reactions or spontaneous fission reactions can be captured on Hydrogen and the 2.2 MeV γs produced in the neutron capture process will contribute as backgrounds in ROI. The major contribution, similar to the external natural radioactive decay γs, comes from the acrylic vessel. The ^{238}U and ^{232}Th contamination of the acrylic vessel are taken from Li et al.[33] and are reproduced in Table 10.3. The ^{235}U contamination is obtained from ^{238}U imposing the natural isotopic ratio. The number of (α, n) neutrons per decay in the acrylic vessel is evaluated for each decay chain with NEUCBOT[34], a tool based on the TALYS simulation package[35].

Table 10.3: Neutron flux from (α, n) reaction and spontaneous fissions and event rates without fiducial volume cut originated from the acrylic sphere (560 ton).

Type	^{238}U	^{235}U	^{232}Th
Concentration (g/g)	10^{-11}	10^{-13}	10^{-11}
(α, n) rate (n/decay)	1.2×10^{-6}	1.2×10^{-6}	1.4×10^{-6}
(α, n) n flux (per day)	7	0.5	3
SF rate $(n \cdot g^{-1} \cdot s^{-1})$	1.36×10^{-2}	3×10^{-4}	$< 1.32 \times 10^{-7}$
SF n flux (per day)	7	< 0.1	< 0.1

The total neutron flux is estimated to be $\mathcal{O}(10)$ per day, so the neutron capture γ event rate is less than $\mathcal{O}(10)$ per day and is negligible compared with external radioactive decay γs. Their contributions to the ROI are summarized in Table 10.4.

Table 10.4: Contribution of $(\alpha,\, n)$ backgrounds to the region of interests.

X	R_X^{tot} (day^{-1})	R_X^{ROI} (day^{-1})
^{8}B ν ES	90.55	13.23
$(\alpha,\, n)$	$\mathcal{O}(10)$	$\mathcal{O}(10)$

10.2.4 *Internal natural radioactive decays*

Decays of radioactive elements dissolved in the liquid scintillator contribute as backgrounds in ROI. In this section, the ^{238}U chain and

---

[32] X.-y. Li et al., "Simulation of natural radioactivity backgrounds in the JUNO central detector", *Chinese Physics C*, vol. 40, no. 2, p. 026001, Feb. 2016, DOI: `10.1088/1674-1137/40/2/026001`, Table 2 and 5.

[33] Ibid., Table 1.

[34] S. Westerdale and P. D. Meyers, "Radiogenic neutron yield calculations for low-background experiments", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 875, no. August, pp. 57–64, 2017, DOI: `10.1016/j.nima.2017.09.007`, arXiv: `1702.02465` See also Github repository shawest/neucbot

[35] See TALYS website.

the ^{232}Th chain are considered. The low energy backgrounds, including ^{40}K, ^{85}Kr, ^{39}Ar and ^{210}Pb whose energy range is below ROI, are also discussed.

The ^{238}U chain and ^{232}Th chain are depicted in Figure 10.8 and 10.9, respectively[36]. Both decay chains are assumed to be in secular equilibrium for simplicity. The feeding rate, or the decay rate of the ^{238}U or ^{232}Th isotopes, is proportional to its concentration and anti-proportional to its lifetime. When the chain is in secular equilibrium, the total event rate is the product of the feeding rate and the number of un-correlated decays. Two sequential decays are considered as un-correlated unless the decay time of the second decay is short enough to be comparable with the 1250 ns event gate. The half-life, reference concentration assumed and the number of un-correlated decays used here and the evaluated feeding rate and the total rate are summarized in Table 10.5. Here the ^{238}U and ^{232}Th concentrations are assumed to be 10^{-17} g/g, which is the "ideal" case in [37] and it is not easy to achieve. As a comparison, the requirement on it in order to determine the neutrino mass ordering is 10^{-15} g/g. The impact of high concentration of ^{238}U and ^{232}Th will be discussed in Section 10.5 on page 182.

Type Unit	$\tau_{1/2}$ years	ref. ϵ_X g/g	$N^{\text{un-corr.}}$	R^{feed} $\text{day}^{-1}\cdot\text{kt}^{-1}$	R^{tot} $\text{day}^{-1}\cdot\text{kt}^{-1}$
^{238}U	4.468×10^9	10^{-17}	14	10.75	150.5
^{232}Th	4.468×10^9	10^{-17}	9.359	3.506	32.81
^{40}K	1.251×10^9	10^{-18}	1	22.84	22.84
^{210}Pb	22.3	10^{-24}	1	234.6	234.6
^{85}Kr					500
^{39}Ar					5

Table 10.5: The parameters used to estimate the internal (liquid scintillator) backgrounds event rates.

Their energy spectra are shown in Figure 10.7. The spectra are normalized to the expected event rates.

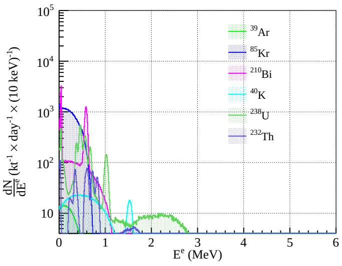


Figure 10.7: Visible energy spectrum of natural radioactive decay backgrounds in the liquid scintillator. Spectra have been normalized to the expected rates.

The corresponding contributions to the ROI are shown in Table 10.6.

X	R_X^{tot} (day^{-1})	R_X^{ROI} (day^{-1})
^{8}B ν ES	90.55	13.23
^{238}U	3009	132.4
^{232}Th	656.3	24.44
other decays	$\mathcal{O}(10^4)$	0

Table 10.6: Contribution of natural radioactive decay backgrounds in the liquid scintillator to ROI.

[36] Malling, D. C. and Fiorucci, S. and Pangilinan, M. and Chapman, J. J. and Faham, C. H. and Verbus, J. R. and Gaitskell, R. J., "Dark Matter Search Backgrounds from Primordial Radionuclide Chain Disequilibrium" May 2013, arXiv: `1305.5183`

[37] An et al., "Neutrino physics with JUNO"

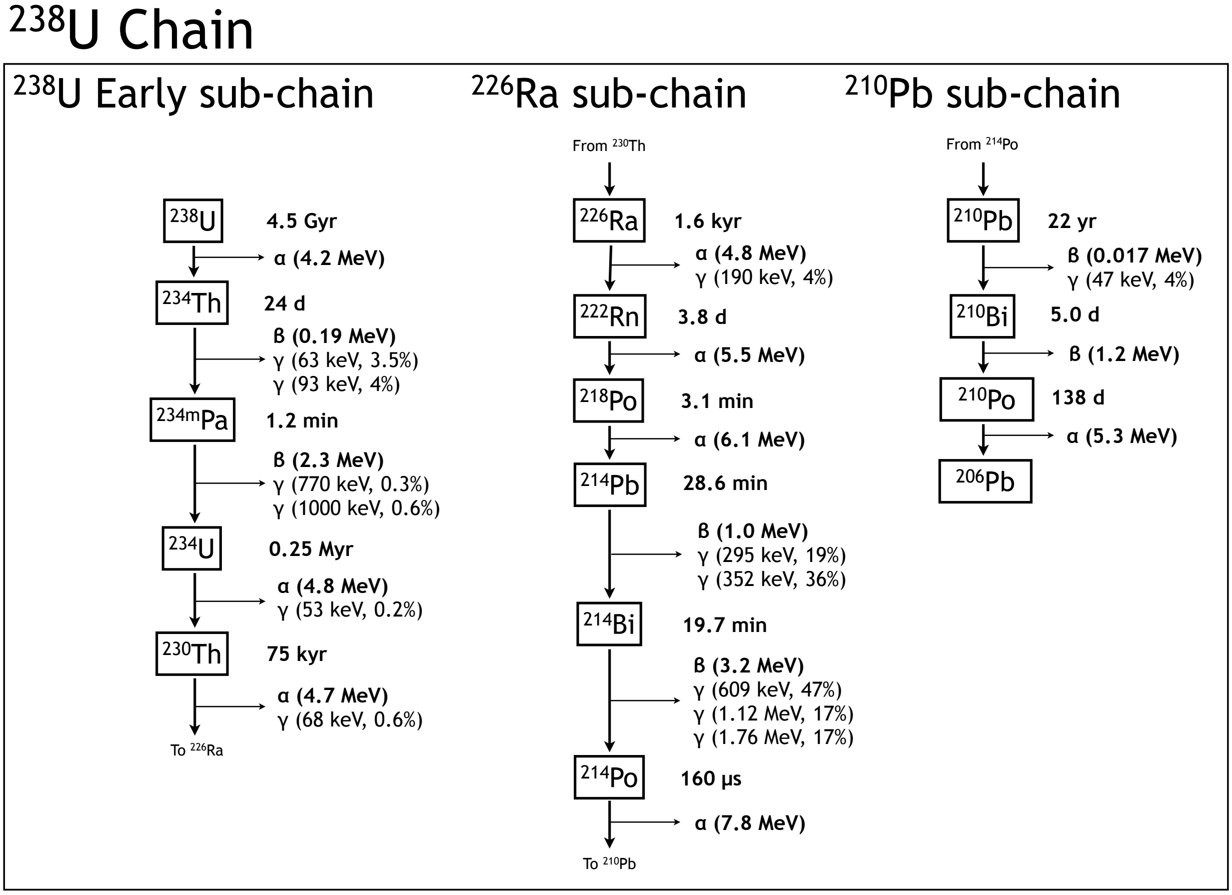


Figure 10.8: ^{238}U chain with Q values and half lives. The figure is reproduced from Malling, D. C. and Fiorucci, S. and Pangilinan, M. and Chapman, J. J. and Faham, C. H. and Verbus, J. R. and Gaitskell, R. J., "Dark Matter Search Backgrounds from Primordial Radionuclide Chain Disequilibrium" May 2013, arXiv: `1305.5183`.

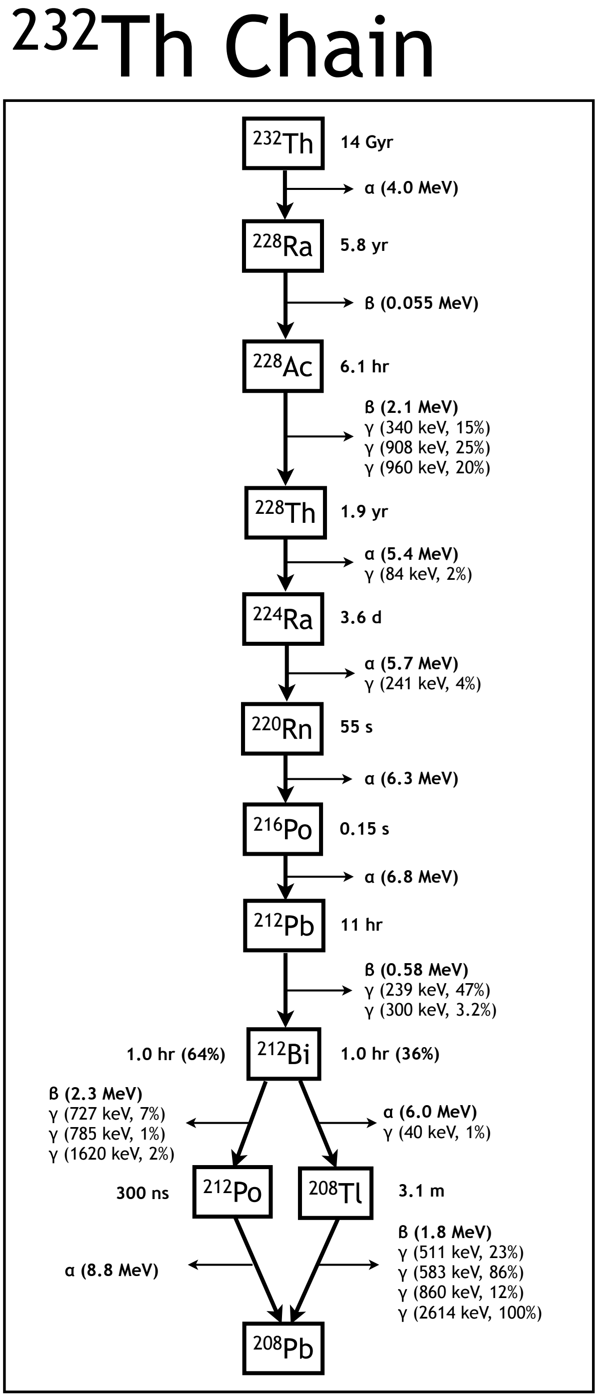


Figure 10.9: ^{232}Th decay chain with *Q* values and half lives. The figure is reproduced from Malling, D. C. and Fiorucci, S. and Pangilinan, M. and Chapman, J. J. and Faham, C. H. and Verbus, J. R. and Gaitskell, R. J., "Dark Matter Search Backgrounds from Primordial Radionuclide Chain Disequilibrium" May 2013, arXiv: `1305.5183`.

10.2.5 Cosmogenic isotopes

The cosmogenic isotopes considered in this section include

1. Long lifetime (> 10 s): ^{10}C, ^{11}Be and ^{16}N
2. 1–1.5 s lifetime: ^{6}He, ^{8}Li and ^{8}B
3. < 300 ms lifetime: ^{9}Li, ^{9}C, ^{8}He, ^{12}B, ^{13}B and ^{12}N
4. Low energy: ^{11}C, ^{7}Be and ^{10}Be.
5. Other low rate cosmogenic isotopes: ^{3}H, ^{11}Li, ^{12}Be, ^{13}B, ^{13}N, ^{14}B, ^{14}C, ^{15}C, ^{16}C, ^{17}N and ^{18}N
6. neutrons

Expected rate The cosmogenic isotope production rate can be predicted using the following scaling formula with respect to a reference experimental measurement:

$$\frac{R_X^{\mathrm{JUNO}}}{R_X^{\mathrm{ref}}} = \frac{\left(E_\mu^{\mathrm{JUNO}}\right)^{\alpha(X)}}{\left(E_\mu^{\mathrm{ref}}\right)^{\alpha(X)}} \cdot \frac{\epsilon_C^{\mathrm{JUNO}}}{\epsilon_C^{\mathrm{ref}}} \cdot \frac{\rho^{\mathrm{JUNO}}}{\rho^{\mathrm{ref}}} \cdot \frac{R_\mu^{\mathrm{JUNO}}}{R_\mu^{\mathrm{ref}}} \cdot \frac{L_\mu^{\mathrm{JUNO}}}{L_\mu^{\mathrm{ref}}} \tag{10.10}$$

where the first term comes from the fact that the cosmogenic production yield is higher when the average muon energy is higher (harder spectrum), and the dependence is verified to be following the power law[38], α is the power law parameter[39], E_μ is the average muon energy, ϵ_C is the mass fraction of Carbon atoms, ρ is the density of the target, R_μ is the muon rate, L_μ is the average μ track length. Their values for JUNO[40], KamLAND[41] and Borexino[42] are summarized in Table 10.7. The average muon energy of two experiments differ because the depths of two two experiments are different. The predicted rates using such method are summarized in Table 10.8.

Experiment	E_μ	ϵ_C	ρ	R_μ	L_μ	m
Unit	GeV		g/cm^3	Hz	m	kton
JUNO	215	0.8792	0.856	3.0	23.6	20
KamLAND	260	0.8568	0.780	0.198	8.78	0.913
Borexino	283	0.9007	0.88	0.00965	4.0	0.0995

Table 10.7: Parameters for predicting cosmogenic isotope production rates using the scaling law

The production rates can also be predicted by FLUKA and Geant4 simulation. According to The Borexino collaboration,[43] the yield predicted by simulation can be biased by up to 50%. We can apply a correction on the prediction for JUNO using the ratio between the measured and simulated yield at the Borexino site according to The Borexino collaboration:[44]

$$R_X^{\mathrm{JUNO}} = R_X^{\mathrm{JUNO}}(\mathrm{MC}) \cdot \frac{R_X^{\mathrm{BX}}(\mathrm{data})}{R_X^{\mathrm{BX}}(\mathrm{MC})} \tag{10.11}$$

where $R_X^Y(Z)$ represents the production rate of isotope X at site Y from Z (experimental measurement or Monte Carlo). The predicted rates by scaling Monte Carlo are also summarized in Table 10.8.

---

[38] The KamLAND collaboration, "Production of radioactive isotopes through cosmic muon spallation in KamLAND", *Physical Review C*, vol. 81, no. 2, p. 025807, Feb. 2010, DOI: `10.1103/PhysRevC.81.025807`, arXiv: `arXiv:0907.0066v2`

[39] ibid. Table IV

[40] An et al., "Neutrino physics with JUNO"

[41] The KamLAND collaboration, "Production of radioactive isotopes through cosmic muon spallation in KamLAND"

[42] The Borexino collaboration, "Cosmogenic backgrounds in borexino at 3800 m water-equivalent depth", *Journal of Cosmology and Astroparticle Physics*, vol. 2013, no. 8 2013, DOI: `10.1088/1475-7516/2013/08/049`, arXiv: `1304.7381`

[43] The Borexino collaboration, "Cosmogenic backgrounds in borexino at 3800 m water-equivalent depth", *Journal of Cosmology and Astroparticle Physics*, vol. 2013, no. 8 2013, DOI: `10.1088/1475-7516/2013/08/049`, arXiv: `1304.7381`, Table 4.

[44] Ibid., Table 4.

At last, the rate predicted by scaling KamLAND data[45] is used, except ^{6}He, ^{11}C and neutron. When evaluating the rate of ^{6}He, ^{11}C and neutron, Borexino[46] is considered as the reference experiment due to better precision. For isotopes without experimental measurements, the rate predicted by FLUKA[47] is used.

Isotope	$R^{\mathrm{BX}}_{\mathrm{scaled}}$	$R^{\mathrm{KL}}_{\mathrm{scaled}}$	$R^{\mathrm{FL}}_{\mathrm{scaled}}$	$R^{\mathrm{G4}}_{\mathrm{scaled}}$	used
Unit	day^{-1}	day^{-1}	day^{-1}	day^{-1}	day^{-1}
^{10}C	$(7 \pm 2) \times 10^2$	$(7.6 \pm 0.9) \times 10^2$	$(6.4 \pm 0.8) \times 10^2$	$(7.4 \pm 0.9) \times 10^2$	760.4
^{11}Be	$< 2.9 \times 10^2$	$(5.1 \pm 0.9) \times 10$	$(5 \pm 2) \times 10$	$(3 \pm 2) \times 10$	51.2
^{16}N	-	-	-	-	13
^{6}He	$(1.6 \pm 0.6) \times 10^2$	-	$(1.2 \pm 0.5) \times 10^3$	$(7 \pm 4) \times 10^2$	1543
^{8}Li	$(3 \pm 3) \times 10^2$	$(6 \pm 1) \times 10^2$	$(4.5 \pm 1.0) \times 10^2$	$(2.6 \pm 0.6) \times 10^2$	560.2
^{8}B	$(6 \pm 3) \times 10^2$	$(4 \pm 1) \times 10^2$	$(3.2 \pm 1.0) \times 10^2$	$(4 \pm 2) \times 10^2$	387.2
^{9}Li	$(1.2 \pm 0.2) \times 10^2$	$(1.02 \pm 0.09) \times 10^2$	$(7 \pm 2) \times 10$	$(2.1 \pm 0.4) \times 10^2$	101.4
^{9}C	$< 6.6 \times 10^2$	$(1.4 \pm 0.6) \times 10^2$	$(1.1 \pm 0.5) \times 10^2$	$(3 \pm 2) \times 10^2$	139.0
^{8}He	< 59	$(4 \pm 2) \times 10$	$(2 \pm 3) \times 10$	< 82	31.83
^{12}B	$(2.3 \pm 0.2) \times 10^3$	$(2.0 \pm 0.2) \times 10^3$	$(1.6 \pm 0.2) \times 10^3$	$(1.3 \pm 0.2) \times 10^3$	1968
^{13}B	-	-	-	-	12
^{12}N	< 44	$(8 \pm 2) \times 10$	$(6 \pm 4) \times 10$	$(9 \pm 3) \times 10$	81.34
^{11}C	$(3.7 \pm 0.5) \times 10^4$	$(4.1 \pm 0.7) \times 10^4$	$(3.1 \pm 0.5) \times 10^4$	$(2.5 \pm 0.4) \times 10^4$	37344
^{7}Be	-	-	-	-	5438
^{10}Be	-	-	-	-	1419
neutron	$(1.28 \pm 0.05) \times 10^5$	$(1.3 \pm 0.2) \times 10^5$	$(2.0 \pm 0.2) \times 10^5$	$(1.87 \pm 0.08) \times 10^5$	127975

Table 10.8: Predicted cosmogenic rates by different scaling methods. Unit: per day

The spectra The spectra of considered cosmogenic isotopes are shown in Figure 10.10 and are normalized to the predicted rates.

---

[45] The KamLAND collaboration, "Production of radioactive isotopes through cosmic muon spallation in KamLAND"

[46] The Borexino collaboration, "Cosmogenic backgrounds in borexino at 3800 m water-equivalent depth"

[47] An et al., "Neutrino physics with JUNO"

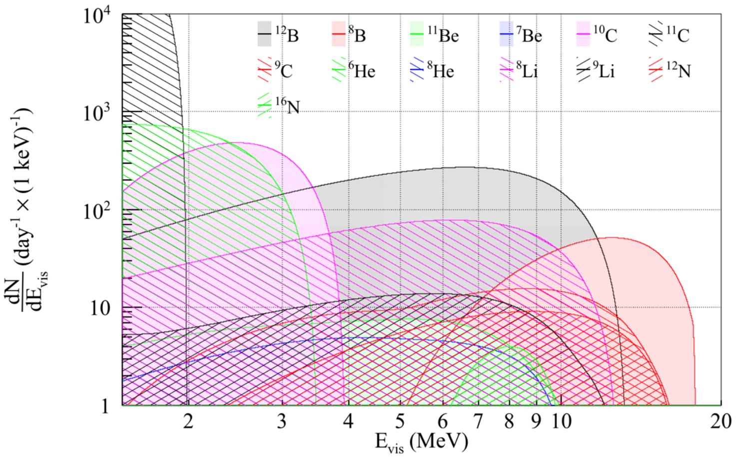


Figure 10.10: The energy spectra of cosmogenic isotope decay products without any cuts. All spectra are normalized to the expected rates.

Contribution from other isotopes According to FLUKA simulation[48], the contribution of other cosmogenic isotopes to ROI, including ^{3}H, ^{11}Li, ^{12}Be, ^{13}N, ^{14}B, ^{14}C, ^{15}C and ^{18}N is 0.10 event per day.

Summary The contribution to the ROI of all cosmogenic isotopes are summarized in Table 10.9:

Table 10.9: Cosmogenic backgrounds

X	R_X^{tot} (day^{-1})	R_X^{ROI} (day^{-1})
^{8}B ν ES	90.55	13.23
^{10}C	760.4	447.85
^{11}Be	51.2	6.10
^{16}N	13[a]	0.39
^{6}He	1543	415.94
^{8}Li	560.2	37.38
^{8}B	387.2	$\ll 0.01$
^{9}Li	101.4	7.76
^{9}C	139.0	5.03
^{8}He	31.83	3.62
^{12}B	1968	112.17
^{13}B	12[a]	1
^{12}N	81.34	1.21
^{11}C	3.734×10^{4}	0
^{7}Be	5438[a]	0
^{10}Be	1419[a]	0
others	1.2×10^{4}[a]	0.10
neutrons	1.2798×10^{5}	1.2798×10^{5}

[a] Based on FLUKA: ibid.

---

[48] An et al., "Neutrino physics with JUNO"

10.2.6 *Other backgrounds*

Following KamLAND ^{8}B solar neutrino analysis[49] the following backgrounds are included:

1. Reactor $\bar{\nu}_e$
2. ^{8}B CC on ^{13}C ground state
3. ^{8}B CC on ^{13}C 3.51 MeV
4. *hep* solar neutrinos ES
5. DSNB and atmospheric neutrinos ES

These backgrounds, except the reactor $\bar{\nu}_e$, are negligible compared with ^{8}B solar neutrino elastic scattering signals. Their contributions to JUNO ROI are estimated assuming uniform energy distribution by scaling the rates from KamLAND[50]. The differences on the mass fraction of Carbon atoms and the electron density have been considered. The results are summarized in Table 10.10.

For the reactor $\bar{\nu}_e$, the inverse beta decay rate is 83 per day[51] and can be removed by coincidence cut. The elastic scattering rate is $\mathcal{O}(10^{-2})$ of the inverse beta decay rate and its contribution to ROI is taken as 1 event per day. Their contributions to ROI are summarized in Table 10.10.

X	R_X^{tot} (day^{-1})	R_X^{ROI} (day^{-1})
^{8}B ν ES	90.55	13.23
^{8}B CC 1	1.06	0.08
^{8}B CC 2	0.20	0.02
hep ES	0.11	0.01
atm ν ES	0.35	0.02
sum	1.72	0.13
Reactor $\bar{\nu}_e$ IBD p	83	12.5
Reactor $\bar{\nu}_e$ IBD d	83	83
Reactor $\bar{\nu}_e$ ES		0.1

Table 10.10: Contribution to ROI of other backgrounds

[49] The KamLAND collaboration, "Measurement of the ^{8}B solar neutrino flux with the KamLAND liquid scintillator detector", *Physical Review C*, vol. 84, no. 3, p. 035804, Sept. 2011, DOI: `10.1103/PhysRevC.84.035804`, arXiv: `1106.0861` Table II

[50] ibid. Table II

[51] An et al., "Neutrino physics with JUNO"

10.2.7 *Summary*

The estimated rates are summarized in Table 10.11.

Type	Name	R_X^{tot}	R_X^{ROI}	Reduction method
signal	^{8}B ν ES	90.55	13.23	
Ext	External γs	3.333×10^7	9.105×10^5	FV cut
	(α, n)	$\mathcal{O}(10)$	$\mathcal{O}(10)$	
LS	^{238}U	3009.26	132.35	Bi-Po tagging
	^{232}Th	656.28	24.44	
	other decay	2×10^4	0	-
cosmogenic	^{10}C	760.4	447.85	veto+TFC
	^{11}Be	51.2	6.10	
	^{16}N	13	0.39	
	^{6}He	1543	415.94	
	^{8}Li	560.2	37.38	
	^{8}B	387.2	0.4×10^{-3}	
	^{9}C	139.0	5.03	
	^{12}B	1968	112.17	
	^{13}B	12	1	
	^{12}N	81.34	1.21	
	^{9}Li	101.4	7.76	IBD + veto
	^{8}He	31.83	3.62	
	others	1.2×10^4	0.10	-
Other	Rea $\bar{\nu}_e$ IBD p	83	12.5	IBD cut
	Rea $\bar{\nu}_e$ IBD d	83	83	
	Rea $\bar{\nu}_e$ ES	-	0.1	
	^{8}B CC 1	1.06	0.08	-
	^{8}B CC 2	0.20	0.02	
	hep ES	0.11	0.01	
	atm ν ES	0.35	0.02	
bkg sum	irreducible	3×10^4	0.33	
	reducible	10^8	10^6	

Table 10.11: Summary of signals and backgrounds. Unit: per day

10.3 Background reductions

In this section, I will describe the technology used to reject background events, the optimization of cuts and the estimation of their efficiencies.

10.3.1 Fiducial volume cut

The fiducial volume cut removes events whose reconstructed event vertex is outside the pre-defined geometrical volume, here $r < r^{\mathrm{FV}}$. The γs from the natural radioactive decay and from the capture process of neutrons produced by (α, n) and spontaneous fissions in components outside the liquid scintillator are rejected.

The spatial distribution of the external γ events is obtained with a Geant4 simulation code. According to the absorption law, the γ intensity and the γ interaction event rates decrease exponentially along the radial direction. Considering that the γ flux along the radius can be described with the exponential function $I(r) = I_0 \exp\left(-\frac{r_0 - r}{\lambda}\right)$, define $v = r^3$, then the external γ event rate per unit volume can be described with

$$\frac{\mathrm{d}R^{\mathrm{ROI}}_{\mathrm{bkg}}}{\mathrm{d}v} = \frac{1}{3 \cdot v^{2/3}} \cdot \frac{r_0^{\mathrm{bkg}}}{\lambda} \cdot \exp\left(-\frac{r_0 - v^{1/3}}{\lambda}\right) \tag{10.12}$$

where r_0 is 17.7 m. The fit result using this model on the simulated distribution is show in Figure 10.11. The best fit parameters are

$$r_0^{\mathrm{bkg}} = (10.7 \pm 0.3)\ \mathrm{Hz} \tag{10.13}$$

$$\lambda = (0.222 \pm 0.006)\ \mathrm{m} \tag{10.14}$$

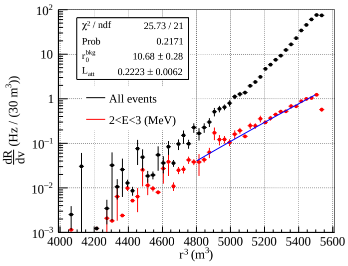


Figure 10.11: Rate of external γ events vs r^3 for all (black) and $2 < E_e < 3$ MeV (red).

The signal to background ratio of external γs with fiducial volume cut $r < r^{\mathrm{FV}}$ is given by

$$\begin{aligned}\frac{R^{\mathrm{ROI}}_{\mathrm{sig}}}{R^{\mathrm{ROI}}_{\mathrm{bkg}}} &= \frac{r_0^{\mathrm{sig}} \cdot \frac{4\pi}{3} \cdot (r^{\mathrm{FV}})^3}{\int_0^{(r^{\mathrm{FV}})^3} \frac{\mathrm{d}R^{\mathrm{ROI}}_{\mathrm{bkg}}}{\mathrm{d}v} \cdot \mathrm{d}v} \\ &= \frac{r_0^{\mathrm{sig}} \cdot \frac{4\pi}{3} \cdot (r^{\mathrm{FV}})^3}{r_0^{\mathrm{bkg}} \cdot \left[\exp\left(-\frac{r_0 - r^{\mathrm{FV}}}{\lambda}\right) - \exp\left(-\frac{r_0}{\lambda}\right)\right]}\end{aligned} \tag{10.15}$$

and it is shown in Figure 10.12.

The optimized value is

$$r^{\mathrm{FV}} = 14\ \mathrm{m} \tag{10.16}$$

where S/B is 118 and the residual rate is 0.11 day^{-1}, and the signal efficiency is 49.48% After such fiducial volume cut, the γ due to neutron from (α, n) reactions or spontaneous fission reactions is negligible.

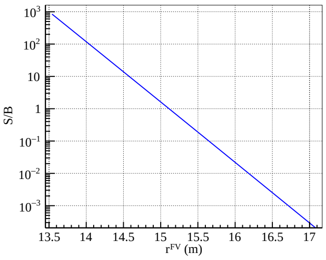


Figure 10.12: Optimization of the FV cut. S/B vs r_{FV}

The rate of affected components are summarized in Table 10.12.

X	R_X^{tot} (day^{-1})	R_X^{ROI} (day^{-1})	FV cut (day^{-1})
^{8}B ν ES	90.55	13.23	6.546
External γs	3.333×10^7	9.105×10^5	0.055
(α, n)	$\mathcal{O}(10)$	$\mathcal{O}(10)$	0

Table 10.12: Rate before and after fiducial volume cut ($r < 14$ m).

10.3.2 *IBD cut*

This cut removes all events that have high energy events within 1 ms and 1.5 meters. The high energy is defined as higher than 0.7 MeV. Because the single event rate above 0.7 MeV is $\mathcal{O}(1)$ Hz, the exposure loss of this cut is negligible. The reactor $\bar{\nu}_e$ IBD prompt and delayed signals, cosmogenic ^{9}Li and ^{8}He decays will be reduced. Besides, the ^{214}Bi–^{214}Po from ^{238}U decay chain will also be reduced. The spectra of ^{238}U decay chain before and after this cut is shown in Figure 10.13. According to a Geant4 simulation, the exposure loss and rejection rate are calculated and reported in Table 10.13.

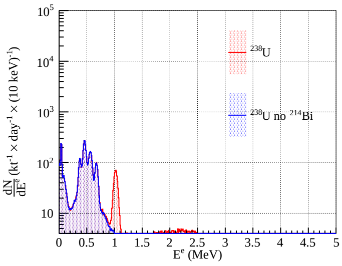


Figure 10.13: ^{238}U decay chain energy spectra before and after IBD cut.

Type	rejected / total (%)
Uncorrelated	0
Reactor $\bar{\nu}_e$ IBD p/d	97.8
cosmogenic ^{9}Li/^{8}He	92.2
LS ^{214}Bi–^{214}Po	99.97

Table 10.13: Efficiency of coincidence cut.

The rates of affected components are summarized in Table 10.14.

X	R_X^{tot}	R_X^{ROI}	FV cut	IBD cut
^{8}B ν ES	90.55	13.23	6.546	6.546
^{9}Li	101.4	7.76	3.84	0.30
^{8}He	31.83	3.62	1.79	0.14
Rea $\bar{\nu}_e$ IBD p	83	12.5	6.19	0.14
Rea $\bar{\nu}_e$ IBD d	83	83	41	0.90
LS ^{238}U	3009	132.4	65.50	0.52

Table 10.14: Rate before and after IBD cut. Unit: per day

10.3.3 *μ veto*

The μ veto removes events that are tagged as μ events and all events after the μ close in space and time. The cut is optimized to be

1. All tagged μ: 1.5 ms, whole detector.

2. tracked μ: $R_{d2\mu} < 3$ m and $T_{d2\mu} < 4.1$ s

$T_{d2\mu}$ is chosen to be 4.1 s so that the inefficiency of time interval cut on ^{6}He is similar to that of distance to μ track cut.

The rejection rates of cosmogenic isotope decay events are evaluated with a Geant4 simulation. We have simulated exposure of around 599.7 days of muon events (1.55×10^8 μ events) for this study. The short lived isotopes are not fully removed because of the leakage from the distance to μ track cut. The distribution of the distance to μ track for cosmogenic ^{6}He decay events are shown in Figure 10.14.

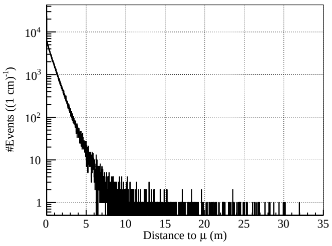


Figure 10.14: Distribution of distance to μ track for cosmogenic ^{6}He decay events. For 3 m cut, the leakage is 2.9%. The Monte Carlo truth of the μ track is used.

Now we calculate the signal efficiency. Consider for any event happed at time t, then its probability of being not rejected by the muon cut is the probability that in the time range from $t - T_{d2\mu}$ to t and in the area of a circle of radius $R_{d2\mu}$ there is no muon. The muon flux is given by

$$\phi_\mu = R_\mu \cdot \frac{L_\mu}{\frac{4\pi}{3}R^3} \tag{10.17}$$

The expected muon event count is $\mu = \phi_\mu \cdot T_{d2\mu} \cdot \pi R_{d2\mu}^2$, and the signal efficiency is $\varepsilon = P(0;\mu)$. So the signal efficiency is given by

$$\varepsilon = \exp\left(-R_\mu \cdot T_{d2\mu} \cdot \frac{\pi \cdot R_{d2\mu}^2 \cdot L_\mu}{\frac{4\pi}{3}R^3}\right) \tag{10.18}$$

where L_μ is the average muon track length, and the result is 70.87%. The rates of affected components are summarized in Table 10.15.

X	R_X^{tot}	R_X^{ROI}	FV cut	IBD cut	μ veto
^{8}B ν ES	90.55	13.23	6.546	6.546	4.639
^{10}C	760.4	447.85	221.69	221.69	186.05
^{11}Be	51.2	6.10	3.02	3.02	2.46
^{16}N	13	0.39	0.39	0.39	0.26
^{6}He	1543	415.94	205.90	205.90	11.99
^{8}Li	560.2	37.38	18.50	18.50	1.22
^{8}B	387.2	0	-	-	-
^{9}C	139.0	5.03	2.49	2.49	0.023
^{12}B	1968	112.17	55.53	55.53	1.58
^{13}B	12	1	0.50	0.50	0
^{12}N	81.34	1.21	0.60	0.60	0.006
^{9}Li	101.4	7.76	3.84	0.30	0.003
^{8}He	31.83	3.62	1.79	0.14	0.001

Table 10.15: Rate before and after muon veto. Unit: per day.

10.3.4 TFC cut

We would like to apply the so-called three-fold-coincidence (TFC) cut[52] to suppress cosmogenic isotope decay events, especially for the ^{10}C and ^{11}Be events.

The cut is defined as the following: for all neutrons that are associated to a μ event, reject all events that are within 2 meters and 111 seconds. The neutron that are within 2 ms with respect to a μ event are considered to be associated to this μ. Due to high luminosity of scintillation photons produced by the μ event, there is around 2 μs dead time for detecting cosmogenic neutrons. Considering the neutron capture time is 220 μs in LAB[53], the overall cosmogenic neutron detection efficiency is 99%.

For comparison, the TFC cut parameters used by Borexino[54,55] and KamLAND[56] are summarized in Table 10.16.

[52] C. Galbiati et al., "Cosmogenic C11 production and sensitivity of organic scintillator detectors to pep and CNO neutrinos", *Physical Review C - Nuclear Physics*, vol. 71, no. 5, pp. 1–11, 2005, DOI: `10.1103/PhysRevC.71.055805`, arXiv: `0411002 [hep-ph]`

[53] An et al., "Neutrino physics with JUNO"

Experiment	$R_{n2\mu}$ (m)	$\Delta T_{n2\mu}$ (μs)	R_{n2x} (m)	ΔT_{n2x} (s)	τ_n (μs)	$\epsilon_\mu^{\text{track}}$	ϵ_n^{tag}	$\epsilon(^{10}\text{C})$ (%)	$\epsilon(^{11}\text{Be})$
Borexino	4.69	[16, 1600]	0.8	[10, 310]	259.7	99.99	94	$92.5^{+7.5}_{-20.0}$	> 71
KamLAND	?	[150,?]	?	?	207.5	97	~ 50	90.7 ± 5.5	2×10

Table 10.16: TFC parameters and efficiencies of Borexino and KamLAND experiments.

For Borexino, considering the ratio of ^{11}C associated with at least one neutron is only 90%, while only 94% neutrons are detected, the tagging efficiency on neutron associated ^{11}C is

$$\frac{N \times 92.5\%}{N \times 90\% \times 94\%} = 109\% \tag{10.19}$$

and is larger than 100%, so the no-neutron associated ^{11}C can also be tagged with such TFC method. This can be explained such that when cosmogenic isotope is produced, it is likely that the μ has induced a shower or has deposited significant amount of energy, and then it is very likely that spallation neutrons are also produced at that point.

We have also performed Geant4 simulation and found this phenomena is reproduced by Geant4 simulation. Consider cosmogenic ^{11}Be isotope, for which the neutron associated ratio is only 11.5%. For each ^{11}Be, if its parent μ event has produced neutrons, we find the nearest neutron and plot the distribution of the distance between the neutron and the ^{11}Be event. For ^{11}Be whose parent μ does not produce any cosmogenic neutrons, such distance is set to infinity. The distribution of such distance is shown in Figure 10.15. As you can see, only 2% of parent μ does not produce any neutrons and more than 97% of nearest neutron are within 2 meters.

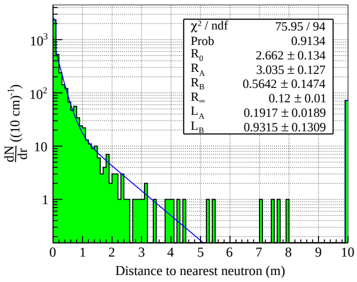


Figure 10.15: Distance to the nearest neutron for cosmogenic ^{11}Be. The distribution can be fitted with an underflow bin, an overflow bin (no neutron ^{11}Be) plus two exponential distributions, whose rates are R_0, R_∞, R_A and R_B, respectively. The unit of R is in per day and that of L is in meters.

The efficiency of TFC cut on rejecting cosmogenic isotope decay events is estimated according to a Geant4 simulation. We have simulated exposure of around 599.7 days of muon events (1.55×10^8 μ events). The signal efficiency, similar to the muon cut, can be evaluated by calculating the probability of finding no cosmogenic neutron in the time interval T_{n2x} before the vent and in the volume of a sphere of radius R_{n2x}. It is given by

$$\epsilon_{\text{TFC}} = \exp\left[-R_n \cdot \Delta T_{n2x} \cdot \left(\frac{R_{n2x}}{R_t}\right)^3\right] \tag{10.20}$$

where R_t is the radius of the target, 17.7 m, R_n is the neutron rate, 1.280×10^5 per day. The residual event rate and the signal-to-background ratio versus various cut conditions are shown in Figure 10.16 and 10.17, respectively. Under the chosen condition, the signal efficiency is 78.88%. The rates of affected components under the chosen cut condition is reported in Table 10.17.

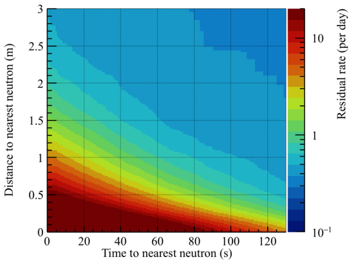


Figure 10.16: Residual cosmogenic isotope decay background rate vs various TFC cut conditions

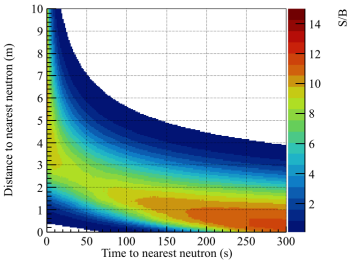


Figure 10.17: $S/B_{\text{cosmogenic}}$ vs various TFC cut conditions.

10.3.5 α cut

The α-like events can be rejected efficiently using pulse shape discrimination technique. When α particles deposit energies in the liquid scintillator, because the unit energy loss $\mathrm{d}E/\mathrm{d}x$ is much larger than that of electrons, LAB molecules are more likely to be excited to triple state, while the decay time of the triple state is longer. As

X	R_X^{tot}	R_X^{ROI}	FV cut	IBD cut	μ veto	TFC veto
^{8}B ν ES	90.55	13.23	6.546	6.546	4.639	3.659
^{10}C	760.4	447.85	221.69	221.69	186.05	0.033
^{11}Be	51.2	6.10	3.02	3.02	2.46	0.046
^{16}N	13	0.39	0.39	0.39	0.26	0.013
^{6}He	1543	415.94	205.90	205.90	11.99	0.212
^{8}Li	560.2	37.38	18.50	18.50	1.22	0.026
^{8}B	387.2	$\ll 0.01$	0	-	-	-
^{9}C	139.0	5.03	2.49	2.49	0.023	0
^{12}B	1968	112.17	55.53	55.53	1.58	0.018
^{13}B	12	1	0.50	0.50	< 0.001	0
^{12}N	81.34	1.21	0.60	0.60	0.006	0
^{9}Li	101.4	7.76	3.84	0.30	0.003	0
^{8}He	31.83	3.62	1.79	0.14	0.001	0
sum	5649	1038	514.3	509.1	203.56	0.347

Table 10.17: Rate before and after TFC veto. Unit: per day.

a result, the scintillation photons produced by α particles have more slow components. According to a simulation, a multi-layer perceptron algorithm can discriminate between β and α particles with almost 100% α efficiency and zero β inefficiency. With this cut, the α decay in liquid scintillator ^{238}U and ^{232}Th decay chain are rejected. The spectra of ^{238}U decay chain and ^{232}Th decay chain after α rejection are shown in Figure 10.18. After such cut the contribution of ^{232}Th to 2–3 MeV region is negligible, because the ^{212}Bi–^{212}Po events are removed.

The rates of affected components under the chosen cut condition are reported in Table 10.18.

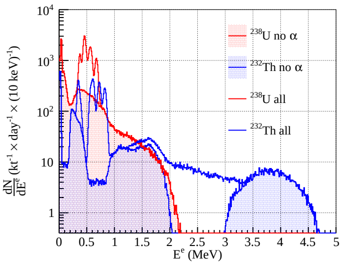


Figure 10.18: Visible energy spectrum of ^{238}U and ^{232}Th chain decay products before and after α cut.

X	R_X^{tot}	R_X^{ROI}	FV cut	IBD cut+μ + TFC	α cut
^{8}B ν ES	90.55	13.23	6.546	3.659	3.659
LS ^{238}U	3009	132.4	65.50	0.291	0.291
LS ^{232}Th	656.28	24.44	12.10	6.76	0.102

Table 10.18: Rate before and after α cut. Unit: per day.

10.3.6 *Summary*

The rates after each cut are summarized in Table 10.19.

Table 10.19: Summary of signals and backgrounds after each cut. Unit: per day.

Name	R_X^{tot}	R_X^{ROI}	FV cut	IBD cut	μ veto	TFC veto	α cut
^{8}B ν ES	90.55	13.23	6.546	6.546	4.639	3.659	3.659
External γs	3.333×10^7	9.105×10^5	0.055	0.055	0.039	0.031	0.031
(α, n)	$\mathcal{O}(10)$	$\mathcal{O}(10)$	0	-	-	-	-
^{238}U	3009.26	132.35	65.50	0.519	0.368	0.291	0.291
^{232}Th	656.28	24.44	12.10	12.10	8.58	6.76	0.102
^{10}C	760.4	447.85	221.69	221.69	186.05	0.033	0.033
^{11}Be	51.2	6.10	3.02	3.02	2.46	0.046	0.046
^{16}N	13	0.39	0.39	0.39	0.26	0.013	0.013
^{6}He	1543	415.94	205.90	205.90	11.99	0.212	0.212
^{8}Li	560.2	37.38	18.50	18.50	1.22	0.026	0.026
^{8}B	387.2	$\ll 0.01$	0	-	-	-	-
^{9}C	139.0	5.03	2.49	2.49	0.023	0	-
^{12}B	1968	112.17	55.53	55.53	1.58	0.018	0.018
^{13}B	12	1	0.50	0.50	0	-	-
^{12}N	81.34	1.21	0.60	0.60	0.006	0	-
^{9}Li	101.4	7.76	3.84	0.30	0.003	0	-
^{8}He	31.83	3.62	1.79	0.14	0.001	0	-
Rea $\bar{\nu}_e$ IBD p	83	12.5	6.19	0.14	0.099	0.078	0.078
Rea $\bar{\nu}_e$ IBD d	83	83	41	0.90	0.638	0.503	0.503
Rea $\bar{\nu}_e$ ES		0.1	0.050	0.050	0.035	0.028	0.028
others	3.2×10^4	0.23	0.114	0.114	0.081	0.064	0.064
bkg sum	3×10^7	9×10^5	639	523	213	8.102	1.444

10.4 *Results*

In this section the estimated final signal to background ratio, the expected precision on the survival probability and the sensitivity to MSW upturn are reported.

Using Monte Carlo simulation the joint probability density function of (E_ν, T) for ^{8}B solar neutrino elastic scattering signals, where T is the kinetic energy of the recoil electron, is obtained. It is shown in Figure 10.19 left. Using such distribution, the energy distribution of electron neutrinos that has contributed to the ROI is obtained and shown in Figure 10.19 right, and the average neutrino energy is 7.09 MeV.

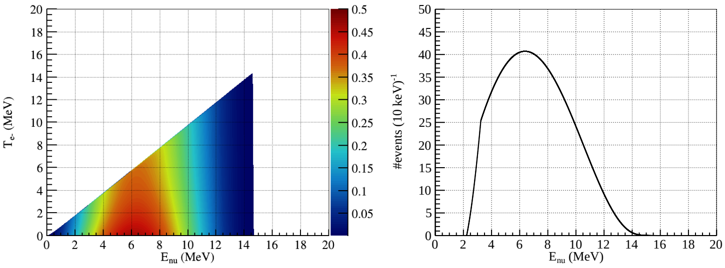


Figure 10.19: Left: Joint distribution of neutrino energy and kinetic energy of the recoil electron for ^{8}B solar neutrino. Right: Distribution of the energy of neutrinos that contribute to the ROI.

With 1 year of data taking and after all event selection criteria, it is expected to observe 1336 ^{8}B solar neutrino elastic scattering events and 527 background events in ROI. Considering that the main contribution of background events are the residual reactor IBD events, ^{238}U decay chain and cosmogenic ^{6}He events, they can be well constrained from independent measurement. If 1863 (1336 + 527) events are observed and the background rate are fixed to be 1.4 per day, then the 90% Feldman-Cousin confidence interval of the signal rate is [3.46, 3.85] per day, and this correspond to $P_{ee} = 0.359 \pm 0.019$ (68% C.L.), where the uncertainty includes the normalization from SNO experiment. JUNO's contribution to the survival probability plot is shown in Figure 10.20 together with the results from Borexino and Super-K/SNO.

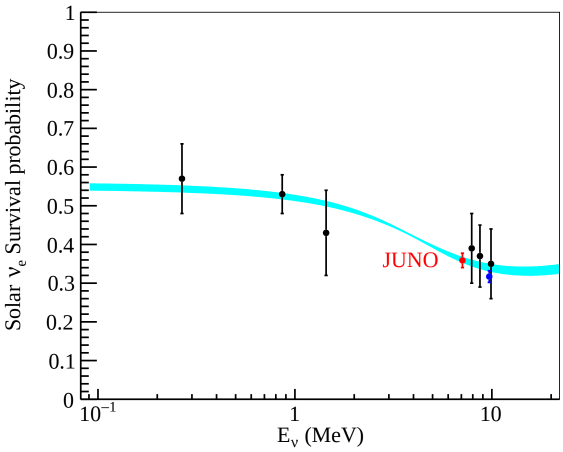


Figure 10.20: The survival probability of solar electron neutrinos. Black: Borexino results. Blue: Super-K + SNO joint anlaysis. Red: this work with the JUNO detector.

The uncertainty of the standard solar model prediction of ^{8}B solar neutrino flux can be eliminated if the ratio variable is used:

$$r^{\mathrm{SK/JUNO}} = \frac{P_{ee}(\text{Super-K})}{P_{ee}(\text{JUNO})} \tag{10.21}$$

where P_{ee}(Super-K) is the average survival probability of ^{8}B solar neutrino with 3.5 MeV detection threshold reported by Super-Kamiokande collaboration[57] and P_{ee}(JUNO) is the survival probability reported in this work. Here the uncertainty from the normalization can be cancelled out. The 99.7% confidence interval of such value versus live time is shown in Figure 10.21. From the figure we can see that with around 261 days of data the statistical sensitivity to reject null hypothesis ($r^{\mathrm{SK/JUNO}} \equiv 1$) is more than 3 σ.

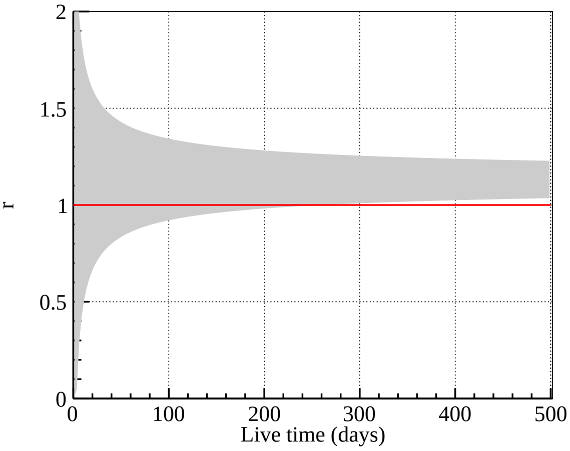


Figure 10.21: The 99.7% confidence interval of the ratio of survival probability of solar electron neutrinos between Super-K and JUNO experiments.

10.5 *Conclusions and discussions*

The background level and analysis methods to suppress backgrounds have been studied in this work, and with 2–3 MeV visible energy range as the region of interests, the final signal to background ratio after all event selection criteria reached 2.5. The ^{8}B solar neutrino elastic scattering signal rate in ROI after all cuts is 3.7 event per day, and with one year of data it is expected that the statistical uncertainty of the measured active ^{8}B solar neutrino flux is 3.2%.

The ratio between the average survival probability estimated in this work and that reported by Super-K is chosen as a criteria to see the MSW up turn, and it is found that with 261 days of data the sensitivity to reject null hypothesis $r \equiv 1$ can reach 3σ.

In this work, the cosmogenic ^{6}He decay events can be further reduced by enlarging the distance to muon track criteria. For reactor anti-electron neutrinos, the IBD events can be further reduced by enlarging the time interval and vertex distance cut while not loosing significant exposure. For ^{238}U and ^{232}Th, their contribution to the ROI can be further suppressed by increasing the low edge of the ROI to, such as, 2.2 MeV, or by performing a shape analysis.

---

[57] The Super-Kamiokande collaboration, "Solar neutrino measurements in Super-Kamiokande-IV"

In this work, the concentration of ^{238}U and ^{232}Th contaminations are assumed to be 10^{-17} g/g. The impact of higher concentration is shown in Figure 10.22. If the ^{238}U and ^{232}Th concentration is as high as 10^{-15} g/g, the time to reach 3σ confidence level increases to around 1200 days.

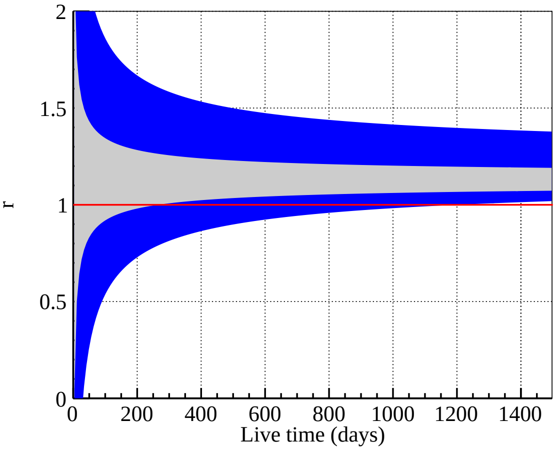


Figure 10.22: The 99.7% confidence interval of the ratio of survival probability of solar electron neutrinos between Super-K and JUNO experiments under different liquid scintillator purity levels. Gray: 10^{-17} g/g. Blue: 10^{-15} g/g.

11
Other studies on Solar neutrinos with JUNO

In this chapter, my further work on solar neutrinos in the JUNO experiment is presented. Solar neutrino generators have been developed in the framework of `Sniper`, which is the Software for Noncollider Physics Experiments. Dedicated reconstruction algorithms are developed for sub-MeV events in order to handle the impact of dark noise. For energy reconstruction, clustering algorithms based on the hit time distribution have been developed, and PMT occupancy based likelihood is developed in the energy reconstruction. For vertex reconstruction, the existing algorithm has been optimized to improve the throughput. Beside these software related development, I have also evaluated the sensitivity of JUNO on other solar neutrino species, including the *pp*, ^{7}Be, *pep*, CNO and *hep* neutrinos.

This chapter is organized as in the following: ff Section 11.1 introduces the software development related to solar neutrinos in JUNO. Section 11.2 presents the evaluated sensitivity of JUNO on low energy solar neutrinos. Section 11.3 presents the evaluated sensitivity on *hep* solar neutrinos.

I have co-worked with Vito Antonelli, Emanuela Meroni, Gioacchino Ranucci, Alessandra Re, Barbara Ricci, Giuseppe Salamanna in this work.

11.1 Softwares for solar neutrinos

In this section I describe the development of softwares related to solar neutrinos in JUNO, including the generator, the vertex and energy reconstruction algorithms and the clustering algorithm.

11.1.1 The solar neutrino generator

The generators are used to provide the kinetic information of the primary particles. Due to the extremely low cross-section, it is unrealistic to simulate the neutrinos then add elastic scattering to the physics process list in GEANT4. Therefore the recoil electrons are chosen as the primary particles. In the generator, first the neutrino energy is sampled according to theoretical shapes, then it is sampled to be transformed into ν_e or $\nu_{\mu,\tau}$. After that the electron energy and momentum is randomly sampled. At last, the sampled energy and momentum are rejected with a probability determined by the differential cross-section. It is important that the sum of acceptance over all electron energies is not 100% given a neutrino energy, because the total cross-section depends on the neutrino energy. This is summarized as the following:

1. Sample E_ν according to the theoretical shapes $f(E)$.
2. Sample a random number. With a probability of $P_{ee}(E_\nu)$, the solar neutrino is still an electron neutrino when detected.
3. Calculate the maximum cross-section over all allowed neutrino energies and all recoil electron energies. Let us denote it as σ_{MM}.
4. Sample the recoil electron energy uniformly between 0 and $T^{\max} = \dfrac{2E_\nu^2}{m_e + 2E_\nu}$, and accept it with a probability of $\dfrac{\sigma(E_\nu, T)}{\sigma_{\mathrm{MM}}}$.
5. If rejected, goes to step 1.

Because the generator is implemented in the framework of JUNO offline software, its usage is similar to all other types of generators. Besides, driver has been implemented so that user can produce samples utilizing the `JunoTest` tool easily. For example, to generate ^{7}Be solar neutrino events , the command is

```
python $JUNOTESTROOT/python/JunoTest/junotest Production --daemon
    --log Be7-dn10k.txt --ini mixing.ini Be7-dn10k
```

and the configuration file reads

```
[Be7-dn10k]
driver = ChainNuSol
scripts = prod_Chain_uniform.py
evtmax = 2000
njobs = 100
seed = 42
elecsim-mode = --enableDarkPulse --darkRate 10e3 --Trigger_FiredPmtNum 300
rec-mode = --elec yes
; tags = Be7
; default tag is Be7
workflow = detsim elecsim calib rec
workDir = Be7-dn10k
```

in which we set the number of events to be simulated, the dark rate and the trigger threshold.

The generated electron kinetic energy spectra are shown in Figure 11.1.

11.1.2 *Throughput optimization of vertex reconstruction*

In JUNO, the main vertex reconstruction algorithm is based on hit time likelihood method. The throughput is not satisfactory. A pack-

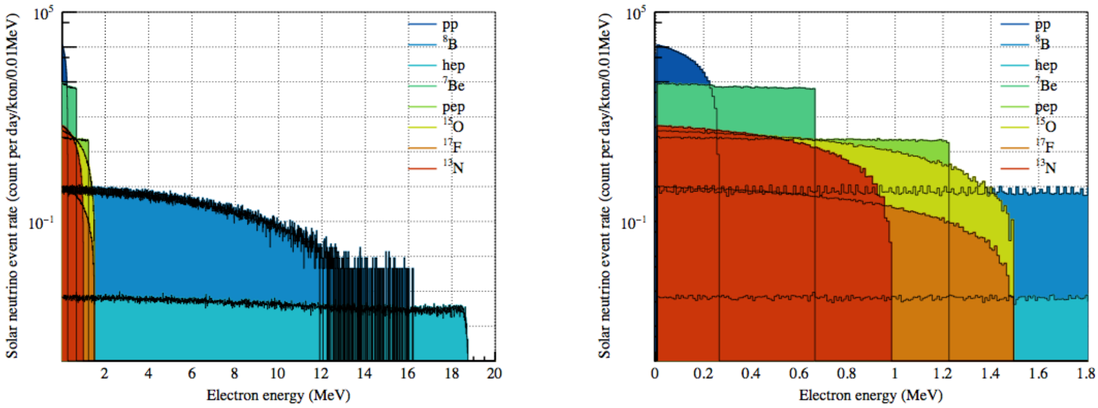


Figure 11.1: Generated recoil electron spectrum produced by the JUNO solar neutrino generator.

age `PosRecTimeLikeAlgV3` based on the same algorithm but optimized on throughput is implemented in my analysis. The package described in this and next two subsections are included in the `SmartRec` package bundle.

First, a caching algorithm has been implemented to save the time for computing likelihood of previously scanned vertex. This is useful, because the grid search algorithm is used in the minimization. The popular optimization algorithm `MINUIT` is not used, because the dependence of the likelihood on fit parameters are not smooth. The caching algorithm is illustrated in Figure 11.2.

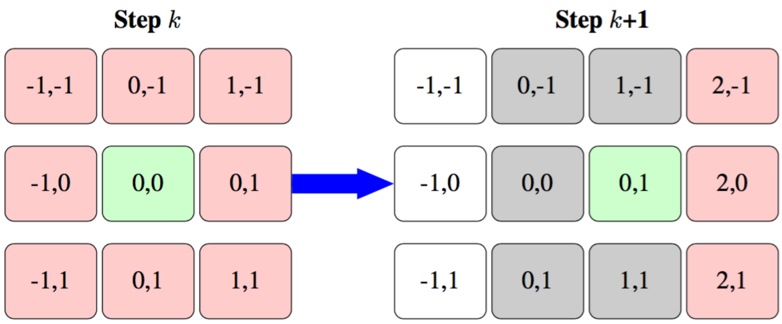


Figure 11.2: A schematic diagram of the caching algorithm. In the grid search optimization algorithm, the pointer will move to the neighbor grid with lowest likelihood. The likelihood of the neighbor grid need to be calculated. As shown, if the likelihood of those grids marked as gray are cached, time needed to compute likelihood on these grids is saved.

Second, CPU profiling tool `perf` was used to identify the bottleneck of the algorithm. The results are summarized in Figure 11.3. From the figure we can see that around 40% of the time is spent on the `acos` and `log` operators. Based on this, the `acos` operator was eliminated using trigonometric formulas, and `log` was replaced by `logf`, which operates on single precision numbers. Considering the `log` operator is needed when calculating the final likelihood, the relative tolerance of the grid search algorithm is around 10^{-4} and the precision of single precision numbers is 10^{-7}, this replacement will not reduce the vertex reconstruction precision.

The final result is shown in Figure 11.4. It can be seen that the throughput of the vertex reconstruction algorithm has been improved from 1.37 ± 0.03 seconds per MeV per event to 0.401 ± 0.005 seconds per MeV per event.

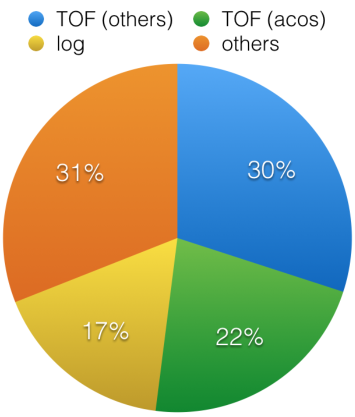


Figure 11.3: The identified bottlenecks and the fraction of time spent on them in the vertex reconstruction before optimization. From the figure we can see that around 40% of the time is spent on the `acos` and `log` operators.

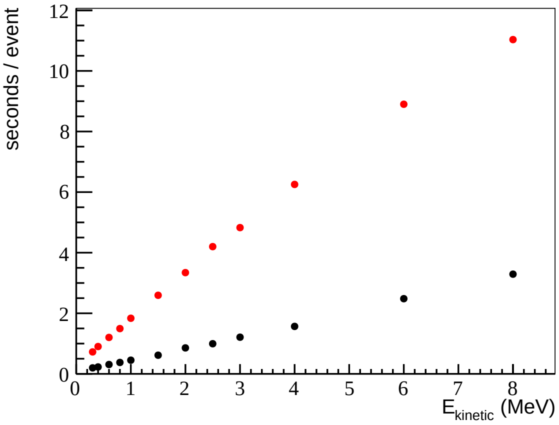


Figure 11.4: Comparison of throughput of vertex reconstruction before (red) and after (black) optimization.

11.1.3 *Occupancy based energy reconstruction*

The deposited energy can be estimated using the number and distribution of the fired PMTs. Compared with the energy reconstructed with the charge distribution, it is not affected by the smearing of single photoelectron charge, and thus has advantage in the low energy region. A likelihood based energy reconstruction algorithm `NpmtLikeRecAlg` using the PMT occupancy information is developed. The likelihood is defined as

$$-\log \mathcal{L} = \sum_{\text{fired}} \mu_i - \sum_{\text{not fired}} \log\left(1 - e^{-\mu_j}\right) \tag{11.1}$$

where μ_i and μ_j are the expected number of p.e. received by i-th and j-th PMT.

11.1.4 *Clustering algorithm*

Clusterization algorithm is like a software based trigger algorithm. It selects hits that are considered to be of the same physical event based on the hit time of hits. This is useful for two reasons: first, hits belonging to the same events comes in a burst, while dark hits distribute uniformly in time, and selecting hits in the burst can suppress contribution of dark hits, especially for low energy events; secondly, sometimes a ^{14}C decay event may fall in the same DAQ window of one physical event, due to its high rate, the hits from these two physical events can be grouped as two bursts and separated. The hit time distribution will be narrower if the time-of-flight of the event is corrected. The hit time distribution with and without time-of-flight correction before and after clusterization is shown in Figure 11.5.

The clusterization algorithm define the start and end of the cluster in the hit time distribution histogram, and selects hits in between. The start of a cluster is easy to be identified. Here we made a scan and find the first time that the number of events in the sliding window is above a certain threshold. The end of a cluster can be defined in two ways, either we fix the length of the cluster, or we make a scan again and find the position where the number of hits in the sliding window is for the first time compatible with no signals.

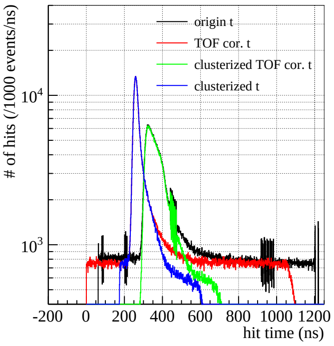


Figure 11.5: Comparison of hit time distribution before/after clusterization with/without time-of-flight correction.

For the fixed length clusters, the efficiency of signal hits and the rejection rate of the dark noise depends on the length of the cluster. The fraction of hits after clusterization is shown in Figure 11.6. Here I have used electrons with kinetic energy 0.515 MeV as the benchmark, and this is the critical energy for ^{7}Be solar neutrinos. From the figure we can see that when the cluster length is set to 400 ns, the signal hits loss is 1% while 10% of dark hits are rejected, and at this cluster length the energy resolution is the best.

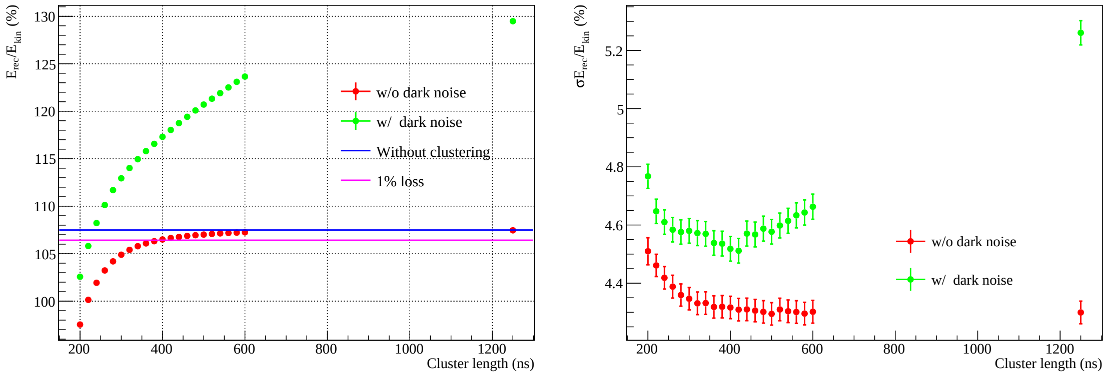


Figure 11.6: Left: The efficiency of clusterization algorithm on accepting signal hits and rejecting dark hits. Right: energy resolution under different cluster length. Here electrons with the kinetic energy of 0.515 MeV are used.

After clusterization, the fraction of dark hits in the event is reduced, and thus the energy resolution is improved. Electrons at detector center are used to compare the energy resolution under various conditions. The result is shown in Figure 11.7. From the figure we can see that for events at the detector center with energy lower than 2 MeV, the PMT occupancy based reconstruction algorithm combined with the fixed length clusterization algorithm gives the best energy resolution.

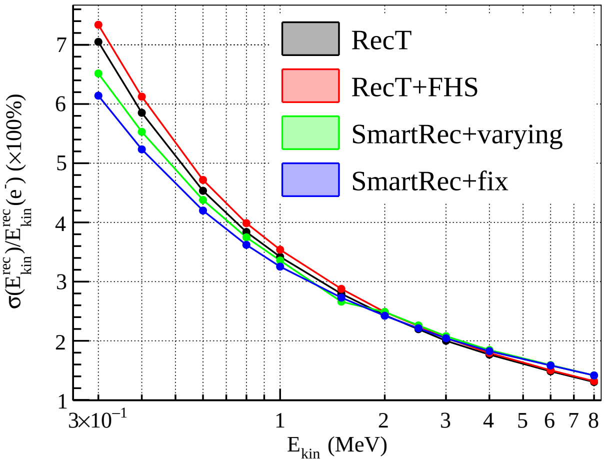


Figure 11.7: Comparison of the energy resolution of different reconstruction algorithms with and without clusterization. Electrons at the detector center are used. The RecT is the original energy reconstruction algorithm based on charge. The SmartRec is the energy reconstruction algorithm based on the occupancy of PMTs developed in this work. FHS is short for first hit selection, and when it is enabled, only the first hit of each PMT is used. SmartRec+Varying and SmartRec+fix correspond to varying or fixed length clusterization algorithms.

11.2 *Potential for solar ^{7}Be and CNO neutrinos*

In this section we present the expected precision of solar neutrino interaction rates using the JUNO detector. As in the case of Borexino analyses, in JUNO it is considered that the solar neutrino rates can be extracted by spectrum fitting. We have assumed the so-called "baseline requirement"[1] for the backgrounds. An exposure of 10^6 ton×years is considered.

11.2.1 *Spectrum fit results*

The analytical response functions described in Chapter 4 are used to build up the probability density functions of all components. The fit range is chosen to be between 0.37 and 1.87 MeV. *pep* solar neutrinos are constrained within 2% precision to break the correlation between *pep* and CNO solar neutrinos. The *pep* solar neutrino interaction rate can be constrained if we apply the solar luminosity constraint which constrains the *pp* solar neutrino fluxes and at the same time apply the constraint on the ratio between the flux of *pp* and *pep* solar neutrinos.

The fit results are shown in Figure 11.8. From the figure we can see that the statistical precision of JUNO on ^{7}Be and CNO solar neutrinos with 10^6 year×ton exposure are 0.03% and 1%, respectively.

To study the systematic uncertainty induced by the light yield, we performed the fit with light yield biased by ±0.5%. The corresponding shift of best fit value is taken as the systematic uncertainties. The results are summarized in Table 11.1. From the table we can see that the uncertainties of these two species will be dominated by the systematic uncertainties. The light yield induced systematic uncertainty, according to our experience, should represent the major contribution of the systematic uncertainties. Other type of detector response related systematic uncertainties, including the non-linearity model and the shape of mono-energetic event energy distribution, will be studied in in the future.

Name	Injected	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$(LY)
$\nu(^7\text{Be})$	483.0	0.2	32
ν(CNO)	49.2	0.5	4.2×10^2

Table 11.1: Fit results of ^{7}Be and CNO solar neutrinos in JUNO. Unit: count per day per kilo-tons

In order to understand if there is any hope to detect CNO solar neutrinos, we studied the impact of the ^{210}Bi decay rate, the precision of the light yield and the effect of three-fold-coincidence, and it is found that the light yield precision is the most important factor among the three. When the light yield precision can be as good as 0.1% and if the cosmogenic ^{11}C can be suppressed by 80%, the systematic uncertainty due to the light yield precision is 13%. The dependence of the light yield induced systematic uncertainty on the precision of the light yield is shown in Figure 11.9. From the figure we can see that when the precision of the light yield is 0.2%, the corresponding systematic uncertainty is 30%. The cosmogenic ^{11}C decay rate can be suppressed with the three-fold-coincidence algorithm. This algorithm rejects events near cosmogenic neutron in space and time. With a naive estimation, 80% suppression is feasible without losing too much exposure. Considering the half life of ^{11}C is 20 minutes, it takes 50 minutes for ^{11}C to decay to be 20%

---

[1] Table 14, page 89 F. An et al., "Neutrino physics with JUNO", *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, no. 3, p. 030401, 2016, DOI: `10.1088/0954-3899/43/3/030401`, arXiv: `1507.05613`

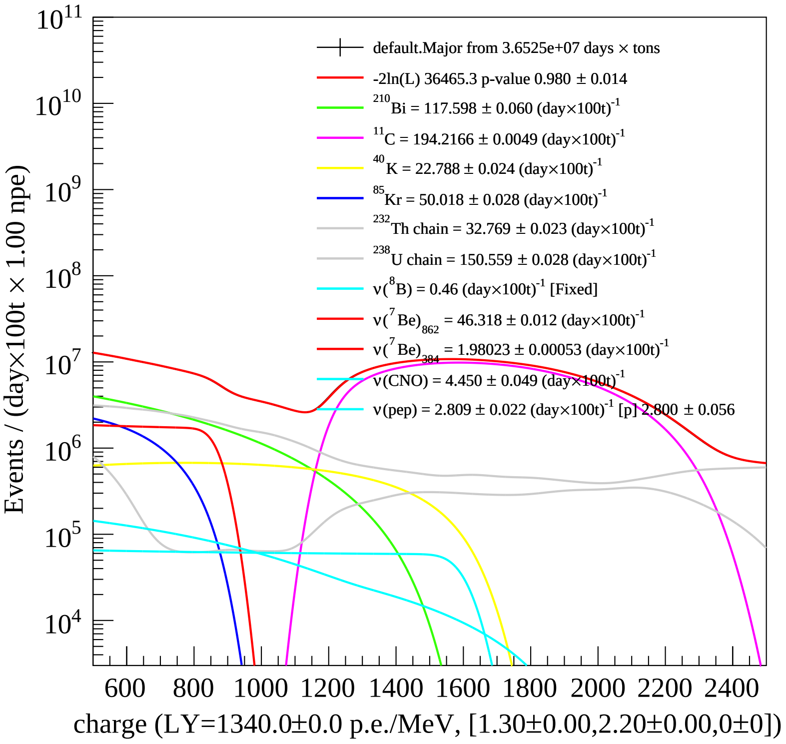


Figure 11.8: Spectral fit result of ^{7}Be and CNO solar neutrinos in JUNO.

left, so the time-to-neutron cut can be set to 1 hour. Considering the precision of the vertex reconstruction of ^{11}C and neutron, the distance-to-neutron cut can be set to 50 centimeters. In this way the exposure loss is around 30%.

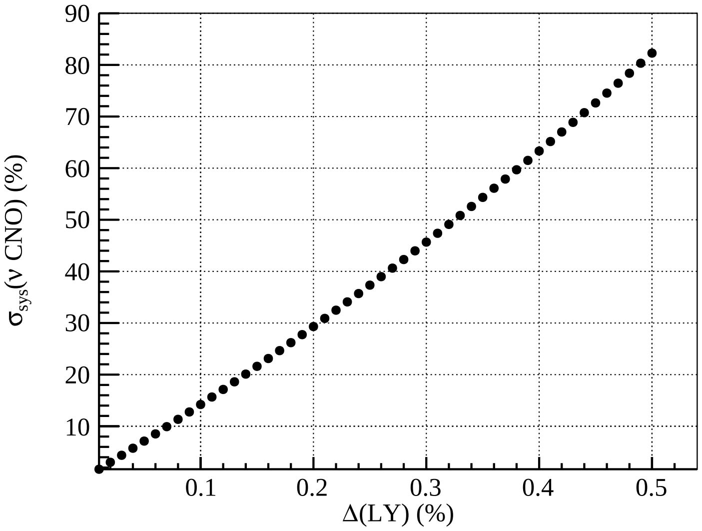


Figure 11.9: Systematic uncertainty of CNO solar neutrinos versus precision of the light yield when ^{11}C rate is suppressed by 80%.

11.3 Potential for *hep* solar neutrino

In this section I described the estimated sensitivity to *hep* solar neutrinos.

The energy spectrum of *hep* solar neutrinos is continuous. The endpoint of it is 18.8 MeV, and its expected flux is even lower than the ^{8}B solar neutrinos, and is $7.98(1 \pm 0.30) \times 10^3$ cm^{-2}s^{-1} (B16-GS98) [2]. This neutrino component was never detected, and currently the global fit upper limit is 47×10^3 cm^{-2}s^{-1} (99% C.L.)[3], which is mainly driven by the result of the D_2O phase of the SNO experiment[4].

In this work, I discuss the detection of the *hep* solar neutrinos via its inelastic scattering on the carbon atoms[5]:

$$\nu_l + ^{12}\mathrm{C}(0^+;\ \mathrm{gnd}) \rightarrow \nu_l' + ^{12}\mathrm{C}^*(1^+;\ 15.1\,\mathrm{MeV}) \tag{11.2}$$

where the 15.11 MeV γ is the signal.

The energy range is chosen to be from 14.4 to 15.4 MeV. Using the GEANT4 based simulation results, 99.7% of γ events will fall in this region. The fiducial volume cut is chosen to be $r < 17.2$ meters, and the efficiency is estimated to be 91.8%. At last, μ and μ daughter cut are applied:

- $\Delta t < 10$ seconds
- $\Delta d < 2$ meters

The efficiency of this cut is evaluated to be 43%. With 7.6 years of data-taking, an exposure of 6×10^4 year×tons can be collected.

The cross-section and *hep* solar neutrino spectrum are shown in Figure 11.10. The estimated signal counts are summarized in Table 11.2.

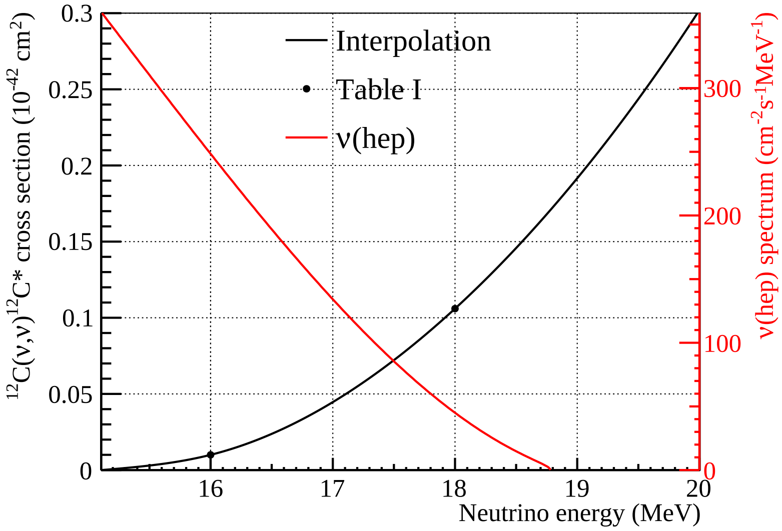


Figure 11.10: Cross-section of inelastic scattering on ^{12}C and energy spectrum of *hep* solar neutrinos. The tabulated cross-section provided in Fukugita, Kohyama, and Kubodera, 1988 is labeled as "Table I", and a spline interpolation of it is shown as the black line.

The ^{8}B solar neutrinos can also conduct such interaction and form backgrounds. The contribution is estimated to be around 0.01 counts

---

[2] Table 6 N. Vinyoles et al., "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, DOI: `10.3847/1538-4357/835/2/202`, arXiv: `1611.09867`

[3] J. Bergström et al., "Updated determination of the solar neutrino fluxes from solar neutrino data", *Journal of High Energy Physics*, vol. 2016, no. 3, p. 132, Mar. 2016, DOI: `10.1007/JHEP03(2016)132`, arXiv: `1601.00972`

[4] The SNO collaboration, "A Search for Neutrinos from the Solar hep Reaction and the Diffuse Supernova Neutrino Background with the Sudbury Neutrino Observatory", *The Astrophysical Journal*, vol. 653, no. 2, pp. 1545–1551, Dec. 2006, DOI: `10.1086/508768`, arXiv: `0607010 [hep-ex]`

[5] M. Fukugita, Y. Kohyama, and K. Kubodera, "Neutrino reaction cross sections on 12C target", *Physics Letters B*, vol. 212, no. 2, pp. 139–144, 1988, DOI: `10.1016/0370-2693(88)90513-8`

Item	Value	Unit	Description
N_{target}	4.38×10^{31}	^{12}C/kton	carbon density
exposure	1.89×10^{9}	kton·seconds	3 years 20 kton
Flux	7.98×10^{3}	cm^{-2}·s^{-1}	B16-GS98
cross-section	2.38×10^{-44}	cm^2	
$E_{\text{th.}}$	7.01%		fraction of above 15.11 MeV
E_{vis}, FV cut	99.71%		
N_{sig}	1.11		counts

Table 11.2: Expected $\nu(hep)$ signal counts.

and thus is negligible. Two other sources of contributing backgrounds include the elastic scattering of ^{8}B solar neutrinos on electrons and the cosmogenic ^{8}B decays. The energy cut $14.4 < E < 15.4$ MeV provides a strong suppression of these two backgrounds. For simplicity, the detector response is simply treated as a Gaussian function with the energy resolution following

$$\frac{\sigma_E}{E} \times 100 = 0.54 + \frac{2.53}{\sqrt{E(\text{MeV})}} \tag{11.3}$$

With such assumption, the efficiencies of the energy cut on these two components are estimated to be 3.3×10^{-6} and 1.9×10^{-3}, respectively. The residual counts of these two backgrounds are summarized in Table 11.3.

Item	Value	Unit	Description
$R_{\nu(^8B)}$	4.5	count/day/kton	
exposure	2.19×10^{7}	kton·days	3 years 20 kton
δ	3.28×10^{-6}		$14.4 < E_{\text{vis}} < 15.4$ MeV
$N_{^8\text{BES}}$	0.32		
$R_{\cos{}^8\text{B}}$	11.25%	count/day/kton	
exposure	2.19×10^{7}	kton·days	3 years 20 kton
ε_μ	1.22×10^{-4}		10s μ efficiency
δ	3.28×10^{-6}		$14.4 < E_{\text{vis}} < 15.4$ MeV
$N_{\cos{}^8\text{B}}$	0.06		

Table 11.3: Expected $\nu(hep)$ background counts.

Considering the expected signal counts is 1.11 and sum of background counts is 0.38, the probability of observing no event is 23% and observing more than 1 event is 42%. If we fix the background counts, in the former case the current upper limit can be improved to 15×10^3 cm^{-2}s^{-1} (90% C.L., Feldman-Cousin), and in the latter case a lower limit might be claimed. The results are summarized in Table 11.4.

Probability	N_{obs}	Lower limit	Upper limit
23%	0	-	14.9
34%	1	-	28.8
25%	2	1.1	40.1
12%	3	5.23	51.1
5%	4	9.0	59.6

Table 11.4: Statistical sensitivity to *hep* neutrino flux. The limit is the 90% Feldman-Cousin confidence interval.

12
Conclusions

This PhD thesis is devoted to the experimental studies of solar neutrinos with Borexino and JUNO. In the context of Borexino with the goal of maximizing the sensitivity to solar neutrinos, we have developed a new fitting method, implemented new models of the detector response functions, evaluated the systematic uncertainties. We have also studied the potential of Borexino on the CNO solar neutrino discovery, and have developed the methods to measure the rate of its main background ^{210}Bi. In JUNO, we have developed the event pile-up rejection algorithm and the energy and vertex reconstruction software optimized for sub-MeV events. The potential of JUNO on the ^{7}Be, CNO, ^{8}B and *hep* solar neutrino flux measurements are evaluated. We have created a multivariate analysis framework `GooStats`. This software shortens the fitting time by a factor of 200 with parallel computing techniques and runs on graphics processing units (GPU). Without this acceleration the analytical multivariate fitting method would have been not feasible. It also makes the method proposed here to evaluate the systematic uncertainties feasible in terms of the computing time. It can be considered as the fundamental tool of the work presented in this thesis. A summary of the main achievement is given in the following, separated for the part of the work done for Borexino and for JUNO.

For Borexino we have tuned the analytical functions used to model the detector energy response against the Monte Carlo software used in Borexino, and have studied the biases of fit results using these functions. It is found that after this tuning the biases on the neutrino interaction rates are negligible compared with the corresponding statistical uncertainties.

We have developed a toy Monte Carlo based method to evaluate the fit model related systematic uncertainties. The systematic uncertainties of measurements are evaluated blindly with this method assuming certain accuracy of detector response models which avoids bias. The fit range is optimized with this method by minimizing the total uncertainties.

We have performed multivariate spectral fits to extract the solar neutrino interaction rates, and have evaluated the solar neutrino fluxes assuming the MSW-LMA model of the neutrino survival probability. The results are summarized in Table 12.1.

Solar ν	Rate [cpd/100 t]	Flux [$\mathrm{cm^{-2}s^{-1}}$]
pp	$134 \pm 10\ ^{+6}_{-10}$	$(6.1 \pm 0.5\ ^{+0.3}_{-0.5}) \times 10^{10}$
^{7}Be	$48.3 \pm 1.1\ ^{+0.4}_{-0.7}$	$(4.99 \pm 0.11\ ^{+0.06}_{-0.08}) \times 10^{9}$
pep (HZ)	$2.43 \pm 0.36\ ^{+0.15}_{-0.22}$	$(1.27 \pm 0.19\ ^{+0.08}_{-0.12}) \times 10^{8}$
pep (LZ)	$2.65 \pm 0.36\ ^{+0.15}_{-0.24}$	$(1.39 \pm 0.19\ ^{+0.08}_{-0.13}) \times 10^{8}$
CNO	< 8.1 (95% C.L.)	$< 7.9 \times 10^{8}$ (95% C.L.)

Table 12.1: Borexino Phase-II results on *pp*, ^{7}Be (862 +384 keV), *pep* and CNO solar ν's: interaction rates and fluxes inferred assuming the MSW-LMA oscillation parameters. The result on *pep* ν's depends on whether we assume HZ or LZ metallicity for CNO ν's.

With these results and the ^{8}B solar neutrino flux measured by Borexino using the data collected in the same period, we have performed phenomenological studies on the solar models and neutrino oscillation models with statistical methods. Using Borexino data only, it is found that, assuming the MSW-LMA model of the neutrino survival probability, the Low Metallicity model (B16-LZ) is rejected with a confidence level of 96.6% (2.2 σ), and the Bayes factor (preference) of the High Metallicity model (B16-HZ) against the Low Metallicity model (B16-LZ) is 4.9. It is also found that, assuming the High Metallicity model (B16-HZ) prediction of the neutrino fluxes, the vacuum oscillation model is rejected with a confidence level of 98.2% (2.4 σ).

We have evaluated the sensitivity of Borexino to CNO solar neutrinos constraining the decay rate of its major background ^{210}Bi. It is found that with ^{210}Bi decay rate constrained within 10% precision, the median sensitivity of Borexino to the CNO solar neutrinos is 3.9 σ assuming the High Metallicity model. This would be the first evidence of the CNO solar neutrinos. The main challenge of this study is to achieve 10% to 20% precision in the ^{210}Bi decay rate measurement.

We have developed methods to measure the ^{210}Bi decay rate with its daughter ^{210}Po. As mentioned before, this study is of fundamental importance for discovering the CNO solar neutrinos in Borexino. The challenge of this method is to get rid of the bias introduced by the ^{210}Po from the inner surface of the nylon vessel migrated into the fiducial volume. We have developed a method to overcome this challenge following the path of the minimum of the spatial distribution of ^{210}Po events.

In order to reduce the systematic uncertainties when testing the fundamental assumptions made in the ^{210}Bi measurement, we have proposed a special dataset enabling only surviving PMTs, and this dataset is found to be useful in various studies, including testing fundamental assumptions made in the method to measure the ^{210}Bi decay rate and improving the light yield monitoring in the Monte Carlo simulation.

For JUNO We have developed fundamental softwares needed for studying the solar neutrinos in JUNO, including the software to generate the energies and momentums of the recoil electrons in the simulation, the algorithms used to suppress event pile-up effect and contribution of dark noise of PMTs and energy/vertex reconstruction software optimized for sub-MeV events.

We have evaluated the signal and backgrounds rates, and the efficiencies of various event selection for detecting the ^{8}B solar neutrinos in JUNO. We demonstrated that JUNO can achieve around 2 to 1 signal-to-background ratio in the visible energy range from 2 to 3 MeV in detecting the ^{8}B solar neutrinos. Combined with Super-Kamiokande results, JUNO can claim a 3 σ evidence of MSW upturn in less than a year of data-taking. This result shows that JUNO has the potential to probe the MSW transition region and would be the unique experiment to do it.

At last, we have evaluated the sensitivity of JUNO to the CNO and ^{7}Be solar neutrinos by performing spectral fits on pseudo-experiment spectra. For ^{7}Be solar neutrinos, when considering only the light yield induced systematic uncertainty and assuming 0.5% precision of the light yield, the expected precision of the ^{7}Be solar neutrino interaction rate is 7% and is dominated by the systematic uncertainty. For CNO solar neutrinos, the measurement is dominated by the systematic uncertainty. When including only the light yield induced systematic uncertainty, if the light yield precision can be as good as 0.1% and if the cosmogenic ^{11}C decay rate can be suppressed by 80%, the total uncertainty of the CNO solar neutrino interaction rate can be as good as 13%. This encouraging result is better than the sensitivity of Borexino. Although this result is encouraging, the requirement on the light yield precision is tight and challenging, and other types of systematic uncertainty such as the energy response non-uniformity statistical uncertainty could also be important. Further studies are needed.

A

Definition of `nhit2`

The motivation of using `nhit2` is to avoid non-uniform binning and bin the data according to the resolution of the detector. Consider we hope when `nhit2` increase by 1 the `nhit` increase by σ, the detector resolution, we have

$$\frac{\mathrm{d\,nhit2}}{1} = \frac{\mathrm{d\,nhit}}{\sigma(\mathrm{nhit})} \tag{A.1}$$

If we parameterize σ(nhit) as

$$\sigma = \sqrt{\mathrm{feq}} \cdot (a \cdot \sqrt{\mathrm{nhit}} + b \cdot \mathrm{nhit}) \tag{A.2}$$

where feq is average of 2000 / (number of live channels).

If we require $\mathrm{nhit2}(0) = 0$, we can solve Eq. (A.1) and get

$$\mathrm{nhit2} = \frac{2}{b} \log\left(1 + \frac{b}{a}\sqrt{\mathrm{nhit}}\right) \tag{A.3}$$

The fit on simulated mono-electron samples from `g4bx2` suggests

$$a = 1.274 \tag{A.4}$$

$$b = -0.01121 \tag{A.5}$$

feq of electrons for the period All (2011 Dec to 2015 May) is 1.289.

Bibliography

Adam, T. et al., "JUNO Conceptual Design Report", p. 328, 2015, arXiv: `1508.07166` (cit. on p. 151).

Agostini, M., *M-STATS: framework for frequentist statistical analysis*, https://github.com/mmatteo/m-stats, 2015 (cit. on pp. 101, 110).

An, F. et al., "Neutrino physics with JUNO", *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, no. 3, p. 030401, 2016, DOI: `10.1088/0954-3899/43/3/030401`, arXiv: `1507.05613` (cit. on pp. 33, 153, 155, 164, 166, 169, 170, 172, 173, 177, 190).

Antonelli, V., L. Miramonti, C. Peña Garay, and A. Serenelli, "Solar Neutrinos", *Advances in High Energy Physics*, vol. 2013, no. supp01a, pp. 1–34, 2013, DOI: `10.1155/2013/351926` (cit. on p. 160).

Asplund, M., N. Grevesse, A. J. Sauval, and P. Scott, "The Chemical Composition of the Sun", *Annual Review of Astronomy and Astrophysics*, vol. 47, no. 1, pp. 481–522, Sept. 2009, DOI: `10.1146/annurev.astro.46.060407.145222`, arXiv: `0909.0948` (cit. on pp. 7, 23, 31).

Bahcall, J. N., "Solar neutrinos. I. theoretical", *Physical Review Letters*, vol. 12, no. 11, pp. 300–302, 1964, DOI: `10.1103/PhysRevLett.12.300` (cit. on pp. 7, 25).

Bahcall, J. N., M. Kamionkowski, and A. Sirlin, "Solar neutrinos: Radiative corrections in neutrino-electron scattering experiments", *Physical Review D*, vol. 51, no. 11, pp. 6146–6158, June 1995, DOI: `10.1103/PhysRevD.51.6146`, arXiv: `arXiv:1011.1669v3` (cit. on pp. 162, 163).

Bahcall, J. N. and C. Peña-Garay, "Solar models and solar neutrino oscillations", *New Journal of Physics*, vol. 6, no. 04, pp. 1–19, 2004, DOI: `10.1088/1367-2630/6/1/063`, arXiv: `0310030 [astro-ph]` (cit. on p. 31).

Bahcall, J. N. and R. K. Ulrich, "Solar models, neutrino experiments, and helioseismology", *Reviews of Modern Physics*, vol. 60, no. 2, pp. 297–372, 1988, DOI: `10.1103/RevModPhys.60.297` (cit. on p. 23).

Bahcall, J., N. Bahcall, and G. Shaviv, "Present Status of the Theoretical Predictions for the ^{37}Cl Solar-Neutrino Experiment", *Physical Review Letters*, vol. 20, no. 21, pp. 1209–1212, 1968, DOI: `10.1103/PhysRevLett.20.1209` (cit. on pp. 23, 25).

Bergström, J., M. C. Gonzalez-Garcia, M. Maltoni, C. Peña-Garay, A. M. Serenelli, and N. Song, "Updated determination of the solar neutrino fluxes from solar neutrino data", *Journal of High En-*

ergy Physics, vol. 2016, no. 3, p. 132, Mar. 2016, DOI: `10.1007/JHEP03(2016)132`, arXiv: `1601.00972` (cit. on pp. 32, 120, 192).

Berlman, I. B., “Efficiency of Energy Transfer in a Solution of PPO in Xylene”, *The Journal of Chemical Physics*, vol. 33, no. 4, pp. 1124–1127, Oct. 1960, DOI: `10.1063/1.1731345` (cit. on p. 64).

Bethe, H. A., “Energy Production in Stars”, *Physical Review*, vol. 55, no. 5, pp. 434–456, Mar. 1939, DOI: `10.1103/PhysRev.55.434` (cit. on p. 120).

Birks, J. B., *The Theory and Practice of Scintillation Counting*, First Edit, Pergamon Press Ltd., 1964, p. 662 (cit. on pp. 64, 66).

Bohm, G. and G. Zech, “Statistics of weighted Poisson events and its applications”, *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 748, pp. 1–6, 2014, DOI: `10.1016/j.nima.2014.02.021`, arXiv: `1309.1287` (cit. on p. 65).

Bravo-Berguño, D., R. Mereu, R. B. Vogelaar, and F. Inzoli, “Fluid-dynamics in the Borexino Neutrino Detector: behavior of a pseudo-stably-stratified, near-equilibrium closed system under asymmetrical, changing boundary conditions”, *Detectors* 2017, arXiv: `1705.09658` (cit. on p. 134).

Capozzi, F., S. W. Li, G. Zhu, and J. F. Beacom, “DUNE as the Next-Generation Solar Neutrino Experiment”, pp. 1–19, Aug. 2018, arXiv: `1808.08232` (cit. on p. 34).

Cerdeño, D. G., J. H. Davis, M. Fairbairn, and A. C. Vincent, “CNO neutrino Grand Prix: The race to solve the solar metallicity problem”, *Journal of Cosmology and Astroparticle Physics*, vol. 2018, no. 4 2018, DOI: `10.1088/1475-7516/2018/04/037` (cit. on p. 120).

Chavarria, A. E., “Study on pep and CNO solar neutrino interaction rates in Borexino”, PhD thesis, Princeton University, 2012 (cit. on pp. 75, 102).

Chen, H. H., “Direct Approach to Resolve the Solar-Neutrino Problem”, *Physical Review Letters*, vol. 55, no. 14, pp. 1534–1536, Sept. 1985, DOI: `10.1103/PhysRevLett.55.1534` (cit. on p. 27).

Christensen-Dalsgaard, J., W. Dappen, S. V. Ajukov, E. R. Anderson, H. M. Antia, S. Basu, V. A. Baturin, G. Berthomieu, B. Chaboyer, S. M. Chitre, A. N. Cox, P. Demarque, J. Donatowicz, W. A. Dziembowski, M. Gabriel, D. O. Gough, D. B. Guenther, J. A. Guzik, J. W. Harvey, F. Hill, G. Houdek, C. A. Iglesias, A. G. Kosovichev, J. W. Leibacher, P. Morel, C. R. Proffitt, J. Provost, J. Reiter, E. J. Rhodes, F. J. Rogers, I. W. Roxburgh, M. J. Thompson, and R. K. Ulrich, “The Current State of Solar Modeling”, *Science*, vol. 272, no. 5266, pp. 1286–1292, May 1996, DOI: `10.1126/science.272.5266.1286` (cit. on pp. 26, 31).

Cleveland, B. T., T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, “Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector”, *The Astrophysical Journal*, vol. 496, no. 1, pp. 505–526, 1998, DOI: `10.1086/305343` (cit. on p. 160).

Cowan, G., K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011, DOI: `10.1140/epjc/s10052-011-1554-0`, arXiv: `1007.1727` (cit. on pp. 59, 114, 126, 154).

D'Angelo, D., "Towards the detection of low energy solar neutrinos in BOREXino: data readout, data reconstruction and background identification", PhD, Munich, Tech. U., 2006, p. 334 (cit. on pp. 64, 102).

Davini, S., "Measurement of the pep and CNO solar neutrino interaction rates in Borexino-I", *The European Physical Journal Plus*, Springer Theses, vol. 128, no. 8, p. 89, Aug. 2013, DOI: `10.1140/epjp/i2013-13089-9` (cit. on pp. 42, 110).

Davis, R., "Solar Neutrinos. II. Experimental", *Physical Review Letters*, vol. 12, no. 11, pp. 303–305, Mar. 1964, DOI: `10.1103/PhysRevLett.12.303` (cit. on pp. 7, 25).

Davis, R., D. S. Harmer, and K. C. Hoffman, "Search for neutrinos from the sun", *Physical Review Letters*, vol. 20, no. 21, pp. 1205–1209, 1968, DOI: `10.1103/PhysRevLett.20.1205`, arXiv: `arXiv:1011.1669v3` (cit. on pp. 7, 23, 25, 160).

Drachnev, I., "New Spectral Analysis of Solar B Neutrino with the Borexino Detector", PhD thesis, Gran Sasso Science Institute, 2016 (cit. on pp. 43, 105, 137, 142, 159, 161).

Fogli, G. L., E. Lisi, A. Marrone, D. Montanino, and A. Palazzo, "Getting the most from the statistical analysis of solar neutrino oscillations", *Physical Review D*, vol. 66, no. 5, pp. 1–30, 2002, DOI: `10.1103/PhysRevD.66.053010`, arXiv: `0206162 [hep-ph]` (cit. on p. 160).

Franco, D., C. Giganti, P. Agnes, L. Agostino, B. Bottino, N. Canci, S. Davini, S. D. Cecco, A. Fan, G. Fiorillo, C. Galbiati, A. M. Goretti, E. V. Hungerford, A. Ianni, A. Ianni, C. Jollet, L. Marini, C. J. Martoff, A. Meregaglia, L. Pagani, M. Pallavicini, E. Pantic, A. Pocar, M. Razeti, A. L. Renshaw, B. Rossi, N. Rossi, Y. Suvorov, G. Testera, A. Tonazzo, H. Wang, and S. Zavatarelli, "Solar neutrino detection in a large volume double-phase liquid argon experiment", *Journal of Cosmology and Astroparticle Physics*, vol. 2016, no. 8 2016, DOI: `10.1088/1475-7516/2016/08/017`, arXiv: `1510.04196` (cit. on p. 34).

Fukuda, Y. et al., "Measurements of the solar neutrino flux from super-kamiokande's first 300 days", *Physical Review Letters*, vol. 81, no. 6, pp. 1158–1162, 1998, DOI: `10.1103/PhysRevLett.81.1158`, arXiv: `9805021 [hep-ex]` (cit. on p. 26).

Fukugita, M., Y. Kohyama, and K. Kubodera, "Neutrino reaction cross sections on 12C target", *Physics Letters B*, vol. 212, no. 2, pp. 139–144, 1988, DOI: `10.1016/0370-2693(88)90513-8` (cit. on p. 192).

Galbiati, C., A. Pocar, D. Franco, A. Ianni, L. Cadonati, and S. Schönert, "Cosmogenic C11 production and sensitivity of organic scintillator detectors to pep and CNO neutrinos", *Physical Review C*

- *Nuclear Physics*, vol. 71, no. 5, pp. 1–11, 2005, DOI: `10.1103/PhysRevC.71.055805`, arXiv: `0411002 [hep-ph]` (cit. on p. 177).

Gallo Rosso, A., C. Mascaretti, A. Palladino, and F. Vissani, "Introduction to neutrino astronomy", *The European Physical Journal Plus*, vol. 133, no. 7, p. 267, 2018, DOI: `10.1140/epjp/i2018-12143-6` (cit. on p. 31).

Gatti, F., V. Lagomarsino, P. Musico, M. Pallavicini, A. Razeto, G. Testera, and S. Vitale, "The Borexino read out electronics and trigger system", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 461, no. 1-3, pp. 474–477, 2001, DOI: `10.1016/S0168-9002(00)01275-4` (cit. on p. 64).

Giunti, C. and C. Kim, *Fundamentals of neutrino physics and astrophysics*, vol. 54, 2, Oxford University Press, USA, 2007, p. 720, DOI: `10.1093/acprof:oso/9780198508717.001.0001` (cit. on pp. 8, 24, 29).

Griffiths, D. J., *Introduction to Elementary Particles*, 2008, DOI: `10.3769/radioisotopes.60.527` (cit. on p. 24).

Guffanti, D., "Measurement of solar neutrino with BOREXINO", PhD thesis, Gran Sasso Science Institute, 2019 (cit. on pp. 91, 114, 117, 119, 121, 147).

Haxton, W., R. Hamish Robertson, and A. M. Serenelli, "Solar Neutrinos: Status and Prospects", *Annual Review of Astronomy and Astrophysics*, vol. 51, no. 1, pp. 21–61, Aug. 2013, DOI: `10.1146/annurev-astro-081811-125539`, arXiv: `1208.5723` (cit. on p. 160).

Herb, S., "The Free Lunch Is Over: A Fundamental Turn Toward Concurrency in Software", *Dr. Dobb's journal*, vol. 30, no. 3, pp. 202–210, 2005 (cit. on p. 51).

Hirata, K., T. Kajita, T. Kifune, K. Kihara, M. Nakahata, K. Nakamura, S. Ohara, Y. Oyama, N. Sato, M. Takita, Y. Totsuka, Y. Yaginuma, M. Mori, A. Suzuki, K. Takahashi, T. Tanimori, M. Yamada, M. Koshiba, T. Suda, K. Miyano, H. Miyata, H. Takei, K. Kaneyuki, H. Nagashima, Y. Suzuki, E. Beier, L. Feldscher, E. Frank, W. Frati, S. Kim, A. Mann, F. Newcomer, Van Berg R, and W. Zhang, "Observation of 8B solar neutrinos in the Kamiokande-II detector.", *Physical review letters*, vol. 63, no. 1, pp. 16–19, 1989, DOI: `10.1103/PhysRevLett.63.16` (cit. on p. 26).

Holanda, P. C. de, W. Liao, and A. Y. Smirnov, "Toward precision measurements in solar neutrinos", *Nuclear Physics B*, vol. 702, no. 1-2, pp. 307–332, 2004, DOI: `10.1016/j.nuclphysb.2004.09.027`, arXiv: `0404042 [hep-ph]` (cit. on p. 29).

Ianni, A., P. Lombardi, G. Ranucci, and O. J. Smirnov, "The measurements of 2200 ETL9351 type photomultipliers for the Borexino experiment with the photomultiplier testing facility at LNGS", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 537, no. 3, pp. 683–697, 2005, DOI: `10.1016/j.nima.2004.07.249`, arXiv: `0406138 [physics]` (cit. on p. 71).

IceCube, T., Fermi-LAT, MAGIC, AGILE, ASAS-SN, HAWC, H. E. S. S, INTEGRAL, Kanata, Kiso, Kapteyn, L. Telescope, Subaru, Swift/NuSTAR, VERITAS, and V.-4. Teams, "Multi-messenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A", *Science*, eaat1378, July 2018, DOI: `10.1126/science.aat1378`, arXiv: `1807.08816` (cit. on pp. 7, 31).

Inc., nVidia, *TESLA K20 GPU ACCELERATOR Board specification*, 2012 (cit. on pp. 51, 56).

Intel, *Intel® Pentium® 4 670 processor with support for HT technology*, 2003 (cit. on p. 51).

– *Intel® Xeon® Processor E5-2630 v3*, 2014 (cit. on p. 51).

King, S. F., "Models of neutrino mass, mixing and CP violation", *Journal of Physics G: Nuclear and Particle Physics*, vol. 42, no. 12 2015, DOI: `10.1088/0954-3899/42/12/123001`, arXiv: `1510.02091` (cit. on p. 24).

Lagomarsino, V. and G. Testera, "A gateless charge integrator for Borexino energy measurement", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 430, no. 2-3, pp. 435–446, July 1999, DOI: `10.1016/S0168-9002(99)00170-9` (cit. on pp. 37, 64).

Li, X.-y., Z.-Y. Deng, L.-J. Wen, W.-D. Li, Z.-Y. You, C.-X. Yu, Y.-M. Zhang, and T. Lin, "Simulation of natural radioactivity backgrounds in the JUNO central detector", *Chinese Physics C*, vol. 40, no. 2, p. 026001, Feb. 2016, DOI: `10.1088/1674-1137/40/2/026001` (cit. on p. 165).

Los Arcos, J. and F. Ortiz, "kB: a code to determine the ionization quenching function Q(E) as a function of the kB parameter", *Computer Physics Communications*, vol. 103, no. 1, pp. 83–94, June 1997, DOI: `10.1016/S0010-4655(97)00012-X` (cit. on p. 66).

Lozza, V. f. t. S. c., "Scintillator phase of the SNO+ experiment" Jan. 2012, DOI: `10.1088/1742-6596/375/1/042050`, arXiv: `1201.6599` (cit. on p. 33).

Maki, Z., M. Nakagawa, and S. Sakata, "Remarks on the Unified Model of Elementary Particles", *Progress of Theoretical Physics*, vol. 28, no. 5, pp. 870–880, Nov. 1962, DOI: `10.1143/PTP.28.870` (cit. on pp. 23, 27).

Malling, D. C. and Fiorucci, S. and Pangilinan, M. and Chapman, J. J. and Faham, C. H. and Verbus, J. R. and Gaitskell, R. J., "Dark Matter Search Backgrounds from Primordial Radionuclide Chain Disequilibrium" May 2013, arXiv: `1305.5183` (cit. on pp. 166, 167, 168).

Maltoni, M. and A. Yu. Smirnov, "Solar neutrinos and neutrino physics", *The European Physical Journal A*, vol. 52, no. 4, p. 87, Apr. 2016, DOI: `10.1140/epja/i2016-16087-0`, arXiv: `1507.05287` (cit. on pp. 8, 30, 160, 162).

Marcocci, S., "Precision Measurement of Solar ν Fluxes with Borexino and Prospects for $0\nu\beta\beta$ Search with Xe -loaded Liquid Scintillators", PhD thesis, Gran Sasso Science Insitute, 2016 (cit. on pp. 42, 76, 93, 103, 111, 133).

Mark Harris, *GPGPU: Beyond Graphics*, Game Developer Conference, 2004 (cit. on p. 51).

McDonald, A. B., "Nobel Lecture: The Sudbury Neutrino Observatory: Observation of flavor change for solar neutrinos", *Reviews of Modern Physics*, vol. 88, no. 3, p. 030502, July 2016, DOI: `10.1103/RevModPhys.88.030502` (cit. on p. 160).

Mikheyev, S. P. and A. Y. Smirnov, "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos", *Sov. J. Nucl. Phys.* Vol. 42, pp. 913–917, 1985 (cit. on pp. 23, 27, 160, 162).

Mosteiro, P., "First measurement of pp neutrinos in real time in the Borexino detector", PhD thesis, Princeton University, 2014 (cit. on p. 103).

Object Management Group, *Unified Modeling Language 2.5.1*, https://www.omg.org/spec/UML/2.5.1/, 2017 (cit. on p. 51).

Particle Data Group, "Review of Particle Physics", *Physical Review D*, vol. 98, no. 3, p. 030001, Aug. 2018, DOI: `10.1103/PhysRevD.98.030001` (cit. on pp. 24, 28, 88, 121, 124).

Parzen, E., "On Estimation of a Probability Density Function and Mode", *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, Sept. 1962, DOI: `10.1214/aoms/1177704472`, arXiv: `arXiv:1011.1669v3` (cit. on p. 147).

Pontecorvo, B., "Inverse beta processes and nonconservation of lepton charge", *Sov.Phys.JETP*, vol. 7, pp. 172–173, 1958 (cit. on pp. 23, 27).

– "Mesonium and anti-mesonium", *Sov.Phys.JETP*, vol. 6, p. 429, 1957 (cit. on pp. 23, 25, 27).

Qian, X., D. A. Dwyer, R. D. McKeown, P. Vogel, W. Wang, and C. Zhang, "Mass hierarchy resolution in reactor anti-neutrino experiments: Parameter degeneracies and detector energy response", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 87, no. 3, pp. 1–7, 2013, DOI: `10.1103/PhysRevD.87.033005`, arXiv: `1208.1551` (cit. on pp. 153, 154).

Raghavan, R. S. and S. Pakvasa, "Probing the nature of the neutrino: The boron solar-neutrino experiment", *Physical Review D*, vol. 37, no. 4, pp. 849–857, Feb. 1988, DOI: `10.1103/PhysRevD.37.849` (cit. on p. 27).

Sakurai, J. J., *Modern quantum mechanics; rev. ed.* Reading, MA: Addison-Wesley, 1994 (cit. on p. 24).

Salas, P. F. de, D. V. Forero, C. A. Ternes, M. Tórtola, and J. W. Valle, "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 782, pp. 633–640, 2018, DOI: `10.1016/j.physletb.2018.06.019`, arXiv: `1708.01186` (cit. on pp. 23, 159, 160, 163).

Saldanha, R. N., "Precision Measurement of the 7 Be Solar Neutrino Interaction Rate in Borexino", PhD thesis, Princeton University, 2012 (cit. on pp. 65, 82, 103).

Seo, S.-h., "Physics Potentials of the Hyper-Kamiokande Second Detector in Korea", pp. 2–6, 2018, arXiv: `1811.06682v1` (cit. on p. 34).

Serenelli, A., "Alive and well: A short review about standard solar models", *The European Physical Journal A*, vol. 52, no. 4, p. 78, Apr. 2016, DOI: `10.1140/epja/i2016-16078-1`, arXiv: `1601.07179` (cit. on p. 31).

Settanta, G., S. M. Mari, C. Martellini, and P. Montini, "e-μ Discrimination at High Energy in the JUNO Detector", no. Cc, pp. 3–6, Jan. 2019, arXiv: `1901.10340` (cit. on p. 156).

Smirnov, O. J., "An approximation of the ideal scintillation detector line shape with a generalized gamma distribution", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 595, no. 2, pp. 410–418, 2008, DOI: `10.1016/j.nima.2008.07.139` (cit. on p. 65).

The Borexino collaboration, "Cosmogenic backgrounds in borexino at 3800 m water-equivalent depth", *Journal of Cosmology and Astroparticle Physics*, vol. 2013, no. 8 2013, DOI: `10.1088/1475-7516/2013/08/049`, arXiv: `1304.7381` (cit. on pp. 169, 170, 177).

– "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 89, no. 11, pp. 1–68, 2014, DOI: `10.1103/PhysRevD.89.112007`, arXiv: `1308.0443` (cit. on pp. 37, 38, 42, 48, 52, 63, 64, 104, 105, 120).

– "First Evidence of *pep* Solar Neutrinos by Direct Detection in Borexino", *Physical Review Letters*, vol. 108, no. 5, p. 051302, Feb. 2012, DOI: `10.1103/PhysRevLett.108.051302`, arXiv: `1110.3230` (cit. on pp. 33, 35, 114).

– "First Simultaneous Precision Spectroscopy of *pp*, ^{7}Be, and *pep* Solar Neutrinos with Borexino Phase-II", pp. 1–8, July 2017, arXiv: `1707.09279` (cit. on pp. 11, 51, 52, 54, 56, 63, 71, 72, 73, 77, 83, 86, 87, 116, 118, 120).

– "Improved measurement of 8B solar neutrinos with 1.5 kt y of Borexino exposure", vol. 016, pp. 1–13, Sept. 2017, arXiv: `1709.00756` (cit. on pp. 8, 35, 38, 116, 117, 159, 161, 177).

– "Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data", *Physical Review D*, vol. 96, no. 9, p. 091103, Nov. 2017, DOI: `10.1103/PhysRevD.96.091103`, arXiv: `1707.09355` (cit. on p. 10).

– "Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data", *Physical Review D*, vol. 96, no. 9, p. 091103, Nov. 2017, DOI: `10.1103/PhysRevD.96.091103`, arXiv: `1707.09355` (cit. on pp. 51, 63).

– "Measurement of the solar ^{8}B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector", *Physical Review D*, vol. 82, no. 3, p. 033006, Aug. 2010, DOI: `10.1103/PhysRevD.82.033006`, arXiv: `0808.2868` (cit. on pp. 33, 35, 38, 159, 161).

– "Neutrinos from the primary proton–proton fusion process in the Sun", *Nature*, vol. 512, no. 7515, pp. 383–386, Aug. 2014, DOI: `10.`

1038/nature13702, arXiv: 1508.05379 (cit. on pp. 32, 33, 35, 76, 108).

The Borexino collaboration, "Observation of geo-neutrinos", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 687, no. 4-5, pp. 299–304, 2010, DOI: 10.1016/j.physletb.2010.03.051, arXiv: 1003.0284 (cit. on p. 33).

– "Precision measurement of the Be7 solar neutrino interaction rate in Borexino", *Physical Review Letters*, vol. 107, no. 14, pp. 1–5, 2011, DOI: 10.1103/PhysRevLett.107.141302, arXiv: 1104.2150 (cit. on pp. 33, 35).

– "The Borexino detector at the Laboratori Nazionali del Gran Sasso", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 600, no. 3, pp. 568–593, 2009, DOI: 10.1016/j.nima.2008.11.076, arXiv: 0806.2400 (cit. on p. 36).

– "The Monte Carlo simulation of the Borexino detector", *Astroparticle Physics*, vol. 97, pp. 136–159, Jan. 2018, DOI: 10.1016/j.astropartphys.2017.10.003, arXiv: 1704.02291 (cit. on pp. 35, 64, 65).

The Daya Bay collaboration, "Measurement of electron antineutrino oscillation based on 1230 days of operation of the Daya Bay experiment", *Physical Review D*, vol. 95, no. 7, p. 072006, Apr. 2017, DOI: 10.1103/PhysRevD.95.072006 (cit. on p. 160).

The GALLEX collaboration, "Solar neutrinos observed by GALLEX at Gran Sasso", *Physics Letters B*, vol. 285, no. 4, pp. 376–389, 1992, DOI: 10.1016/0370-2693(92)91521-A (cit. on p. 26).

The Kamiokande collaboration, "Solar neutrino data covering solar cycle 22", *Physical Review Letters*, vol. 77, no. 9, pp. 1683–1686, 1996, DOI: 10.1103/PhysRevLett.77.1683 (cit. on p. 26).

The KamLAND collaboration, "First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance", *Physical Review Letters*, vol. 90, no. 2, p. 6, 2003, DOI: 10.1103/PhysRevLett.90.021802, arXiv: 0212021 [hep-ex] (cit. on pp. 27, 160).

– "Measurement of the ^{8}B solar neutrino flux with the KamLAND liquid scintillator detector", *Physical Review C*, vol. 84, no. 3, p. 035804, Sept. 2011, DOI: 10.1103/PhysRevC.84.035804, arXiv: 1106.0861 (cit. on p. 173).

– "Production of radioactive isotopes through cosmic muon spallation in KamLAND", *Physical Review C*, vol. 81, no. 2, p. 025807, Feb. 2010, DOI: 10.1103/PhysRevC.81.025807, arXiv: arXiv:0907.0066v2 (cit. on pp. 169, 170, 177).

– "Reactor on-off antineutrino measurement with KamLAND", *Physical Review D - Particles, Fields, Gravitation and Cosmology*, vol. 88, no. 3, pp. 1–10, 2013, DOI: 10.1103/PhysRevD.88.033001, arXiv: 1303.4667 [hep-ex] (cit. on pp. 8, 30, 160, 161).

The MARCO collaboration, "Measurement of the atmospheric neutrino-induced upgoing muon flux using MACRO", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 434, no. 3-4, pp. 451–457, 1998, doi: 10.1016/S0370-2693(98)00885-5 (cit. on p. 26).

The SAGE collaboration, "Measurement of the solar neutrino capture rate with gallium metal", *Physical Review C - Nuclear Physics*, vol. 60, no. 5, p. 32, 1999, doi: 10.1103/PhysRevC.60.055801, arXiv: 0901.2200 (cit. on p. 26).

The SNO collaboration, "A Search for Neutrinos from the Solar hep Reaction and the Diffuse Supernova Neutrino Background with the Sudbury Neutrino Observatory", *The Astrophysical Journal*, vol. 653, no. 2, pp. 1545–1551, Dec. 2006, doi: 10.1086/508768, arXiv: 0607010 [hep-ex] (cit. on p. 192).

– "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory", *Physical Review Letters*, vol. 89, no. 1, pp. 1–6, 2002, doi: 10.1103/PhysRevLett.89.011301, arXiv: 0204008 [nucl-ex] (cit. on pp. 23, 27).

– "Measurement of the Rate of $\nu_e + d \top + p + e^-$ Interactions Produced by ^{8}B Solar Neutrinos at the Sudbury Neutrino Observatory", *Physical Review Letters*, vol. 87, no. 7, p. 071301, July 2001, doi: 10.1103/PhysRevLett.87.071301 (cit. on pp. 7, 23).

The Super-Kamiokande collaboration, "Evidence for oscillation of atmospheric neutrinos", *Physical Review Letters*, vol. 81, no. 8, pp. 1562–1567, 1998, doi: 10.1103/PhysRevLett.81.1562, arXiv: 9807003 [hep-ex] (cit. on p. 26).

– "Solar neutrino measurements in Super-Kamiokande-IV", *Physical Review D*, vol. 94, no. 5, pp. 1–33, 2016, doi: 10.1103/PhysRevD.94.052010, arXiv: 1606.07538 (cit. on pp. 7, 8, 30, 159, 160, 161, 182).

The T2K collaboration, "Measurement of neutrino and antineutrino oscillations by the T2K experiment including a new additional sample of veinteractions at the far detector", *Physical Review D*, vol. 96, no. 9, pp. 1–49, 2017, doi: 10.1103/PhysRevD.96.092006, arXiv: 1707.01048 (cit. on p. 160).

Villante, F. L., A. Ianni, F. Lombardi, G. Pagliaroli, and F. Vissani, "A step toward CNO solar neutrino detection in liquid scintillators", *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, vol. 701, no. 3, pp. 336–341, 2011, doi: 10.1016/j.physletb.2011.05.068, arXiv: 1104.1335 (cit. on pp. 120, 131, 132).

Vinyoles, N., A. M. Serenelli, F. L. Villante, S. Basu, J. Bergström, M. C. Gonzalez-Garcia, M. Maltoni, C. Peña-Garay, and N. Song, "A New Generation of Standard Solar Models", *The Astrophysical Journal*, vol. 835, no. 2, p. 202, Jan. 2017, doi: 10.3847/1538-4357/835/2/202, arXiv: 1611.09867 (cit. on pp. 31, 32, 35, 115, 124, 132, 162, 192).

Vissani, F., "Joint analysis of Borexino and SNO solar neutrino data and reconstruction of the survival probability", pp. 1–12, Sept. 2017, arXiv: 1709.05813 (cit. on pp. 30, 163).

Vissani, F., "Joint analysis of borexino and sno solar neutrino data and reconstruction of the survival probability", *Nuclear Physics and Atomic Energy*, vol. 18, no. 4, pp. 303–312, 2017, DOI: `10.1207/s15327868ms0402_3`, arXiv: `1709.05813` (cit. on p. 35).

– "Solar neutrino physics on the beginning of 2017", *Nuclear Physics and Atomic Energy*, vol. 18, no. 1, pp. 5–12, Mar. 2017, DOI: `10.15407/jnpae2017.01.005` (cit. on p. 163).

Westerdale, S. and P. D. Meyers, "Radiogenic neutron yield calculations for low-background experiments", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 875, no. August, pp. 57–64, 2017, DOI: `10.1016/j.nima.2017.09.007`, arXiv: `1702.02465` (cit. on p. 165).

Winkelmann, J., "Diffusion coefficient of acetic acid in benzene", in: *Diffusion in Gases, Liquids and Electrolytes: Nonelectrolyte Liquids and Liquid Mixtures - Part 1: Pure Liquids and Solute in Solvent Systems*, ed. by M. D. Lechner, Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 623–623, DOI: `10.1007/978-3-540-73735-3_404` (cit. on p. 144).

Winter, W. T., S. J. Freedman, K. E. Rehm, and J. P. Schiffer, "The B8 neutrino spectrum", *Physical Review C - Nuclear Physics*, vol. 73, no. 2, pp. 1–15, 2006, DOI: `10.1103/PhysRevC.73.025503`, arXiv: `0406019 [nucl-ex]` (cit. on p. 162).

Wolfenstein, L., "Neutrino oscillations in matter", *Physical Review D*, vol. 17, no. 9, pp. 2369–2374, May 1978, DOI: `10.1103/PhysRevD.17.2369` (cit. on pp. 23, 27, 160, 162).

Yang, M. and Y. Cheng, "Energy Non-linearity study at Daya Bay", *International Journal of Modern Physics: Conference Series*, vol. 31, p. 1460312, Mar. 2014, DOI: `10.1142/S2010194514603123`, arXiv: `1403.1550` (cit. on p. 75).

Zeyuan, Y., "Study of the calibration, reconstruction and neutrino disappearance at Daya Bay neutrino experiment", PhD thesis, Chinese Acamedy of Science University, 2013 (cit. on p. 75).

Acknowledgement

In the end, I would like to thank everybody who has accompanied me for these three and half years.

I would like to thank my supervisor Nicola Rossi. You have trained me to think in a scientist with your insights. You have also given me enough freedom to let me grow. Staying with you reminds me the joy of research and the desire to become a scientist. You have also given me strength when I feel depressed.

I would like to thank my supervisor Gioacchino Ranucci. You are always supporting me strongly without any doubt. Without these support, I could not imagine how I can finish the JUNO ^{8}B studies.

I would like to thank my supervisor Matteo Agostini. Your understanding of physics is the wonderland to me. You are the man who lighten my life.

I would like to thank Ilia Drachnev, who was my roommate. I learned Borexino from you, however you left before I realized one year was so short.

I would like to thank Daniele Guffanti. It is hard to find a friend like you who could understand my non-sense jokes.. And, we are real friends who fight the same battle in the same trenches. You understand that!

I would like to thank Giovanni Ciani. I remember on my first night of L'Aquila, we were in the GSSI hall, you were the first one who walked towards me and spoke with me.

I would like to thank Irene Nutini. In Neutrino 2016 when we were in London, it was the best trip I ever made. Not only because of the beautiful city, but also the memory of staying with friends.

I would like to thank Gialex Anastasi. I spoke many jokes, but I always remember the joke about the detector size when we were having the beer near the piazza duomo LOL!

I would like to thank Lorenzo Aiello, Silvia Celli, Valerio Gentile, Imran Khan, Vsevolod Orekhov, Andrea Gallo Rosso and Akshat Singhal. When I first came to Italy, I felt lonely. After a few times of dinner together, I felt we are like families. It was so good to be with you cool physicists. I would like to also thank all GSSI staffs.

I would like to thank Chiara Ghiano, Paolo Saggese, Lee Stokes, Lidio Pietrofaccia, Valentino Di Marcello and Antonio Di Ludovico. We had lots of good memories at LNGS and at Milan. I would like to thank Aldo Ianni, Andrea Ianni and all my LNGS colleagues. I would like to thank also all my Borexino and JUNO colleagues. It is

a pleasure to work with you.

I would like to thank my referees Francesco Villante, Gemma Testera and Alberto Garfagnini for devoting your time reading and commenting my thesis.

I would like to thank Francesco Vissani. I came to GSSI because of you. In 2016 when I was experiencing the cultural shock, you helped me out.

I would like to thank my relatives, especially my father, who is supporting me from six thousand kilometers away.

I would like to thank Le. I would like to thank also Le's mom, Penny, Léa and Eltonsphowe for helping me when I was depressed .

I feel sorry not able to list everyone whom I would like to express my gratitude to. Many thanks to you.