UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics

Final Dissertation

Probing the atmospheric mass splitting Δm^2_{31} with reactor antineutrino oscillations at JUNO

Thesis supervisor

Prof. Alberto Garfagnini

Thesis co-supervisor

Dr. Andrea Serafini

Candidate

Vanessa Cerrone

Academic Year 2022/2023

Abstract

Neutrinos, elementary particles that interact only weakly with matter, constitute the first evidence of physics beyond the Standard Model of particle physics. The study of their properties, both on experimental and theoretical grounds, is one of the most active directions within particle physics. In the last few decades, many experiments provided evidence of neutrinos' existence and revealed peculiar properties, such as their non-zero but extremely small masses and the possibility of oscillation among different flavors. To date, almost all neutrino data collected with accelerator, solar, atmospheric, and reactor neutrinos can be explained within the standard three-neutrino oscillation paradigm, where a total of six parameters are needed to fully describe neutrino oscillations: three mixing angles (θ_{12}, θ_{23} and θ_{13}), one Dirac CP phase ($\delta_{\rm CP}$), and two independent mass squared differences (Δm^2_{21} and Δm^2_{31}, or equivalently Δm^2_{32}). Despite significant advancements in neutrino experiments and their precision in recent years, many properties of neutrinos still remain unknown, including the nature of neutrinos (Dirac or Majorana), the existence of CP violation in the leptonic sector, and the scale of the neutrino mass eigenstates, commonly referred to as neutrino Mass Ordering (MO).

The Jiangmen Underground Neutrino Observatory (JUNO), situated in South China, is a forthcoming multi-purpose neutrino experiment. With a substantial active mass of 20 kton, it is foreseen to become the World's largest liquid scintillator-based neutrino detector in the next decade. JUNO's primary goal is to determine the neutrino MO using reactor antineutrinos ($\overline{\nu}_e$) emitted from two adjacent nuclear power plants at a 52.5 km baseline from the experimental site. The oscillation pattern observed in JUNO exhibits subtle variations depending on the neutrino MO, thus providing sensitivity to this parameter. In JUNO's location, the energy spectrum will be distorted by a slow (low frequency) oscillation driven by Δm^2_{21} and modulated by $\sin^2(2\theta_{12})$, as well as by a fast (high frequency) oscillation regulated by Δm^2_{31} and modulated by $\sin^2(2\theta_{13})$. Moreover, it will be the first experiment to simultaneously observe neutrino oscillations from two different frequencies, and multiple oscillation cycles of the atmospheric mass splitting Δm^2_{31}. In addition, JUNO's data-taking period of less than one year is sufficient to establish its dominance in the global precision of three of these parameters. Notably, this is particularly evident in the case of the atmospheric mass splitting Δm^2_{31}, for which sub-percent precision can be attained within just 100 days of data taking.

This thesis focuses on investigating JUNO's sensitivity to oscillation parameters, with a specific emphasis on Δm^2_{31}. The initial part involves generating the expected energy spectrum in JUNO using Monte Carlo simulations. Subsequently, by fitting a pseudo-dataset, we evaluate the impact of statistical and systematic uncertainties on the estimation of oscillation parameters.

Chapter 1 is dedicated to introducing the phenomenology of neutrino mixing and oscillations, covering the three-flavor framework in both vacuum and matter. Among the several open questions in the field, a brief overview of the current status on mass ordering measurements will be given. A concise overview of the present state-of-the-art understanding of oscillation parameters will conclude this chapter.

Chapter 2 provides a comprehensive overview of the JUNO experiment, offering in-depth details regarding its main detector components. Emphasis is placed on the physics-driven requirements and how they are effectively addressed through the experiment's design. Subsequently, the formalism described earlier is applied to the specific case of reactor antineutrino oscillations, highlighting the key aspects that are most relevant to this study.

Chapter 3 thoroughly outlines the entire process involving reactor antineutrinos, spanning from their production at the source to their interaction and detection in the liquid scintillator. This chapter

begins by presenting information on reactor antineutrino flux models and the necessary corrections involved in their estimation. It then focuses on describing the Inverse Beta Decay (IBD) reaction, which serves as the "golden channel" for electron antineutrinos detection at the MeV energy scale, thus playing a crucial role in JUNO. Furthermore, the chapter inquires the detector response, leading to the derivation of the expected spectrum in JUNO.
Chapter 4 describes how the main spectral components are obtained through the official simulation software and how the residual event count after all section cuts is calculated. A part of this work was devoted to the development of an algorithm to implement time correlation analysis, a process that involves selecting two events that occur within a defined time and spatial window. In this section, the distinctive temporal and spatial signature yielded by the IBD reaction is leveraged to differentiate IBD events from the dominant background. Then, by incorporating all the relevant information gathered in the preceding chapters, an Asimov pseudo-dataset is constructed to simulate the nominal energy spectrum at JUNO, thus modeling the antineutrino incoming flux, the oscillation probability, the detector response, and all spectral components. Finally, Chapter 5 is dedicated to the investigation of JUNO's sensitivity to oscillation parameters, placing a particular emphasis on the atmospheric mass splitting. The latter serves as a crucial input for global fits and analyses that combine data from various neutrino experiments, including those involving solar, reactor, and accelerator neutrinos. Furthermore, the chapter presents the results of the analysis, highlighting JUNO's remarkable sensitivity in measuring three oscillation parameters with an unprecedented level of precision.

Contents

Abstract **iii**

1 Introduction **1**
1.1 Neutrinos in the Standard Model of Particle Physics . . . 1
1.2 Phenomenology of neutrino mixing and oscillations . . . 2
1.2.1 Neutrino oscillations in matter . . . 6
1.2.2 Neutrino Mass Ordering . . . 8
1.2.3 Current knowledge on oscillation parameters . . . 8

2 The JUNO experiment **11**
2.1 JUNO detector design . . . 12
2.1.1 JUNO main detector . . . 13
2.1.2 JUNO liquid scintillator . . . 14
2.1.3 JUNO calibration system . . . 15
2.2 JUNO simulation and event reconstruction framework . . . 16
2.2.1 Event and detector simulation . . . 18
2.2.2 Event reconstruction . . . 18
2.3 High precision neutrino oscillation physics in JUNO . . . 19
2.3.1 Antineutrino survival probability . . . 20
2.3.2 Matter effects in JUNO . . . 20

3 Reactor antineutrino production and detection **23**
3.1 Reactor antineutrino flux . . . 23
3.2 IBD reaction and detection . . . 25
3.2.1 Neutron recoil effects . . . 27
3.3 Detector response . . . 28
3.3.1 Energy non-linearity . . . 28
3.3.2 Energy resolution . . . 31

4 Reactor antineutrino selection: signal and backgrounds **35**
4.1 Time correlation analysis . . . 35
4.2 IBD selection . . . 37
4.2.1 IBD signal selection efficiency . . . 38
4.3 Background estimation . . . 41
4.3.1 Accidental background . . . 41
4.3.1.1 Accidental coincidences rate . . . 42
4.3.2 Cosmogenic isotopes . . . 44
4.3.3 Geo-neutrinos . . . 46
4.3.4 Other background sources . . . 48
4.4 Expected energy spectrum . . . 49

5 JUNO's sensitivity to oscillation parameters **51**
5.1 Analysis strategy . . . 51
5.2 Likelihood function and spectral fit . . . 53
5.2.1 Relationship between chi-squared and log-likelihood . . . 54
5.2.2 Correlation among parameters . . . 56
5.3 Impact of nuisance parameters on Δm^2_{31} . . . 57
5.3.1 Nuisances affecting the precision . . . 59
5.3.2 Nuisances affecting the accuracy . . . 61
5.3.3 Profile likelihood scans . . . 63
5.4 Systematic uncertainties . . . 63
5.4.1 Reactor flux-related uncertainties . . . 65
5.4.2 Observed spectrum and detector response uncertainties . . . 67
5.5 Sensitivity determination . . . 69
5.5.1 Impact of the reference spectrum on Δm^2_{31} . . . 72
5.5.2 Neutrino mass ordering hypothesis impact . . . 73

6 Conclusions **75**

Bibliography **77**

A Electron antineutrino survival probability in matter **83**

B Detailed IBD kinematics **85**
B.1 Cross section . . . 86

C Bayesian inference **87**

D Likelihood profile scans results **89**

E Sensitivity results **91**

F Neutrino MO synergy between JUNO and LBL experiments **95**

List of Figures

1.1 Feynman diagrams of the coherent forward elastic scattering processes that generate the charged current potential through W-boson exchange and the NC potential through Z-boson exchange [1]. 7
1.2 The two possible neutrino mass ordering schemes. 8

2.1 Location of the JUNO experiment in South China [21]. The main 20 kton LS detector is indicated in blue, at a baseline of $\sim 52.5\,\text{km}$ from six 2.9 GW_{th} reactor cores in the Yangjiang NPP and two 4.6 GW_{th} cores in the Taishan NPP. The 2.8 ton JUNO-TAO detector [31], indicated in red, is located about 30 meters away from one of the Taishan reactor cores. 11
2.2 Top: Expected antineutrino spectrum at JUNO in the ideal configuration (without detector-related effects) with and without oscillations (using global parameters in Table 1.2). Bottom: $\overline{\nu}_e$ survival probability with a 52.5 km baseline. To enhance visual clarity, the oscillation induced by the solar term is depicted using $\sin^2\theta_{12} = 0.28$ instead of $\sin^2\theta_{12} = 0.303$. The mass splittings and mixing angles modulating the spectrum are indicated as well. 12
2.3 Schematic representation of the main JUNO detector [21]. The 20 kton acrylic sphere is submerged in a 44-meters-deep water pool and instrumented with 17612 20-inch and 25600 3-inch photomultiplier tubes. The upper part of the detector consists of a Top Tracker system with three layers of plastic scintillator and a calibration house. . 13
2.4 Schematic representation of the calibration system (not to scale), taken from [42]. The different subsystems are shown: the Automatic Calibration Unit (ACU), two Cable Loop Systems (CLSs), the Guide Tube (GT), and the Remotely Operated Vehicle (ROV). The red points represent a source assembly. The AURORA is an auxiliary laser diode system to monitor the attenuation and scattering length of the LS. 17
2.5 Detector simulation scheme in SNiPER [49], from the generator of the physics process to the number of photoelectrons released at the photocatode. 18
2.6 Absolute and relative (in the inset) difference between $\mathcal{P}(\overline{\nu}_e \to \overline{\nu}_e)$ in matter and vacuum, for a fixed baseline $L = 52.5\,\text{km}$. 20
2.7 Antineutrino survival probability according to Equation 2.4 at fixed energy $E_{\overline{\nu}_e} = 3\,\text{MeV}$ and variable baseline. To enhance visual clarity, the oscillation induced by the solar term is illustrated using $\sin^2\theta_{12} = 0.28$ instead of $\sin^2\theta_{12} = 0.303$. Two region of interest for reactor antineutrino experiments are highlighted, corresponding to short baseline and medium baseline facilities, e.g., Daya Baya [14] and JUNO [30], respectively. 21

3.1 (a) Contribution from the single reactor cores to the expected reconstructed energy spectrum, for 6 years exposure and 11/12 duty cycle power. Yangjiang is indicated with "YJ", Taishan with "TS" and Daya Bay with "DYB". The inset plot shows the cumulative contribution and the resulting total one (black line). (b) Additional contributions to the reconstructed energy spectrum due to non-equilibrium effect and Spent Nuclear Fuel (SNF) [60]. 25
3.2 Schematic illustration of an IBD reaction in the LS. The electron antineutrino interacts with a proton (p) in the LS, creating a positron (e^+) and a neutron (n). The positron deposits its energy and annihilates into two 0.511 MeV photons (γ), producing a prompt signal. After a random walk, the neutron is captured mainly by a free proton in the LS, emitting a 2.22 MeV γ-ray, and giving rise to a delayed signal. The time-charge diagram is not in scale. 26
3.3 (a) Reactor antineutrinos yield per fission and total IBD cross section. (b) Expected $\overline{\nu}_e$ energy spectrum, normalized to one, obtained using the Huber-Mueller model without any additional corrections. 27
3.4 (a) Fractional recoil effects in terms of antineutrino energy, in three different cases: no recoil, mid recoil approximation, and full recoil. (b) Energy spread due to nucleon recoil compared with energy resolution expected in JUNO. The definition of $\Delta E_{\rm dep}$ can be found in Equation 3.7. 28
3.5 Energy conversion scheme. The energy of antineutrinos, denoted as $E_{\bar{\nu}_e}$, undergoes a conversion to positron deposited energy, represented as $E_{\rm dep}$, through the IBD reaction. Subsequently, the energy spectrum experiences distortion due to the non-linearity of the liquid scintillator (LSNL), leading to the measurement of visible energy $E_{\rm vis}$ in terms of detected photoelectrons. Finally, the finite energy resolution introduces additional smearing, resulting in the reconstructed energy $E_{\rm rec}$. 29
3.6 The expected prompt energy spectra of JUNO with and without the different detector response effects, i.e., liquid scintillator non-linearity (NL) and energy resolution (Res). The detector response curves used in this analysis, i.e., LSNL and resolution, are shown in the insets as a function of deposited and visible energy, respectively. 30

4.1 (1) Pure IBD event in two trigger windows. (2) IBD pair interposed by a background event. (3) IBD event interposed by another IBD pair. 36
4.2 Bi-dimensional distribution of prompt (left) and delayed (right) events energy and associated distance from the center of the detector (r^3). The radius corresponding to the fiducial volume cut is indicated by the dashed line. 39
4.3 Prompt (left) and delayed (right) events energy distributions, with selection areas indicated by dashed lines. 39
4.4 Δt (left) and Δr (right) distribution between prompt and delayed events. The corresponding PDFs are indicated by the red lines. 40
4.5 Δt and Δr distribution between prompt and delayed events. 40
4.6 Radiogenic *singles* energy spectrum (left) and distribution of events as a function of volume and energy (right). 42
4.7 3-dimensional distribution of radioactive events inside the full detector volume, without (left) and with (right) energy cut at 0.7 MeV. The color scale indicates the distance from the detector center. 43
4.8 Distribution of time intervals Δt between two consecutive radiogenic events above the 0.7 MeV threshold. 44

4.9 (a) Accidental coincidences and reactor IBD daily rates after all selection criteria, as a function of different values of FV radii. (b) IBD daily rate as a function of FV volume radii; both the simulated data (dashed line) and the cubic interpolation curve (solid line) are shown. 45
4.10 Δt, Δr (a) and energy distributions (b) for ^{8}He-induced prompt-delayed correlated pairs. 46
4.11 Δt, Δr (a) and energy distributions (b) for ^{9}Li-induced prompt-delayed correlated pairs. 47
4.12 Δt, Δr (a) and energy distributions (b) for ^{238}U geo-neutrinos prompt-delayed correlated pairs. 47
4.13 Δt, Δr (a) and energy distributions (b) for ^{232}Th geo-neutrinos prompt-delayed correlated pairs. 48
4.14 Reconstructed energy spectra at JUNO for normal ordering, 6 years exposure at 11/12 duty cycle reactor power. The background spectra are stacked in the filled histogram and shown separately in the inset. The background rates were calculated after applying the selection criteria, as described in Table 4.2. 50

5.1 Reactor and terrestrial antineutrino MC-driven energy spectra. The corresponding smoothed PDFs are reported in the insets. 52
5.2 Cosmogenic ^{9}Li/^{8}He (a) and accidental coincidences (b) MC-driven energy spectra. The corresponding smoothed PDFs are reported in the insets. 53
5.3 $\Delta\chi^2$ curves for different test statistics: log-likelihood, χ^2_{Pearson}, χ^2_{Neyman}, and combined χ^2_{CNP}, obtained with 400 pseudo-experiments. The dashed line indicates the true value of the parameter of interest. 55
5.4 Correlation matrix obtained keeping all parameters free. 56
5.5 Examples of Asimov (a) and pseudo-experiment (b) datasets for 6 years exposure. The best-fit line is also reported. 58
5.6 Two-dimensional 68.27%, 95.45%, and 99.73% confidence level contours for one parameter of interest ($\theta = \Delta m^2_{21}$) and one nuisance ($\eta = N_{\text{U}}$). In the secondary axis, the conditional and profile likelihoods are compared: the former (in orange) corresponds to the narrower profile, while the latter is significantly wider due to the known correlation between the solar mass splitting and the number of Uranium geoneutrinos N_{U}. 59
5.7 Conditional likelihood profile considering $\theta = \Delta m^2_{31}$ as parameter of interest and fixing the nuisance parameters related to the energy resolution, i.e., $\eta = (a, b, c)$. . . 60
5.8 Relative precision on Δm^2_{31} as a function of the bias introduced on resolution parameters, obtained from the conditional likelihood profiles. 60
5.9 (a) Conditional likelihood profile considering $\theta = \Delta m^2_{31}$ as parameter of interest and fixing the number of reactor antineutrino events. (b) Relative precision on Δm^2_{31} for different values of N_{rea}, i.e., as a function of the bias introduced on N_{rea}. 61
5.10 Conditional likelihood profiles considering $\theta = \Delta m^2_{31}$ as parameter of interest and fixing the nuisance parameters related to the energy non-linearity, i.e., $\eta = (p_0, p_1, p_2, p_3)$. 62
5.11 $\Delta\chi^2$ profiles for $\theta = \Delta m^2_{31}$, showing the conditional likelihood and the profile likelihood. In (a) the conditional likelihood is compared with the profile likelihood with constrained non-linearity (NL) parameters; (b) reports the profiles for constrained resolution parameters. 64
5.12 Antineutrino spectrum b2b uncertainty from different models, rescaled for a bin width of 10 keV, according to Equation 5.20: the satellite detector TAO-based estimation is taken from [31], while the Daya Bay (DYB) is based on [88]. 66

5.13 (a) Nominal pull curves for liquid scintillator non-linearity (LSNL). (b) Ratio of the LSNL curves to the nominal LSNL curve, as a function of deposited energy. 68
5.14 Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). The empty boxes represent the statistical uncertainty. The impact of each source of systematic error, represented by the filled boxes, is assessed by enabling the corresponding uncertainty together with the statistical uncertainty, and removing the latter. The last row of each table, displayed in a lighter color, represents the total uncertainty obtained by simultaneously considering all sources of statistical and systematic error. The presented uncertainties correspond to six years of JUNO data with 11/12 duty cycle. 70
5.15 Relative precision of the oscillation parameters as a function of JUNO exposure. The dotted and solid lines represent the statistical and total uncertainties, respectively. . 71
5.16 Comparison of $\Delta\chi^2$ profiles of oscillation parameters: the dashed curve indicates the current results provided by NuFIT 2022 [27], while the solid ones represent JUNO projection with 100 days (in blue) and 6 years (in orange) of data-taking. 72
5.17 Bi-dimensional 68.27%, 95.45%, and 99.73% confidence level (C.L.) contours for all oscillation parameters, corresponding to 1σ, 2σ, and 3σ C.L. regions, for 6 years exposure. The black star and the dotted lines represent the best-fit values as well as the input values of the oscillation parameters. The correlation coefficient ρ between pair of parameters is also indicated. 73
5.18 Relative precision on Δm^2_{31}, for 100 days (a) and 20 years (b) exposures, with 11/12 duty cycle. The bars with different colors correspond to two reference antineutrino spectrum models, Daya Bay (gray) and TAO (black), and corresponding bin-to-bin (b2b) uncertainties. 74

6.1 1-dimensional $\Delta\chi^2$ profiles for Δm^2_{31}: (a) for JUNO (60 days) with NO Asimov data fitted with the correct MO hypothesis (NO), LBL with NO hypothesis, and combining the two. (b) for JUNO (60 days) with NO Asimov data fitted with the wrong MO hypothesis (IO), LBL with IO hypothesis, and combining the two. Details on the procedure employed to include the LBL external information on $\Delta m^2_{3\ell}$ (with $\ell = 1$ for NO and $\ell = 2$ for IO hypotheses) in JUNO's analysis is provided in Appendix F. 76

D.1 Conditional likelihood scans for $\theta = \Delta m^2_{31}$ and all nuisance parameters. The profiles are obtained using Equation 5.14. 89
D.2 Conditional likelihood scans for $\theta = \Delta m^2_{31}$ and all nuisance parameters. The profiles are obtained using Equation 5.15. 90

E.1 Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). All uncertainties correspond to 100 days of JUNO data with 11/12 duty cycle. 91
E.2 Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). All uncertainties correspond to 1 year of JUNO data with 11/12 duty cycle. 92
E.3 Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). All uncertainties correspond to 20 years of JUNO data with 11/12 duty cycle. 93

List of Tables

1.1 Types of neutrino oscillation experiments with their typical source–detector distance, energy, and sensitivity to Δm^2. The acronyms SBL and LBL stand for Short Baseline and Long Baseline, respectively. 5

1.2 Current best fit values on oscillation parameters from NuFIT 2022 global analysis [27], without Super-Kamiokande atmospheric data. Note that $\Delta m^2_{3\ell} = \Delta m^2_{31} > 0$ for Normal Ordering and $\Delta m^2_{3\ell} = \Delta m^2_{32} < 0$ for Inverted Ordering. 9

2.1 List of calibration sources and corresponding radioactive processes considered by the JUNO collaboration [42]. 16

3.1 Characteristics of NPPs and their reactor cores considered in this analysis: the two closest ones to JUNO, Taishan and Yangjiang, at an approximate distance of 52.5 km, and the next closest, Daya Bay. The average distance and the total average thermal power of each NPP are indicated in the corresponding row. The expected daily IBD rates are estimated considering the baselines, full thermal power of the reactors, and neutrino oscillation parameters in Table 1.2. 24

3.2 Parameters related to the normalization of the number of detected IBDs. 31

4.1 Summary of cumulative IBD selection efficiencies. The reported IBD rates refer to the expected events per day after the selection criteria are progressively applied. These rates are calculated for nominal reactor power rate. 41

4.2 Summary of background rates after all selection cuts. Rate and shape (bin-to-bin uncorrelated) uncertainties are reported. All shape uncertainties are given for 36 keV bins, with the exception of global reactors (20 keV bins). 49

5.1 Oscillated spectrum fit parameters. The cyan highlighted entries are associated with detector-related parameters, therefore efficiency, energy resolution, and non-linearity. The α_i parameters are used to model the non-linearity response, and their meaning will be explained in section 5.4. The oscillation parameters and the matter density ρ are indicated in orange. In green, we denote the quantities associated with the reactor antineutrino flux: N_{rea} is the total predicted number of reactor antineutrinos, with entries from 24 to 32 representing the weight of each core; N_{SNF} and N_{noneq} indicate the corrections due to spent nuclear fuel and non-equilibrium. The background components are hued in purple. 54

5.2 Summary of systematic uncertainties which impact the JUNO detector and their bin-to-bin correlation considered in this analysis. 68

5.3 Current relative precision on oscillation parameters from global analysis (NuFIT 5.2) [27], and JUNO (different exposures) in this analysis, assuming NO. 72

5.4 Relative uncertainty and best fit values of the four oscillation parameters after 6 years of data taking at 11/12 duty cycle, for different mass ordering hypotheses. The best fit values for the right mass ordering are given by the input parameters in Table 1.2. 74

Chapter 1

Introduction

Neutrinos, elementary particles that interact only weakly with matter, have played a pivotal role in unveiling the limitations of the Standard Model (SM) of particle physics [1]. They were first proposed by Wolfgang Pauli in 1930 to explain the apparent energy and momentum conservation violation observed in beta decay. In the following decades, many experiments provided evidence of neutrinos' existence and revealed peculiar properties, such as their non-zero but extremely small masses and the possibility of oscillation among different flavors.
The study of neutrinos has become one of the most active directions within particle physics. Neutrino experiments, both on the experimental and theoretical fronts, have been devised to investigate their fundamental properties, such as mass hierarchy, CP violation, and the nature of their masses (whether they are Dirac or Majorana particles).

1.1 Neutrinos in the Standard Model of Particle Physics

The Standard Model of particle physics [1–3] is a quantum field theory that describes the fundamental particles and their interactions. It is based on the gauge symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where the subscripts C, L, and Y represent the color charge, left-handed chirality, and hypercharge, respectively.
The gauge symmetry group uniquely determines the types of interactions and the number of vector gauge bosons associated with the generators of the group. In the SM, there are eight massless gluons, corresponding to the eight generators of the $SU(3)_C$ group, which mediate strong interactions. Additionally, there are four gauge bosons associated with the electroweak interactions. Among them, three are massive ($W^{\pm}$ and Z^0 bosons), while one is massless, representing the photon (γ). These gauge bosons arise from the three generators of the $SU(2)_L$ group and one generator of the $U(1)_Y$ group. The electroweak part of the SM Lagrangian is determined by the symmetry group $SU(2)_L \otimes U(1)_Y$, which governs the interactions of various particles, including neutrinos [1, 3]. In the original formulation of the SM, which was based on the data available at the time, neutrinos were treated as massless chiral[1] (Weyl) fermions. The model considered only the left-handed component ν_L and its corresponding right-handed antineutrino state $\overline{\nu}_R$, where L and R represent the left-handed and right-handed chirality, respectively. The left-handed neutrinos, ν_L, are arranged into weak isospin doublets together with their corresponding charged leptons:

$$L_{eL} = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \quad L_{\mu L} = \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \quad L_{\tau L} = \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix},$$

[1] Chiral fermions are particles whose left-handed and right-handed components have distinct transformation properties under the gauge symmetries of the SM.

while the right-handed antineutrinos, $\overline{\nu}_R$, are singlets. Active left handed neutrinos interact via weak charged current (CC) and neutral current (NC) interactions, described by the following Lagrangians [3] in the interaction basis:

$$\begin{aligned} -\mathcal{L}_{\mathrm{CC}} =& \frac{g}{\sqrt{2}} \sum_{\alpha} \overline{\nu}_{\alpha L} \gamma^{\mu} \ell_{\alpha L} W_{\mu}^{+} + \mathrm{h.c.} \\ -\mathcal{L}_{\mathrm{NC}} =& \frac{g}{2\cos\theta_W} \sum_{\alpha} \overline{\nu}_{\alpha L} \gamma^{\mu} \nu_{\alpha L} Z_{\mu}^{0}, \end{aligned} \tag{1.1}$$

where g is the coupling constant associated to $SU(2)_L$, θ_W is the Weinberg angle, and $\alpha = e, \mu, \tau$. Charged current interactions allow us to define a neutrino of a definite flavor α as the one that takes part in a CC process with the charged lepton ℓ_α.

In the SM, neutrinos were initially considered to be massless due to the restricted particle content and gauge symmetry of the model [1, 3]. The SM, as formulated, does not allow for a renormalizable mass term for neutrinos with its fermionic content and gauge symmetry. However, subsequent experimental evidence, such as the observations from the Super-Kamiokande experiment [4], has unequivocally demonstrated that neutrinos do possess non-zero masses and can therefore undergo flavor oscillations [5, 6], transitioning between different flavor states as they propagate. This discovery required an extension of the original SM framework to accommodate neutrino masses and mixing. Various approaches have been explored, including extending the particle content of the model, departing from strict gauge invariance, or relaxing the requirement of renormalizability [3]. In the mass basis, the Lagrangian for leptonic CC interactions includes the mixing of neutrino flavor eigenstates $(\nu_e, \nu_\mu, \nu_\tau)$ with the mass eigenstates (ν_1, ν_2, ν_3) and is expressed as:

$$-\mathcal{L}_{\mathrm{CC}} = \frac{g}{\sqrt{2}} \sum_{i,\alpha} \overline{\nu}_{iL} U_{i\alpha} \gamma^{\mu} \ell_{\alpha L} W_{\mu}^{+} + \mathrm{h.c.} \ ,$$

where U is the unitary *Pontecorvo-Maki-Nakagawa-Sakata* (PMNS) [5] mixing matrix, describing the mixing of neutrino flavors, which will be addressed in the following section.

1.2 Phenomenology of neutrino mixing and oscillations

As outlined in the previous section, neutrinos are known to exist in three distinct flavors, namely electron neutrinos (ν_e), muon neutrinos (ν_μ), and tau neutrinos (ν_τ). Each flavor is linked to a corresponding charged lepton: the electron, muon, and tau, respectively. However, a notable phenomenon known as neutrino oscillation was observed [4, 7], where neutrinos can alter their flavor as they travel through space. This suggests that the neutrino flavors produced in a particular interaction may not match the ones detected in a different location.
The reason behind this phenomenon lies in the fact that the three neutrino flavors are not equivalent to the three neutrino mass eigenstates (ν_1, ν_2, ν_3). Rather, the flavor eigenstates are a combination of the mass eigenstates. The standard three-flavor neutrino framework establishes a relationship between the neutrino flavor eigenstates (ν_e, ν_μ, ν_τ) and the mass eigenstates (ν_1, ν_2, ν_3) through the lepton mixing matrix,

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U_{\mathrm{PMNS}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{1.2}$$

where U_{PMNS} is a 3x3 unitary matrix, also called *Pontecorvo-Maki-Nakagawa-Sakata* (PMNS) [5] matrix. The standard parametrization for U_{PMNS} is given by [3],

$$U_{\text{PMNS}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\eta_1} & 0 & 0 \\ 0 & e^{i\eta_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1.3)$$

where the notation $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$ is used. The Dirac CP phase, δ_{CP}, is to denote the magnitude of charge conjugation parity (CP) symmetry violation. On the other hand, the Majorana CP phases, represented by η_i $(i = 1, 2)$, are physical only if neutrinos are Majorana particles. However, since they appear in a diagonal matrix, they do not affect neutrino oscillations and are not accessible in oscillation experiments [8].
Let ν_k be a mass eigenstate[2] and ν_α a flavor eigenstate. Neutrinos are produced through charged current interactions in a given flavor eigenstate that is the superposition of different mass eigenstates, namely:

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i}|\nu_i\rangle, \quad (1.4)$$

where $U_{\alpha k}$ are the elements of the PMNS matrix in Eq. (1.3). Since the massive neutrino states $|\nu_k\rangle$ are eigenstates of the Hamiltonian, their time evolution $|\nu_k(t)\rangle$ can be written as:

$$|\nu_k(t)\rangle = e^{-iE_k t}|\nu_k(0)\rangle. \quad (1.5)$$

where the energy associated with the k-th mass eigenstate, assuming a monochromatic beam, is

$$E_k = \sqrt{|\vec{p}|^2 + m_k^2} = |\vec{p}|\sqrt{1 + \frac{m_k^2}{|\vec{p}|^2}} \simeq |\vec{p}| + \frac{m_k^2}{2|\vec{p}|} \simeq E + \frac{m_k^2}{2E}, \quad (1.6)$$

with $E \simeq |\vec{p}|$. Using the expression in Eq. (1.4) for the flavor states in terms of the mass eigenstates, we can write:

$$|\nu_\alpha(t)\rangle = \sum_{i=1}^{3} U_{\alpha i}e^{-iE_i t}|\nu_i\rangle. \quad (1.7)$$

The probability of a neutrino of flavor α (at $t = 0$) being detected at the time t as a neutrino of flavor β is given by the absolute square of the transition amplitude:

$$P(\nu_\alpha \to \nu_\beta) = |\langle\nu_\beta|\nu_\alpha(t)\rangle|^2 = \sum_{i,j} U^\star_{\beta i}U_{\beta j}U_{\alpha i}U^\star_{\alpha j}e^{-i(E_i - E_j)\cdot t} \quad (1.8)$$

Given their extremely small masses, neutrinos can be assumed to be ultra-relativistic, thereby allowing to identify the distance L between source and detector with the time of flight t[3]. Moreover, given Eq. (1.6),

$$\Delta E_{ij} = E_i - E_j \simeq \frac{\Delta m_{ij}^2}{2E}, \quad (1.9)$$

[2]The neutrino mass eigenstates are numbered with descending electron neutrino content: ν_1 has the highest electron flavor content and ν_3 is the neutrino with the lower electron flavor content.
[3]We use natural units, i.e., $c = 1$.

hence the standard 3-flavor neutrino model (νSM) is further characterized by two distinct mass-squared differences, which are defined as follows:

$$\Delta m^2_{ij} \equiv m^2_i - m^2_j \qquad \text{with} \qquad i,j = 1,2,3 \qquad (i > j). \tag{1.10}$$

An important prediction of the νSM is the so called *mass sum rule*, establishing that

$$\Delta m^2_{21} - \Delta m^2_{31} + \Delta m^2_{32} = 0. \tag{1.11}$$

Therefore, the oscillation probability in vacuum is given by

$$\begin{aligned} \mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) &= \left| \sum_{i=1}^{3} U^\star_{\beta i} e^{-iE_i t} U_{\alpha i} \right|^2 = \sum_{i,j} U^\star_{\beta i} U_{\beta j} U_{\alpha i} U^\star_{\alpha j} \exp\left\{ -i \frac{\Delta m^2_{ij} L}{2E} \right\} \\ &= \sum_{i=j} |U_{\beta i}|^2 \, |U_{\alpha i}|^2 + 2\,\mathrm{Re} \sum_{i>j} U^\star_{\beta i} U_{\beta j} U_{\alpha i} U^\star_{\alpha j} \exp\left\{ -i \frac{\Delta m^2_{ij} L}{2E} \right\}, \end{aligned} \tag{1.12}$$

where in the last equality we have decomposed the sum with $i \neq j$ into two sums that are complex conjugate of each other (one with $i > j$ for which $\Delta m^2_{ij} > 0$ and the other with $i < j$ for which $\Delta m^2_{ij} < 0$) and then we used the property $z + z^\star = 2\,\mathrm{Re}\, z$. By performing the necessary calculations and exploiting the fact that $\sum_i U^\star_{\beta i} U_{\alpha i} = \delta_{\alpha\beta}$, we obtain the expression:

$$\begin{aligned} \mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) &= \delta_{\alpha\beta} - 2\,\mathrm{Re} \left[\sum_{i>j} U^\star_{\beta i} U_{\beta j} U_{\alpha i} U^\star_{\alpha j} \left(1 - \exp\left\{ -i \frac{\Delta m^2_{ij} L}{2E} \right\} \right) \right] \\ &= \delta_{\alpha\beta} - 2\,\mathrm{Re} \left[\sum_{i>j} U^\star_{\beta i} U_{\beta j} U_{\alpha i} U^\star_{\alpha j} \left(1 - \cos\left\{ \frac{\Delta m^2_{ij} L}{2E} \right\} \right) \right] \\ &\quad + 2\,\mathrm{Im} \left[\sum_{i>j} U^\star_{\beta i} U_{\beta j} U_{\alpha i} U^\star_{\alpha j} \sin\left\{ \frac{\Delta m^2_{ij} L}{2E} \right\} \right], \end{aligned} \tag{1.13}$$

where in the last equality we exploited the identity $\mathrm{Re}(ab) = \mathrm{Re}\, a\, \mathrm{Re}\, b - \mathrm{Im}\, a\, \mathrm{Im}\, b$, $\delta_{\alpha\beta}$ is the Kronecker delta and L is the distance traveled by the neutrino. The following expression is also commonly used:

$$\begin{aligned} \mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) &= \delta_{\alpha\beta} - 4\,\mathrm{Re} \left[\sum_{i>j} U^\star_{\beta i} U_{\beta j} U_{\alpha i} U^\star_{\alpha j} \sin^2\left\{ \frac{\Delta m^2_{ij} L}{4E} \right\} \right] \\ &\quad + 2\,\mathrm{Im} \left[\sum_{i>j} U^\star_{\beta i} U_{\beta j} U_{\alpha i} U^\star_{\alpha j} \sin\left\{ \frac{\Delta m^2_{ij} L}{2E} \right\} \right]. \end{aligned} \tag{1.14}$$

Consequently, from Eq. (1.14), it can be inferred that the oscillation phase is driven by the neutrino energy E, the source-detector distance L and the squared mass splitting Δm^2_{ij}. The first two quantities depend on the characteristics of the experiment (hence they can be, in some cases[4], optimized in order to be sensitive to a specific mass splitting), while the latter is a physical constant.

[4] This applies to artificial sources of neutrinos, such as those generated by reactors and accelerators. For natural sources, such as solar and atmospheric neutrinos, the baseline and energy are predetermined by the underlying physical processes.

On the other hand, the amplitude of the oscillation is driven only by the elements of the mixing matrix [1], thus by the mixing angles θ_{ij}, as specified in Eq. (1.3).
The so-called sensitivity to a given mass splitting Δm^2_{ij} of an experiment is the value of Δm^2_{ij} for which

$$\frac{\Delta m^2 L}{4E} \simeq 1.27 \cdot \left(\frac{\Delta m^2}{1\ \mathrm{eV}^2}\right)\left(\frac{1\,\mathrm{GeV}}{E}\right)\left(\frac{L}{1\,\mathrm{km}}\right) \overset{!}{\simeq} 1, \tag{1.15}$$

where without the loss of generality, we use the notation $\Delta m^2 = \Delta m^2_{ij}$. Different types of neutrino oscillation experiments are traditionally classified also based on the average value of L/E, which determines their sensitivity to a given Δm^2 [1]. Typical values of baselines, energy ranges, and corresponding sensitivity to Δm^2 are reported in Table 1.1.

Type of experiment	Baseline L	Energy E	Δm^2 sensitivity [eV^2]
Reactor SBL	$\sim 10\,\mathrm{m}$	$\sim 1\,\mathrm{MeV}$	~ 0.1
Reactor LBL	$\sim 1 - 100$ km	$\sim 1\,\mathrm{MeV}$	$\sim 10^{-3} - 10^{-5}$
Accelerator SBL	$\sim 0.1 - 1$ km	$\geq 1\,\mathrm{GeV}$	$\gtrsim 1$
Accelerator LBL	$\sim 100 - 1000$ km	$\geq 1\,\mathrm{GeV}$	$\sim 10^{-2} - 10^{-3}$
Atmospheric	$\sim 10 - 10^4$ km	$0.5 - 10^2$ GeV	$\sim 10^{-3} - 10^{-4}$
Solar	$\sim 10^{11}$ km	$0.2 - 15$ MeV	$\sim 10^{-12}$

TABLE 1.1: Types of neutrino oscillation experiments with their typical source–detector distance, energy, and sensitivity to Δm^2. The acronyms SBL and LBL stand for Short Baseline and Long Baseline, respectively.

The oscillation probabilities associated with channels where $\alpha \neq \beta$ are commonly referred to as *transition probabilities*. Conversely, the oscillation probabilities linked to channels where $\alpha = \beta$ are typically known as *survival probabilities*. Oscillation experiments follow a similar distinction [1], outlined in the following paragraphs.

***Appearance* experiments.** They measure transitions between distinct neutrino flavors. If the final flavor to be detected is absent in the initial beam, the background noise can be effectively minimized. The primary signal sought in such experiments is the observation of a non-zero count of this different neutrino flavor. This configuration enables experiments to exhibit high sensitivity even to relatively small values of the mixing angle. Neutrino appearance experiments typically employ beams predominantly consisting of a single type of neutrino and aim to detect neutrinos of different flavors. However, their sensitivity is often hindered by uncertainties in the knowledge of beam contamination at the source. For instance, accelerator-generated muon-neutrino beams are not completely pure and contain a small fraction of electron neutrinos, typically around 1%, and particle identification in detectors may not be 100% efficient. On the other hand, the presence of tau neutrino contamination at the source is negligible [9]. However, the sensitivity of muon-to-tau neutrino appearance experiments is limited either by statistical constraints or by the ability to accurately identify tau particles. In practice, appearance experiments using an accelerator-generated muon-neutrino beam are always disappearance experiments as well: this is due to the similarity between the squared mass differences Δm^2_{32} and Δm^2_{31}. Consequently, an experiment designed to observe the appearance of electron or tau neutrinos will also possess sensitivity to the disappearance of muon neutrinos. For example, the OPERA experiment [9] used a muon-neutrino beam produced by CERN Neutrinos to Gran Sasso (CNGS) facility. Despite the appearance of ν_τ being its primary goal, OPERA was also sensitive to the disappearance of ν_μ during their propagation. Accelerator experiments that produced crucial results by studying muon (anti)neutrino

disappearance and electron (anti)neutrino appearance are, for example, the MINOS (Main Injector Neutrino Oscillation Search) [10] and the T2K (Tokai to Kamioka) [11] experiments.

Disappearance experiments. They are designed to measure the survival probability of a neutrino flavor. The distinctive oscillation signal manifests as a deficit in the observed event count. However, due to statistical fluctuations inherent in the number of detected events, even in the absence of oscillations, uncovering a subtle disappearance becomes challenging. Reactor neutrino experiments are always disappearance experiments, because the energies of the produced electron antineutrinos, typically of the order of few MeV, are not sufficient to create muons or taus, and therefore ν_μ and ν_τ cannot be seen in charged-current reactions. For low-energy neutrinos, the only viable approach for conducting appearance experiments is through a direct comparison between neutral-current and charged-current reactions, as done by the Sudbury Neutrino Observatory (SNO) experiment [7]. Disappearance experiments have yielded significant results in the study of neutrino oscillations. SNO provided compelling evidence for neutrino oscillations and confirmed the existence of the phenomenon. It observed a deficit in the number of solar electron neutrinos compared to expectations based on solar models. This deficit indicated that electron neutrinos from the Sun had oscillated into other neutrino flavors during their journey to Earth [7]. The KamLAND (Kamioka Liquid Scintillator Antineutrino Detector) experiment [12] similarly observed a deficit in the number of electron antineutrinos emitted by nuclear reactors, providing further confirmation of neutrino oscillations. The Borexino experiment played a key role in the field by performing direct measurements of solar electron neutrinos survival probability [13]. The Daya Bay [14] and RENO (REactor Neutrino Oscillation) [15] experiments provided precise measurements of the mixing angle θ_{13}, determining its non-zero value.

1.2.1 Neutrino oscillations in matter

The formalism presented thus far applies to neutrino oscillations in vacuum. However, when considering neutrino propagation through a dense medium, the free-particle approximation used in the dispersion relation in Eq. (1.6) breaks down [1, 16]. In this scenario, the evolution equation for active flavor neutrinos is modified due to the presence of effective potentials arising from their interactions with matter. These interactions primarily involve coherent forward elastic weak charged current (CC) and neutral current (NC) scatterings. In the context of ordinary matter composed of electrons, the CC channel is accessible exclusively to electron (anti)neutrinos, while the NC channel is flavor blind, affecting all (anti)neutrino flavors equally. The Feynman diagrams of CC and NC-induced scatterings are shown in Figure 1.1.
The dispersion relation for neutrinos in the flavor basis in a medium can be expressed as follows:

$$E_m^\alpha = \underbrace{E_0^\alpha}_{\text{vacuum}} + \underbrace{E_{\text{CC}}^{\text{int}}\,\delta_e^\alpha + E_{\text{NC}}^{\text{int}}}_{\text{matter}}. \tag{1.16}$$

where E_0^α represents the energy of neutrino α in vacuum, whereas $E_{\text{NC}}^{\text{int}}$ and $E_{\text{CC}}^{\text{int}}$ account for the effects of NC and CC interactions with matter. It is worth noting that the contributions from neutral current (NC) scatterings cancel out in the neutrino oscillation probability due to their flavor universality. Thus, for simplicity, $E_{\text{NC}}^{\text{int}}$ can be neglected. On the other hand, it is demonstrated [1] that $E_{\text{CC}}^{\text{int}}$ is proportional to the Fermi constant G_F and the number density of electrons N_e, yielding:

$$E_{\text{CC}}^{\text{int}} = \sqrt{2} G_F N_e. \tag{1.17}$$

The corresponding matter potential term, indicated as V_{CC}, induces a shift in the effective masses of the neutrino mass eigenstates, leading to the Mikheyev-Smirnov-Wolfenstein (MSW) effect [16, 17];

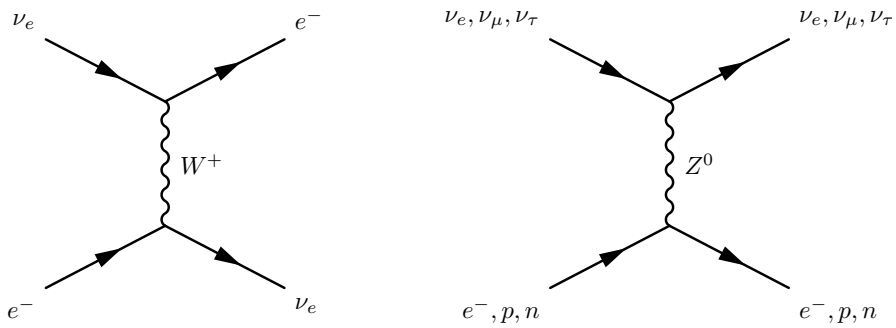


FIGURE 1.1: Feynman diagrams of the coherent forward elastic scattering processes that generate the charged current potential through W-boson exchange and the NC potential through Z-boson exchange [1].

this phenomenon, occurring under specific conditions, leads to resonant maximal mixing between different neutrino flavors.

Since E_m^α in Eq. (1.16) are the eigenvalues of the Hamiltonian in the flavor basis, one can express the total effective Hamiltonian in matter[5], neglecting NC contributions, as:

$$\widetilde{\mathcal{H}}_{\text{eff}}^f = \mathcal{H}_0^f + \mathcal{H}_\mathcal{I}, \qquad \text{with} \qquad \mathcal{H}_\mathcal{I}|\nu_\alpha\rangle = V_{\text{CC}}\delta_\alpha^e|\nu_\alpha\rangle = E_{\text{CC}}^{\text{int}}\delta_\alpha^e|\nu_\alpha\rangle, \tag{1.18}$$

where $\mathcal{H}_0^f$ is the vacuum Hamiltonian in the flavor basis.
Detailed calculations can be found in [1] and lead to the following expression:

$$\begin{aligned}\widetilde{\mathcal{H}}_{\text{eff}}^f &= \frac{1}{2E}\left[U\begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}U^\dagger - \begin{pmatrix} A & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right] \\ &= \frac{1}{2E}\left[\widetilde{U}\begin{pmatrix} \widetilde{m}_1^2 & 0 & 0 \\ 0 & \widetilde{m}_2^2 & 0 \\ 0 & 0 & \widetilde{m}_3^2 \end{pmatrix}\widetilde{U}^\dagger\right],\end{aligned}$$

where $\widetilde{U}$ and $\widetilde{m}_i$ stand for the effective neutrino mixing matrix and the i-th neutrino mass in matter, respectively. The matter parameter A can be expressed as

$$A = \pm\, 2\sqrt{2}G_{\text{F}}N_e E, \tag{1.19}$$

where the $+(-)$ sign refers to the neutrino (antineutrino) channel.
By accounting for matter-induced potentials and their influence on neutrino propagation, the formalism can be extended to describe neutrino oscillations in dense media, enabling the study of oscillation phenomena in astrophysical environments, terrestrial experiments, and even within the Earth itself. In Section 2.3, the role of matter effects in the JUNO experiment is further explained.

[5]The oscillation in vacuum is described by Schroedinger's equation in the mass basis, with Hamiltonian in the mass basis $\mathcal{H}_0^m$. To describe the evolution of flavor eigenstates in vacuum, one needs to pass to the flavor basis by applying a rotation, namely $\mathcal{H}_0^f = U\mathcal{H}_0^m U^\dagger$, with U the PMNS mixing matrix.

1.2.2 Neutrino Mass Ordering

Despite significant advancements in neutrino experiments and their precision in recent years, many properties of neutrinos still remain unknown, including the nature of neutrinos (Dirac or Majorana), the existence of CP violation in the leptonic sector, and the scale of the neutrino mass eigenstates, commonly referred to as neutrino Mass Ordering (MO) [18]. This section will address the last of the aforementioned unknowns.

Neutrino oscillation experiments are only sensitive to the squared mass differences Δm^2_{ij} in Eq. (1.10), but not to the absolute mass scale. Thanks to matter effects in the Sun [19], we know that $\Delta m^2_{21} > 0$. Contrariwise, the sign of Δm^2_{31}, or equivalently Δm^2_{32}, still remains undetermined. As a consequence, there are two possible scenarios: normal ordering (NO, $\Delta m^2_{31} > 0$) or inverted ordering (IO, $\Delta m^2_{31} < 0$). The two mentioned hierarchical neutrino mass spectra are shown in Figure 1.2, providing a graphical representation of the neutrino flavor content of each mass eigenstate[6].

Its determination thus relies on the precise measurement of Δm^2_{31} and Δm^2_{32} and can be experimentally addressed via different techniques [20]: (1) long baseline (LBL) accelerator experiments, (2) atmospheric neutrino experiments, and (3) medium baseline reactor experiments (i.e., JUNO [21]). Matter effects introduce discrepancies in the oscillation probabilities for neutrinos and antineutrinos: in the case of NO, the electron neutrino appearance is enhanced and the electron anti-neutrino appearance is suppressed (while for IO the configuration is opposite), thus manifesting as an effective *pseudo* CP-violating effect. In LBL beam experiments, the sensitivity to the matter effect relies on the appearance channel, specifically the transitions $\nu_\mu \to \nu_e$ and $\overline{\nu}_\mu \to \overline{\nu}_e$. However, it is important to note that LBL experiments are also sensitive to the genuine CP-violating phase, making it nontrivial to distinguish the asymmetry caused by CP violation and the different mass orderings. Among the accelerator experiments, NOνA [22] currently holds the most relevance in terms of intrinsic MO sensitivity due to its relatively large matter effects compared to T2K [11], and they will be followed by the next generation DUNE [23]. A similar approach, based on the MSW effect [19], is pursued by atmospheric neutrino experiments (e.g., the planned PINGU [24]), which exploit (anti)neutrinos with a wide range of baselines ($\sim 10 - 10^4$ km) and energies. In contrast, JUNO stands out as the only experiment currently capable of resolving the mass ordering through dominant vacuum oscillations[7]. This unique characteristic grants JUNO unparalleled insight and capability in the study of mass ordering within the neutrino oscillation framework [20]. JUNO's MO sensitivity relies on high-precision spectral analysis of reactor antineutrinos.

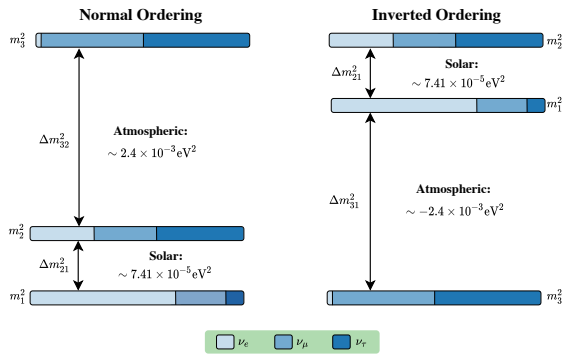


FIGURE 1.2: The two possible neutrino mass ordering schemes.

1.2.3 Current knowledge on oscillation parameters

The current knowledge on three-flavor oscillation parameters is summarized in Table 1.2: the results are obtained from a global fit (NuFIT 5.2 [27]) as of November 2022, without the inclusion of Super-Kamiokande tabulated $\Delta\chi^2$ data [28]. The numbers on the first (second) column are obtained

[6] It is worth noting that the probability of finding a neutrino of flavor α in the i-th mass eigenstate also depends on the value of the CP violating phase [1, 18]. This feature is not shown in Figure 1.2, where all mixing parameters are fixed, and $\delta_{\text{CP}} = 0$.

[7] Matter effects play a minor role in JUNO [25, 26], as explained in Section 2.3.2.

assuming NO (IO). The following convention is adopted:

$$\Delta m^2_{3\ell} \qquad \text{with} \begin{cases} \ell = 1 & \text{for } \Delta m^2_{3\ell} > 0 \qquad \text{Normal Ordering (NO)} \\ \ell = 2 & \text{for } \Delta m^2_{3\ell} < 0 \qquad \text{Inverted Ordering (IO)} \end{cases} \tag{1.20}$$

Following Eq. (1.20), $\Delta m^2_{3\ell}$ stands for the mass splitting with the largest absolute value, thus Δm^2_{31} for NO, Δm^2_{32} for IO.
The significant difference in magnitude between $\Delta m^2_{3\ell}$ and Δm^2_{21} is evident in Table 1.2. This feature allows many experiments to employ a simplified two-flavor neutrino oscillation framework as a reasonable approximation to the full three-flavor framework [1]. Within this scheme, the parameters Δm^2_{21} and θ_{12} are commonly referred to as the *solar parameters* as they govern the disappearance of solar neutrinos. Similarly, the parameters $\Delta m^2_{3\ell}$ and θ_{23} are known as the *atmospheric parameters* since they were first measured by the Super-Kamiokande (SK) collaboration through the detection of atmospheric neutrinos [4]. The smallest mixing angle, θ_{13}, was experimentally determined to be non-zero mainly through reactor experiments, such as Daya Bay [14], RENO [15], and Double Chooz [29].

	Normal Ordering (best fit)		Inverted ordering ($\Delta\chi^2 = 2.3$)	
	Best fit $\pm$ 1σ	3σ range	Best fit $\pm$ 1σ	3σ range
$\frac{\Delta m^2_{21}}{10^{-5}\text{eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$
$\frac{\Delta m^2_{3\ell}}{10^{-3}\text{eV}^2}$	$+2.511^{+0.028}_{-0.027}$	$+2.428 \to +2.597$	$-2.498^{+0.032}_{-0.025}$	$-2.581 \to -2.408$
$\sin^2\theta_{12}$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$
$\sin^2\theta_{13}$	$0.02203^{+0.00056}_{-0.00059}$	$0.02029 \to 0.02391$	$0.02219^{+0.00060}_{-0.00057}$	$0.02047 \to 0.02396$
$\sin^2\theta_{23}$	$0.572^{+0.018}_{-0.023}$	$0.406 \to 0.620$	$0.572^{+0.018}_{-0.023}$	$0.412 \to 0.623$
δ_{CP} / $^\circ$	197^{+42}_{-25}	$108 \to 404$	286^{+27}_{-32}	$192 \to 360$

TABLE 1.2: Current best fit values on oscillation parameters from NuFIT 2022 global analysis [27], without Super-Kamiokande atmospheric data. Note that $\Delta m^2_{3\ell} = \Delta m^2_{31} > 0$ for Normal Ordering and $\Delta m^2_{3\ell} = \Delta m^2_{32} < 0$ for Inverted Ordering.

Chapter 2

The JUNO experiment

The Jiangmen Underground Neutrino Observatory (JUNO) [21, 30], is a multi-purpose liquid scintillator (LS) experiment currently under construction in South China. JUNO is designed primarily for the determination of the neutrino mass ordering with reactor antineutrinos ($\overline{\nu}_e$), emitted from the Taishan and Yangjiang nuclear power plants (NPPs), both located at a baseline of about 52.5 km from the experimental site. In order to achieve accurate results, JUNO relies on precise knowledge of the unoscillated reactor antineutrino spectrum shape. To accomplish this, a dedicated small satellite detector called Taishan Antineutrino Observatory [31] (TAO or JUNO-TAO) will be installed at a distance of around 30 m from one of the Taishan reactors. TAO will measure the spectrum with sub-percent energy resolution, serving as a data-driven reference to constrain the spectra of the other reactor cores. A schematic illustrating the location of both JUNO and TAO is shown in Figure 2.1.

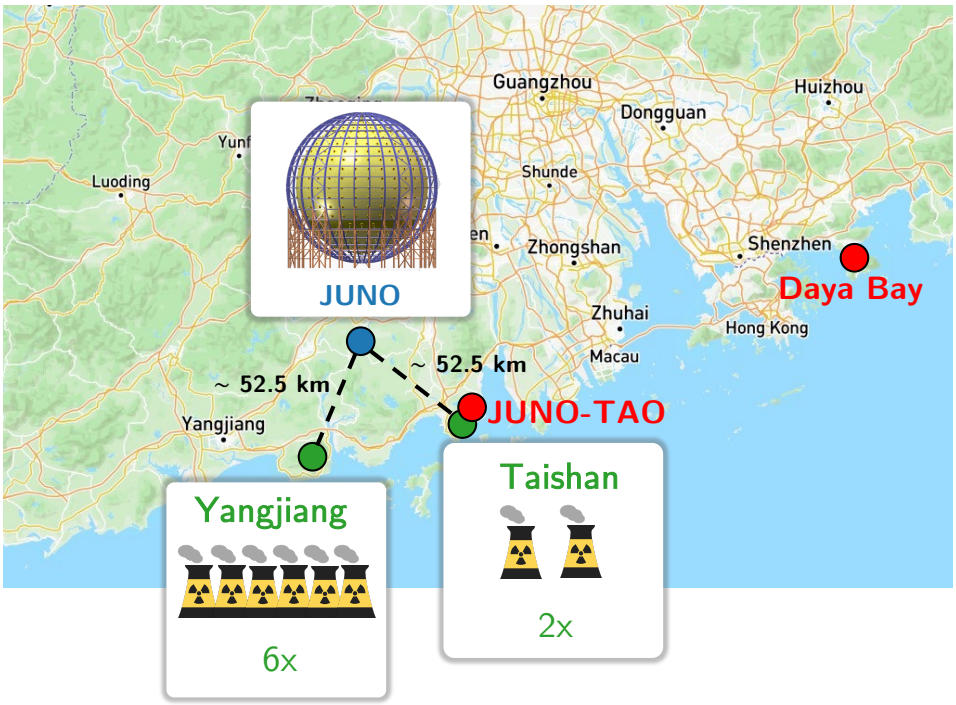


FIGURE 2.1: Location of the JUNO experiment in South China [21]. The main 20 kton LS detector is indicated in blue, at a baseline of $\sim$ 52.5 km from six 2.9 GW_{th} reactor cores in the Yangjiang NPP and two 4.6 GW_{th} cores in the Taishan NPP. The 2.8 ton JUNO-TAO detector [31], indicated in red, is located about 30 meters away from one of the Taishan reactor cores.

In JUNO's location, the energy spectrum will be distorted by a slow (low frequency) oscillation driven by Δm^2_{21} and modulated by $\sin^2(2\theta_{12})$, as well as by a fast (high frequency) oscillation regulated by Δm^2_{31} and modulated by $\sin^2(2\theta_{13})$, as shown in Figure 2.2. The oscillated spectrum in JUNO changes subtly depending on the neutrino mass ordering, thus providing sensitivity to this parameter.

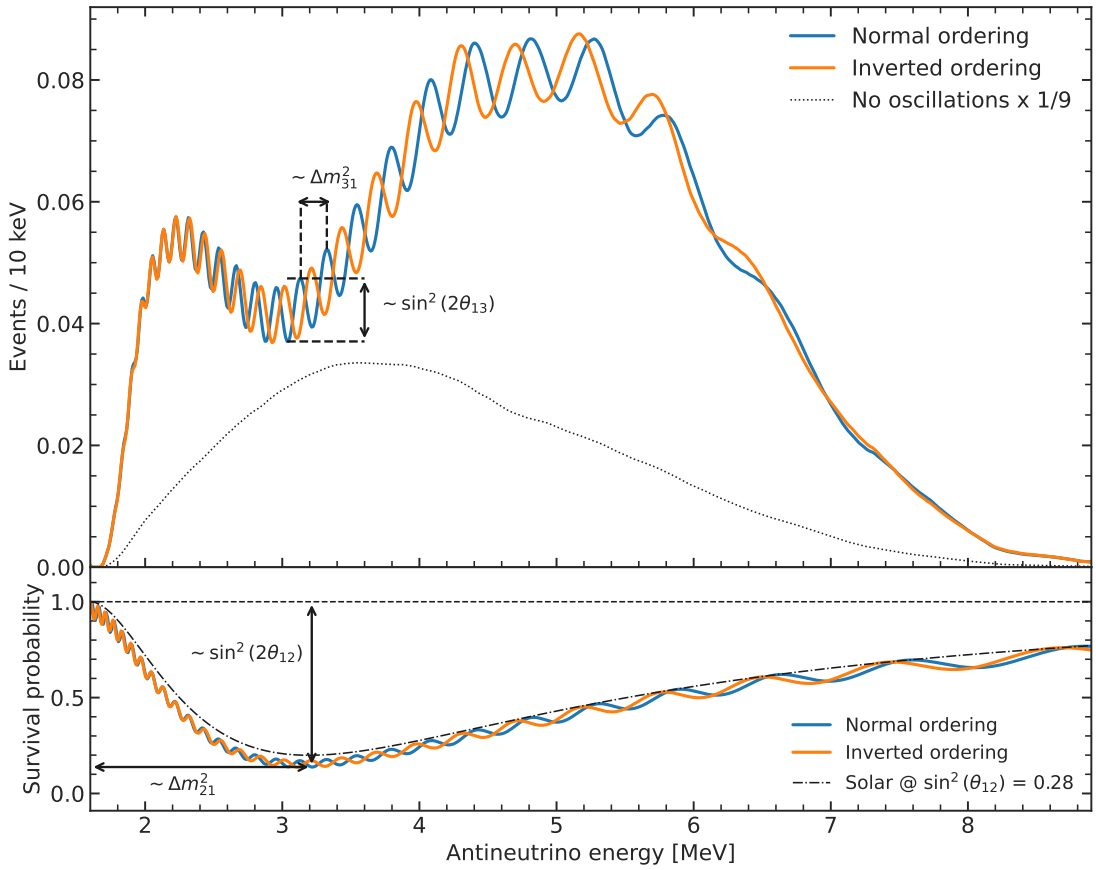


FIGURE 2.2: Top: Expected antineutrino spectrum at JUNO in the ideal configuration (without detector-related effects) with and without oscillations (using global parameters in Table 1.2). Bottom: $\overline{\nu}_e$ survival probability with a 52.5 km baseline. To enhance visual clarity, the oscillation induced by the solar term is depicted using $\sin^2\theta_{12} = 0.28$ instead of $\sin^2\theta_{12} = 0.303$. The mass splittings and mixing angles modulating the spectrum are indicated as well.

2.1 JUNO detector design

The JUNO detector is located in an underground laboratory with approximately 690 m of rock overburden, i.e., 1800 meters-water-equivalent (m.w.e.), at an average baseline of 52.5 km from two nuclear power plants (NPPs). The detector consists of a 20 kton liquid scintillator (LS) target, contained inside a 35.4-meter-diameter spherical acrylic vessel. It will be the largest liquid scintillator-based detector in the World for the next decade [21], exceeding comparable existing detectors, e.g. KamLAND [12], SNO+ [7], or Borexino [13], by a factor of approximately twenty. Both the determination of the neutrino MO and the precision measurement of oscillation parameters are challenging and imply strict requirements, including:

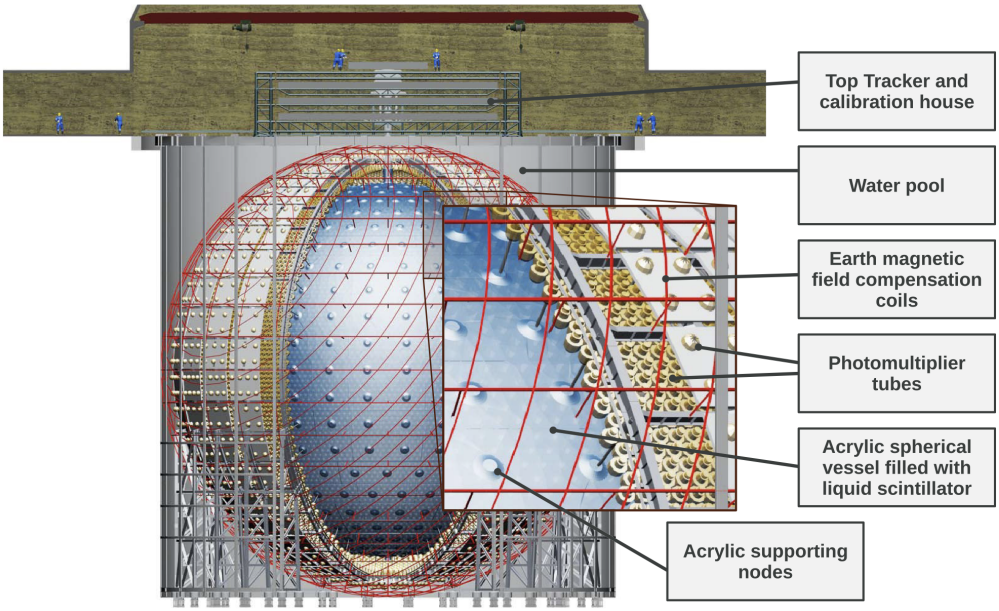


FIGURE 2.3: Schematic representation of the main JUNO detector [21]. The 20 kton acrylic sphere is submerged in a 44-meters-deep water pool and instrumented with 17612 20-inch and 25600 3-inch photomultiplier tubes. The upper part of the detector consists of a Top Tracker system with three layers of plastic scintillator and a calibration house.

- high light yield to efficiently detect and measure neutrino interactions, and achieve high energy resolution;
- control of the energy scale to accurately resolve the high-frequency oscillations in the antineutrino spectrum;
- a large number of detected antineutrino events to achieve robust statistical significance;
- minimization of background contamination to enhance the signal-to-background ratio;
- thorough control of systematic uncertainties to mitigate potential biases.

These constraints are addressed by careful detector design, as outlined in the subsequent sections.

2.1.1 JUNO main detector

The main detector of JUNO can be divided into several sub-detector systems: these are the *Central Detector* (CD), the *Water Pool* (WP), and the *Top Tracker* (TT). A schematic view of the main detector is shown in Figure 2.3.

The CD is the key component of the JUNO detector, serving the purpose of detecting physics events covering the experiment's comprehensive physics program [21, 30], including reactor, solar, atmospheric and supernovae neutrinos and antineutrinos. The driving requirements for its design are the energy resolution within 3% at 1 MeV, a precise control of the energy scale (overall non-linearity effects below 1%), and a substantial antineutrino statistics [21, 30]. The statistics constraint is secured by the sizeable fiducial volume, with a 20 kton LS target, contained in an acrylic sphere of 35.4 m. On the other hand, to achieve the unprecedented energy resolution (for a LS-based experiment), the detector incorporates a sophisticated photo-detection system consisting of 17612

20-inch (or Large) photomultiplier tubes (PMTs) and 25600 3-inch (or Small) PMTs; these PMTs point inward and are attached to the surrounding Stainless Steel (SS) structure. This configuration provides an extensive total photo-coverage of $\sim 78\%$, granting high photoelectron (PE) statistics. Moreover, to mitigate the impact of the Earth's magnetic field on the PMTs photoelectron collection efficiency, compensation coils are mounted on the SS structure. These coils generate magnetic fields that counteract the Earth's magnetic field within the detector, minimizing any potential detrimental effects on the PMTs performance.

Due to the extremely small cross sections of neutrino weak interactions, neutrino events are inherently rare. As a result, reducing background noise, mainly due to cosmic rays and natural radioactivity, becomes of utmost importance in JUNO.
The radiogenic contamination is efficiently contained thanks to proper detector design, meticulous environmental control, and careful selection and screening of the employed materials [32]. Further details on the contribution of radioactive backgrounds will be discussed in Section 4.3.1.
To decrease muon-induced backgrounds, the CD is located in an experimental hall with approximately 690 m of rock overburden (1800 m.w.e.), which provides natural shielding from cosmic rays. Nevertheless, due to the large size of the detector, the residual muon rate is still high, amounting to approximately $0.004\,\mathrm{Hz/m^2}$ [21, 33]. For this reason, the experiment will be equipped with two veto systems acting as efficient taggers for cosmic muons[8]. The CD is housed within a cylindrical water pool (WP), with a height of 44 m and a diameter of 43.5 m, filled with 35 kton of ultra-pure water. The WP is instrumented with 2400 20-inch PMTs attached to the outer part of the SS shell and pointing outwards, thereby acting as a Cherenkov veto detector for cosmic muons. Furthermore, a water buffer connected to the outer Cherenkov detector acts as a shielding layer for the natural radioactivity of the surrounding rock. On top of the water pool, a muon tracker, called TT [34], will be installed. The system comprises scintillator strips decommissioned from the Target Tracker of the OPERA experiment [35]. As the TT is foreseen to have a very high rate of background events from the radioactivity of the rock and detector material, it will be built in a 3-layer design, requiring the coincident detection of traversing muons. The TT is centrally placed on top of the WP, covering only about 60% of the top surface of the JUNO CD [21, 34].
A chimney connects the CD to the outside and will be used to deploy calibration sources into the detector. The calibration operations, which are presented in Section 2.1.3 are carried out in the Calibration House.

2.1.2 JUNO liquid scintillator

The preparation of the liquid scintillator (LS) formula requires high quality of all the chemicals to satisfy JUNO's strict requirements:

- One crucial aspect is achieving a high light yield, which directly impacts the energy resolution of the detector. State-of-the-art simulations indicate that thanks to the high LS light yield and the large photo-coverage (see Section 2.1.1), JUNO aims for a light level exceeding 1350 photoelectrons per MeV [36], more than double the values reached in previous LS-based detectors like Daya Bay ($\sim$ 160 PEs/MeV [37]) and Borexino ($\sim$ 500 PEs/MeV [38]).

- Transparency of the LS to its emitted light is another essential criterion. The LS must have an attenuation length greater than 20 m [21, 30] at a wavelength of 430 nm. This ensures that photons emitted in the inner regions of the CD can effectively reach the PMTs for detection.

[8] Given the rather long lifetime of the cosmogenic isotopes and the muon rate of approximately 3-4 Hz, a veto of the full detector volume is not viable. Instead, a dedicated veto strategy aims at reconstructing muons tracks inside the CD [33].

- The LS must also comply with strict cleanliness and radiopurity standards.

Extensive studies, including dedicated investigations with a Daya Bay detector [39], were aimed at optimizing the LS composition. The chosen recipe contains *linear alkylbenzene* (LAB) as a solvent, due to its excellent transparency, high flash point, low chemical reactivity, and good light yield [21]. To attain transparency to the emitted light, an organic flour is usually added to the mixture. In JUNO, a concentration of 2.5 g/L of 2,5-diphenyloxazole (PPO) is used as the first wavelength shifter. The flour absorbs the photons emitted by the solvent and re-emits them at a higher wavelength ($\sim$ 390 nm), enabling their propagation through the LS. Additionally, p-bis-(o-methylstyryl)-benzene (bis-MSB) is incorporated (at a concentration of 3 mg/L) as an additional wavelength shifter. This choice ensures compatibility with the region of highest detection efficiency of the PMTs, which is approximately at 430 nm [40].

On the other hand, the targeted LS radiopurity is set at 10^{-15} g/g of ^{238}U and ^{232}Th for reactor antineutrinos analysis, while for the solar neutrino analysis concentrations below 10^{-17} g/g are required [32]. To ensure proper cleanliness of the LS, a combined system of purification plants, employing different techniques [21], is planned: the raw solvent will pass through alumina columns and distillation plants in the experimental surface area, before being mixed with the wavelength shifters. Then, the LS will be sent underground to undergo water extraction and stripping. A final quality check is performed by the Online Scintillator Internal Radioactivity Investigation System (OSIRIS) [41] stand-alone detector, to monitor the radiopurity of the LS while the CD is filled and to confirm the proper operation of the purification plants. More details on the aforementioned process can be found in [21].

2.1.3 JUNO calibration system

To achieve a precise determination of the MO, it is crucial to control the uncertainty in the positron kinetic energy scale within 1% [21]. To meet this requirement, a comprehensive calibration program is envisaged, which focuses on measuring and understanding the intrinsic non-linear behavior of the scintillation and Cherenkov light emitting mechanisms [42]. In addition to addressing energy non-linearity (NL), the calibration program also accounts for the position-dependence of the total number of PEs collected by the Large and Small PMT systems. This position non-uniformity, which is primarily independent of energy, needs to be corrected through a multi-positional calibration approach to optimize the energy resolution. The details of the NL model will be discussed in Section 3.3. However, in this section, the primary focus is on describing the calibration systems and the program established to ensure accurate energy scale calibration.
A comprehensive calibration will be carried out at the beginning of the experiment to gain a fundamental understanding of the CD performance, and then periodically throughout the lifespan of JUNO. Multiple sources will be deployed at the CD center: in particular, a UV laser diffuser ball, two neutron sources, and several gamma sources. The radioactive sources selected for calibration in JUNO, along with the types of emitted radiation, are listed in Table 2.1. To monitor and track significant changes in the detector properties, such as variations in the light yield of the LS, PMT gains, and electronics, weekly and monthly calibrations will be performed using the laser and the ^{241}Am-^{13}C source. These calibrations will be conducted at a limited number of positions within the detector. For a detailed overview of the calibration program, refer to [42].
The hardware design of the calibration system consists of several independent subsystems, shown in Figure 2.4.

- **Automatic Calibration Unit (ACU).** The ACU is developed to perform calibrations along the central vertical axis of the CD. Three sources can be regularly deployed, including a

Source/process	Type	Emitted radiation
^{137}Cs	γ	0.662 MeV
^{54}Mn	γ	0.835 MeV
^{60}Co	γ	1.173 MeV + 1.333 MeV
^{40}K	γ	1.461 MeV
^{68}Ge	e^+	2×0.511 MeV
^{241}Am-Be	n, γ	neutron + 4.43 MeV (^{12}C*)
^{241}Am-^{13}C	n, γ	neutron + 6.13 MeV (^{16}O*)
$(n, \gamma)p$	γ	2.22 MeV
$(n, \gamma)^{12}$C	γ	4.94 MeV or 3.68 MeV + 1.26 MeV

TABLE 2.1: List of calibration sources and corresponding radioactive processes considered by the JUNO collaboration [42].

neutron source (^{241}Am-^{13}C), a gamma source (^{40}K), and a pulsed UV laser source carried by an optical fiber with a diffuser ball attached to the end. The ACU is expected to be used frequently during data-taking to monitor the stability of the energy scale and, to some extent, the position non-uniformity.

- **Guide Tube system (GT).** The GT consists of a tube that runs outside the acrylic sphere of the CD, forming a loop along a longitudinal circle. This tube serves as a pathway for a radioactive source, which is connected to cables attached to both ends of the tube. The GT subsystem's ability to position the sources with a 3 cm precision the and its position make it an effective tool for characterizing the non-uniformity at the boundary of the CD.
- **Cable Loop System (CLS).** Two CLSs will be installed in the two opposite half-planes of the detector to ensure comprehensive coverage. The central cable of the CLS extends upwards toward the north pole of the CD, while the side cable winds through an anchor on the inner surface of the acrylic sphere. Both cables ultimately lead towards the north pole.
- **Remotely Operated Vehicle (ROV).** A ROV is foreseen to deploy radioactive sources in almost the entire LS volume, spanning areas outside the CLS plane.

Moreover, the response of each individual channel of the Large PMTs (LPMTs) might a priori be different, and also influenced by its position within the detector, resulting in a position-dependent channel-level non-linearity. This non-linearity introduces a discrepancy between the actual number of photons detected and the measured charge. To address this issue and ensure accurate measurements, the LPMT system undergoes a process specifically designed to account for channel-level non-linearity. This is performed thanks to the dual calorimetry technique [43], which involves comparing the response of the LPMT system to that of the Small PMT (SPMT) system. A tunable laser light source covering the entire energy range of interest is employed [42, 44]. Using this approach, the residual event-level non-linearity can be potentially reduced to less than 0.3%.

2.2 JUNO simulation and event reconstruction framework

The JUNO experiment relies on a software package based on the *Software for Non-collider Physics Experiments* (SNiPER) framework [45] for its Monte Carlo (MC) simulations of events. SNiPER is a modular software framework designed to meet the requirements of both reactor electron anti-neutrino experiments and cosmic ray experiments. It is implemented using a combination of C++

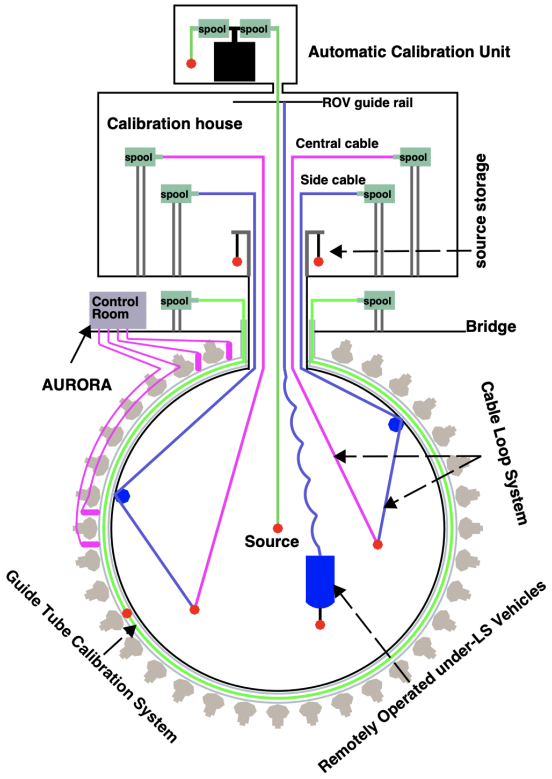


FIGURE 2.4: Schematic representation of the calibration system (not to scale), taken from [42]. The different subsystems are shown: the Automatic Calibration Unit (ACU), two Cable Loop Systems (CLSs), the Guide Tube (GT), and the Remotely Operated Vehicle (ROV). The red points represent a source assembly. The AURORA is an auxiliary laser diode system to monitor the attenuation and scattering length of the LS.

and Python, with C++ providing the main body for efficient execution and Python serving as the user interface. In addition to SNiPER, JUNO also uses the Geant4 [46] software for detector simulations and the ROOT data analysis framework [47]. Geant4 is a widely used toolkit for simulating the behavior of particles as they pass through matter, including their interactions with materials and the resulting energy loss, scattering, and production of secondary particles. It also models the response of the detector materials and electronics, such as scintillation light production and propagation, photoelectric effect, Compton scattering, and pair production. The simulation process consists of several steps, which are briefly described in Section 2.2.1 and Section 2.2.2.

The integration of SNiPER and Geant4 creates a comprehensive simulation and reconstruction pipeline for JUNO, a crucial component in optimizing detector design and developing necessary techniques for future real-data analysis. In the scope of this thesis, the SNiPER software has been extensively utilized to simulate the JUNO detector's response to various types of events, such as reactor antineutrinos interactions, geo-neutrinos, and cosmogenic backgrounds. A detailed description of this step is provided in Chapter 4.

Additionally, SNiPER provides a framework for developing and testing new algorithms for event reconstruction and analysis, making it an important tool for JUNO's ongoing development and optimization efforts. Within this work, an algorithm to perform event correlation analysis is developed to efficiently distinguish signals from backgrounds (see Section 4.1).

2.2.1 Event and detector simulation

The JUNO simulation software implements three stages: the physics generator, the detector simulation (`DetSim`), and the electronics simulation (`ElecSim`). The physics generator produces a list of particles to be simulated, including a simple particle gun and various generators for different studies, such as positron-neutron pairs from an Inverse Beta Decay reaction following the expected JUNO spectrum, as well as generators for atmospheric and solar neutrinos.
The Geant4-based `DetSim` stage is dedicated to simulating the detector's geometry and the particles' interactions inside the target material. It comprises physics processes such as the quenching effect in liquid scintillators, light production and propagation, and the PMT response. The output of this stage is a set of PMT hits representing the number of PEs released at the photocathode before amp ction
Effi f the
dete

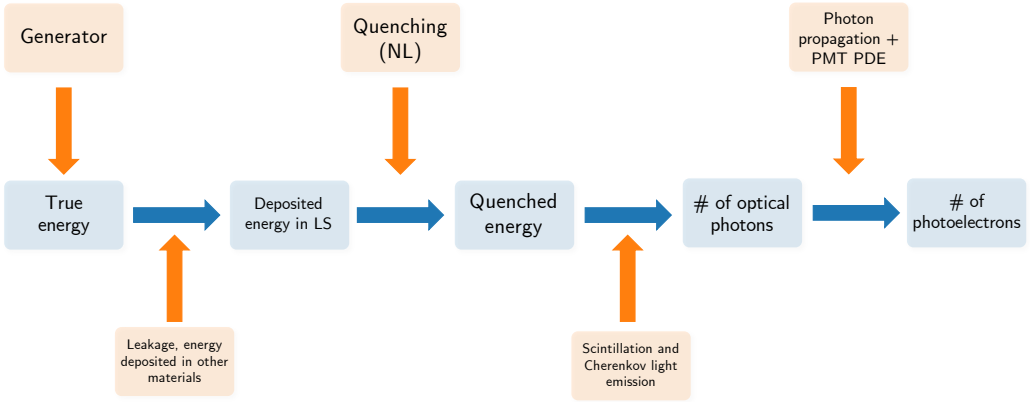


FIGURE 2.5: Detector simulation scheme in SNiPER [49], from the generator of the physics process to the number of photoelectrons released at the photocatode.

The `ElecSim` stage in JUNO software is responsible for modeling the response of the PMTs to hits from the `DetSim` package and for simulating the behavior of the readout electronics of the CD, WP, and TT, with their corresponding trigger logic. The hits are first merged, and then the detected waveforms, or PMT signals, are simulated with a digitization sampling time of 1 ns. The multiplicity trigger in JUNO [50] has a default trigger condition of 200 fired large PMTs within a 80 ns window. If the trigger condition is met, a positive trigger decision is made, and all waveforms in a window of 1 µs, starting from −100 ns before the trigger time, are bundled into an event.
Furthermore, this package [49] incorporates an event mixing procedure that combines events from different `DetSim` sources, and arranges the corresponding hits according to predefined event rates. This feature allows for the accounting and handling of pileup events, where multiple events occur within the same time window, thus creating MC samples that simulate data-like features.

2.2.2 Event reconstruction

The reconstruction part of the JUNO software consists of two additional stages, comprising calibration and event reconstruction. The calibration stage is aimed at reconstructing the PMT waveforms that are provided as output by either the `ElecSim` simulation or, in the future, by the data acquisition system. In particular, when particles deposit energy in the LS, they generate optical photons that result in PMT signals. The LPMT readout electronics [51] employs Flash Analog-to-Digital

[9] PDE = QE · CE, and it is of the order of 28-30% for JUNO 20-inch PMTs [48].

converters (FADCs) to convert these PMT signals into 1 µs Data Acquisition (DAQ) windows with a sampling frequency of 1 GHz. These raw PMT waveforms may contain overshoots and systematic electronics noise, which have to be carefully addressed in order to achieve accurate reconstruction. To that end, different algorithms are implemented in the JUNO software and have been tested to perform waveform reconstruction [52]. The default approach is based upon a Fast-Fourier Transform (FFT): waveforms undergo a filtering procedure to eliminate high-frequency white noise, then they are deconvoluted with the single PE PMT response. This process disentangles the high-frequency electronics noise from the low-frequency signal, thereby yielding charge and time information for each photon hit. Subsequently, these reconstructed charge and time values serve as inputs for reconstructing the event vertex and energy, employing various algorithms [53, 54]. One of the simplest methods involves estimating the vertex position by summing the charge-weighted positions of the PMTs [53]. This approach provides a rough estimation of the vertex location based on the distribution of light detected by the PMTs. A more complex data-driven method [54] develops a simultaneous vertex and energy reconstruction: the algorithm combines the charge and time information of PMTs to enhance the reconstruction accuracy of positron events in JUNO.

2.3 High precision neutrino oscillation physics in JUNO

One of the main goals of JUNO [21, 30] is to determine the neutrino MO at high statistical significance ($\simeq 3\sigma$ in approximately 6 years of data taking). The MO information is encoded in the high-frequency oscillation peaks driven by $\Delta m^2_{3\ell}$ in the oscillated antineutrino spectrum, as illustrated in Figure 2.2. Therefore, a precise measurement of the oscillated antineutrino spectrum is essential for JUNO to determine the neutrino MO and requires a detector with an unprecedented effective energy resolution. JUNO can obtain valuable insights into the neutrino MO, independently of the unknown CP-violating phase and the octant of θ_{23}. This distinctive feature provides unique and complementary information when combined with other neutrino experiments.

JUNO's remarkable capability to measure the oscillated reactor antineutrino spectrum opens up the opportunity for an independent determination of the oscillation parameters Δm^2_{31}, Δm^2_{21}, $\sin^2\theta_{12}$, and $\sin^2\theta_{13}$, which constitute the main focus of this thesis work. Among these parameters, JUNO will attain a level of precision significantly better than 1% for the determination of the first three [55]. According to the latest global results in Table 1.2, the relative uncertainties on the parameters Δm^2_{21}, Δm^2_{31}, and $\sin^2\theta_{12}$ are estimated to be $\approx 2.7\%$, $\approx 1.1\%$, and $\approx 4\%$, respectively. The latest published study on JUNO's sensitivity to oscillation parameters [55] shows that the experiment will be able to estimate these parameters to a precision of 0.2%, 0.3%, and 0.5%, with 6 years exposure, and reaching sub-percent precision within the first year of data taking. This groundbreaking result heralds a new era of unparalleled precision in neutrino oscillation measurements. The foreseen precision will have far-reaching implications across various research fields, including particle physics, astrophysics, and cosmology. It will enable rigorous tests of the νSM, such as investigating the unitarity of the PMNS matrix [56], thereby offering the potential to uncover physics beyond the Standard Model. Moreover, this enhanced precision will significantly impact other experimental endeavors, reducing the parameter space in the search for leptonic CP violation [21, 30] and neutrinoless double beta decay [57], for example. This cutting-edge precision will serve as a powerful discriminator of neutrino masses and mixing models, provide constraints to other experiments, and strong handles to probe the three-flavor neutrino paradigm well beyond current limits.

2.3.1 Antineutrino survival probability

JUNO will detect electron antineutrinos from nuclear reactors at energies of the order of few MeV. Since the energy is not sufficiently high to allow the production of muons and taus in the LS target, only the $\overline{\nu}_e$ disappearance channel is accessible. In particular, the case where $\alpha = \beta$ in Eq. (1.14) yields the ν_α survival probability, namely

$$\mathcal{P}(\nu_\alpha \to \nu_\alpha) = 1 - 4 \sum_{i>j} |U_{\alpha i}|^2 \, |U_{\alpha j}|^2 \sin^2 \left(\frac{\Delta m^2_{ij} L}{4E} \right). \tag{2.1}$$

Moreover, CPT (Charge-Parity-Time reversal) is a symmetry of the oscillation probabilities [1], hence $\mathcal{P}(\nu_\alpha \to \nu_\alpha) = \mathcal{P}(\overline{\nu}_\alpha \to \overline{\nu}_\alpha)$. Consequently, the $\overline{\nu}_e$ survival probability reads:

$$\mathcal{P}(\overline{\nu}_e \to \overline{\nu}_e) \;=\; 1 - \sin^2 2\theta_{12} \, c^4_{13} \, \sin^2 \Delta_{21} - \sin^2 2\theta_{13} \left(c^2_{12} \sin^2 \Delta_{31} + s^2_{12} \sin^2 \Delta_{32} \right), \tag{2.2}$$

where the notation $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$, and $\Delta_{ij} = \Delta m^2_{ij} L/4E$ is used. There is no dependence on $\sin^2\theta_{23}$ nor on δ_{CP}.

2.3.2 Matter effects in JUNO

Despite the relatively minor nature of matter effects in JUNO compared to long-baseline oscillation experiments, their inclusion remains crucial for accurately determining the mixing parameters. Overlooking these effects would primarily impact the solar-induced oscillation, resulting in notable biases in Δm^2_{21} and $\sin^2\theta_{12}$, estimated to be approximately 1% and 0.2% respectively [25, 26]. Figure 2.6 shows the absolute and relative (in the inset) difference between $\mathcal{P}(\overline{\nu}_e \to \overline{\nu}_e)$ in matter and vacuum. The absolute difference reaches a maximum of approximately 0.0065 in the proximity of the first low-frequency oscillation peak, consequently yielding a relative correction of roughly 3.7%. As a consequence, the matter-induced effect would hinder the foreseen accuracy in the estimation of solar parameters driving the slow oscillation pattern. This clearly highlights the significance of accounting for sub-percent effects arising from matter propagation, which are not negligible compared with the foreseen accuracy on the parameters of interest [58]. In order to include matter effects, one needs to consider the effective Hamiltonian in Eq. (1.19).

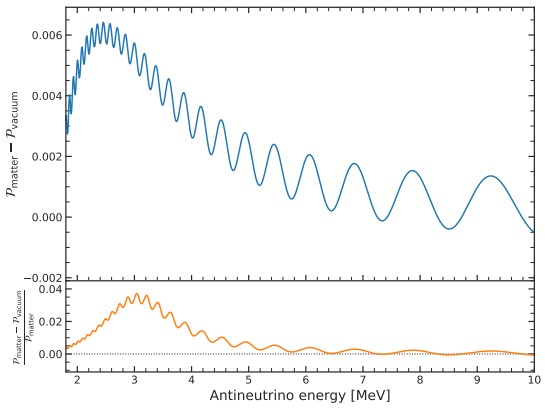


FIGURE 2.6: Absolute and relative (in the inset) difference between $\mathcal{P}(\overline{\nu}_e \to \overline{\nu}_e)$ in matter and vacuum, for a fixed baseline $L = 52.5\,\mathrm{km}$.

For medium baseline experiments like JUNO, reactor $\overline{\nu}_e$ mostly propagate within the upper part of the crust and the sediment of the Earth. Thus, for JUNO case, the parameter A in Eq. (1.19) can be expressed as:

$$A = \pm\, 2\sqrt{2} G_{\mathrm{F}} N_e E \simeq \pm\, 1.52 \times 10^{-4} \mathrm{eV}^2 \cdot Y_e \cdot \frac{\rho}{\mathrm{g/cm^3}} \cdot \frac{E}{\mathrm{GeV}}, \tag{2.3}$$

where G_{F} is the Fermi constant, N_e is the number density of electrons, $Y_e \simeq 0.5$ is the electron fraction, and $\rho = (2.45 \pm 0.15)\ \mathrm{g/cm^3}$ is the estimated average matter density [59] with its associated uncertainty. For a constant matter profile, it is sufficient to replace the mixing angles

and mass splittings in Eq. (2.2) by the corresponding matter parameters (indicated with a tilde). Consequently, the reactor antineutrino survival probability can be expressed as

$$\begin{aligned}\mathcal{P}(\overline{\nu}_e \to \overline{\nu}_e) &= 1 - \sin^2 2\widetilde{\theta}_{12}\,\tilde{c}_{13}^4\,\sin^2\widetilde{\Delta}_{21} - \sin^2 2\widetilde{\theta}_{13}\left(\tilde{c}_{12}^2\sin^2\widetilde{\Delta}_{31} + \tilde{s}_{12}^2\sin^2\widetilde{\Delta}_{32}\right) \\ &= 1 - \sin^2 2\widetilde{\theta}_{12}\tilde{c}_{13}^4\sin^2\widetilde{\Delta}_{21} - \frac{1}{2}\sin^2 2\widetilde{\theta}_{13}\left(\sin^2\widetilde{\Delta}_{31} + \sin^2\widetilde{\Delta}_{32}\right) \\ &\quad - \frac{1}{2}\cos 2\widetilde{\theta}_{12}\sin^2 2\widetilde{\theta}_{13}\sin\widetilde{\Delta}_{21}\sin(\widetilde{\Delta}_{31} + \widetilde{\Delta}_{32}). \end{aligned} \tag{2.4}$$

The effective matter-induced mixing angles in $\widetilde{U}$ are denoted by $\widetilde{\theta}_{ij}, (i,j = 1,2,3,, i < j)$, and are parameterized as $\tilde{c}_{ij} \equiv \cos\widetilde{\theta}_{ij}$ and $\tilde{s}_{ij} \equiv \sin\widetilde{\theta}_{ij}$, keeping the same structure as Eq. (1.3). Here, $\Delta\tilde{m}^2_{ij}$ represents the effective matter-induced mass squared difference, and $\widetilde{\Delta}_{ij} = \Delta\tilde{m}^2_{ij}L/4E$, where L stands for the baseline, and E is the neutrino energy. The full expression of the oscillation parameters appearing in Eq. (2.4) can be found in Appendix A.
In the second and third lines of Eq. (2.4), the survival probability is reformulated to factor out the solar-dominated, atmospheric-dominated, and MO-sensitive terms, respectively.
A graphical representation of the electron antineutrino survival probability is shown in Figure 2.7, with a fixed energy $E_{\overline{\nu}_e} = 3\,\text{MeV}$ and variable baseline.

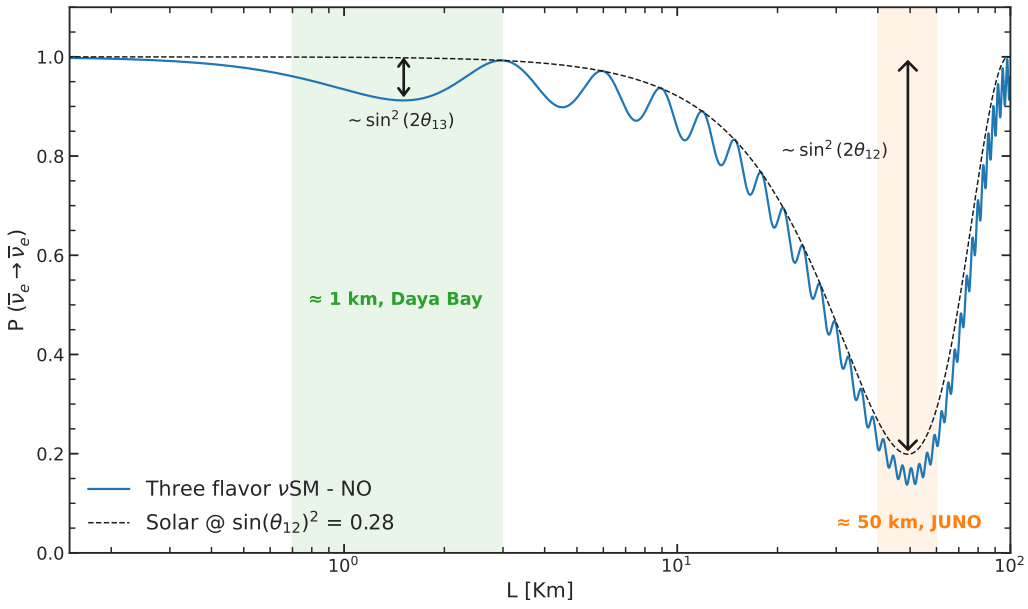


FIGURE 2.7: Antineutrino survival probability according to Eq. (2.4) at fixed energy $E_{\overline{\nu}_e} = 3\,\text{MeV}$ and variable baseline. To enhance visual clarity, the oscillation induced by the solar term is illustrated using $\sin^2\theta_{12} = 0.28$ instead of $\sin^2\theta_{12} = 0.303$. Two region of interest for reactor antineutrino experiments are highlighted, corresponding to short baseline and medium baseline facilities, e.g., Daya Baya [14] and JUNO [30], respectively.

Chapter 3

Reactor antineutrino production and detection

To accurately predict the observed energy spectrum, several factors must be taken into consideration, including the source, propagation, interaction, and subsequent detection of the involved particles. The source of interest in this context is the reactor antineutrino flux, which is discussed in detail in Section 3.1. The propagation aspect relates to the survival probability of electron antineutrinos ($\overline{\nu}_e$), as described by Eq. (2.4). Then, a comprehensive overview of the kinematics and cross section of the Inverse Beta Decay (IBD) is provided in Section 3.2. This reaction serves as the primary interaction mechanism between antineutrinos and the LS target. Finally, the detection of the antineutrinos requires understanding the response of the detector, which is described in Section 3.3, considering the latest knowledge and advancements specific to the JUNO experiment.

3.1 Reactor antineutrino flux

Reactor antineutrinos detected by JUNO originate from the nearby Taishan and Yangjiang NPPs, which consist of two and six reactor cores, respectively. The average distance from the experiment to all reactor cores is 52.5 km. The next closest reactor complex is Daya Bay, at a distance of 215 km. Due to its longer baseline, its antineutrino flux produces a reduced sensitivity to oscillation parameters [55]. These NPPs operate commercial pressurized water reactors (PWRs), where electron antineutrinos are produced by the β decay of fission products of four major isotopes: ^{235}U, ^{238}U, ^{239}Pu, and ^{241}Pu. These isotopes contribute over 99% of the thermal power and total neutrino flux [30] above the 1.8 MeV reaction threshold in JUNO (see Section 3.2).
Table 3.1 provides a summary of the average reactor power, baselines, and expected IBD rates, taking into account oscillation effects, for each core of the Taishan, Yangjiang, and Daya Bay NPPs. In Figure 3.1a, the breakdown of each reactor core's contribution to the expected reconstructed energy spectrum at JUNO is illustrated; the Daya Bay-induced oscillation pattern notably stands out in the figure.
The oscillated antineutrino flux at time t is predicted as

$$\Phi(E_{\overline{\nu}_e}, t) = \sum_r \frac{\mathcal{P}_{\overline{\nu}_e \to \overline{\nu}_e}(E_{\overline{\nu}_e}, L_r)}{4\pi L_r^2} \frac{W_r(t)}{\sum_i f_{ir}(t) e_i} \sum_i f_{ir}(t) s_i(E_{\overline{\nu}_e}), \tag{3.1}$$

where r is the reactor index, $\mathcal{P}_{\overline{\nu}_e \to \overline{\nu}_e}(E_{\overline{\nu}_e}, L_r)$ is the $\overline{\nu}_e$ survival probability in Eq. (2.4) at a distance L_r from the reactor r, $W_r(t)$ is the reactor thermal power, $f_{ir}(t)$ is the fission fraction of one isotope i among the four, e_i is the mean energy released per fission for isotope i, and $s_i(E_{\overline{\nu}_e})$ is the antineutrino energy spectrum per fission for each isotope. The reactor thermal power and fission fractions are time-dependent and will be provided by the NPPs during the data-taking period; for this reason, in current studies [55] the reactor thermal power and fission fraction are assumed to

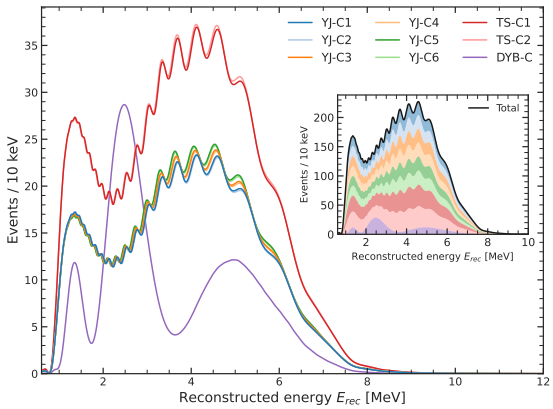
Reactor	Power [GW_{th}]	Baseline [km]	IBD rate [events/day]
Taishan	9.2	52.71	Total: 18.4
Core 1	4.6	52.77	9.2
Core 2	4.6	52.64	9.3
Yangjiang	17.4	52.46	Total: 35.3
Core 1	2.9	52.74	5.8
Core 2	2.9	52.82	5.7
Core 3	2.9	52.41	5.9
Core 4	2.9	52.49	5.9
Core 5	2.9	52.11	6.0
Core 6	2.9	52.19	6.0
Daya Bay	17.4	215	3.7
Total flux			57.4

TABLE 3.1: Characteristics of NPPs and their reactor cores considered in this analysis: the two closest ones to JUNO, Taishan and Yangjiang, at an approximate distance of 52.5 km, and the next closest, Daya Bay. The average distance and the total average thermal power of each NPP are indicated in the corresponding row. The expected daily IBD rates are estimated considering the baselines, full thermal power of the reactors, and neutrino oscillation parameters in Table 1.2.

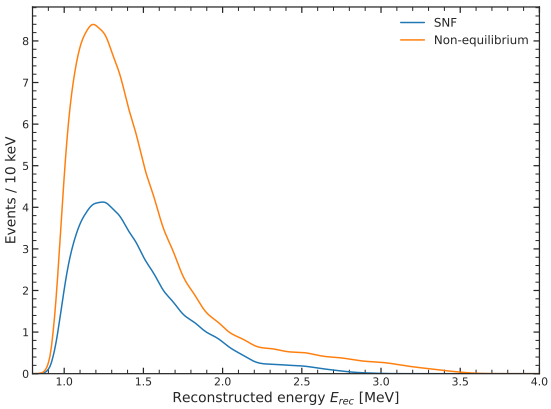
be stable at averaged values. To account for the refueling period during which the reactors are temporarily shut down, typically lasting for one month each year, a duty cycle factor of 11/12 is applied. This factor is used to scale down the thermal power, resulting in a reduction of the antineutrino flux.
In this work, the energy spectrum of $\overline{\nu}_e$ per fission for ^{235}U, ^{239}Pu, and ^{241}Pu is obtained from the Huber model [61], and for ^{238}U from Mueller's [62]. The reactor antineutrinos yield per fission is shown in Figure 3.3a.
Furthermore, additional corrections can be applied to these models to account for experimental observations by other reactor experiments, as well as effects related to the working principle of the NPPs.

Spent nuclear fuel and non-equilibrium effect In the Huber-Mueller model, the beta decay rates of some long-lived fission fragments do not reach equilibrium with their production rates, while in the reactor core these fission fragments will accumulate and reach equilibrium. As a result, this so-called *non-equilibrium* effect produces a 0.6% increased antineutrino flux with respect to the evaluation in [62]. Moreover, the Spent (burned) Nuclear Fuel (SNF) is removed to cooling pools near the reactor core but still emits antineutrinos, contributing to an additional 0.3% in the flux (based on calculations in [60]). Thus, the antineutrino energy spectrum has to be modified to cover the contributions of SNF and non-equilibrium: the resulting contributions to the reconstructed visible energy spectrum are reported in Figure 3.1b.

Reactor antineutrino anomaly from Daya Bay The observed antineutrino yield per fission shows a $\sim$ 5% deficit compared to the model predictions, which is referred to as the reactor antineutrino anomaly [63]. Furthermore, recent reactor antineutrino experiments, e.g., Daya Bay [64], RENO [15], etc., also observed an additional discrepancy in the antineutrino energy spectrum. Specifically, an excess of neutrino events around 5 MeV, the so-called *reactor bump*, is observed [64]. In this analysis, the $\overline{\nu}_e$ antineutrino flux is obtained considering the Huber-Mueller model with



(A) Contribution from the single reactor cores to the expected reconstructed energy spectrum at JUNO.



(B) Non-equilibrium effect and Spent Nuclear Fuel contributions to JUNO visible energy spectrum.

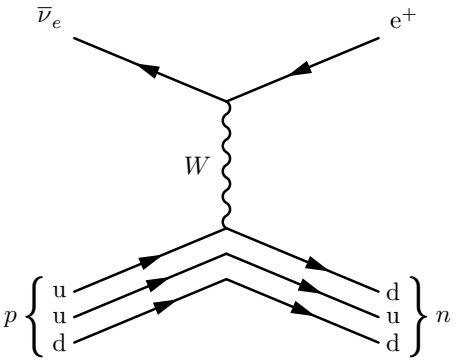
FIGURE 3.1: (a) Contribution from the single reactor cores to the expected reconstructed energy spectrum, for 6 years exposure and 11/12 duty cycle power. Yangjiang is indicated with "YJ", Taishan with "TS" and Daya Bay with "DYB". The inset plot shows the cumulative contribution and the resulting total one (black line). (b) Additional contributions to the reconstructed energy spectrum due to non-equilibrium effect and Spent Nuclear Fuel (SNF) [60].

a correction on both the rate and spectral shape based on the measurements in the Daya Bay experiment [64].

3.2 IBD reaction and detection

Reactor antineutrinos are detected in JUNO through the IBD reaction

$$\overline{\nu}_e + p \rightarrow e^+ + n \quad :$$



specifically by measuring the produced positron energy spectrum, since it retains almost all of the incoming antineutrino kinetic energy. The electron antineutrino interacts with a proton (p) in the LS, creating a positron (e^+) and a neutron (n). The positron quickly deposits its energy and annihilates into two 0.511 MeV photons, producing a *prompt* signal containing both the positron kinetic energy T_{e^+} and the 1.022 MeV annihilation energy. While the emission of positrons is approximately isotropic, with a slight predilection for the backward direction [65], the neutron absorbs most of the momentum of the antineutrino and its initial direction is largely parallel to that of the incoming antineutrino. The neutron then makes a random walk and is thermalized in the detector medium by successive scattering events, causing a deflection from its original path, after an average time of 220 µs[10]. It is finally captured either by a free proton in the LS ($\sim$ 99% probability)

[10]The duration of this process depends on the average neutron capture cross section, hence on the target material.

or by carbon ($\sim 1\%$), subsequently emitting a 2.22 MeV or 4.95 MeV γ-ray, respectively, and giving rise to a *delayed* signal. A sketch, depicting the IBD kinematics from the emission of both particles to their annihilation or capture, is shown in Figure 3.2.

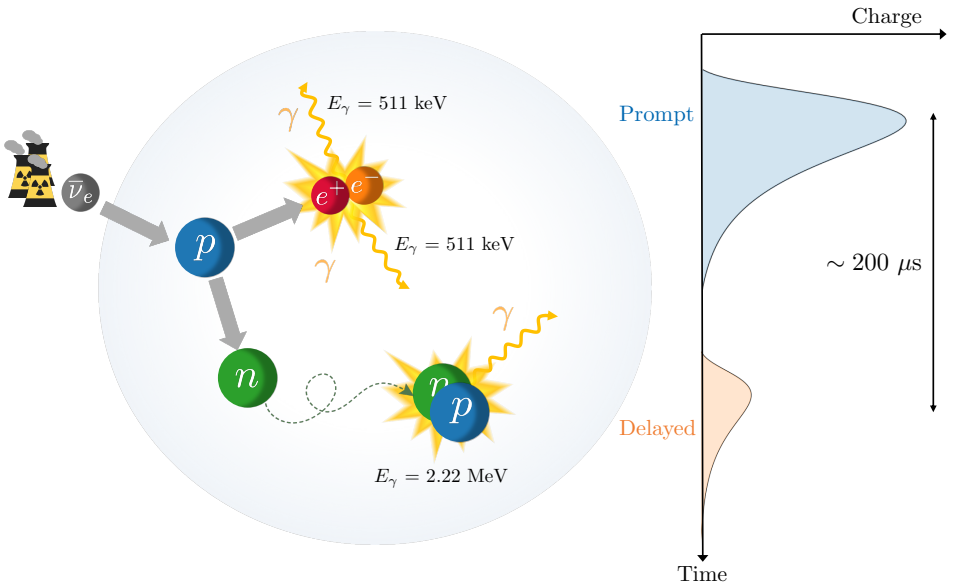


FIGURE 3.2: Schematic illustration of an IBD reaction in the LS. The electron antineutrino interacts with a proton (p) in the LS, creating a positron (e^+) and a neutron (n). The positron deposits its energy and annihilates into two 0.511 MeV photons (γ), producing a prompt signal. After a random walk, the neutron is captured mainly by a free proton in the LS, emitting a 2.22 MeV γ-ray, and giving rise to a delayed signal. The time-charge diagram is not in scale.

In the IBD reaction the incident antineutrino energy $E_{\bar{\nu}_e}$ is given by [66]

$$E_{\bar{\nu}_e} \simeq T_{e^+} + T_n + \Delta_{np} + m_{e^+}, \tag{3.2}$$

where T_{e^+} and T_n are the positron and neutron kinetic energies, $\Delta_{np} = m_n - m_p = 1.293$ MeV is the mass difference between neutron and proton, and m_{e^+} is the positron mass. T_n ranges from 0 to tens of keV, thus the neutron recoil energy is often considered to be fairly negligible with respect to the $O(1\,\text{MeV})$ energy scales of the experiment. Under this *recoilless* approximation, the positron energy is directly related to the antineutrino energy, i.e., $E_{\bar{\nu}_e} \simeq E_{e^+} + \Delta_{np}$ [58], with $E_{e^+} = T_{e^+} + m_{e^+}$. The IBD kinematical threshold is roughly 1.8 MeV (as explained in Appendix B).
The deposited energy (E_{dep}) for the IBD prompt signal is thus defined as the sum of the positron kinetic energy T_{e^+} and the annihilation energy generating two 0.511 MeV photons, as expressed in Eq. (3.3).

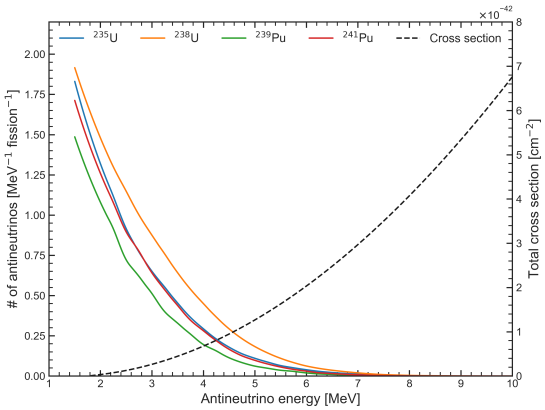
$$E_{\text{dep}} = T_{e^+} + 2 \times 0.511\,\text{MeV} = E_{e^+} + m_{e^+}. \tag{3.3}$$

As a result, $E_{e^+} + m_{e^+} \sim E_{\bar{\nu}_e} - 0.782$ MeV, or equivalently

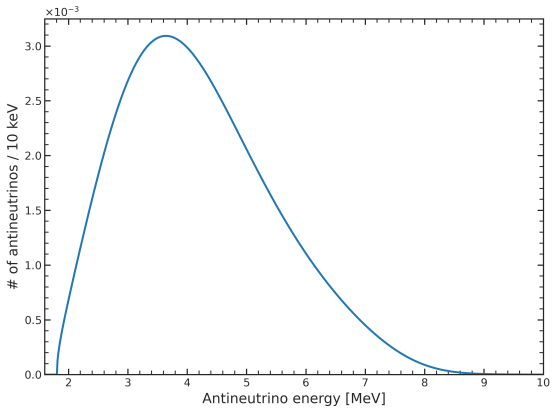
$$E_{\text{dep}} \simeq E_{\bar{\nu}_e} - 0.782\,\text{MeV}. \tag{3.4}$$

E_{dep} is assumed to be fully deposited in the detector since energy losses due to escaping secondary particles generated by Compton scattering and pair production only affect less than 1% of the IBD events. Therefore, the measured positron spectrum corresponds to the antineutrino spectrum,

shifted in energy and weighted by the IBD differential cross section $\frac{d\sigma}{d\cos\theta}(E_{\bar{\nu}_e}, \cos\theta)$, which is a function of the neutrino energy and the scattering angle [67], and accounts for the interaction probability. There are two widely used formulas for evaluating the IBD cross-section: one by Vogel and Beacom [65], and the other by Strumia and Vissani [67]. For this analysis, the Strumia-Vissani model shown in Figure 3.3a (dashed line) was used, as it includes radiative corrections and is valid in most of the sub-GeV energy range. Further details on the cross section are provided in Appendix B. The expected $\overline{\nu}_e$ energy spectrum, based on the Huber-Muller model and without anomaly corrections, obtained as a product of the IBD cross section and the $\overline{\nu}_e$ flux, is shown in Figure 3.3b.



(A) Reactor antineutrinos flux per fission and total IBD cross section calculated by Strumia and Vissani [67].



(B) $\overline{\nu}_e$ energy spectrum obtained as a product of the IBD cross section and the $\overline{\nu}_e$ flux.

FIGURE 3.3: (a) Reactor antineutrinos yield per fission and total IBD cross section. (b) Expected $\overline{\nu}_e$ energy spectrum, normalized to one, obtained using the Huber-Mueller model without any additional corrections.

3.2.1 Neutron recoil effects

Recoil effects, although often neglected in calculations for the sake of simplicity, have been shown to have a non-negligible impact in high-precision experiments like JUNO [58]. When a neutrino scatters off a nucleus, a small fraction of energy on the order of $O(E_{\bar{\nu}_e}/m_p)$ is transferred to the recoiling nucleon. Consequently, the estimation provided in Eq. (3.4) only serves as an approximate upper bound. Indeed, the recoil of the neutron induces an angle-dependent deficit in E_{e^+}, which falls inside a defined kinematical range, namely

$$E_{e^+} \in [E_1, E_2], \tag{3.5}$$

where the explicit expressions of $E_{1,2}$ are taken from [67] and reported in Appendix B. In this analysis, we employ a less accurate approximation, denoted as "mid-recoil" [58, 68], where the mid-point of Eq. (3.5) is considered as a proxy for E_{e^+}:

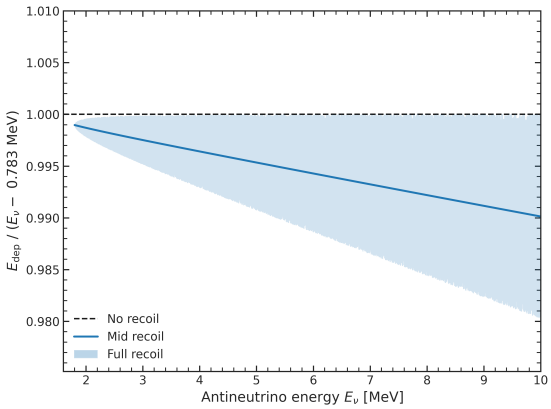
$$E_{e^+} \simeq E_{e^+}^{\text{mid}} \equiv \frac{E_1 + E_2}{2} \quad \rightarrow \quad E_{\text{dep}} = E_{e^+}^{\text{mid}} + m_{e^+}. \tag{3.6}$$

Figure 3.4a shows the recoil correction as a function of the antineutrino energy, in terms of deviation from the relation in Eq. (3.4): the solid line marks the mid-point average effect, while the filled area corresponds to the full recoil-induced energy deficit. It is worth noting that at high antineutrino energy, the deposited prompt energy results to be both shifted and smeared. Hence, while the mid-recoil recipe provides a reasonable estimate of the average shift, it does not account for the energy spread induced by recoil effects. To further investigate this aspect, Figure 3.4b reports the fractional

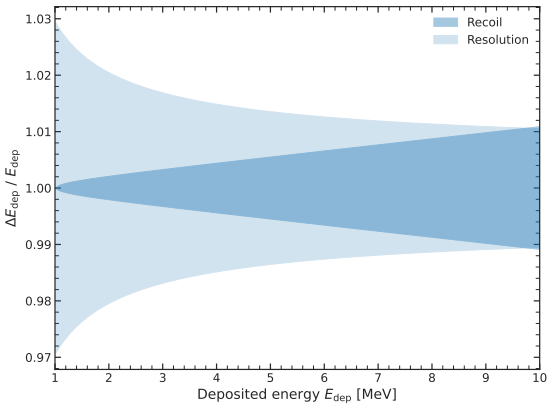
energy spread $\Delta_{E_{\rm dep}}/E_{\rm dep}$ due to recoil effects and intrinsic detector resolution (see Section 3.3 for the details). Specifically, by defining $\delta = \frac{E_2 - E_1}{2}$, the fractional energy spread can be expressed as:

$$\Delta E_{\rm dep} = \begin{cases} E_{\rm dep} \pm \delta & \text{for recoil effects} \\ E_{\rm dep} \pm \sigma_{E_{\rm dep}} & \text{for intrinsic resolution} \end{cases} \tag{3.7}$$

where $\sigma_{E_{\rm dep}}$ is set at an effective 3% energy resolution. Figure 3.4b shows that in JUNO, the intrinsic energy smearing of the IBD reaction caused by the neutron recoil is up to some extent relatively small compared to the limited energy resolution of the detector; this observation holds true at least for $E_{\rm dep} \leq 4\,\text{MeV}$, i.e., the region of highest sensitivity to the neutrino MO. The two effects can be simultaneously taken into account in the calculation of reactor antineutrino spectra by incorporating suitable modifications to the energy resolution function. This correction has been explored in the literature, as demonstrated in [58], but it is not investigated in the present work.



(A) Fractional recoil as a function of antineutrino energy [68].



(B) Energy smearing due to recoil compared with 1σ JUNO energy resolution.

FIGURE 3.4: (a) Fractional recoil effects in terms of antineutrino energy, in three different cases: no recoil, mid recoil approximation, and full recoil. (b) Energy spread due to nucleon recoil compared with energy resolution expected in JUNO. The definition of $\Delta E_{\rm dep}$ can be found in Eq. (3.7).

3.3 Detector response

The detector energy response model is encoded in a function mapping the antineutrino energy $E_{\bar{\nu}_e}$ to the reconstructed energy of the prompt IBD signal $E_{\rm rec}$, namely $R(E_{\bar{\nu}_e}, E_{\rm rec})$. The development of $R(E_{\bar{\nu}_e}, E_{\rm rec})$ takes into account three effects: the energy transfer in the IBD reaction (as discussed in Section 3.2), energy non-linearity brought on by scintillation and Cherenkov processes, and energy resolution. The energy conversion scheme depicted in Figure 3.5 elucidates the conventions employed for denoting the different forms of energy that are involved.
Figure 3.6 shows the energy spectrum in terms of deposited, visible, and reconstructed energy, for normal mass ordering and with oscillation parameters fixed to the global best fit values reported in Table 1.2.

3.3.1 Energy non-linearity

When positrons interact with a scintillator detector, they generate photons through scintillation and sub-dominant Cherenkov radiation mechanisms. However, the correlation between the deposited

FIGURE 3.5: Energy conversion scheme. The energy of antineutrinos, denoted as $E_{\bar{\nu}_e}$, undergoes a conversion to positron deposited energy, represented as E_{dep}, through the IBD reaction. Subsequently, the energy spectrum experiences distortion due to the non-linearity of the liquid scintillator (LSNL), leading to the measurement of visible energy E_{vis} in terms of detected photoelectrons. Finally, the finite energy resolution introduces additional smearing, resulting in the reconstructed energy E_{rec}.

energy and the number of scintillation photons detected by the photomultiplier tubes (PMTs) is not strictly linear, primarily due to the quenching effect. This effect is conventionally characterized by Birks' formula [69], an empirical parameterization that establishes the relationship between the measured light yield and the energy deposited in the scintillator. Birks' formula is expressed as follows:

$$\frac{dL}{dx} = \frac{S\frac{dE}{dx}}{1 + k_B\frac{dE}{dx}}, \tag{3.8}$$

where L is the scintillation light yield, E is the energy deposited per unit path length, S is the scintillation efficiency, k_B is the Birks' constant, and $\frac{dE}{dx}$ is the energy loss per path length. Moreover, to address the non-linear behavior of the quenching factor at high energy depositions, a modified version of Birks' formula was introduced. This extended model, known as the Birks-Chou model [70], provides a more accurate representation and is given by:

$$\frac{dL}{dx} = \frac{S\frac{dE}{dx}}{1 + k_B\frac{dE}{dx} + k_C\left(\frac{dE}{dx}\right)^2}, \tag{3.9}$$

where k_C is a new constant introduced to account for the nonlinear behavior of the quenching factor at high energies.
Furthermore, Cherenkov light is typically emitted at shorter photon wavelengths and is predominantly absorbed and re-emitted within the scintillator material. Consequently, distinguishing Cherenkov light from scintillation light becomes challenging, requiring the consideration of both components in the non-linear response of the detector.
In contrast, the *instrumental* non-linearity stemming from the charge response of JUNO PMTs and electronics can be considered quite negligible, with an estimated magnitude of less than 0.3%. This can be achieved thanks to the implementation of the dual-calorimetry calibration technique [42].

To describe deviations from linearity, the liquid scintillator non-linearity (LSNL) is defined as

$$E_{\mathrm{vis}} = f_{\mathrm{LSNL}}(E_{\mathrm{dep}}) \cdot E_{\mathrm{dep}}, \tag{3.10}$$

where E_{dep} is the deposited energy, E_{vis} is the visible energy assuming perfect energy resolution and $f_{\mathrm{LSNL}}(E_{\mathrm{dep}})$ is the LSNL function. Since a gamma deposits its energy into the LS via secondary electrons, it will be possible to achieve a rigorous determination of the LSNL function $f_{\mathrm{LSNL}}(E_{\mathrm{dep}})$ using gamma calibration data. According to the JUNO calibration strategy [42], the visible energy E_{vis} is estimated from the total number of detected photoelectrons with a scale constant set by 2.22 MeV gammas, which are radiated by neutron capture on hydrogen (n-H) in the detector. In

other words, the visible energy E_{vis} is defined as the energy in terms of detected photoelectrons, i.e.,

$$E_{\text{vis}} = \text{PE}/Y_0, \tag{3.11}$$

where Y_0 is a constant light yield obtained by considering neutron captures on hydrogen at the detector center and dividing the mean number of detected PEs by the 2.22 MeV expected capture energy.

In this analysis, $f_{\text{LSNL}}(E_{\text{dep}})$ is modeled as a four-parameter function, as used in the first Daya Bay analysis [14], namely

$$f_{\text{LSNL}}(E_{\text{dep}}) = \frac{p_0 + \frac{p_3}{E_{\text{dep}}}}{1 + p_1 e^{-p_2 E_{\text{dep}}}}. \tag{3.12}$$

Since the composition of the LS in Daya Bay and JUNO is similar, it is reasonable to adopt the Daya Bay non-linearity curves [71], rescaled with a constant obtained from the gamma LSNL curve at 2.22 MeV [42] in JUNO. The inset of Figure 3.6 displays the energy non-linearity response as a function of the deposited IBD prompt energy. It is worth noting that the curve intersects $\frac{E_{\text{vis}}}{E_{\text{dep}}} = 1$ at approximately 3.5 MeV, deviating from the 2.2 MeV anchoring point, owing to the differing non-linearity between positrons and gammas.

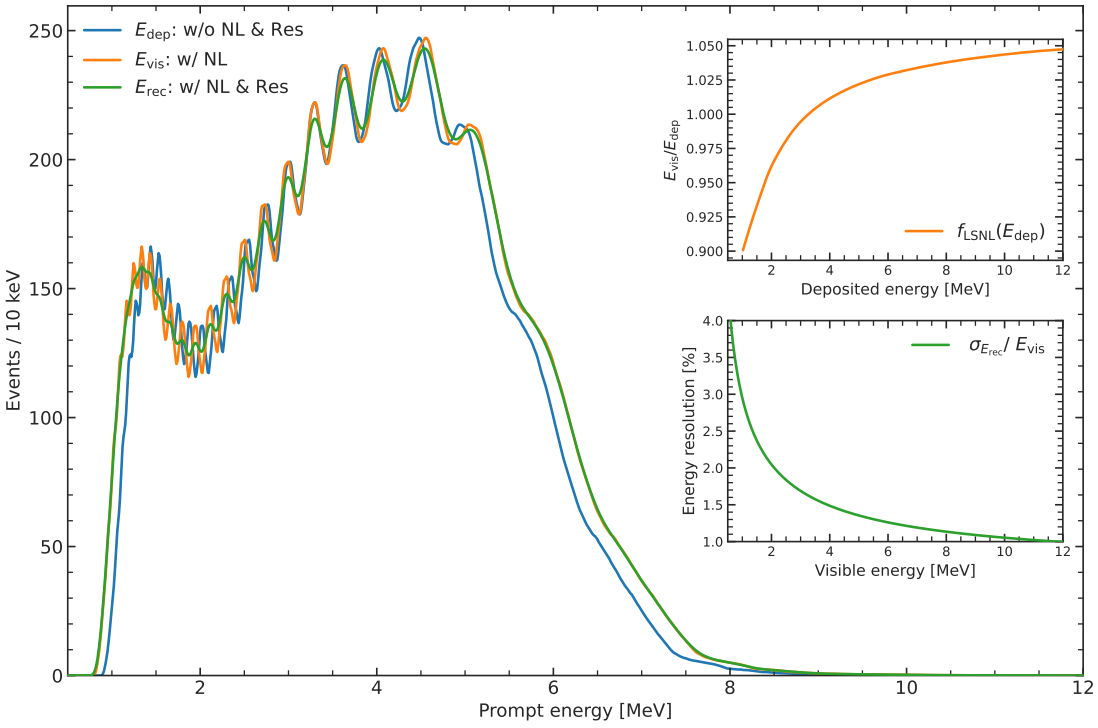


FIGURE 3.6: The expected prompt energy spectra of JUNO with and without the different detector response effects, i.e., liquid scintillator non-linearity (NL) and energy resolution (Res). The detector response curves used in this analysis, i.e., LSNL and resolution, are shown in the insets as a function of deposited and visible energy, respectively.

3.3.2 Energy resolution

The visible energy E_{vis} is further smeared because of the finite energy resolution of the detector; neglecting detector leakage effects[11], the smeared reconstructed energy E_{rec} is assumed to follow a Gaussian distribution with mean E_{vis} and standard deviation $\sigma_{E_{\text{rec}}}$. A standard parameterization for the fractional energy resolution is considered in Eq. (3.13), in order to include systematic effects.

$$\frac{\sigma_{E_{\text{rec}}}}{E_{\text{vis}}} = \sqrt{\left(\frac{a}{\sqrt{E_{\text{vis}}}}\right)^2 + b^2 + \left(\frac{c}{E_{\text{vis}}}\right)^2}, \tag{3.13}$$

where a is the term driven by the Poisson statistics of the number of detected photoelectrons, b is independent of energy and is dominated by the detector's spatial and temporal non-uniformity, and c accounts for the PMT dark noise and for the two annihilation gammas. The updated energy resolution curve is shown in the inset of Figure 3.6 as a function of reconstructed IBD prompt energy for events uniformly distributed inside the fiducial volume. The obtained parameters [72] are: $a = (2.614 \pm 0.005) \times 10^{-2}$ $\sqrt{\text{MeV}}$, $b = (0.640 \pm 0.003) \times 10^{-2}$, $c = (1.20 \pm 0.01) \times 10^{-2}$ MeV. Thus, considering both LSNL and resolution, the final formula of the response function $R(E_{\bar{\nu}_e}, E_{\text{rec}})$ can be expressed with Eq. (3.14).

$$R\left(E_{\bar{\nu}_e}, E_{\text{rec}}\right) \equiv \frac{d\sigma}{d\cos\theta}\left(E_{\bar{\nu}_e}, \cos\theta\right) \cdot N_p \cdot \varepsilon \cdot \frac{1}{\sigma_{E_{\text{rec}}}\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{E_{\text{rec}} - E_{\text{vis}}}{\sigma_{E_{\text{rec}}}}\right)^2\right). \tag{3.14}$$

E_{dep} is calculated with Eq. (3.6), E_{vis} is evaluated accounting for LS non-linearity with Eq. (3.10), and E_{rec} refers to the smeared visible (i.e., reconstructed) energy of the prompt signal.
$N_p = 1.44 \times 10^{33}$ is the number of free protons in the 20-kton liquid scintillator target [21], and it is calculated as follows:

$$N_p = \frac{M_{\text{target}} f_{\text{H}} \alpha_{^1\text{H}}}{m_{^1\text{H}}}, \tag{3.15}$$

where M_{target} is the mass of the LS in the target, f_{H} is the hydrogen mass fraction in JUNO LS, $\alpha_{^1\text{H}}$ is the isotope abundance (in terms of mass fraction) of ^{1}H in natural hydrogen, and $m_{^1\text{H}}$ is the atomic mass. $\varepsilon = 82.1\%$ is the IBD event selection efficiency calculated later in Section 4.2.1. These parameters and the corresponding values are reported in Table 3.2.

Parameter	Value
M_{target}	20 kton
f_{H}	0.1201
$\alpha_{^1\text{H}}$	0.999885
$m_{^1\text{H}}$	1.008 Da [73]
ε	82.1%
N_p	1.43512×10^{33}

TABLE 3.2: Parameters related to the normalization of the number of detected IBDs.

[11] This approximation is highly accurate, considering the substantial size of the detector and the implementation of a fiducial volume cut. The impact of energy losses caused by escaping secondary particles is minimal, affecting less than 1% of the IBD events.

Additionally, it can be advantageous to isolate the cross-section contribution from Eq. (3.14) by introducing a separate definition,

$$R'\left(E_{\text{vis}}, E_{\text{rec}}\right) \equiv N_p \cdot \varepsilon \cdot \frac{1}{\sigma_{E_{\text{rec}}}\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{E_{\text{rec}} - E_{\text{vis}}}{\sigma_{E_{\text{rec}}}}\right)^2\right), \tag{3.16}$$

therefore

$$R\left(E_{\bar{\nu}_e}, E_{\text{rec}}\right) = \frac{d\sigma}{d\cos\theta}\left(E_{\bar{\nu}_e}, \cos\theta\right) \cdot R'\left(E_{\text{vis}}, E_{\text{rec}}\right). \tag{3.17}$$

Putting together the expected reactor antineutrino flux in Eq. (3.1), the detector response function in Eq. (3.17) and integrating over the scattering angle, the expected IBD signal reconstructed energy spectrum is obtained:

$$S(E_{\text{rec}}) = \int_{T_{\text{DAQ}}} dt \int_{1.8\,\text{MeV}}^{15\,\text{MeV}} dE_{\bar{\nu}_e} \cdot \Phi(E_{\bar{\nu}_e}, t) \cdot \sigma(E_{\bar{\nu}_e}) \cdot R'\left(E_{\text{vis}}, E_{\text{rec}}\right), \tag{3.18}$$

T_{DAQ} is the data taking period, $\Phi(E_{\bar{\nu}_e}, t)$ the oscillated reactor antineutrino flux in at time t (from Eq. (3.1)) and $\sigma(E_{\bar{\nu}_e})$ the total IBD cross section,

$$\sigma(E_{\bar{\nu}_e}) = \int_{-1}^{+1} d\cos\theta \frac{d\sigma}{d\cos\theta}\left(E_{\bar{\nu}_e}, \cos\theta\right). \tag{3.19}$$

The integration interval of the antineutrino energy starts from the IBD reaction threshold at 1.8 MeV and ends at 15 MeV, where the reactor antineutrino flux becomes negligible. The time evolution of the reactors and consequently of the antineutrino spectrum will be crucial in real data analysis, whereas it is currently unrealistic to precisely predict it. Hence, the neutrino flux is assumed to be time-independent in this analysis, and the time dependence in Eq. (3.18) is dropped.
Considering the expression of the expected IBD reconstructed prompt signal, it is possible to calculate the expected number of events s_i^{IBD} for the i-th energy bin, which in this analysis are obtained with the formula in Eq. (3.20).

$$s_i^{\text{IBD}} = \int_{E_{\text{rec}}^i}^{E_{\text{rec}}^{i+1}} dE_{\text{rec}} \int_{0.8\,\text{MeV}}^{12\,\text{MeV}} dE_{\text{vis}} \cdot \Phi(E_{\bar{\nu}_e}) \cdot \sigma(E_{\bar{\nu}_e}) \cdot R'\left(E_{\text{vis}}, E_{\text{rec}}\right) \cdot \frac{dE_{\bar{\nu}_e}}{dE_{\text{dep}}} \cdot \frac{dE_{\text{dep}}}{dE_{\text{vis}}}. \tag{3.20}$$

To compute a binned PDF we need to convolve the antineutrino spectrum $\Phi(E_{\bar{\nu}_e})$ with the energy smearing function $R'\left(E_{\text{vis}}, E_{\text{rec}}\right)$. In order to do that, both functions have to be expressed in terms of the same coordinates. Firstly the deposited spectrum $S_{\text{dep}}(E_{\text{dep}})$ is calculated from the antineutrino flux:

$$S_{\text{dep}}(E_{\text{dep}}) = \Phi(E_{\bar{\nu}_e}(E_{\text{dep}})) \cdot \sigma(E_{\bar{\nu}_e}) \cdot \frac{dE_{\bar{\nu}_e}}{dE_{\text{dep}}}, \tag{3.21}$$

where $E_{\bar{\nu}_e}$ (E_{dep}) represents the antineutrino energy at given deposited energy, and $\frac{dE_{\bar{\nu}_e}}{dE_{\text{dep}}} = \frac{1}{\mathcal{J}_{\text{dep}}}$, with $\mathcal{J}_{\text{dep}}$ energy conversion Jacobian[12], defined as follows:

$$\mathcal{J}_{\text{dep}} = \frac{E_{\text{dep}}(E^{upper}_{\bar{\nu}_e}) - E_{\text{dep}}(E^{lower}_{\bar{\nu}_e})}{E^{upper}_{\bar{\nu}_e} - E^{lower}_{\bar{\nu}_e}}. \tag{3.22}$$

$E^{upper}_{\bar{\nu}_e}$ and $E^{lower}_{\bar{\nu}_e}$ are the antineutrino energy bins lower and upper edges, while $E_{\text{dep}}(E^{upper}_{\bar{\nu}_e})$ and $E_{\text{dep}}(E^{lower}_{\bar{\nu}_e})$ the corresponding deposited energies. E_{dep} is obtained from Eq. (3.6), considering neutron recoil.
Then, the visible energy spectrum $S_{\text{vis}}(E_{\text{vis}})$ is calculated from $S_{\text{dep}}(E_{\text{dep}})$:

$$S_{\text{vis}}(E_{\text{vis}}) = S_{\text{dep}}(E_{\text{dep}}(E_{\text{vis}})) \cdot \frac{dE_{\text{dep}}}{dE_{\text{vis}}}. \tag{3.23}$$

The corresponding conversion Jacobian is used, namely $\frac{dE_{\text{dep}}}{dE_{\text{vis}}} = \frac{1}{\mathcal{J}_{\text{vis}}}$, with

$$\mathcal{J}_{\text{vis}} = \frac{E_{\text{vis}}(E^{upper}_{\text{dep}}) - E_{\text{vis}}(E^{lower}_{\text{dep}})}{E^{upper}_{\text{dep}} - E^{lower}_{\text{dep}}}, \tag{3.24}$$

where a similar notation is used to indicate the upper and lower energy bins edges, and $E_{\text{vis}}(E_{\text{dep}}) = f_{\text{LSNL}}(E_{\text{dep}}) \cdot E_{\text{dep}}$ from Eq. (3.10). Therefore, Eq. (3.20) reduces to

$$s_i^{\text{IBD}} = \int_{E^i_{\text{rec}}}^{E^{i+1}_{\text{rec}}} dE_{\text{rec}} \int_{0.8\,\text{MeV}}^{12\,\text{MeV}} dE_{\text{vis}} \cdot S_{\text{vis}}(E_{\text{vis}}) \cdot R'\left(E_{\text{vis}}, E_{\text{rec}}\right) \tag{3.25}$$

Therefore, the number of reactor IBD events in the i-th bin is obtained by eventually performing a convolution between the visible spectrum and the gaussian energy resolution model in Eq. (3.13). The resulting smeared spectrum is a histogram binned with 10 keV bins between 0.8 MeV and 12 MeV.

[12] Since the bin width is kept fixed for all the different energy spectra, the Jacobian is a scale factor that modifies the height of each bin. This ensures event number conservation.

Chapter 4

Reactor antineutrino selection: signal and backgrounds

As described in Section 3.2, reactor antineutrinos interact with protons inside the LS, producing a pair of prompt-delayed events. The detection of neutron capture in a delayed coincidence with the positron is the distinguishing hallmark of the antineutrino interaction with protons, allowing signal identification and background suppression. This temporal and spatial signature can be mimicked by other events depositing energy inside the detector, giving rise to non-negligible backgrounds. For this reason, several selection criteria are designed to distinguish signal from background.
In this study, the JUNO offline software framework [49] is used for Monte Carlo simulations of the spectral components and to perform signal selection, through the analysis of time and space correlations.

4.1 Time correlation analysis

In contrast to collider experiments, where the signal of interest is often provided by a single event, the identification of a neutrino interaction in JUNO relies on the detection of a time-correlated pair of events. For this reason, a *time correlation analysis* is to be performed: the process typically involves selecting two events that occur within a certain time and spatial window. The time window is based on the known lifetime of the neutron, which is on the order of hundreds of microseconds: it has to be long enough to allow for neutron capture by the nucleus, but short enough to avoid contamination from other types of events that may occur later in time. The spatial window is instead determined by the expected radial distribution of the emitted neutrons. Figure 4.1 shows three possible configurations for an IBD event. In the ideal scenario (1), a prompt-delayed pair gives rise to two positive triggers without any other events occurring between them. In the most likely scenario (2), one or more background events occur between the pair of interest. In rare cases, an IBD event may be interposed by another IBD pair (3), but this is highly unlikely given the reactor antineutrino rate of approximately 60 events per day and the time selection window of the order of milliseconds.

A time correlation analysis is therefore critical for isolating IBD events and achieving high precision in the measurement of the reactor antineutrino flux. To fulfill this goal, the JUNO software [49] will offer various tools, some of which are still in the developmental phase, designed to effectively perform offline data analysis. To date, no existing tool is already available to conduct data-like analysis of SNiPER-based simulations. Consequently, a part of this work was devoted to the development of an algorithm to implement a time correlation analysis. The main body of the code is written in C++ and uses methods from the ROOT framework [47] while the configuration script is in Python[13].

[13] More information about the implementation and execution of SNiPER algorithms can be found at https://sniper-framework.github.io/Algorithms/1.CppCode.html

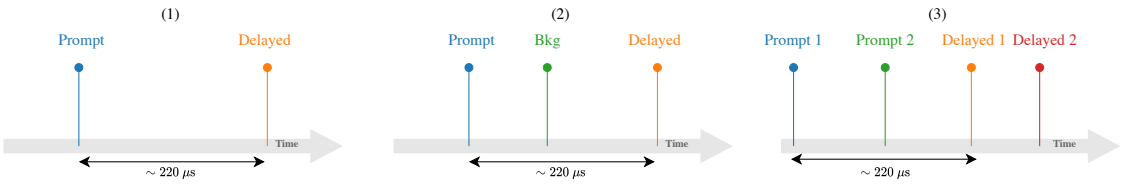


FIGURE 4.1: (1) Pure IBD event in two trigger windows. (2) IBD pair interposed by a background event. (3) IBD event interposed by another IBD pair.

The algorithm rationale is based on a sliding time window acting on a buffer of readout events, which maintains a chronological arrangement. Initially, an object is instantiated to store different properties associated with each event, including timestamp, reconstructed vertex, and energy. As the algorithm progresses through the loop, pairs of events located at the beginning and end of the buffer are analyzed to determine whether they satisfy the prescribed time and vertex IBD selection criteria. If these criteria are met, the pair is recognized as a prompt-delayed candidate and subsequently removed from the buffer. Following this, the sliding time window is updated, and the next event in the sequence is added to the buffer to undergo further processing. If two events in the buffer satisfy the time cut but fail to meet the vertex cut, the first event is not immediately discarded. This allows for the possibility of combining the next event added to the buffer with the first event, creating an IBD candidate that fulfills both the time and vertex cuts. In contrast, if a pair of events in the buffer does not satisfy the time cut, it is guaranteed that no IBD candidate can be formed from those events due to their chronological arrangement. In this case, the first event in the buffer is removed, and the process continues with the next event in the sequence.
The algorithm continues adding events to the buffer until a valid IBD candidate is identified or until the time window reaches its maximum capacity, at which point the oldest event is removed from the buffer. Once all events in the data stream have been processed, the loop is exited, and specific information regarding the IBD candidates, such as event ID, timestamp, reconstructed energy, and vertex, is stored for further offline analysis. The main steps of the process are summarized hereafter:

1. Initialize an empty buffer.
2. Retrieve the next event from the data stream and add it to the buffer.
3. Check the buffer for the following conditions:
   (a) If the buffer is empty, go to step 2.
   (b) If the buffer contains only one event, go to step 2.
   (c) If the time difference between the first and last events in the buffer exceeds the specified time cut threshold, remove the first event from the buffer and return to step 3.
   (d) If only the distance between the vertices of the first and last events in the buffer exceeds the specified vertex cut threshold (while satisfying the time cut), return to step 2.
4. If both the time and vertex cuts are satisfied, consider the events to be a candidate IBD prompt-delayed pair and remove them from the buffer.
5. If the time cut is satisfied but the vertex cut is not, continue adding events to the buffer until either a valid IBD candidate is found or the buffer reaches the maximum size allowed by the time window. If the buffer reaches the maximum size allowed by the time window, remove the oldest event from the buffer and go back to step 3.

6. If there are no more events in the data stream, end the process.

Due to limited computational resources, it was not feasible to analyze a data-like mixed dataset, comprising the IBD signal and the major background sources in realistic proportions. Consequently, the study focused solely on the ideal configuration (1) depicted in Figure 4.1 (for reactor antineutrino signal, geo-neutrinos and cosmogenic ^{9}Li/^{8}He). However, some tests were conducted using a smaller-scale mixed sample containing reactor IBD events and radiogenic events. These tests confirmed two important findings: (1) the algorithm effectively identified and tagged IBD prompt-delayed pairs, and (2) no spurious correlations were observed between a prompt/delayed pair and an event induced by radioactivity. It should be noted that while these tests were performed on a limited scale, they provided valuable insights into the algorithm's performance and its capability to distinguish genuine IBD events from background sources.
In addition to its effectiveness in the main time-correlation analysis, the algorithm also demonstrated its versatility for off-window analyses. One specific application involved the study of the radiogenic background, as discussed in Section 4.3.1. Most radioactive events contribute to the uncorrelated background. However, there are cases where certain decays are correlated, specifically the fast coincidence of ^{214}Bi-^{214}Po and ^{212}Bi-^{212}Po present in the decay chains of Uranium and Thorium, respectively. The algorithm proved to be valuable in identifying and mitigating these correlated decays within the radiogenic background. By leveraging the double signature characteristic of the Bismuth-Polonium decays, the algorithm successfully differentiated them from the uncorrelated background events. By implementing this approach, a significant reduction in the overall radiogenic background was achieved, specifically narrowing it down to only accidental coincidences.

4.2 IBD selection

The study of reactor antineutrino oscillations is primarily hindered by the dominant background signal, which far exceeds the actual IBD signal. With an expected event rate of $O(100\,\mathrm{Hz})$, only a small fraction corresponds to IBD-induced processes. However, the distinctive signature yielded by the IBD reaction can be leveraged to differentiate it from the background.
The accidental background, cosmogenic ^{8}He/^{9}Li, geo-neutrinos, fast neutrons, and (α, n) are the major backgrounds for the oscillation analysis. IBD selection criteria are designed to suppress the background contribution while keeping a high efficiency for true reactor antineutrino IBD events. For example, the fiducial volume cut can significantly reduce the accidental and the (α, n) backgrounds. To further mitigate the impact of accidental coincidences, the selection is based upon temporal coincidence, vertex correlation of the prompt and delayed signals, and energy selection. To reject the cosmogenic backgrounds such as ^{8}He/^{9}Li and fast neutrons, muon veto cuts are employed. A detailed discussion of the calculation of the residual rates will be presented in Section 4.3.
The IBD standard selection cuts are the following [21, 55]:

- **Fiducial volume (FV) cut:** prompt or delayed candidates are discarded if their vertices are more than 17.2 m away from the detector center, since the external background rate is exponentially increasing at the edges of the acrylic sphere. However, it is worth noting that this fiducial volume cut is subject to modification and optimization during the data-taking phase.

- **Energy cut:** prompt and delayed candidate events are restricted to the energy windows $E_p \in (0.7, 12.0)$ MeV and $E_d \in (1.9, 2.5) \cup (4.4, 5.5)$ MeV, respectively. The prompt energy range is the dominant one for IBD prompt events, as it can be observed in Figure 3.6. On the other hand, the delayed signal selection windows are defined to be centered around 2.2 MeV and 4.9 MeV, which correspond to neutron capture on hydrogen and carbon, respectively.

- **Time cut:** the surviving pairs are required to fall in a time coincidence window of 1 ms, corresponding to approximately 5× the neutron capture time; hence, the time interval between delayed and prompt signals $\Delta t < 1$ ms.
- **Vertex cut:** the distance between prompt and delayed events vertices has to be smaller than 1.5 m, hence $\Delta r < 1.5$ m.
- **Muon veto criteria:** cosmic muon veto cuts are crucial to suppress the cosmogenic backgrounds, most of which meet the aforementioned IBD coincidence selection criteria and can easily mimic the signal signature. Indeed, time and space correlations are not enough to fully eliminate longer-lived isotopes, such as ^{8}He/^{9}Li produced along muon tracks by cosmic ray spallation. The current state-of-the-art muon veto strategy [33, 55] involves the following steps:
  - for muons tagged by either the external water pool Cherenkov detector or by the Central Detector, a veto of 1 ms after each muon is applied over the whole FV. This strategy is aimed at suppressing spallation neutrons and short-lived radio-isotopes;
  - for well-reconstructed tracks inside the CD, produced by single or two far-apart muons, a veto of 0.6 s, 0.4 s, and 0.1 s is applied when candidate events have reconstructed vertices smaller than 1 m, 2 m, and 4 m, respectively;
  - for events containing two close and parallel muons (within 3 m, constituting $\sim 0.6\%$ of muon-related events), a single track is often reconstructed. The veto is then applied around this track as described by the previous point, but the cylinder radii are increased according to their separation;
  - for events whose track cannot be properly reconstructed (about 2% of all muon-related events, and mostly dominated by more than two muons going through the detector simultaneously), a 0.5 s veto is applied over the whole fiducial volume.
  - a 1.2 s veto is applied on any candidate event detected inside a 3 m-radius sphere around spallation neutron capture events.

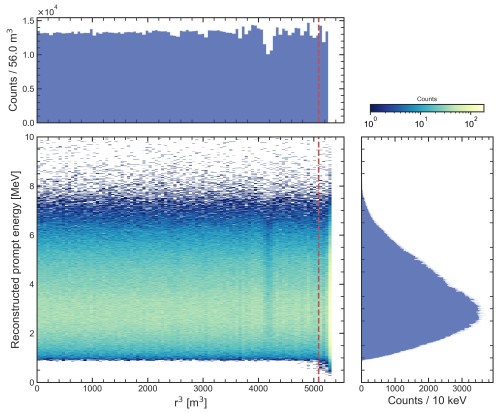
4.2.1 IBD signal selection efficiency

To estimate the selection efficiency of the IBD signal, 1.5 million unoscillated reactor antineutrino events are generated with the JUNO offline software [49], using the full chain detector and electronics simulation and reconstruction. For each cut, the *selection efficiency* ε is obtained as the ratio of the number of selected events, i.e., meeting that specific criterion, to the total number of reconstructed events before applying that cut, namely

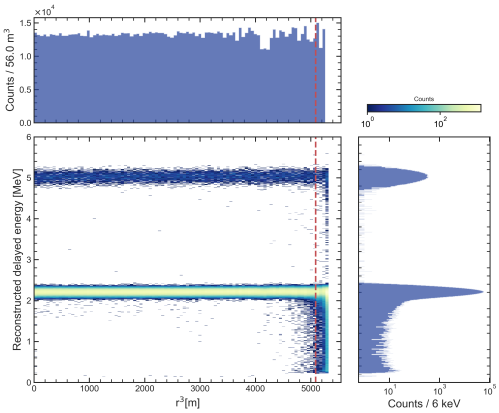
$$\varepsilon = \frac{N_{selected}}{N_{tot}}. \tag{4.1}$$

The subsequent paragraphs provide a comprehensive breakdown of the calculations performed to determine the individual efficiencies for the reactor antineutrino sample.

Fiducial volume cut: after the application of the fiducial volume (FV) cut, 91.3% of the initial events are retained. A crucial aspect of IBD prompt-delayed pairs is that they exhibit a uniform distribution within the entire detector volume. This feature can be clearly seen in both Figure 4.2a and Figure 4.2b, which illustrate the prompt and delayed event energies as a function of the radial distance between their vertices (in terms of r^3) and the detector center.



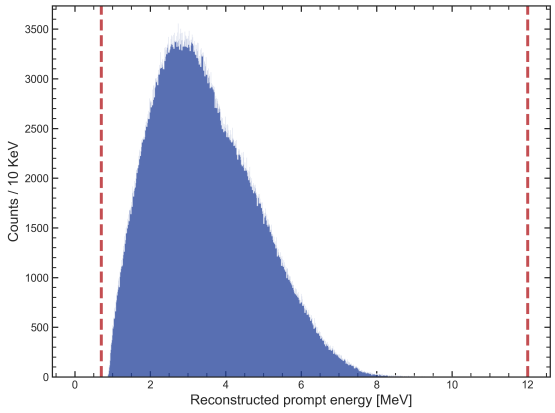
(A) Bi-dimensional distribution of prompt events energy and associated vertex.



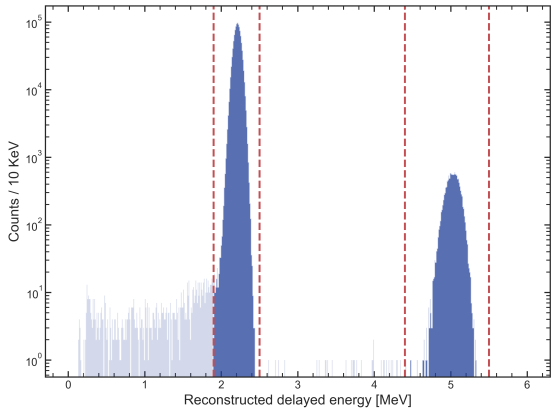
(B) Bi-dimensional distribution of delayed events energy and associated vertex.

FIGURE 4.2: Bi-dimensional distribution of prompt (left) and delayed (right) events energy and associated distance from the center of the detector (r^3). The radius corresponding to the fiducial volume cut is indicated by the dashed line.

Energy cut: all prompt events (99.999%) lie in the energy selection region while the efficiency is 99.9% for delayed events. The energy cut is applied only over the fiducial volume. The resulting energy spectra after selection cuts are reported in Figure 4.3a and Figure 4.3b.



(A) Reactor $\overline{\nu}_e$ IBD prompt events energy distribution.

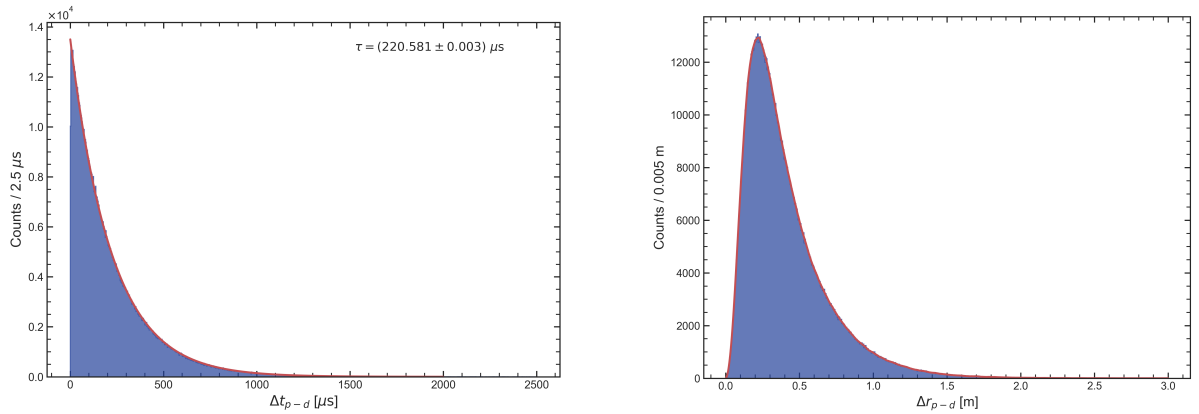


(B) Reactor $\overline{\nu}_e$ IBD delayed events energy distribution.

FIGURE 4.3: Prompt (left) and delayed (right) events energy distributions, with selection areas indicated by dashed lines.

Vertex and time cut: it is important to note that vertex and time cuts can be applied over the entire fiducial volume, independently of the energy selection. When it comes to temporal selection, it is reasonable to assume that the time interval between two candidate events does not depend on the energy deposited by the two particles. As for the relative position cut, it can be argued that positrons are uniformly distributed in energy inside the detector, as one can observe in Figure 4.2a. Similarly, the same assumption can be made for the photons emitted as a result of neutron capture, but only after applying the FV cut. Any slight deviation from uniformity is only due to photons being produced at the edge of the detector where energy is not fully deposited, resulting in the

shoulder in the delayed spectrum shown in Figure 4.2b. To apply this criterion, a semi-analytical approach is employed. Firstly, the differences Δt and Δr between prompt and delayed events are calculated and presented in Figure 4.4a and Figure 4.4b, respectively. Based on the data, we can derive the Probability Distribution Functions (PDFs) $F_t(\Delta t)$ and $F_r(\Delta r)$. $F_t(\Delta t)$ is modeled by an exponential distribution, and by fitting the data, we determine a decay time of approximately 220 µs, which is aligned with the expected neutron decay time in the LS. $F_r(\Delta r)$ depends on the spatial distribution of the emitted neutrons; the latter can be modeled as a random walk process, where the neutrons undergo multiple elastic and inelastic collisions with the target medium, losing energy and changing direction each time.



(A) Δt distribution between prompt and delayed events.

(B) Δr distribution between prompt and delayed events.

FIGURE 4.4: Δt (left) and Δr (right) distribution between prompt and delayed events. The corresponding PDFs are indicated by the red lines.

Two different methods are used to evaluate the efficiency of the time and vertex cuts. In the first method, the area under the two PDFs inside the selection region is calculated and divided by the total area. This yields an efficiency of 99.0% and 99.3% for the time and vertex selections, respectively. In the second method, the joint probability distribution is constructed by multiplying $F_t(\Delta t)$ and $F_r(\Delta r)$, which is shown in Figure 4.5 along with the marginal densities. This approach enables the calculation of the probability that the two variables simultaneously belong to a given 2d interval of values, by integrating the joint density over the specified ranges. Using this approach, a combined efficiency of 98.3% is obtained, which is equivalent to the product of the individual efficiencies. This confirms the independence of the Δt and Δr cuts.

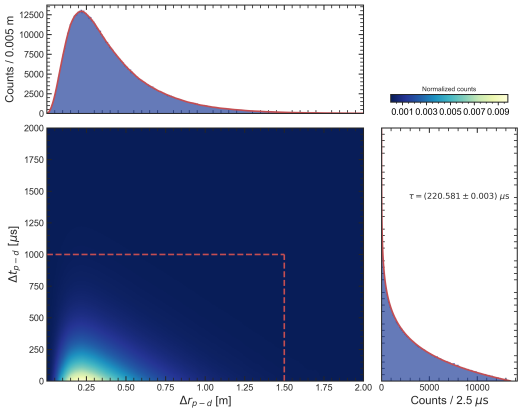


FIGURE 4.5: Δt and Δr distribution between prompt and delayed events.

Muon veto cut: the muon veto strategy is taken from [55] and yields a selection efficiency of 91.6% for IBD events.

The combined antineutrino detection efficiency after all selection cuts is found to be 82.1%, with a resulting IBD rate of 47.1 events/day with nominal reactor power (hence $\sim$ 43.2 events per day with 11/12 duty cycle). The breakdown of the selection efficiencies is summarised in Table 4.1.

Selection criterion	Efficiency [%]	IBD rate [events/day]
All IBDs	100	57.4 [55]
Fiducial Volume	91.3	52.4
Energy Range	99.9	52.35
Time cut Δt	99.0	51.8
Vertex cut Δr	99.3	51.5
Muon veto	91.6	47.1
Combined selection	82.1	47.1

TABLE 4.1: Summary of cumulative IBD selection efficiencies. The reported IBD rates refer to the expected events per day after the selection criteria are progressively applied. These rates are calculated for nominal reactor power rate.

4.3 Background estimation

This section investigates the contribution of the main background sources, which can be categorized as *correlated* or *uncorrelated*. The rates of these backgrounds are determined using either the official simulation framework [49], which accurately models the detector response and background sources, or by using updated information from the JUNO common analysis framework [74].

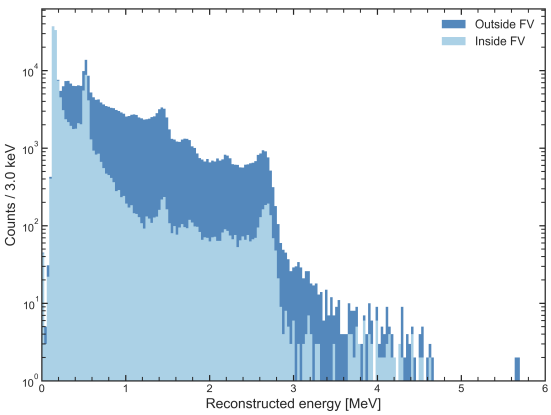
4.3.1 Accidental background

Correlated backgrounds are by definition produced by a single physics process and yield both a prompt and a delayed signal. The coincidence of two otherwise uncorrelated events, typically of radiogenic origin, gives rise to the so-called *accidental* background. Natural radioactivity is a significant source of accidental background that cannot be entirely discriminated through software-based cuts because it satisfies the IBD selection criteria in the same energy range. Therefore, it is of paramount importance (1) to maintain the natural radioactivity background at extremely low levels [32], by implementing strict reduction strategies during detector design and construction; (2) to precisely characterize the shape and rate of the residual components. The accidental background consists of mainly three types of random coincidences: (radioactivity, radioactivity), (radioactivity, cosmogenic isotope), and (radioactivity, spallation neutrons). Studies [32] have shown that radioactivity-radioactivity coincidences are the primary contribution, while the other two potential random coincidences are negligible. This background dominates the low energy part of the spectrum, providing a deposited energy up to 5 MeV, thus overlapping with the IBD selection region. The radioactivity of the materials employed in the construction of the JUNO detector is one of the main sources of accidental background. The main contaminants, depositing energy after their radioactive decays are [32]:

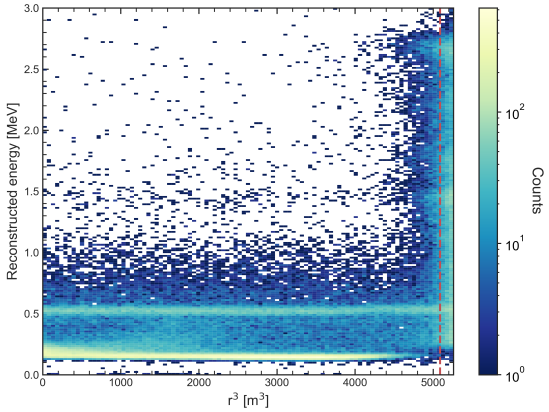
- natural long-lived radionuclides, ^{238}U, ^{232}Th, and ^{40}K;
- natural medium-lived radionuclides ^{226}Ra, ^{210}Pb-^{210}Bi, and ^{210}Po when secular equilibrium is broken in the ^{238}U chain;
- the gaseous radionuclide ^{222}Rn and the anthropogenic radionuclide ^{60}Co.

In addition, the background can be classified as either *internal* or *external*, depending on whether the source of the background is located within the LS or in other detector components, respectively. The internal background is caused by radionuclides, such as those from the Uranium and Thorium chains, that directly release energy inside the target volume regardless of the type of emitted particles

(e.g., α, β, γ). These events are uniformly distributed within the detector. On the other hand, only a portion of the α and β particles will contribute to the external background due to their relatively short ranges in solid media. In contrast, high-energy photons can traverse the solid materials and reach the LS, depositing energy and making a more substantial contribution to the external background.
A full detector simulation is conducted to assess the background count rate, or the so-called *singles* rate, resulting from natural radioactive contamination. Figure 4.6a presents the singles energy spectrum, distinguishing between events inside and outside the fiducial volume. Meanwhile, Figure 4.6b shows the singles event rate as a function of r^3 and reconstructed energy. Both plots reveal two major peaks below 1 MeV corresponding to events that are uniformly distributed within the LS. The first peak is caused by ^{14}C with a β-decay end-point energy of approximately 156 keV, while the second peak is produced by quenched α particles stemming from decays of single particles, resulting in a well-defined structure at around 530 keV. Additionally, higher energy events dominate the outer region of the detector, as expected.



(A) Reconstructed energy spectrum of radiogenic events. Contributions from inside and outside the FV are separately indicated.



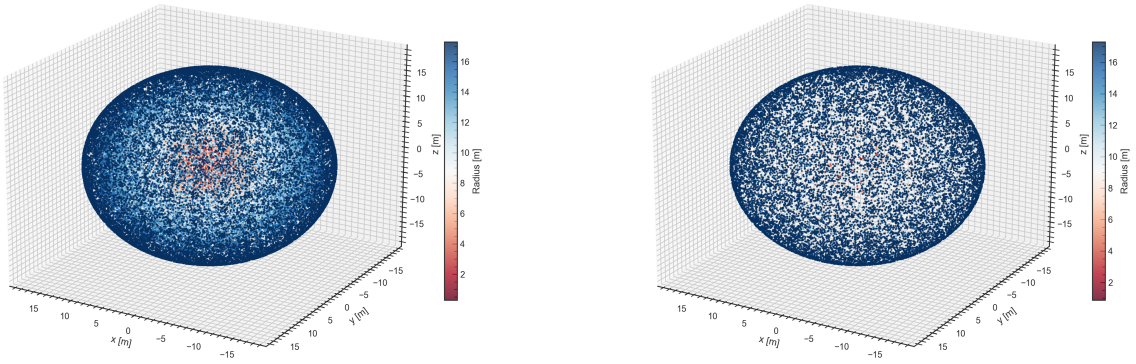
(B) Singles event rate as a function of volume (r^3) and reconstructed energy. The standard FV cut is indicated by the dashed line.

FIGURE 4.6: Radiogenic *singles* energy spectrum (left) and distribution of events as a function of volume and energy (right).

Furthermore, it is important to note that a strong correlation exists between energy and radial distributions, which renders the IBD-like approach inapplicable to this scenario. Indeed, selection efficiency terms cannot be computed independently for the radiogenic background, i.e., vertex and energy selection efficiencies cannot be calculated individually and then combined.
Hereinafter, a preliminary energy cut at 0.7 MeV is applied, by discarding all events with an energy lower than this threshold.

4.3.1.1 Accidental coincidences rate

The accidental coincidences rate is estimated using the full detector simulation and reconstruction. Specifically, a 1-hour radioactivity simulation dataset is used. Since the accidental coincidences rate is expected to be of the order of 1 count per day (cpd) [21, 32, 55], a long exposure dataset would be needed to directly retrieve it from data. For this reason, a toy Monte Carlo is exploited to estimate it, using the default IBD selection cuts presented in Section 4.2. The accidental coincidences rate can be evaluated as



(A) 3-dimensional distribution of singles without any energy threshold inside the detector.

(B) 3-dimensional distribution of singles above the 0.7 MeV threshold inside the detector.

FIGURE 4.7: 3-dimensional distribution of radioactive events inside the full detector volume, without (left) and with (right) energy cut at 0.7 MeV. The color scale indicates the distance from the detector center.

$$R_{\text{acc}} = R_d \cdot R_p \cdot \Delta t \cdot \varepsilon_{\text{acc}}, \tag{4.2}$$

where R_p and R_d represent the prompt-like and delayed-like event rates, respectively, Δt is the applied time cut, and ε_{acc} is an efficiency scale factor, evaluated from the MC sample. The following procedure is carried out:

1. As mentioned in Section 4.1, the correlated contribution present in the radioactivity sample is preemptively eliminated, thereby isolating the uncorrelated component responsible for the accidental background.

2. An energy cut $E > 0.7\,\text{MeV}$ is implemented, resulting in a sample of N events over threshold. The distribution of events within the detector is displayed in Figure 4.7a and Figure 4.7b, before and after applying this energy threshold, respectively.

3. All possible prompt-delayed pairs are then constructed, giving rise to a number of combinations $N_{\text{comb}} = N(N-1)$. IBD selection cuts are applied simultaneously to each pair of events (i, j), where the indices (i, j) are referred to potential prompt and delayed candidates, respectively, with $i \neq j$. Subsequently:

   - The spatial distance between the two events vertices, $\Delta\vec{r} = \vec{r_j} - \vec{r_i}$, is evaluated.
   - Fiducial volume, energy, and Δr selection criteria are then implemented, thereby obtaining a certain number of events N_{sel} meeting all the requirements.
   - The efficiency factor is computed using Eq. (4.1), providing

$$\varepsilon_{\text{acc}} = \frac{N_{\text{sel}}}{N_{\text{comb}}} = (1.41 \pm 0.01) \times 10^{-6} \quad \text{with} \quad \sigma_{\varepsilon_{\text{acc}}} = \sqrt{\frac{\varepsilon_{\text{acc}}(1-\varepsilon_{\text{acc}})}{N_{\text{comb}}}}$$

where the uncertainty is obtained assuming a binomial distribution for the number of events.

4. The singles event rate above 0.7 MeV is then inferred from the MC sample by constructing the distribution of time intervals Δt between two consecutive events; the latter is shown in Figure 4.8 and fitted with an exponential function. The selection efficiency related to the FV and the energy ranges is already accounted for in the MC scaling factor in the previous point. Since events are assumed to be uncorrelated by definition, the rate of prompt-like and delayed-like candidates is the same, namely from the fit, $R_d = R_p = (39.4 \pm 0.1)$ Hz.

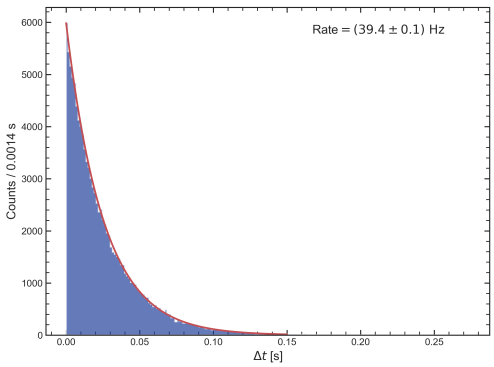


FIGURE 4.8: Distribution of time intervals Δt between two consecutive radiogenic events above the 0.7 MeV threshold.

Finally, using Eq. (4.2) and considering the additional muon veto efficiency (91.6% from Table 4.1) the accidental coincidence rate is found to be:

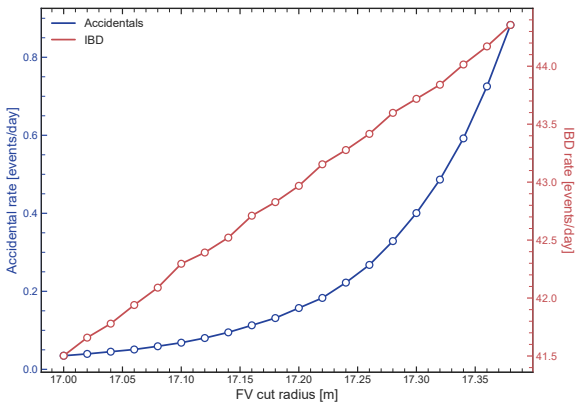
$$R_{\text{acc}} = (0.172 \pm 0.002) \text{ events/day},$$

with a relative uncertainty of the order of 1%. This rate is highly sensitive to the choice of the FV radius with more than a factor 2 variation when shifting the FV cut by 0.1 m around 17.2 m; indeed an exponential increase is observed. This sensitivity is evident in Figure 4.9a, which displays the daily rate of accidental coincidences after applying all selection criteria for different FV radii ranging from 17 to 17.4 meters. Notably, a clear exponential increase can be observed. Similarly, the study is extended to the case of IBD events, and a linear increase is observed as the FV radius is varied. It is important to mention that IBD events are uniformly distributed within the detector. Consequently, one would expect the rate to increase with the volume following a cubic dependence (i.e., r^3) rather than a linear one (i.e., r). However, due to the relatively small range of FV radii considered in Figure 4.9a, the cubic dependence is well approximated by a linear function within this limited interval. To further investigate the relationship between the IBD daily rate and the FV radius, Figure 4.9b presents the IBD daily rate calculated across the full range of possible FV radii. As expected, a clear cubic trend emerges, confirming the uniform distribution of events throughout the entire CD.
It is assumed that the accidental background, benefiting from a significant statistics, carries negligible uncertainty on the spectral shape PDF[14]. Additionally, the rate uncertainty can be controlled within 1%.

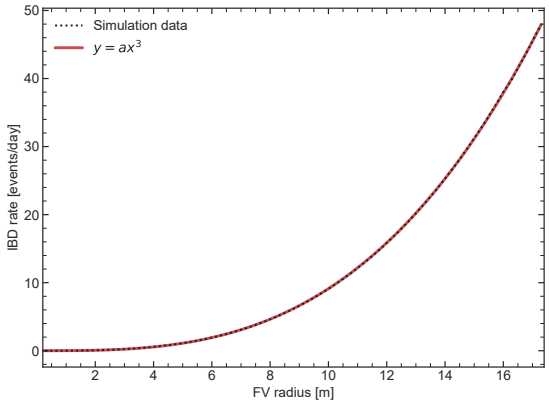
4.3.2 Cosmogenic isotopes

For underground neutrino observatories, a sufficient amount of overburden above the detector is the most effective approach to suppress the cosmogenic backgrounds. The JUNO experimental site is located under a 240-meters high hill, and the detector will be at 450 m depth, resulting in a rock overburden of about 690 m [21, 30]. The estimated muon flux in the JUNO detector is about 0.004 Hz/m^2 with a mean energy of 207 GeV [33]. In the liquid scintillator, energetic cosmic muons

[14] The definition of the uncertainties on the spectral shape (i.e., bin-to-bin uncorrelated uncertainties, here denoted as "shape" uncertainties) is given in Section 5.4, when dealing with the systematics.



(A) Accidental coincidences and reactor IBD daily rates as a function of variable FV radii.



(B) Reactor IBD daily rate as a function of variable FV volume radii.

FIGURE 4.9: (a) Accidental coincidences and reactor IBD daily rates after all selection criteria, as a function of different values of FV radii. (b) IBD daily rate as a function of FV volume radii; both the simulated data (dashed line) and the cubic interpolation curve (solid line) are shown.

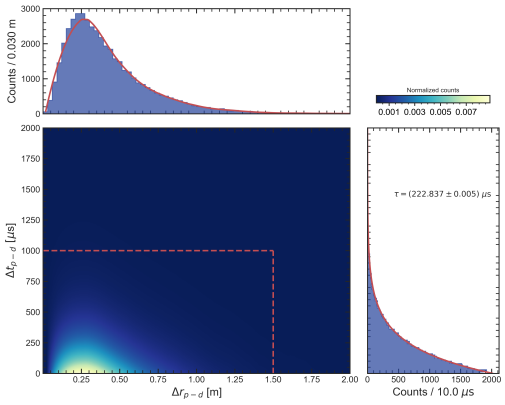
and the subsequent showers can interact with ^{12}C and produce radioactive isotopes with $Z \leq 6$ by electromagnetic or hadronic processes. Among them, ^{9}Li and ^{8}He, with half-lives of 0.178 s and 0.119 s, respectively, are the major correlated background source, because they can decay by emitting both a β^- and a neutron, thereby mimicking a reactor IBD signal. ^{9}Li and ^{8}He isotopes rates are extrapolated using KamLAND and Borexino measured yields [75, 76]: specifically, the expected yield is 127.1 events/day and 40.4 events/day, respectively. Other isotopes are also β-neutron emitters but are negligible with respect to ^{9}Li/^{8}He, whereas some long-lived cosmogenic isotopes undergo beta decay without an accompanying neutron. These cannot give rise to correlated backgrounds by themselves but can contribute to the accidental neutron-like signal (if the energy is in the proper range). The β^--decay of ^{8}He populates a neutron-unstable excited state in ^{8}Li, with a 16% branching ratio. For ^{9}Li, the branching ratio to a neutron unstable state in ^{9}Be is 51%. The subsequently emitted neutron is captured mainly on hydrogen with a mean capture time of 220 µs. The triple coincidence of a muon, a β-emission, and a delayed neutron capture provides a precise signature, that allows to efficiently decrease this background. To that end, the recently updated muon veto strategy [33] reduces the residual ^{9}Li/^{8}He background to 0.81 events/day and 0.09 events/day, respectively. Consequently, the total background level amounts to 0.9 events/day in the full LS volume. Subsequently, the IBD standard selection cuts are applied to this remaining component, with the same procedure employed for reactor antineutrinos.
We assume the rate relative uncertainty to be 20% and assign 10% bin-to-bin uncorrelated spectrum shape uncertainty for 36 keV bins, which is the estimation reported in Ref. [30].

Cosmogenic ^{8}He To estimate the number of ^{8}He-induced correlated pairs passing the IBD selection criteria, 50k events were generated using the offline software and the full detector simulation. Then, as previously described in Section 4.2, selection efficiencies are calculated and combined. The results of these calculations are summarized hereafter:

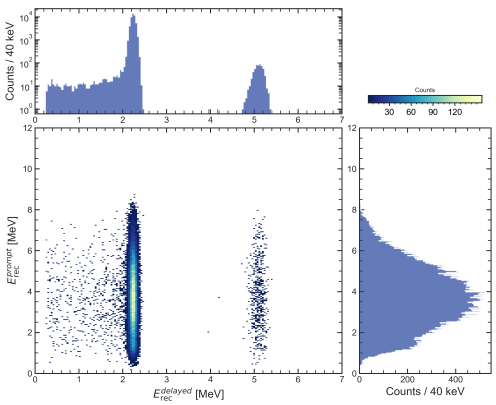
- **Fiducial volume cut:** 91.4% of the events are retained by applying the FV cut.
- **Energy cut:** 99.8% of prompt events lie in the energy selection region while the efficiency is 99.9% for delayed events. The energy cut is applied only over the fiducial volume. The resulting energy spectra are reported in Figure 4.10b.

- **Vertex and time cut:** From the joint probability distribution, shown in Figure 4.10a, together with the marginal densities, a combined time-vertex efficiency of 98.3% is found.

The resulting ^{8}He residual contribution after IBD selection is therefore 0.08 events/day.



(A) ^{8}He: Δt and Δr distribution between prompt and delayed events. The selection cut range is indicated by the red dashed line.



(B) ^{8}He: prompt and delayed energy distributions, both the bi-dimensional and uni-dimensional distributions are reported.

FIGURE 4.10: Δt, Δr (a) and energy distributions (b) for ^{8}He-induced prompt-delayed correlated pairs.

Cosmogenic ^{9}Li The same procedure is carried out for the ^{9}Li sample, obtaining:

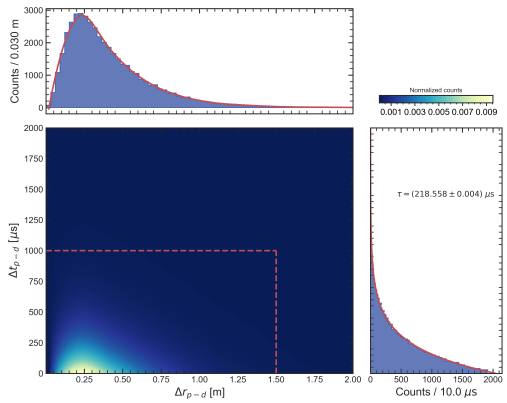
- **Fiducial volume cut:** 91.2% of the events is retained by applying the FV cut.
- **Energy cut:** The energy selection region contains 99.6% of prompt events, while the efficiency for delayed events is 99.9%. The energy cut is applied only over the fiducial volume. The resulting energy spectra can be seen in Figure 4.11b.
- **Vertex and time cut:**. The combined time-vertex efficiency of 98.4% is obtained from the joint probability distribution, as illustrated in Figure 4.11a, along with the 1-d probability densities.

The resulting ^{9}Li residual contribution after IBD selection is 0.72 events/day.

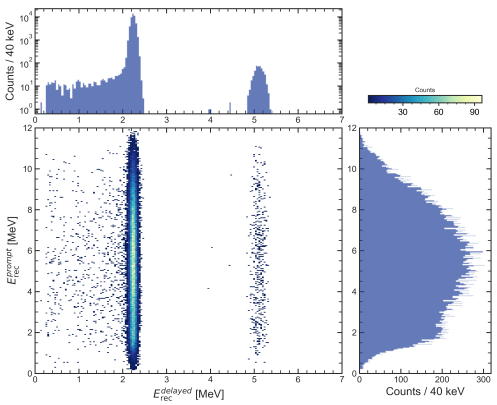
4.3.3 Geo-neutrinos

Geo-neutrinos (geo-ν) are electron antineutrinos produced by the radioactive decay chains of Uranium and Thorium inside the Earth's mantle and crust. They are produced via beta decay and have similar energy spectra and detection signatures as reactor antineutrinos. The total rate of geo-neutrinos at the JUNO site is estimated to be 1.5 events per day [21], with the ^{238}U decay chain contributing approximately 77% and the ^{232}Th decay chain contributing approximately 23% [21]. The residual geo-ν rate after applying all selection cuts is calculated in the following paragraphs. The rate and shape relative uncertainties are set to 30% and 5%, respectively [30].

Geo-neutrinos from Uranium chain A 500k events MC sample is used to calculate selection efficiencies, following the same procedure as in earlier examples.



(A) ^{9}Li: Δt and Δr distribution between prompt and delayed events. The selection cut range is indicated by the red dashed line.

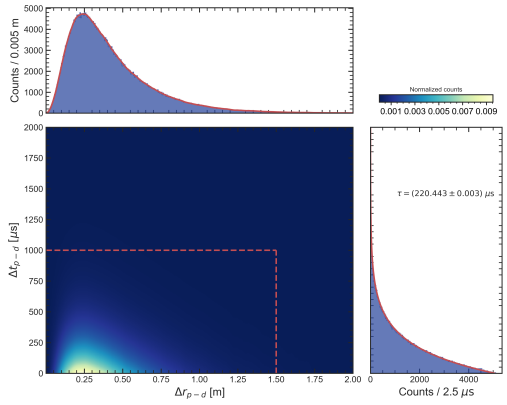


(B) ^{9}Li: prompt and delayed energy distributions, both the bi-dimensional and uni-dimensional distributions are reported.

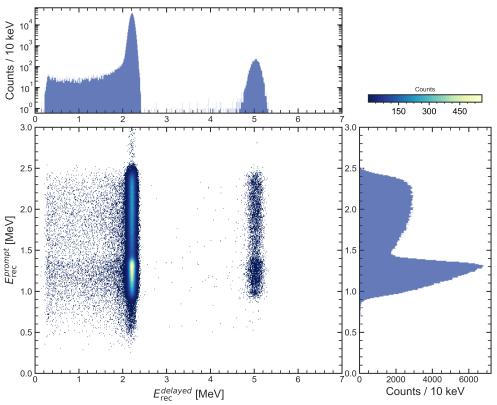
FIGURE 4.11: Δt, Δr (a) and energy distributions (b) for ^{9}Li-induced prompt-delayed correlated pairs.

- **Fiducial volume cut:** 91.0% of the events is retained by applying the FV cut.
- **Energy cut:** All prompt events lie in the energy selection region while the efficiency is 99.9% for delayed events. The energy cut is applied exclusively within the fiducial volume, and the resulting energy spectra are depicted in Figure 4.12b.
- **Vertex and time cut:** The joint probability distribution in Figure 4.12a, along with the 1-d probability densities, was used to determine a combined time-vertex efficiency of 98.2%.

The combined antineutrino detection efficiency for geo-ν from the ^{238}U chain after all selection cuts is 81.7%, with a resulting rate of 0.94 events/day.



(A) Geo-ν from ^{238}U chain: Δt and Δr distribution between prompt and delayed events. The selection cut range is indicated by the red dashed line.



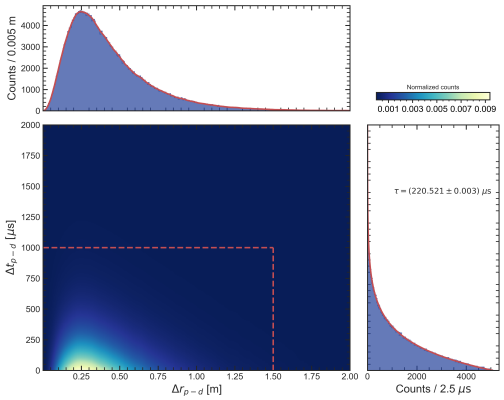
(B) Geo-ν from ^{238}U chain: prompt and delayed energy distributions, both the bi-dimensional and uni-dimensional distributions are reported.

FIGURE 4.12: Δt, Δr (a) and energy distributions (b) for ^{238}U geo-neutrinos prompt-delayed correlated pairs.

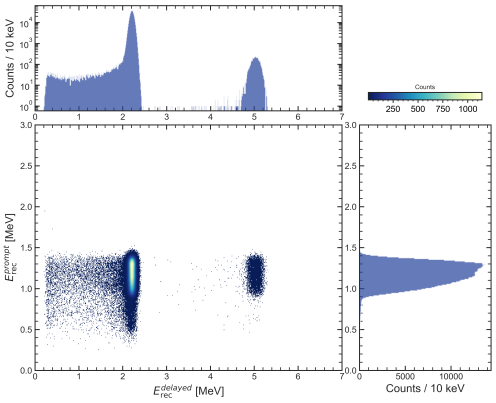
Geo-neutrinos from Thorium chain Similarly, a 500k events MC sample is used to calculate selection efficiencies.

- **Fiducial volume cut:** After applying the FV cut to the reconstructed events, 91.2% of the events are retained.
- **Energy cut:** The energy selection criteria retains 99.99% of prompt events and 99.89% of delayed events, applied exclusively within the fiducial volume. The resulting energy spectra are presented in Figure 4.13b.
- **Vertex and time cut:** The joint probability distribution displayed in Figure 4.13a leads to a combined time-vertex efficiency of 98.3%.

The total detection efficiency for geo-neutrinos from the ^{232}Th chain after all selection cuts is found to be approximately 82%. Consequently, the expected rate of events is 0.28 events/day.



(A) Geo-ν from ^{232}Th chain: Δt and Δr distribution between prompt and delayed events. The selection cut range is indicated by the red dashed line.



(B) Geo-ν from ^{232}Th chain: prompt and delayed energy distributions, both the bi-dimensional and uni-dimensional distributions are reported.

FIGURE 4.13: Δt, Δr (a) and energy distributions (b) for ^{232}Th geo-neutrinos prompt-delayed correlated pairs.

4.3.4 Other background sources

Other sources of correlated background contribute significantly to the expected number of events detected in JUNO. To estimate their rates and spectral shape, we rely on previous studies and experiments.

Fast neutrons Cosmic muons produced in the rock surrounding the water pool or those that corner-clip, i.e., pass through the CD but leave only a short track, are not effectively tagged. This results in fast neutrons produced by these muons scattering off a proton and being subsequently captured in the LS, creating an irreducible background that mimics the signal signature. The rate of fast neutrons is estimated to be approximately 0.1 events per day (with 100% rate uncertainty), based on a full simulation that does not consider optical processes [55]. The prompt energy spectrum of the fast neutrons is found to be consistent with a flat distribution, with 20% shape uncertainty for 36 keV bins.

^{13}C$(\alpha, n)^{16}$O The ^{13}C$(\alpha, n)^{16}$O background [77] is primarily caused by alpha particles that are produced by the decay of naturally occurring radioactive isotopes along the chains of ^{238}U, ^{235}U, and ^{232}Th [78]. These alpha particles can interact with the surrounding materials, leading to the production of neutrons via the ^{13}C$(\alpha, n)^{16}$O reaction. The total rate of this background is estimated to be 0.05/day with a 50% uncertainty in both the rate and shape of the spectrum. In JUNO, this background is effectively mitigated by using materials that contain low levels of ^{13}C.

Atmospheric neutrinos Atmospheric neutrinos can interact with nuclei in the JUNO detector through charged current (CC) or neutral current (NC) interactions. According to simulations and estimates in [79], the number of events from CC interactions is negligible after the IBD selection, with a rate of $\sim 3.2 \times 10^{-6}$ per day per kiloton. However, the final states of NC interactions can induce correlated events, and the total event rate after applying the IBD selection criteria mentioned above is about 0.1-0.2 per day. The dominant contribution to this background is from atmospheric neutrinos interacting via NC interactions with ^{12}C nuclei. In these interactions, the neutrino only transfers a fraction of its energy to the final products and can eject a neutron from the carbon nucleus, leaving it in an excited state with multiple decay modes [80]. The rate is fixed at 0.16 events/day, with 50% uncertainty in both rate and shape.

Global reactor neutrinos The JUNO analysis considers electron antineutrinos produced by reactors beyond 300 km-distance as background. The energy spectrum and event rate were estimated by taking into account the global reactor distribution and load factor in 2019 [81]. The estimated rate amounts to 1 event/day, with 2% rate and 5% shape uncertainties.

4.4 Expected energy spectrum

The residual background rates, following the application of all physics selection criteria, are presented in Table 4.2, along with their corresponding relative input uncertainties. These combined rates amount to a total of 3.5 events per day, which can be compared to the expected 47.1 reactor antineutrinos per day (at full reactor power).

Background	Rate [events/day]	Rate unc. $\delta_{B,R}$ [%]	Shape unc. $\delta_{B,S}$ [%]
Geo-neutrinos	1.22	30%	5%
Accidentals	0.172	1%	negligible
Fast-n	0.1	100%	20%
^{13}C$(\alpha, n)^{16}$O	0.05	50%	50%
^{9}Li/^{8}He	0.8	20%	10%
Atmospheric neutrinos	0.16	50%	50%
World reactors	1.0	2%	5%
Total	3.5	-	-

TABLE 4.2: Summary of background rates after all selection cuts. Rate and shape (bin-to-bin uncorrelated) uncertainties are reported. All shape uncertainties are given for 36 keV bins, with the exception of global reactors (20 keV bins).

Compared to other underground liquid scintillator experiments, the impact of backgrounds on the precision of oscillation parameter measurements is relatively limited, as discussed in Ref. [55]. This favorable outcome is achieved by exploiting the significant spectral shape distortion observed in the prompt spectrum as the primary means for extracting oscillation parameters. This aspect will be

further analyzed in Chapter 5.
Finally, Figure 4.14 depicts the total reconstructed energy spectra at JUNO, both with and without the background contribution.

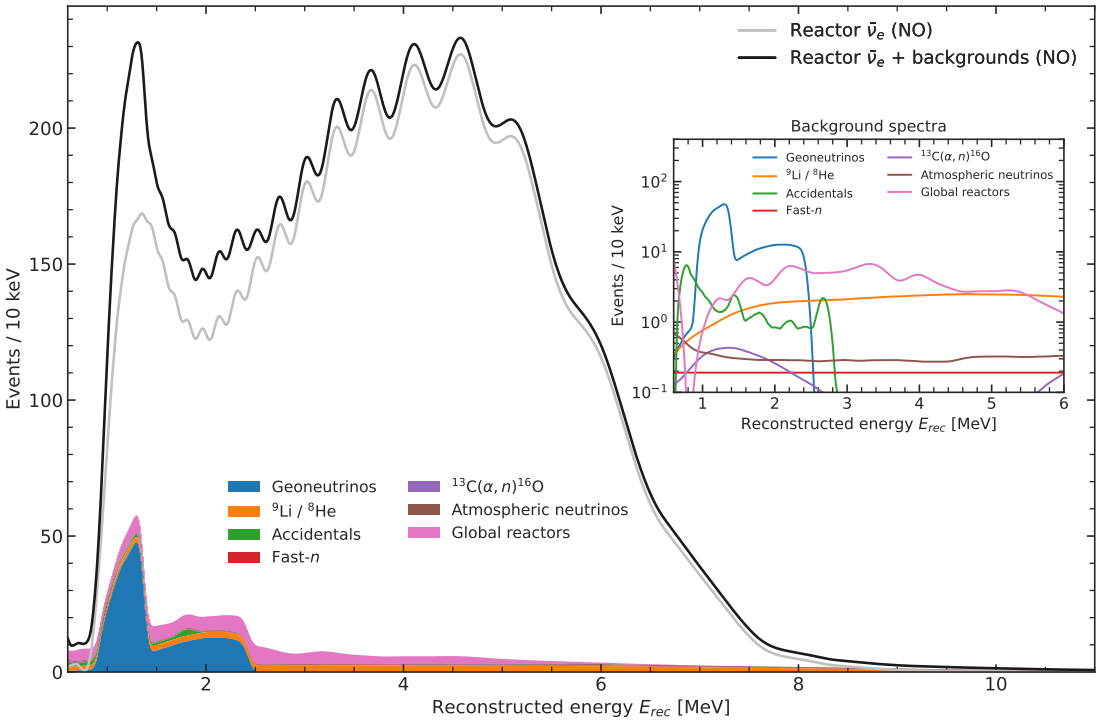


FIGURE 4.14: Reconstructed energy spectra at JUNO for normal ordering, 6 years exposure at 11/12 duty cycle reactor power. The background spectra are stacked in the filled histogram and shown separately in the inset. The background rates were calculated after applying the selection criteria, as described in Table 4.2.

Chapter 5

JUNO's sensitivity to oscillation parameters

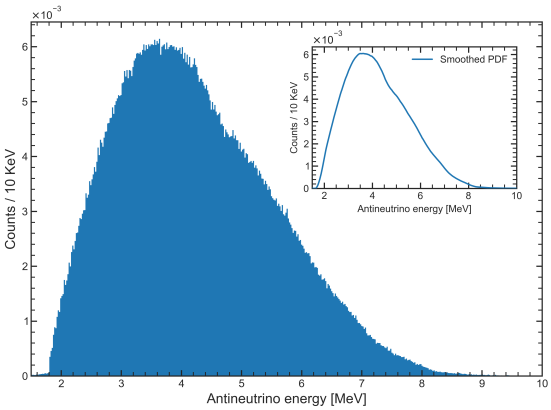
This chapter is dedicated to the investigation of JUNO's sensitivity to oscillation parameters, placing a particular emphasis on the atmospheric mass splitting Δm^2_{31}. The analysis workflow is presented in Section 5.1. The employed cost function and the details on the spectral fit are outlined in Section 5.2. In Section 5.3, a preliminary evaluation of the most influencing parameters on the estimation of Δm^2_{31} is performed, and the results are then integrated in the sensitivities studies in Section 5.4, where results are reported for all oscillation parameters JUNO is sensitive to.

5.1 Analysis strategy

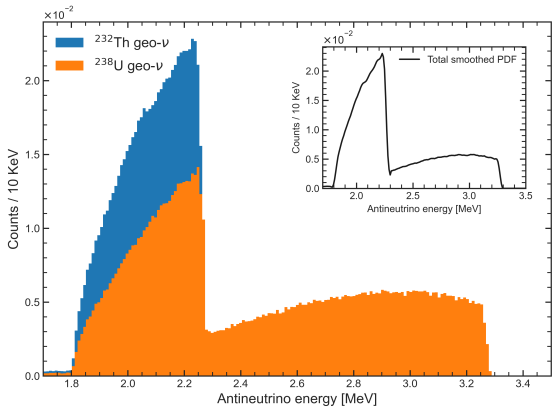
As outlined in previous sections, the information about neutrino oscillation parameters is conveyed through the oscillated reactor antineutrino spectrum. Notwithstanding, as discussed in Section 4.3, there exist several sources of background that are indistinguishable from the signal. Nevertheless, with the help of MC simulations and a thorough comprehension of the detector response, it is possible to predict the relative contributions of the signal and background based on their differing energy distributions. The main steps of the analysis workflow are detailed hereafter:

Monte Carlo production. Both the IBD signal produced by reactor antineutrinos and background components are simulated using the JUNO offline software [49], as mentioned in Chapter 4.

Spectral components PDF production. A Probability Density Function (PDF) for the energy variable of interest can be retrieved directly from the corresponding Monte Carlo sample. To model each component, a specific MC simulation is employed, and the PDF is obtained taking into account the accessible energy scale in a real data acquisition. IBD-like signals, induced from reactor and terrestrial antineutrinos, can be modeled knowing the expected flux, hence as a function of the antineutrino kinetic energy. However, while the energy scale of interest for expressing neutrino oscillations is the antineutrino kinetic energy, this scale is not experimentally measurable. The experimental observable is the total charge collected for a given event, which is the reconstructed energy accounting for factors such as the interaction cross section, LSNL, and finite energy resolution (refer to Figure 3.5). For IBD and geo-neutrino events, the PDFs are generated by starting from the MC truth prompt deposited energy and retrieving the corresponding antineutrino energy. This process is accomplished by considering four-momentum conservation [58]. The obtained antineutrino energy distributions are reported in Figure 5.1.
For ^{9}Li/^{8}He cosmogenic decays, a different approach is taken. The muon veto strategy allows for the isolation of this background component and the measurement of its effective spectral shape using real data. Consequently, the ^{9}Li/^{8}He energy spectra are expressed in terms of reconstructed



(A) Reactor antineutrino energy spectrum (main plot), with entries normalized to unity. The smoothed probability distribution function is reported in the inset.



(B) Geo-ν energy spectra before and after the smoothing process. Entries are normalized to unity, taking into account the Uranium-Thorium relative contribution.

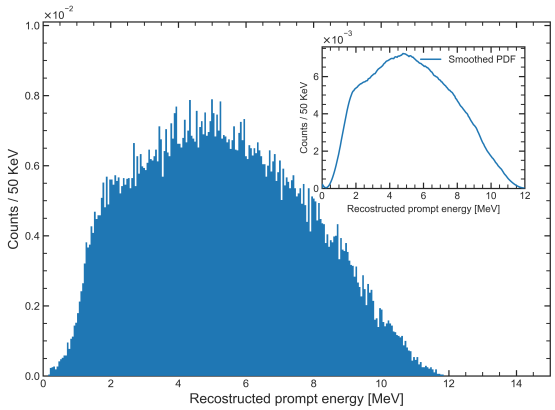
FIGURE 5.1: Reactor and terrestrial antineutrino MC-driven energy spectra. The corresponding smoothed PDFs are reported in the insets.

energy (see Figure 5.2a). Moreover, uncorrelated backgrounds (i.e., accidental coincidences) can be directly inferred from data by employing off-windows techniques[15]: hence, this component is expressed in reconstructed energy as well and reported in Figure 5.2b.
The other backgrounds (fast-n, atmospheric ν, $^{13}C(\alpha, n)^{16}O$ and global reactors) constitute irreducible contributions that cannot be experimentally distinguished. Their spectral shape is taken from JUNO common inputs [74], and they are obtained either through other experiments' data or from dedicated simulations of involved physical processes: they all are reported in terms of reconstructed energy, and are shown in the inset of Figure 4.14. The single spectral components are finally normalized according to the residual rates after all selection cuts (see Table 4.2).

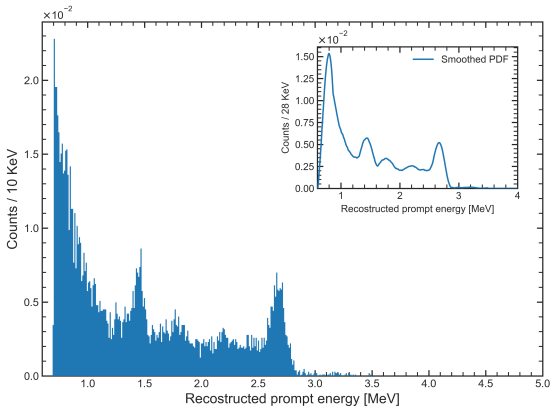
Pseudo-dataset production. Asimov datasets [82] are produced by sampling the aforementioned MC PDFs, based on the expected number of events for a given exposure. Prior to this, a smoothing procedure is applied to the PDFs to reduce statistical fluctuations between adjacent bins while preserving the spectral features. This is achieved using a *Savitzky Golay filter* [83], which fits low-degree polynomials to subsets of adjacent data points and uses the best-fit parameters to estimate the signal value at the center of the window. The filter is applied recursively to the entire data range, effectively removing high-frequency noise and maintaining the original PDF shape [84]. Finally, the datasets are constructed by summing up the contributions from each spectral component, in terms of reconstructed energy.

Spectral fit. To extract the parameters of interest, a binned likelihood function is maximized as part of the fitting strategy. The produced PDFs serve both as a model for the fit and to create several pseudo-datasets. Further details on the likelihood function and spectral fit are provided in Section 5.2.

[15] The contribution of the accidental background spectrum can be determined by searching for coincidences (with the same energy and vertex cuts) in an off-time window, typically much larger than the IBD time cut.



(A) Cosmogenic ^{9}Li/^{8}He energy spectrum (main plot), with entries normalized to unity and taking into account the Lithium-Helium relative contribution. The smoothed probability distribution function is reported in the inset.



(B) Accidental coincidences energy spectrum after fiducial volume and 0.7 MeV energy threshold cuts (see Section 4.3.1), with entries normalized to unity. The smoothed probability distribution function is reported in the inset.

FIGURE 5.2: Cosmogenic ^{9}Li/^{8}He (a) and accidental coincidences (b) MC-driven energy spectra. The corresponding smoothed PDFs are reported in the insets.

5.2 Likelihood function and spectral fit

A multivariate binned likelihood is used as a cost function for the analysis. Each dataset is represented as a histogram with N energy bins, where the i-th bin contains M_i measured events. The likelihood reads

$$\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M}) = \prod_{i=1}^{N} \frac{(s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{M_i}}{M_i!} e^{-(s_i(\boldsymbol{\theta})+b_i(\boldsymbol{\theta}))}, \tag{5.1}$$

where s_i and b_i represent the expected number of signal and background events in the i-th bin, respectively, and $\mathbf{M} = (M_0, M_1, ..., M_i, ..., M_N)$. Hereinafter, the notation $\mu_i = s_i + b_i$ is employed to indicate the total predicted number of events per bin. The likelihood function is the product of Poisson probability density functions corresponding to each bin with the number of entries given by M_i. The expected number of entries in each bin depends on a set of unknown parameters $\boldsymbol{\theta} = (\theta_0, ..., \theta_m)$. For the JUNO oscillated spectrum use case, we consider 33 parameters, i.e., $m = 32$, which are listed in Table 5.1. An in-depth description of all parameters will be provided in Section 5.4, when addressing the source of systematic uncertainties.
The best-fit parameters $\hat{\boldsymbol{\theta}} = (\hat{\theta}_0, ..., \hat{\theta}_m)$ can be determined by performing a *maximum likelihood* (ML) estimation. To accomplish this, it is common to take the natural logarithm of the likelihood function, which is known as log-likelihood. Typically, the optimization problem involves minimizing the negative log-likelihood function, which can be expressed as:

$$\begin{aligned} -2\ln\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M}) &= -2\ln\prod_{i=1}^{N} \frac{\mu_i(\boldsymbol{\theta})^{M_i} e^{-\mu_i(\boldsymbol{\theta})}}{M_i!} \\ &= -2\sum_{i=1}^{N} \ln \frac{\mu_i(\boldsymbol{\theta})^{M_i} e^{-\mu_i(\boldsymbol{\theta})}}{M_i!} \\ &= 2\sum_{i=1}^{N} (\mu_i(\boldsymbol{\theta}) - M_i \ln \mu_i(\boldsymbol{\theta}) + \ln M_i!) \end{aligned} \tag{5.2}$$

Index	0	1	2	3	4	5	6	7	8	9	10
Parameter	$\mathcal{E}$	a	b	c	α_1	α_2	α_3	α_4	ρ	Δm^2_{21}	Δm^2_{31}
Index	11	12	13	14	15	16	17	18	19	20	21
Parameter	$\sin^2\theta_{12}$	$\sin^2\theta_{13}$	N_{rea}	N_{U}	N_{Th}	N_{acc}	N_{fn}	N_{LiHe}	N_{α}	N_{glob}	N_{atm}
Index	22	23	24	25	26	27	28	29	30	31	32
Parameter	N_{SNF}	N_{noneq}	$w^{\text{c1}}_{\text{YJ}}$	$w^{\text{c2}}_{\text{YJ}}$	$w^{\text{c3}}_{\text{YJ}}$	$w^{\text{c4}}_{\text{YJ}}$	$w^{\text{c5}}_{\text{YJ}}$	$w^{\text{c6}}_{\text{YJ}}$	$w^{\text{c1}}_{\text{TS}}$	$w^{\text{c2}}_{\text{TS}}$	w_{DYB}

TABLE 5.1: Oscillated spectrum fit parameters. The cyan highlighted entries are associated with detector-related parameters, therefore efficiency, energy resolution, and non-linearity. The α_i parameters are used to model the non-linearity response, and their meaning will be explained in Section 5.4. The oscillation parameters and the matter density ρ are indicated in orange. In green, we denote the quantities associated with the reactor antineutrino flux: N_{rea} is the total predicted number of reactor antineutrinos, with entries from 24 to 32 representing the weight of each core; N_{SNF} and N_{noneq} indicate the corrections due to spent nuclear fuel and non-equilibrium. The background components are hued in purple.

The expression in Eq. (5.2) is rarely used, while a different construction based on the ratio of two Poisson functions is preferred

$$\lambda(\boldsymbol{\theta}) = \frac{\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M})}{\mathcal{L}(\mathbf{M} \mid \mathbf{M})}, \tag{5.3}$$

where the denominator is a model-independent constant that maximizes the likelihood of the data without any restriction on the model. Maximizing this likelihood ratio is equivalent to minimizing the Poisson-likelihood chi-square function [85]

$$\chi^2_{\text{Poisson}} = -2\ln\lambda(\boldsymbol{\theta}) = 2\sum_{i=1}^{N}\left(\mu_i(\boldsymbol{\theta}) - M_i + M_i\ln\frac{M_i}{\mu_i(\boldsymbol{\theta})}\right). \tag{5.4}$$

The χ^2_{Poisson} function is equal to zero when data is perfectly in agreement with data and is close to the regular χ^2 around the minimum.

It is also possible to relate this formalism to Bayesian parameter estimation (see Appendix C), where the posterior distribution $P(\boldsymbol{\theta} \mid \mathbf{M})$ is proportional to the product of the likelihood function $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M})$ and the prior distribution $P(\boldsymbol{\theta})$:

$$P(\boldsymbol{\theta} \mid \mathbf{M}) \propto \mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M})P(\boldsymbol{\theta}). \tag{5.5}$$

Here, the likelihood function $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M})$ is the probability of observing the data $\mathbf{M}$ given the model parameters $\boldsymbol{\theta}$, and the prior distribution $P(\boldsymbol{\theta})$ expresses our knowledge or beliefs about the model parameters before observing the data. The mode of the posterior distribution corresponds to the maximum a posteriori (MAP) estimate of the parameters, which can be obtained by maximizing the logarithm of the posterior distribution, i.e., finding the values of $\boldsymbol{\theta}$ that maximize

$$\ln P(\boldsymbol{\theta} \mid \mathbf{M}) = \ln\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M}) + \ln P(\boldsymbol{\theta}) + \text{const.} \tag{5.6}$$

5.2.1 Relationship between chi-squared and log-likelihood

At large statistics, the Poisson distribution can be approximated by a normal distribution with mean μ_i and variance $\sigma_i^2 = \mu_i$ [85]. Therefore, under certain regularity conditions, and for a sufficiently

large sample size, the log-likelihood function can be approximated by a quadratic form, leading to a chi-square distribution. Specifically, for Gaussian distributed variables

$$\ln \lambda(\boldsymbol{\theta}) \simeq -\frac{1}{2}\chi^2 + k,$$

hence,

$$\chi^2 \simeq -2 \ln \lambda(\boldsymbol{\theta}) + k. \tag{5.7}$$

Assuming the statistical variance to be represented by the expected number of events in each bin, i.e., μ_i we are considering the so-called Pearson's chi-square; on the other hand, if the variance is approximated by the number of observed events, thus $\sigma_i^2 = M_i$, we obtain the Neyman chi-square [85].

$$\chi^2_{\text{Pearson}} = \sum_i \frac{(\mu_i(\boldsymbol{\theta}) - M_i)^2}{\mu_i(\boldsymbol{\theta})} \quad , \quad \chi^2_{\text{Neyman}} = \sum_i \frac{(\mu_i(\boldsymbol{\theta}) - M_i)^2}{M_i}. \tag{5.8}$$

The advantage of χ^2-like functions is related to their close connection to the covariance matrix formalism, given by

$$\chi^2_{\text{cov}} = (\mathbf{M} - \boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{\eta}))^T \cdot V^{-1} \cdot (\mathbf{M} - \boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{\eta})) \,. \tag{5.9}$$

Hereinafter, we adopt the following notation: $\boldsymbol{\theta}$ is a set of free parameters of interest, while $\boldsymbol{\eta}$ are referred to as *nuisance parameters*, i.e. the additional parameters incorporated into the model to account for uncertainties or external effects that are not the primary focus but are essential for accurately analyzing the data. Within this context, all parameters, except for the oscillation parameters, are classified as nuisances. In Eq. (5.9), V is the covariance matrix of the prediction.

While the ML estimate from the Poisson-likelihood is proven to be an asymptotically unbiased estimator [3], it is known that the estimator of model parameters constructed from Pearson's or Neyman's chi-square leads to biases[16] especially when the large-statistics condition is not met [85]. One way to overcome the possible bias is to consider a combined Neyman-Pearson χ^2 [85], defined as follows

$$\begin{aligned} \chi^2_{\text{CNP}} &= \frac{1}{3}\chi^2_{\text{Neyman}} + \frac{2}{3}\chi^2_{\text{Pearson}} \\ &= \sum_i \frac{(\mu_i(\boldsymbol{\theta}, \boldsymbol{\eta}) - M_i)^2}{3\left(\frac{1}{M_i} + \frac{2}{\mu_i(\boldsymbol{\theta}, \boldsymbol{\eta})}\right)^{-1}} \end{aligned} \tag{5.10}$$

It can be demonstrated [85] that the combined chi-square is closer to the Poisson-likelihood in Eq. (5.4). To investigate this behavior, a study was conducted using 400 pseudo-experiments to construct the $\Delta\chi^2$ profiles shown in Figure 5.3, where $\Delta\chi^2 \equiv \chi^2 - \chi^2_{\text{min}}$. The results clearly illustrate that the Pearson and Neyman chi-square statistics exhibit biases in opposite directions. However, the combined version χ^2_{CNP} effectively compensates for this shift, resulting in a much closer alignment with the log-likelihood estimation.

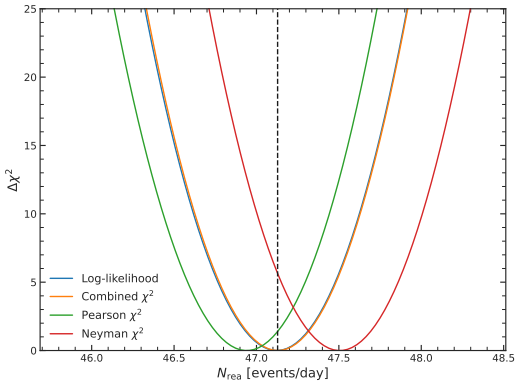


FIGURE 5.3: $\Delta\chi^2$ curves for different test statistics: log-likelihood, χ^2_{Pearson}, χ^2_{Neyman}, and combined χ^2_{CNP}, obtained with 400 pseudo-experiments. The dashed line indicates the true value of the parameter of interest.

[16] The χ^2 best fit estimate is unbiased on Asimov datasets, while it introduces a bias for real data or pseudo-experiments, i.e., in presence of statistical fluctuations.

Furthermore, from the definition in Eq. (5.3), one can see that $0 \leq \lambda \leq 1$, with λ near 1 implying good agreement between the data and the hypothesized value of the model parameters. Equivalently, it is convenient to use the test statistic [82] in Eq. (5.11).

$$t_\theta = -2 \ln \lambda(\theta) \simeq \Delta\chi^2. \tag{5.11}$$

In the context of maximum likelihood estimation, the likelihood ratio can also be related to the chi-square statistic. If the sample size is large and under certain assumptions, the logarithm of the likelihood ratio is approximately equal to half of the chi-square statistic. Therefore, the test statistic t follows a chi-square distribution, as suggested by Eq. (5.7). $\Delta\chi^2$ is defined as the difference in the chi-square statistic between the best-fit model and a reference point, typically the minimum chi-square value.

5.2.2 Correlation among parameters

In order to examine the correlation among various parameters, we carry out an initial fitting process where the parameters are allowed to vary without any prior assumptions. During this stage, for the sake of simplicity, we do not take into account the individual contributions from the reactor cores. Instead, we calculate the reactor antineutrino flux using average power and baselines.

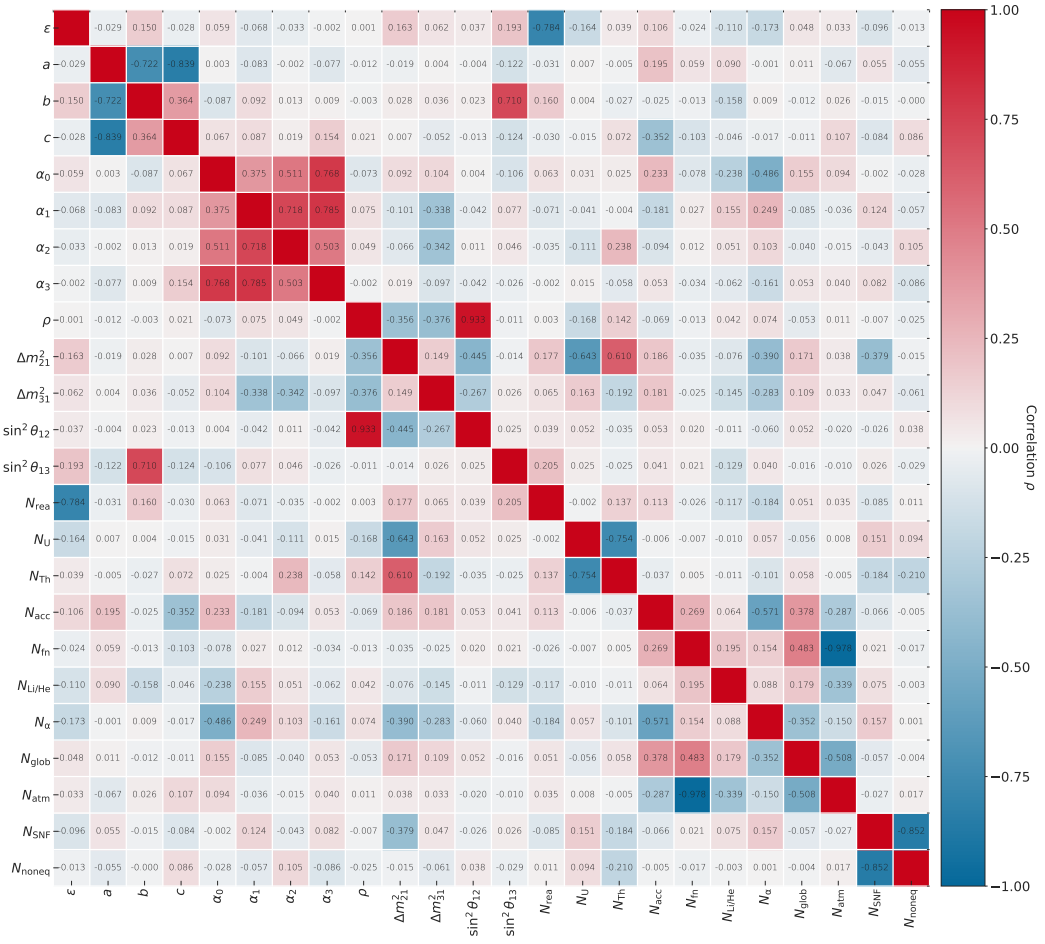


FIGURE 5.4: Correlation matrix obtained keeping all parameters free.

In this work, unless otherwise specified, the minimization is performed using the `Minuit2` [86] library via the Python interface `iminuit` [87]. The correlation matrix, depicted in Figure 5.4, is derived by calculating the covariance matrix provided by the `Minuit` algorithm. This is achieved by inverting the matrix of second derivatives (Hessian matrix) at the minimum point [87]. Specifically, each element (i, j) of the correlation matrix is determined using the following equation:

$$c_{ij} = \frac{\mathrm{cov}(\theta_i, \theta_j)}{\sqrt{\sigma_i^2 \sigma_j^2}},$$

where $\mathrm{cov}(\theta_i, \theta_j)$ denotes the covariance between parameters θ_i and θ_j, and σ_i^2 and σ_j^2 represent the variances[17] of parameters θ_i and θ_j respectively.

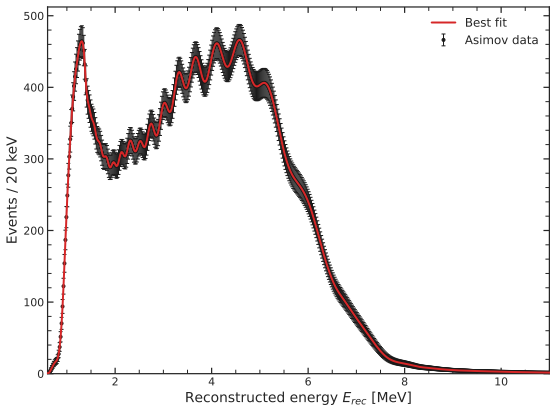
Upon examining the correlation matrix in Figure 5.4, the nuisance parameters that significantly influence the estimation of the two mass splittings and mixing angles can be deduced. Notably, Δm^2_{21} exhibits a pronounced correlation with the geo-neutrino background component,i.e., N_{U} and N_{Th}. This correlation arises due to the dominance of the geo-neutrino background in the low energy region, potentially causing a bias in the position of the first minimum in the oscillated spectrum. As discussed in Section 2.3.2, matter effects also impact the solar oscillation pattern [25, 26], leading to a correlation with the matter density parameter, ρ. Generally speaking, two parameters are expected to be correlated if they produce a similar effect on the energy spectrum. This is the case, e.g., of the atmospheric and fast neutrons backgrounds, since they are both uniform in the prompt signal energy range. Regarding the atmospheric mass splitting, the resolution and non-linearity play the most significant role, suggesting that the precision of its estimation primarily relies on spectral shape distortions rather than the absolute number of events. This foreshadows the importance of accurately characterizing the shape of the spectrum for precise determination of the atmospheric mass splitting.

5.3 Impact of nuisance parameters on Δm^2_{31}

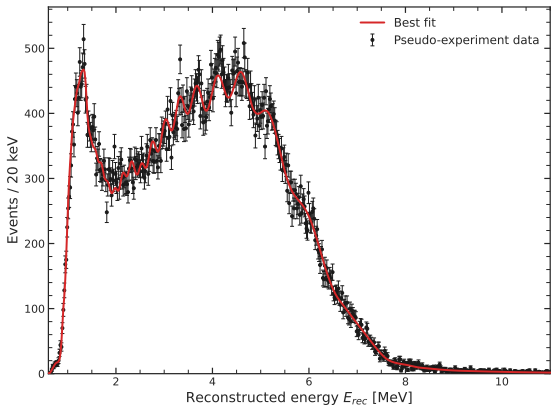
In this section, a preliminary analysis is conducted to qualitatively identify the nuisance parameters that have the greatest impact on the estimation of the atmospheric mass splitting. The objective of our analysis is to find the best-fit values for both the parameters of interest and the nuisance parameters, aiming to minimize the discrepancy between the model predictions and the observed (simulated) data. To accomplish this, we employ an Asimov dataset [82], which is constructed to perfectly align with the expected outcomes predicted by a given model. In other words, the model's parameters are set to their true values, and no random variation is introduced in the generated data. The Asimov dataset represents a "best-case scenario" for the model, serving as a benchmark for evaluating its performance considering the median effect of statistical fluctuations. When employed to estimate all parameters, the Asimov dataset yields the true parameter values without introducing biases. Figure 5.5a illustrates an example of an Asimov dataset, while Figure 5.5b showcases a pseudo-experiment where each bin is subject to fluctuations following a Poisson distribution.
In this section, several alternative techniques are investigated to handle nuisance parameters. They all involve evaluating the likelihood as a function of only the parameter (or parameters) of interest: the main approaches are the profile and conditional likelihoods.

Profile likelihood. The profile likelihood is obtained by maximizing the likelihood function with respect to the nuisance parameters, effectively "profiling out" their influence. It is constructed by

[17] The variance is given by the diagonal entries of the covariance matrix.



(A) Fit performed on Asimov dataset, for 6 years exposure.



(B) Fit performed on pseudo-experiment dataset, for 6 years exposure.

FIGURE 5.5: Examples of Asimov (a) and pseudo-experiment (b) datasets for 6 years exposure. The best-fit line is also reported.

maximizing the likelihood function for each value of the parameter of interest θ, while allowing the nuisance parameters $\boldsymbol{\eta}$ to vary freely.
To test a hypothesized value of a parameter of interest θ we consider the so called *profile likelihood ratio* [82] as test statistic:

$$\lambda(\theta) = \frac{\mathcal{L}(\theta, \hat{\hat{\eta}})}{\mathcal{L}(\hat{\theta}, \hat{\eta})}, \tag{5.12}$$

where θ indicates the parameter of interest and η one nuisance parameter. The numerator represents the maximized likelihood function for the parameter of interest θ with a given nuisance parameter maximized at its conditional maximum likelihood estimate, denoted as $\hat{\hat{\eta}}$. The denominator is the maximized (unconditional) likelihood, i.e., both $\hat{\theta}$ and $\hat{\eta}$ are ML estimates.
Therefore, the profile likelihood function represents the variation of the likelihood with respect to the parameter of interest, taking into account the best-fitting values of the nuisance parameters for each value of the parameter of interest. By maximizing over the nuisance parameters, the profile likelihood implicitly takes into account their correlations with the parameter of interest.

Conditional likelihood. The conditional likelihood is obtained by conditioning the likelihood function on specific values or ranges of the nuisance parameters. The conditional likelihood is expressed as a function of the parameter of interest, with the nuisance parameters fixed to given values. This approach is commonly used to study the effect of fixing a nuisance parameter on the estimation of the parameter of interest. Mathematically, the conditional likelihood ratio with a fixed nuisance parameter can be written as

$$\lambda(\theta \mid \eta_0) = \frac{\mathcal{L}(\theta, \eta = \eta_0)}{\mathcal{L}(\hat{\theta}, \eta = \eta_0)}. \tag{5.13}$$

By fixing the nuisance parameters, the conditional likelihood focuses solely on the parameter of interest, disregarding their correlations.

A comparison between the two aforementioned methods is reported in Figure 5.6, where we report the 2-dimensional 68.27%, 95.45%, and 99.73% confidence level contours for one parameter of interest (Δm^2_{21}) and one nuisance (N_{U}), with the corresponding 1-d $\Delta\chi^2$ profiles.

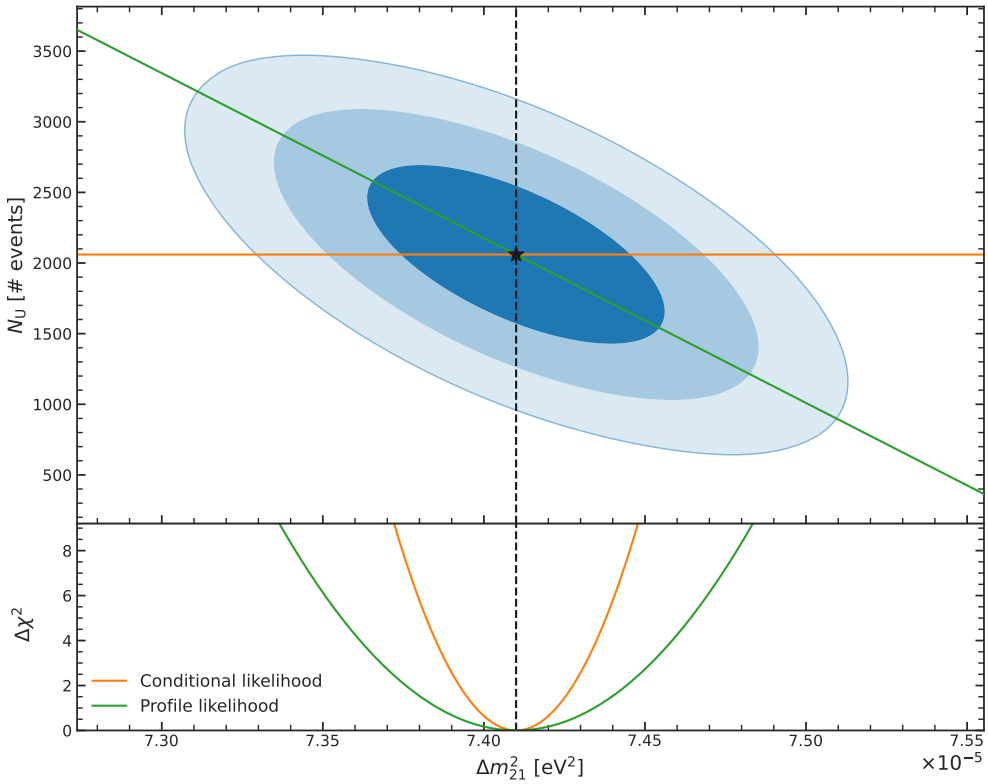


FIGURE 5.6: Two-dimensional 68.27%, 95.45%, and 99.73% confidence level contours for one parameter of interest ($\theta = \Delta m^2_{21}$) and one nuisance ($\eta = N_{\mathrm{U}}$). In the secondary axis, the conditional and profile likelihoods are compared: the former (in orange) corresponds to the narrower profile, while the latter is significantly wider due to the known correlation between the solar mass splitting and the number of Uranium geoneutrinos N_{U}.

5.3.1 Nuisances affecting the precision

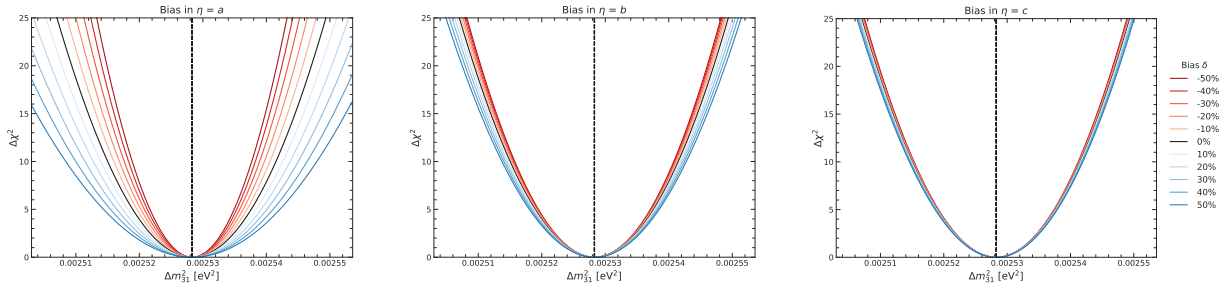
The initial set of preliminary tests focuses on the conditional likelihood given in Eq. (5.13). The objective is to evaluate the impact of individual nuisance parameters on the estimation of Δm^2_{31}. In this analysis, we assume that each nuisance parameter deviates from its initial expected value within a well-defined range, while still being accurately measurable. This deviation is denoted as δ, and we assess the likelihood ratio as follows:

$$\lambda(\theta \mid \eta_0 + \delta) = \frac{\mathcal{L}(\boldsymbol{\mu}(\theta, \eta_0(1+\delta)), \mathbf{M}(\theta, \eta_0(1+\delta)))}{\mathcal{L}(\boldsymbol{\mu}(\hat{\theta}, \eta_0(1+\delta)), \mathbf{M}(\theta, \eta_0(1+\delta)))}, \tag{5.14}$$

where δ is a bias introduced on each nuisance, such that $\delta \in [-0.5, 0.5]$, and η_0 is the input value, i.e. the expected one from the current knowledge of the experiment. Here, we explicitly express the dependence on the model, represented by the expected number of events $\boldsymbol{\mu}$, and on the data sample $\mathbf{M}$ to emphasize that the parameters are accurately estimated through likelihood maximization. It should be noted that we are working with Asimov datasets; as a result, the ML estimate obtained from these datasets will not exhibit a bias in its central value. However, the width of the profile,

which is a proxy of its uncertainty, may be affected by variations in the nuisance parameters. For a Gaussian distribution and one degree of freedom (e.g., when considering a single parameter), a change of $\Delta\chi^2 = 1$ corresponds to a one-sigma deviation from the minimum chi-square value, representing approximately a 68% confidence interval[18]. It has to be underlined that the 1σ interval obtained from the conditional likelihood provides an underestimated result, since it may not fully capture the true uncertainty due to the neglect of correlations between parameters.
The main sources of uncertainty in the measurement of Δm^2_{31} are the statistics of reactor antineutrinos and all parameters that can distort the spectral shape. Indeed the statistical precision is by definition improved with a higher number of events, while the energy resolution has to be sufficiently high to be able to resolve the fast oscillation pattern. On the other hand, factors related to rate normalization have little impact on the determination of Δm^2_{31}. Therefore, while the results for all nuisance parameters are reported in Appendix D, we hereafter concentrate on a limited number of nuisance parameters that are known to influence the spectral shape and analyze their effects in detail.

Resolution parameters. Figure 5.7 shows the conditional likelihood curves for the parameters a, b, c describing the energy resolution, as in Eq. (3.13).



(A) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = a$.

(B) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = b$.

(C) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = c$.

FIGURE 5.7: Conditional likelihood profile considering $\theta = \Delta m^2_{31}$ as parameter of interest and fixing the nuisance parameters related to the energy resolution, i.e., $\eta = (a, b, c)$.

The color scale indicates the magnitude of the bias δ, such that for each curve the nuisance parameter is fixed at $\eta = \eta_0(1 + \delta)$. Firstly, it can be observed that, as expected for an Asimov sample, the ML estimate is not affected, whereas the width of the profiles turns out to be either enlarged or narrowed. Looking at the parameterization in Eq. (3.13), it is straightforward to infer that higher (lower) values of a, b, c correspond to worse (better) energy resolution. For instance, this is exactly what can be seen in Figure 5.7, where the width of the profiles varies with η_i. For each value of η, we calculate the 1σ uncertainty on

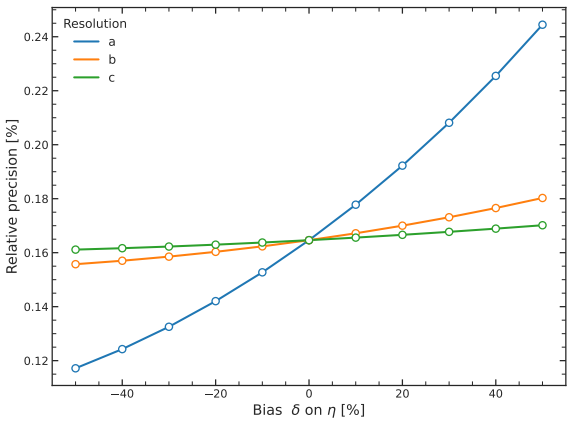
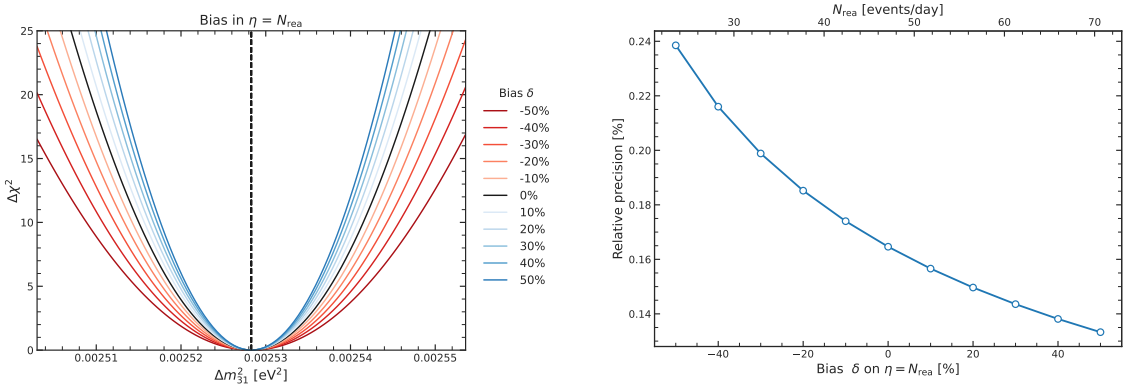


FIGURE 5.8: Relative precision on Δm^2_{31} as a function of the bias introduced on resolution parameters, obtained from the conditional likelihood profiles.

[18] These approximations are valid in certain asymptotic limits and may not hold exactly for small sample sizes or in the presence of non-Gaussian distributions.

Δm^2_{13} by determining the range where $\Delta\chi^2 = 1$, relative to the minimum chi-squared value. It is important to note that the conditional likelihood approach tends to underestimate the uncertainty because it does not take into account correlations with the nuisance parameters being considered. A careful evaluation of uncertainties comprehensive of correlations will be performed in Section 5.4. The results for the resolution parameters are presented in Figure 5.8, which illustrates the relative precision on Δm^2_{31} as a function of δ. As expected, the term influenced by the Poisson statistics of the detected photoelectrons (i.e., a) has the most significant impact since it contributes the most to the energy resolution.

Reactor antineutrino statistics. It is evident that the statistical contribution to the overall uncertainty will primarily be influenced by the reactor antineutrino statistics. This observation becomes apparent when examining Figure 5.9a, which illustrates that a higher number of events (referred to as N_{rea}) leads to narrower profiles, resulting in smaller errors. This trend is further supported by Figure 5.9b, where the relative precision is depicted as a function of N_{rea}. These findings underscore the importance of maximizing the number of signal events through the optimization of selection cuts, among other strategies.



(A) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = N_{\text{rea}}$.

(B) Relative precision on Δm^2_{31} obtained from the conditional likelihood profiles.

FIGURE 5.9: (a) Conditional likelihood profile considering $\theta = \Delta m^2_{31}$ as parameter of interest and fixing the number of reactor antineutrino events. (b) Relative precision on Δm^2_{31} for different values of N_{rea}, i.e., as a function of the bias introduced on N_{rea}.

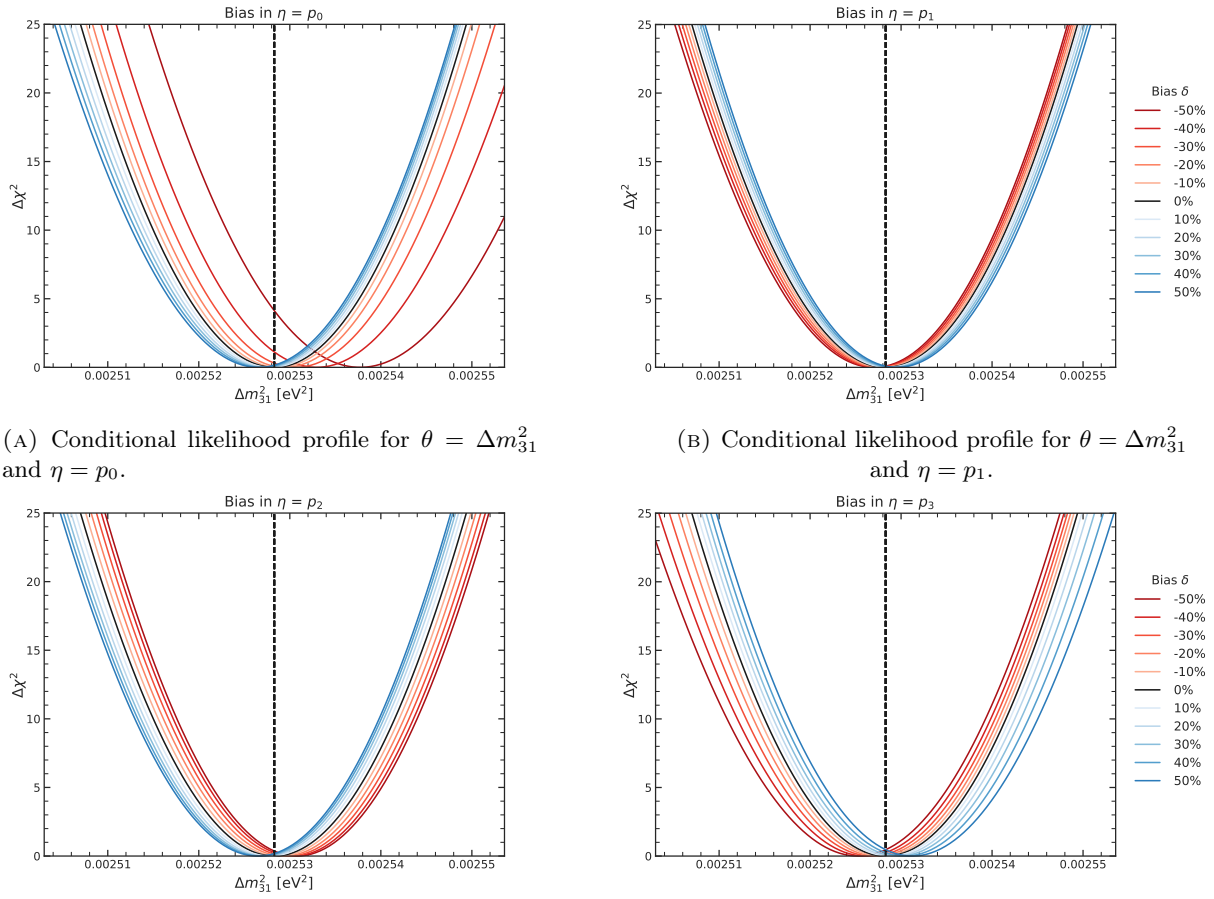
Looking at the results with all parameters in Appendix D, it can be seen that the efficiency factor has a comparable effect with what is shown in Figure 5.9. This similarity arises from the fact that the efficiency factor influences the antineutrino statistics in a comparable manner.

5.3.2 Nuisances affecting the accuracy

The other nuisance parameters are not expected to significantly affect the uncertainty on Δm^2_{31}. Rather, it is important to understand what happens if, for a fixed data sample, we fail to accurately estimate a given η. This scenario involves estimating the nuisance parameter with a bias, resulting in a value of $\eta = \eta_0(1+\delta)$, while still using the same fixed data sample, dubbed $\mathbf{M}(\theta, \eta_0)$. Hence, the following expression is used:

$$\lambda(\theta \mid \eta_0 + \delta) = \frac{\mathcal{L}(\boldsymbol{\mu}(\theta, \eta_0(1+\delta)), \mathbf{M}(\theta, \eta_0))}{\mathcal{L}((\boldsymbol{\mu}(\hat{\theta}, \eta_0(1+\delta)), \mathbf{M}(\theta, \eta_0))}, \tag{5.15}$$

where the likelihood function is evaluated for fixed data, denoted as $\mathbf{M}(\theta, \eta_0)$, but with an incorrectly estimated nuisance parameter η_i given by $\eta_0(1+\delta)$. Specifically, this procedure provides a qualitative assessment of how sensitive the estimation of θ is to inaccuracies in measuring η within the given fixed data sample.
This aspect holds significant relevance for the LSNL, since it induces a shift in the energy spectrum that can potentially bias the estimation of the oscillation parameters. Indeed, this shift, in turn, affects the oscillation pattern and introduces distortions in the observed data. To effectively enhance the effect, we use the LSNL Daya Bay parameterization provided by Eq. (3.12), which offers a more realistic representation based on experimental data, with respect to using the pull curves α_i in Figure 5.13b. Therefore the group of nuisance parameters will be given by $\eta = (p_0, p_1, p_2, p_3)$.



(A) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = p_0$.

(B) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = p_1$.

(C) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = p_2$.

(D) Conditional likelihood profile for $\theta = \Delta m^2_{31}$ and $\eta = p_3$.

FIGURE 5.10: Conditional likelihood profiles considering $\theta = \Delta m^2_{31}$ as parameter of interest and fixing the nuisance parameters related to the energy non-linearity, i.e., $\eta = (p_0, p_1, p_2, p_3)$.

Figure 5.10 shows the conditional likelihood curves obtained by varying all non-linearity parameters of a certain quantity δ, as previously mentioned. It is evident that the Δm^2_{31} ML estimate is skewed in a particular direction. The direction of the systematic bias depends on the specific combination of NL parameters and their impact on the oscillation pattern. In some cases, the distortions introduced by the NL parameters may align in a way that pushes the ML estimate of Δm^2_{31} towards higher values, while in other cases, it may adjust it towards lower values.
This simple study serves as a valuable reminder of the importance of properly accounting for the complex interplay that exists between the non-linearity phenomenon and the underlying oscillation

pattern, since they might induce subtle changes in the spectrum which are not easily attributed solely to either of these factors. It indicates that failure to accurately model non-linear effects can lead to biased estimations of the oscillation parameters, such as Δm^2_{31}.

5.3.3 Profile likelihood scans

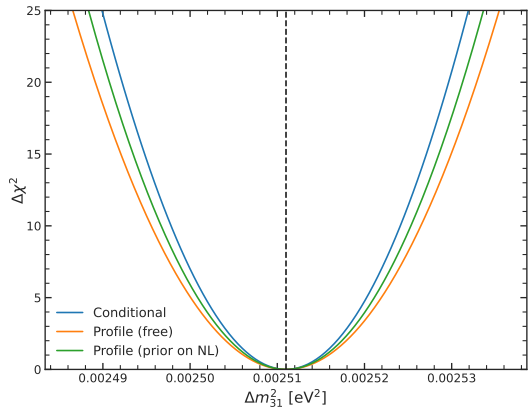
The conditional likelihood method employed so far proved to be an essential tool to grasp the relationship between θ and individual nuisances. Nonetheless, this approach is known to provide the narrowest $\Delta\chi^2$ profiles, as it scans the conditional distribution of one parameter of interest given fixed values of the other nuisances. Hence, by construction, it does not account for parameter correlations, leading to underestimated uncertainties. Quantifying these parameters' correlations is crucial for effectively prioritizing which nuisances, and their corresponding physics phenomena or detector-related properties, to focus on during calibration campaigns and data analysis. In order to do that, the profile likelihood approach introduced in Section 5.3 will be used. Specifically, we need to perform a fit on Asimov datasets, using nominal parameters, and analyze the profile likelihood using different constraints on the nuisances.
From Figure 5.4, it is evident that there is nearly no correlation between Δm^2_{31} and all nuisance parameters. The dominant contribution to the correlation arises from non-linearity effects. In Figure 5.11a, we compare the conditional likelihood profile with the profile likelihood in two different configurations: one with all free nuisances and the other with NL parameters constrained based on the input priors. The conditional and "all free" cases represent the two extreme scenarios, yielding the highest and lowest possible precision, respectively. Placing a prior on the NL parameters makes the profile notably narrower, due to the known correlation with Δm^2_{31}. The procedure described above was also applied to the resolution parameters, and the corresponding results are depicted in Figure 5.11a: no observable effect is produced by the introduction of a constraint on a, b, c. All other nuisances affecting the rate normalization are found to have little to no impact on the precision of Δm^2_{31}, which is driven by uncertainties affecting the shape of the reactor $\overline{\nu}_e$ spectrum. This aspect will be further investigated in Section 5.5. This study allowed us to verify that the presence of the nuisance parameters broadens the profile likelihood as a function of θ compared to a scenario where their values are fixed. This reflects the loss of information about θ due to correlations and systematic uncertainties.

5.4 Systematic uncertainties

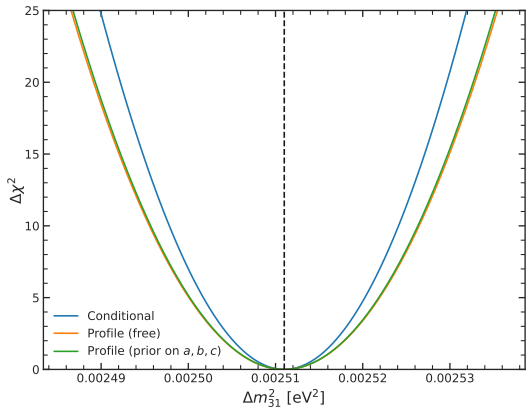
Systematic effects can be categorized into two distinct groups: bin-to-bin correlated uncertainties and bin-to-bin uncorrelated uncertainties. The former can be further divided into fully correlated, relating to factors that influence the overall event count or normalization, often referred to as *rate uncertainties*, and partially correlated. The latter is the case of non-linearity, resolution, and matter effects uncertainties, since they do not act as a rate normalization but modify the spectral shape. On the other hand, bin-to-bin uncorrelated uncertainties refer to the (statistical and systematic) uncertainties resulting from the estimation of reference spectra PDFs, and they are uncorrelated among individual energy bins. The systematic uncertainties, summarized in Table 5.2, are implemented with two different methods:

1. External constraints on bin-to-bin correlated uncertainties are propagated through Gaussian penalty (or pull) terms on χ^2-like functions, which correspond to adding a gaussian prior in a likelihood-based cost function. The penalty term is defined as:

$$\chi^2_{\text{penalty}}(\boldsymbol{\eta}) = (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})^T V_{\boldsymbol{\eta}}^{-1} (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}), \tag{5.16}$$



(A) $\Delta\chi^2$ profiles for $\theta = \Delta m^2_{31}$, showing the conditional likelihood and the profile likelihood (with all free parameters and with constrained non-linearity parameters).



(B) $\Delta\chi^2$ profiles for $\theta = \Delta m^2_{31}$, showing the conditional likelihood and the profile likelihood (with all free parameters and with constrained resolution parameters).

FIGURE 5.11: $\Delta\chi^2$ profiles for $\theta = \Delta m^2_{31}$, showing the conditional likelihood and the profile likelihood. In (a) the conditional likelihood is compared with the profile likelihood with constrained non-linearity (NL) parameters; (b) reports the profiles for constrained resolution parameters.

where $V_{\boldsymbol{\eta}}$ is the covariance matrix for the nuisance parameters. For a set of uncorrelated $\boldsymbol{\eta}$, the expression reduces to

$$\chi^2_{\text{penalty}}(\boldsymbol{\eta}) = \sum_i \frac{(\eta_i - \hat{\eta}_i)^2}{\sigma^2(\eta_i)}. \tag{5.17}$$

$\hat{\eta}_i$ is the central value, $\sigma(\eta_i)$ is the uncertainty of i-th nuisance parameter. The nuisances $\boldsymbol{\eta}$ act on the energy spectrum through their dependence on $\mu = \mu(\boldsymbol{\theta}, \boldsymbol{\eta})$. In our case all nuisance parameters are assumed to be uncorrelated, hence the formalism in Eq. (5.17) is used.

2. External constraints on bin-to-bin uncorrelated uncertainties are added via the extension of the covariance matrix. In this case, using penalty terms would require a significant increase in the number of minimization parameters, hence it is more convenient to use a χ^2 in the covariance matrix formalism. Specifically, the covariance matrix V in Eq. (5.9) is extended to include both statistical and systematical bin-to-bin (b2b[19]) uncertainties, namely

$$V = V^{\text{stat}} + V^{\text{b2b}} \tag{5.18}$$

V^{b2b} is constructed as follows:

$$V^{\text{b2b}}_{ij} = \delta^2_{\text{b2b}} \cdot \mu_i \mu_j \delta_{ij}, \tag{5.19}$$

where δ_{b2b} is the b2b relative uncertainty, μ_k represents the expected number of events in the k-th bin, and δ_{ij} is the Kronecker delta. Therefore, V^{b2b} is a diagonal matrix such that $V^{\text{b2b}}_{ii} = \delta^2_{\text{b2b}} \cdot \mu^2_i$. These uncertainties are defined for a specific bin width, denoted as a. To scale the relative uncertainty δ_a from bin width a to b, a multiplication factor is required, as shown in Eq. (5.20):

$$k = \frac{\delta_b}{\delta_a} = \sqrt{\frac{a}{b}}. \tag{5.20}$$

[19] Hereinafter with the notation *b2b* we indicate the bin-to-bin uncorrelated uncertainties.

Since there is no convenient way to introduce b2b uncertainties within a likelihood function, the following approach is used:

- When dealing with purely statistical uncertainties, the multivariate binned likelihood (Eq. (5.4)) is employed to obtain the best fit parameters.
- Considering the equivalence in Eq. (5.7), it is reasonable to switch to a combined χ^2, as in Eq. (5.10), to quantify the impact of systematical uncertainties. This choice is motivated by the fact that the method approximates a log-Poisson ratio, ensuring unbiasedness for both Asimov and pseudo-experiments [85]. Additionally, this approach facilitates the incorporation of systematic effects through pull terms and extended covariance.

Therefore, using the combined Neyman-Pearson prescription, the final cost function reads

$$\chi^2_{\mathrm{CNP}} = (\mathbf{M} - \boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{\eta}))^T \cdot \left(V^{\mathrm{stat}}_{\mathrm{CNP}} + V^{\mathrm{b2b}}\right)^{-1} \cdot (\mathbf{M} - \boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{\eta})) + \sum_i \frac{(\eta_i - \hat{\eta}_i)^2}{\sigma^2(\eta_i)} \tag{5.21}$$

with

$$V^{\mathrm{stat}}_{\mathrm{CNP}}(\boldsymbol{\theta})_{ij} \equiv 3\left(\frac{1}{M_i} + \frac{2}{\mu_i(\boldsymbol{\theta})}\right)^{-1}\delta_{ij} \qquad \text{and} \qquad \mathrm{V}^{\mathrm{b2b}}_{\mathrm{ij}} = \mathrm{k_i k_j}\cdot\left(\delta^2_{\mathrm{b2b,rea}}\cdot \mathrm{s_i s_j}\delta_{\mathrm{ij}} + \sum_{\mathrm{B}} \delta^2_{\mathrm{b2b,B}}\cdot \mathrm{b_i b_j}\delta_{\mathrm{ij}}\right).$$

The notation with s_i and b_i is intended: s_i refers to the reactor antineutrino spectrum and b_i to the expected number of events for the background component B. The different sources of rate and b2b systematic uncertainties are summarized in the following sections.

5.4.1 Reactor flux-related uncertainties

The uncertainty associated with the reactor flux comprises both rate effects, which are correlated across all energy bins, and the uncertainty arising from different flux models. The systematic uncertainties related to the predicted reactor antineutrino spectrum can be classified as either correlated or uncorrelated among different reactor cores. Additionally, the energy-dependent uncertainties can be further categorized as either correlated or uncorrelated between different energy bins.

Reactor correlated uncertainty. This type of uncertainty accounts for factors that are common to all reactor cores, including the energy released per fission, the antineutrino yield per fission, and the spectra of the fuel isotopes. It is correlated across all reactors, thereby producing an overall shift in the number of reactor antineutrinos (N_{rea} in Table 5.1), and has been estimated to contribute a total relative uncertainty of $\delta_C = 2\%$, as calculated in [60]. The corresponding pull term in Eq. (5.21) is given by

$$\chi^2_{\mathrm{rea,corr}}(N_{\mathrm{rea}}) = \frac{(N_{\mathrm{rea}} - \hat{N}_{\mathrm{rea}})^2}{\sigma^2_C}. \tag{5.22}$$

Reactor uncorrelated uncertainty. This uncertainty arises from variations in normalization that occur independently from one reactor core r to another, due to the reactor power, the burn-up calculation, and the fission fractions. The total contribution of reactor uncorrelated uncertainty amounts to a relative uncertainty of $\delta_r = 0.8\%$. This systematic effect is applied by rescaling the weight of each reactor core w_{r} (with indices from 24 to 32 in Table 5.1) with an independent Gaussian centered at 1 with a width of 0.008, to account for the 0.8% relative uncertainty, resulting in a penalty term

$$\chi^2_{\text{rea,uncorr}}(w_\text{r}) = \sum_r^{\text{cores}} \frac{(w_\text{r} - \hat{w}_\text{r})^2}{\sigma_r^2}. \tag{5.23}$$

Non-equilibrium and SNF uncertainties. The normalization rate uncertainty of non-equilibrium and spent nuclear fuel is assumed to be 30% with negligible b2b uncertainty. The pull term for both contributions is given by

$$\chi^2_{\text{SNF,noneq}}(N_\text{SNF}, N_\text{noneq}) = \frac{(N_\text{SNF} - \hat{N}_\text{SNF})^2}{\sigma^2_\text{SNF}} + \frac{(N_\text{noneq} - \hat{N}_\text{noneq})^2}{\sigma^2_\text{noneq}}. \tag{5.24}$$

Given the differences in the NPPs characteristics, it is likely that single cores will exhibit diverse contributions from both SNF and non-equilibrium effects. Consequently, the systematic uncertainties associated with these two factors are expected to be uncorrelated among different reactors. However, the contributions from these corrections will be averaged over time, leading to a high degree of correlation among reactors. For the purpose of this analysis, we thus consider these uncertainties to be common to all cores, being a reasonable approximation over an extended period of data taking.

Signal spectrum bin-to-bin uncertainty. This uncertainty arises from the calculation of the antineutrino energy spectra, the fission fractions, and the spectra of the fuel isotopes. The spectrum b2b uncertainty is assumed to be correlated among all reactors while being uncorrelated between different energy bins. The JUNO collaboration has examined various models, but for this analysis, we focus on the two most recent ones [21, 55], based on Daya Bay and TAO, which are shown in Figure 5.12:

- **Daya Bay based flux model:** The Daya Bay measurement [88] can be used as a reference spectrum. The uncertainty of the Daya Bay antineutrino energy spectrum accounts for both statistical and systematical contributions and is given for a bin width of 250 keV.
- **TAO based flux model:** The primary objective of the Taishan Antineutrino Observatory (TAO) satellite experiment is to establish a model-independent and highly accurate reference antineutrino spectrum for JUNO. TAO spectrum uncertainties are calculated based on a thorough detector simulation described in [31]. By integrating various sources of uncertainties, an energy-dependent uncertainty is derived, which is reported in Figure 5.12.

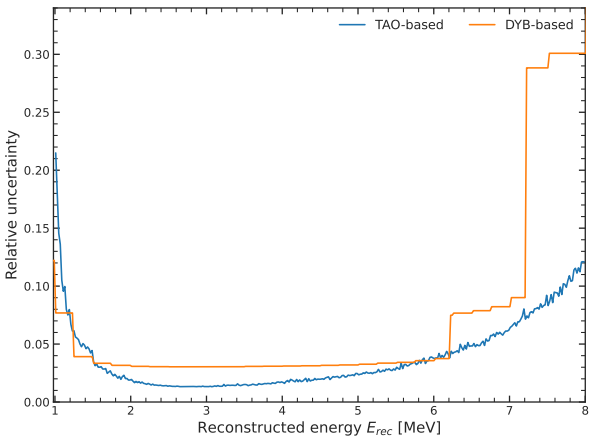


FIGURE 5.12: Antineutrino spectrum b2b uncertainty from different models, rescaled for a bin width of 10 keV, according to Eq. (5.20): the satellite detector TAO-based estimation is taken from [31], while the Daya Bay (DYB) is based on [88].

The systematic variance induced by the reactor spectrum b2b uncertainty for the i-th energy bin is embedded in the chi-square via a covariance matrix, namely

$$V_{ij}^{\text{b2b,rea}} = k_i k_j \cdot \delta^2_{\text{b2b,rea}} \cdot s_i s_j \cdot \delta_{ij} \tag{5.25}$$

5.4.2 Observed spectrum and detector response uncertainties

Detector efficiency uncertainty. An overall rate uncertainty of $\delta_D = 1\%$ is assigned to the detection efficiency $\mathcal{E}$. This uncertainty accounts for all effects that impact the $\overline{\nu}_e$ interaction probability, including the number of target protons and the H/C ratio (as shown in Eq. (3.15)) and the selection efficiency. The normalization factor associated with this uncertainty is applied to all spectral components with PDFs expressed in terms of antineutrino energy, such as geo-neutrinos and reactor antineutrinos. This effect is assigned with a pull term:

$$\chi^2_{\mathrm{D}}(\mathcal{E}) = \frac{(\mathcal{E} - \hat{\mathcal{E}})^2}{\sigma_D^2}. \tag{5.26}$$

Energy resolution. The parameters a, b and c in Eq. (3.13) are pulled considering Gaussian functions of width described by the corresponding uncertainties, resulting in the following nuisance term

$$\chi^2_{\mathrm{res}}(a,b,c) = \sum_{\eta=a,b,c} \frac{(\eta - \hat{\eta})^2}{\sigma_\eta^2}, \tag{5.27}$$

where the relative uncertainties are $\delta_a = 0.19\%$, $\delta_b = 0.47\%$ and $\delta_c = 0.89\%$ [72].

Non-linearity. The relative energy scale distortion due to LSNL is parameterized as:

$$\frac{E_{\mathrm{vis}}}{E_{\mathrm{dep}}} = \left(f_{\mathrm{nom}}\left(E_{\mathrm{dep}}\right) + \sum_l \alpha_l \left(f_l\left(E_{\mathrm{dep}}\right) - f_{\mathrm{nom}}\left(E_{\mathrm{dep}}\right)\right)\right) \cdot \frac{f_{\mathrm{nom}}^{\mathrm{SNiPER}}\left(E_{\mathrm{dep}}\right)}{f_{\mathrm{nom}}^{\mathrm{Daya\ Bay}}\left(E_{\mathrm{dep}}\right)}. \tag{5.28}$$

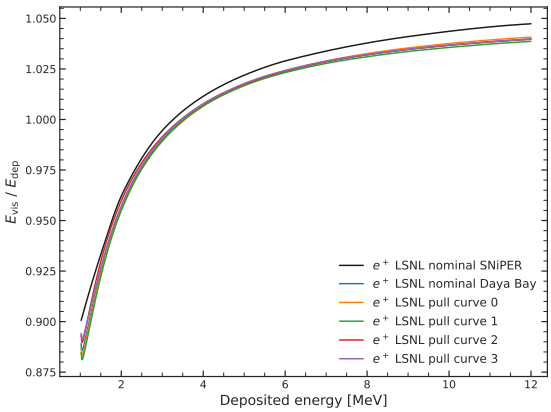
f_{nom} is the nominal LSNL curve from Daya Bay [71] and f_l with $l = 0, ..., 3$ are four pull curves to capture the corresponding uncertainty, all shown in Figure 5.13a. The variation of the pull curves with respect to the nominal value is expressed in Figure 5.13b. Moreover, the Daya Bay LSNL curve is corrected so that the LSNL with nominal parameters curve matches the LSNL curve from the SNiPER simulation $f_{\mathrm{nom}}^{\mathrm{SNiPER}}$, grasping the most recent knowledge of the JUNO detector [74]. α_l are the nuisance parameters associated with each pull curve, assumed to follow a Gaussian distribution centered in zero and with a standard deviation of 1; consequently, a 100% relative uncertainty is assumed for each of the four non-linearity nuisance parameters.

$$\chi^2_{\mathrm{LSNL}}(\boldsymbol{\alpha}) = \sum_{l=0}^{3} \frac{(\alpha_l - \hat{\alpha}_l)^2}{\sigma_l^2}, \tag{5.29}$$

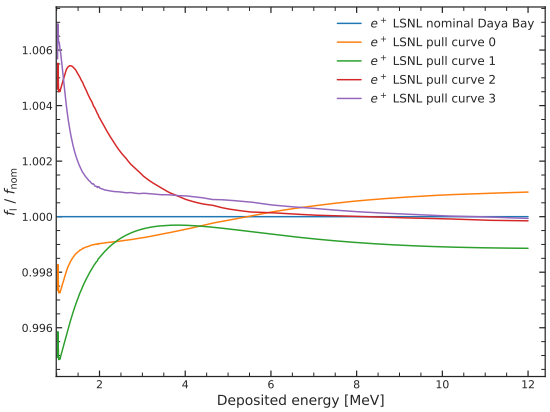
Backgrounds uncertainties. Both rate and b2b uncertainties related to the main backgrounds (here indexed with B) are summarized in Table 5.2. The rate uncertainty is added as a regular pull term, hence

$$\chi^2_{\mathrm{bkg}}(N_{\mathrm{B}}) = \sum_{\mathrm{B}}^{\mathrm{bkgs}} \frac{(N_{\mathrm{B}} - \hat{N}_{\mathrm{B}})^2}{\sigma_{\mathrm{B}}^2}. \tag{5.30}$$

The systematic variance induced by the backgrounds b2b uncertainty for the i-th energy bin is embedded in the chi-square via a covariance matrix, namely



(A) Nominal and pull LSNL curves from Daya Bay and LSNL curve from SNiPER.



(B) Ratio of the LSNL pull curves from Daya Bay to the nominal LSNL curve..

FIGURE 5.13: (a) Nominal pull curves for liquid scintillator non-linearity (LSNL). (b) Ratio of the LSNL curves to the nominal LSNL curve, as a function of deposited energy.

$$V_{ij}^{\text{b2b,bkg}} = k_i k_j \cdot \sum_{\text{B}}^{\text{bkgs}} \delta_{\text{b2b,B}}^2 \cdot b_i b_j \delta_{ij}. \tag{5.31}$$

Matter density uncertainty. Oscillation probabilities are calculated with an average matter density $\rho = (2.45 \pm 0.15)$ g/cm^3, hence assuming a 6% relative uncertainty [55, 59]. The pull term is defined as follows:

$$\chi_\rho^2(\rho) = \frac{(\rho - \hat{\rho})^2}{\sigma_\rho^2}. \tag{5.32}$$

Systematic effect	Relative uncertainty [%]	Bin-to-bin correlation
Reactors		
Core to core uncorrelated	0.8	Fully correlated
Core to core correlated	2	Fully correlated
Reference spectrum	Daya Bay or TAO based	Uncorrelated
Spent Nuclear Fuel rate	30	Fully correlated
Non-equilibrium rate	30	Fully correlated
Detector		
Normalization/efficiency	1	Fully correlated
Energy resolution	0.19, 0.47, 0.83	Correlated
LSNL (pull terms)	100	Correlated
Backgrounds rate	see Table 4.2	Fully correlated
Backgrounds b2b	see Table 4.2	Uncorrelated
Matter density	6	Correlated

TABLE 5.2: Summary of systematic uncertainties which impact the JUNO detector and their bin-to-bin correlation considered in this analysis.

5.5 Sensitivity determination

The sensitivity to oscillation parameters is determined by χ^2 minimization, which is performed using the `Minuit2` [86] library via the Python interface `iminuit` [87].
The measured spectrum is calculated with the nominal setup (Normal Ordering), as an Asimov data sample. Then the data is fitted by minimizing the combined chi-square in Eq. (5.21). The 1σ uncertainty for Δm^2_{21}, Δm^2_{31}, $\sin^2\theta_{12}$, and $\sin^2\theta_{13}$ is calculated with all rate and b2b systematic uncertainties in four different regimes of data-taking time: 100 days (statistics-dominated regime), 1 year, 6 years (nominal), and 20 years (systematics-dominated regime).

In order to estimate the uncertainties we employ the `MINOS` method provided by `Minuit` [86], which is based on the likelihood profile scan, and estimates the 1σ interval for each parameter, according to the condition:

$$\ln\lambda(\theta) = \frac{\mathcal{L}(\theta,\hat{\hat{\eta}})}{\mathcal{L}(\hat{\theta},\hat{\eta})} \stackrel{!}{=} -\frac{1}{2}. \tag{5.33}$$

Generally speaking, the obtained 1σ interval can be asymmetric. However, in the gaussian regime $-2\ln\lambda(\theta)$ has a parabolic shape and the uncertainty of the parameter is sufficiently symmetric. The parabolic approximation given by `HESSE` [86, 87] was verified to yield the same results[20]. Since this assumption is valid in most cases, Eq. (5.33) can be related to the expression in terms of $\Delta\chi^2$, namely the $N\sigma$ interval is given by

$$N\sigma = \sqrt{\Delta\chi^2(\theta)} = -2\ln\lambda(\theta). \tag{5.34}$$

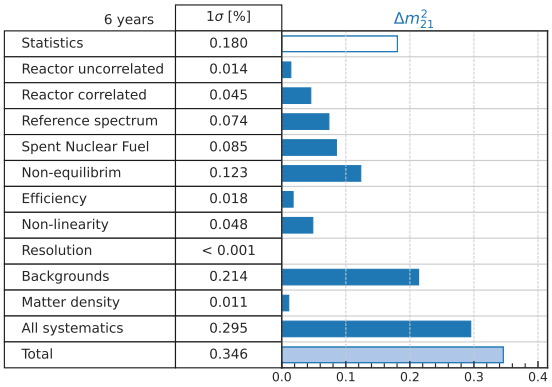
Therefore, the $N\sigma$ interval provides a measure of the deviation from the expected value in terms of the number of standard deviations, based on the change in the chi-square statistic or the likelihood ratio.
To evaluate the impact of statistical contributions, we consider only the statistical covariance matrix V_{stat} in Eq. (5.21). In particular, the cost function has the following expression:

$$\chi^2_{\text{stat}} = (\mathbf{M} - \boldsymbol{\mu}(\boldsymbol{\theta},\hat{\boldsymbol{\eta}}))^T \cdot \left(V^{\text{stat}}_{\text{CNP}}\right)^{-1} \cdot (\mathbf{M} - \boldsymbol{\mu}(\boldsymbol{\theta},\hat{\boldsymbol{\eta}})) \tag{5.35}$$

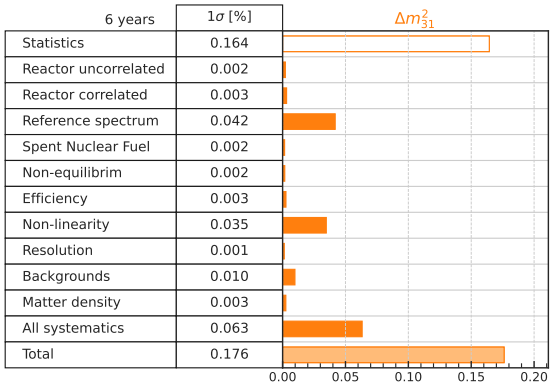
In this assessment, all nuisance parameters are fixed at their nominal values $\hat{\boldsymbol{\eta}}$, while the oscillation parameters are left free to vary. This allows us to isolate and quantify the effect of statistical uncertainties, thus driven by the Poisson statistics of the number of events, without the influence of systematic effects. Then, we consider only one systematic effect at a time: this is done by enabling the corresponding pull term or additional covariance matrix in the expression V_{b2b} of Eq. (5.21). The list of these considered effects, along with the explicit distinction between bin-to-bin correlated and uncorrelated contributions, is provided in Table 5.2.
The breakdown of statistical and systematic uncertainties on each parameter is shown in Figure 5.14 for a nominal exposure of 6 years and 11/12 duty cycle, corresponding to approximately 2010 days of data taking with full reactor power. Additional tables for 100 days, 1 year, and 20 years exposure are reported in Appendix E. The impact of each source of systematic error is assessed by enabling the corresponding effect together with the statistical uncertainty and removing the latter, namely:

$$\sigma_{syst,i} = \sqrt{\sigma^2_{tot,i} - \sigma^2_{stat}}, \tag{5.36}$$

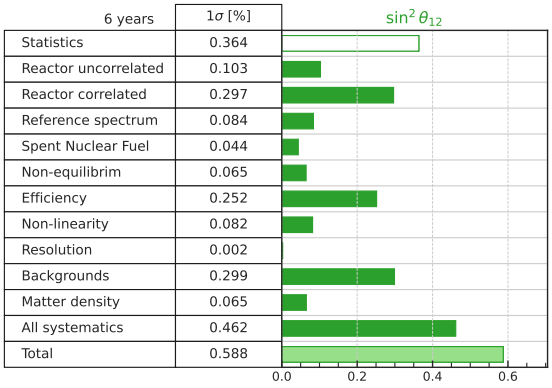
[20] Minor deviations from symmetric uncertainties may be observed for certain parameters when dealing with low statistics, such as an exposure of less than one year. Nevertheless, they are generally found to be fairly negligible.



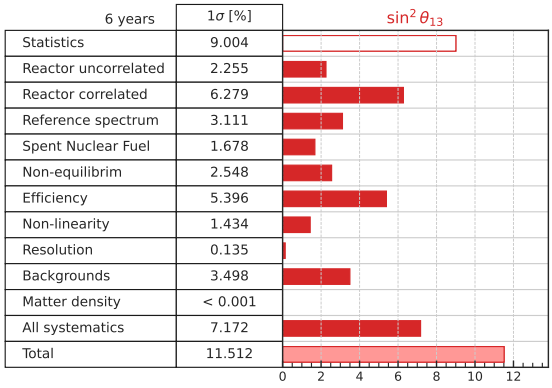
(A) Relative uncertainty contributions to the precision of Δm^2_{21}.



(B) Relative uncertainty contributions to the precision of Δm^2_{31}.



(C) Relative uncertainty contributions to the precision of $\sin^2\theta_{12}$.



(D) Relative uncertainty contributions to the precision of $\sin^2\theta_{13}$.

FIGURE 5.14: Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). The empty boxes represent the statistical uncertainty. The impact of each source of systematic error, represented by the filled boxes, is assessed by enabling the corresponding uncertainty together with the statistical uncertainty, and removing the latter. The last row of each table, displayed in a lighter color, represents the total uncertainty obtained by simultaneously considering all sources of statistical and systematic error. The presented uncertainties correspond to six years of JUNO data with 11/12 duty cycle.

where $\sigma^2_{tot,i}$ indicates the total uncertainty obtained by enabling the i-th systematic contribution. The removal is done by assuming that the statistical and systematic uncertainties add in quadrature, which allows to isolate the systematic components and identify the most significant ones[21].
The "Reactor uncorrelated" entry refers to the bin-to-bin uncorrelated uncertainty arising from the reference spectrum. Since TAO might not be ready to provide JUNO with a reference spectrum in the first months of data-taking, the nominal results in Figure 5.14 are obtained using the Daya Bay data [88]. The entry labeled "All systematics" is obtained using the same methodology but by simultaneously considering all sources of systematic uncertainty rather than individually. The statistics-only sensitivity is also provided in the empty boxes in the first row of the tables in Figure 5.14. Finally, the last row displayed in a lighter color, represents the total uncertainty.

[21]The quadratic combination of statistical and systematic uncertainty assumed in Eq. (5.36) can be used to estimate the impact of single systematics, since the two terms are uncorrelated. Nevertheless, multiple systematic uncertainties cannot be combined with a simple quadratic sum, because of their non-zero correlation.

By examining Figure 5.14c and Figure 5.14d, we can deduce that the precision of the two mixing angles, $\sin^2\theta_{12}$ and $\sin^2\theta_{13}$, is predominantly influenced by rate systematic uncertainties. These uncertainties primarily arise from two sources: the reactor flux normalization (referred to as "Reactor correlated" in the tables) and the detector efficiency, both impacting the expected spectrum in a similar way. Furthermore, despite the reactor correlated uncertainty being twice as large as the efficiency contribution (2% and 1% according to Table 5.2), the relative difference between the effects is only roughly 15%. This observation suggests that the impact on the precision of the oscillation parameters does not scale linearly with these systematic uncertainties. Even with a substantial increase in them, the resulting changes in the precision are not significant. Furthermore, the backgrounds related uncertainties have a non-negligible effect on the estimation of $\sin^2\theta_{12}$ and $\sin^2\theta_{13}$.
On the other hand, when considering the two mass splittings, reported in Figure 5.14a and Figure 5.14b, the systematic uncertainties that have the greatest impact are those that influence the spectral shape. Specifically, these uncertainties are associated with the reference reactor antineutrino spectrum and detector non-linearity. The former uncertainties can be efficiently mitigated by comparing near (unoscillated) and far (oscillated) event spectra [68], and this aspect will be investigated in Section 5.5.1. Moreover, a comprehensive understanding of the detector response is essential. Calibration campaigns [42] are envisaged to characterize the non-linearity features of the detector, aiming to minimize biases in the estimation of Δm^2_{31} and other oscillation parameters. The spent nuclear fuel, non-equilibrium, and background systematic effects also distort the spectral shape, particularly in the low energy region, impacting the precision of Δm^2_{21}.
Figure 5.15 shows the time evolution of JUNO precision sensitivity for all four oscillation parameters, with the dotted and solid lines representing the statistical only and total uncertainties, respectively. The precision of the atmospheric parameters remains statistics-dominated even after 20 years of data-taking, as indicated by the relatively small contribution of the systematic uncertainties. In contrast, the precision of the solar parameters is notably influenced by systematic effects.

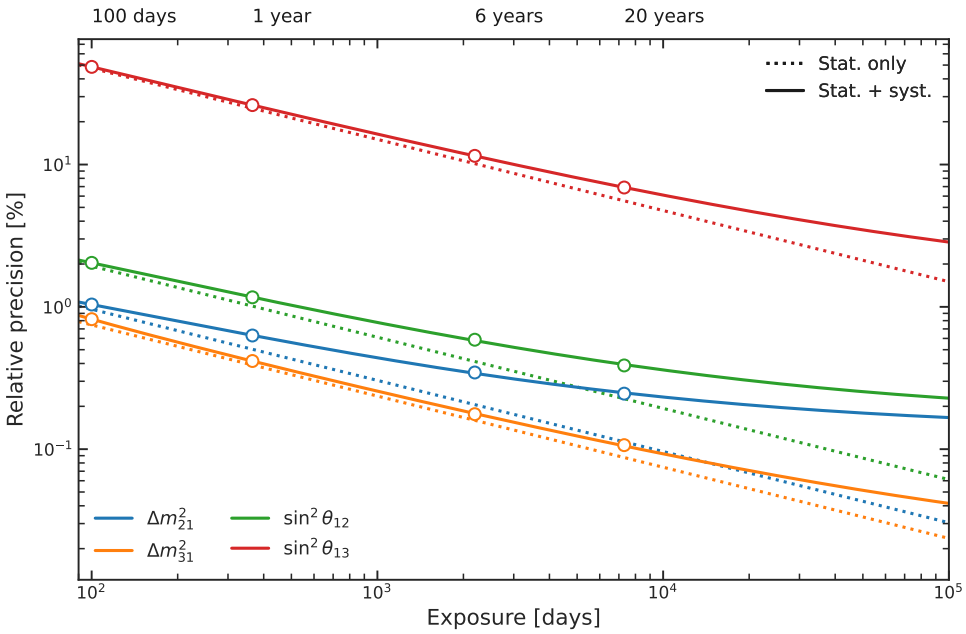


Figure 5.15: Relative precision of the oscillation parameters as a function of JUNO exposure. The dotted and solid lines represent the statistical and total uncertainties, respectively.

The obtained relative precision on the oscillation parameters is reported in Table 5.3, and compared with state-of-the-art knowledge, namely the NuFIT 5.2 2022 analysis [27]. Since the uncertainties on the NuFIT results (in Table 1.2) are asymmetric, the precision in Table 5.3 is calculated considering the higher absolute value and compared with JUNO predictions for different exposures. Consistently with previous studies [55], JUNO is projected to surpass global precision on three parameters within the initial 100 days of data acquisition. Moreover, Figure 5.16 reports the $\Delta\chi^2$ profiles of JUNO (at 100 days and 6 years exposure), compared to combined results from state-of-the-art measurements. JUNO is expected to improve the current precision of almost one order of magnitude in 6 years for Δm^2_{21}, Δm^2_{31}, and $\sin^2\theta_{12}$, while it has a very limited ability to determine $\sin^2\theta_{13}$ beyond today's global precision.

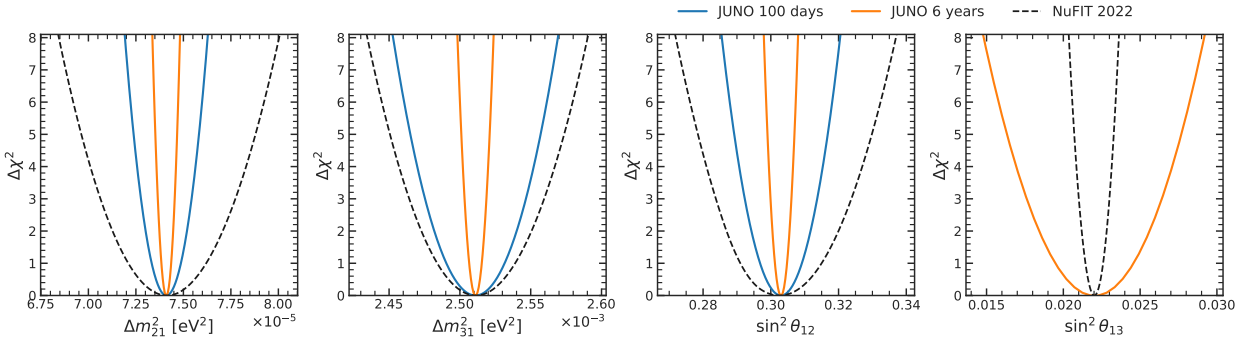


FIGURE 5.16: Comparison of $\Delta\chi^2$ profiles of oscillation parameters: the dashed curve indicates the current results provided by NuFIT 2022 [27], while the solid ones represent JUNO projection with 100 days (in blue) and 6 years (in orange) of data-taking.

	Central value	NuFIT 5.2 [27]	100 days	1 year	6 years	20 years
Δm^2_{21} [eV^2]	7.410×10^{-5}	2.7%	1.0%	0.6%	0.3%	0.2%
Δm^2_{31} [eV^2]	2.511×10^{-3}	1.2%	0.8%	0.4%	0.2%	0.1%
$\sin^2\theta_{12}$	0.303	4.0%	2.0%	1.2%	0.6%	0.4%
$\sin^2\theta_{13}$	0.02203	2.7%	48.6%	26.2%	11.5%	6.9%

TABLE 5.3: Current relative precision on oscillation parameters from global analysis (NuFIT 5.2) [27], and JUNO (different exposures) in this analysis, assuming NO.

Finally, Figure 5.17, depicts the bi-dimensional 1σ, 2σ, and 3σ confidence level (C.L.) regions for all oscillation parameters and 6 years exposure. It shows how these parameters are nearly uncorrelated, highlighting the abundant information available in JUNO's high-resolution measurement of the reactor antineutrino spectrum. Correlations are indeed mitigated by JUNO's rich spectral shape information, since each parameter has a specific effect on the spectral shape, acting with minimal interference with the others.

5.5.1 Impact of the reference spectrum on Δm^2_{31}

The reference spectrum uncertainty is of particular importance as it introduces a distortion in the spectral shape and significantly impacts the precision of the measurement, especially for Δm^2_{31}. Current reactor flux models (e.g., based on conversion and summation methods [89]) exhibit discrepancies when compared to experimental data from the Daya Bay experiment [60], making them

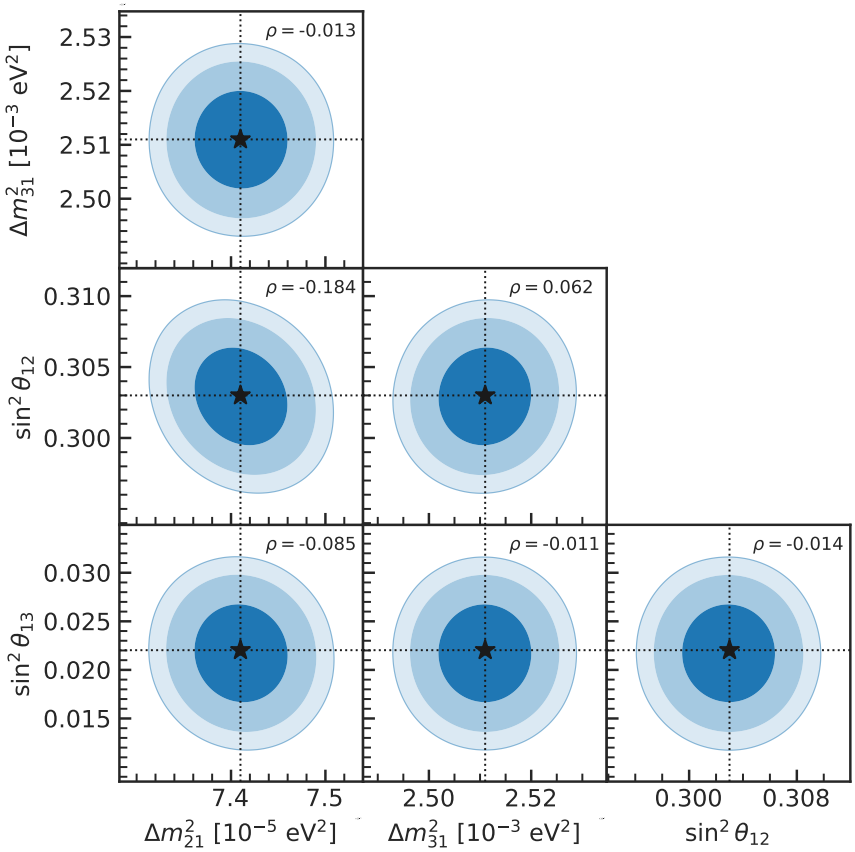
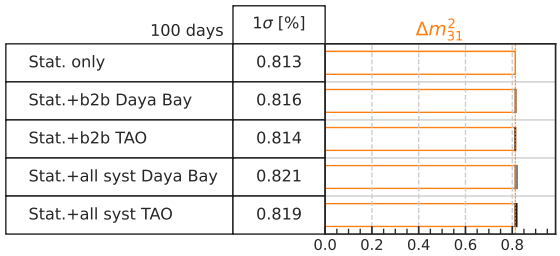


FIGURE 5.17: Bi-dimensional 68.27%, 95.45%, and 99.73% confidence level (C.L.) contours for all oscillation parameters, corresponding to 1σ, 2σ, and 3σ C.L. regions, for 6 years exposure. The black star and the dotted lines represent the best-fit values as well as the input values of the oscillation parameters. The correlation coefficient ρ between pair of parameters is also indicated.

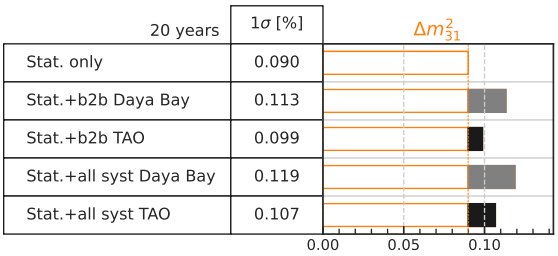
unreliable as a reference spectrum for JUNO. To address this issue, the Taishan Antineutrino Observatory (TAO) [21, 31], a satellite experiment, is expected to monitor the unoscillated spectrum with a sub-percent energy resolution and a 30× gain in event statistics with respect to JUNO oscillated flux [21, 31]. It will play a crucial role in providing JUNO with a model-independent reference spectrum for reactor antineutrinos; this will be possible by unfolding the isotopic spectra, i.e., the extracted prompt energy spectra for specific fission isotopes, of ^{235}U and ^{239}Pu, as done in the Daya Bay experiment [88, 89]. Figure 5.18 illustrates the systematic contribution to the precision of the atmospheric mass splitting, comparing the bin-to-bin uncorrelated uncertainties based on both the Daya Bay and TAO models. The results are shown for 100 days and 20 years exposures: it is evident that in the statistics-dominated regime, the spectrum shape uncertainty has a minor impact on the precision, suggesting that there is no substantial difference between the two models. However, as the exposure increases, the discrepancy becomes progressively more significant, thereby emphasizing the critical role of the future constraint provided by TAO measurements.

5.5.2 Neutrino mass ordering hypothesis impact

The results presented in the previous section were obtained by fitting the Asimov data under the assumption of the correct MO hypothesis. To explore the impact of this choice, we compare four different configurations: fitting NO (Normal Ordering) and IO (Inverted Ordering) data with the



(A) Relative uncertainty contributions to the precision of Δm^2_{31}, for different reference spectrum models and 100 days exposure.



(B) Relative uncertainty contributions to the precision of Δm^2_{31}, for different reference spectrum models and 20 years exposure.

FIGURE 5.18: Relative precision on Δm^2_{31}, for 100 days (a) and 20 years (b) exposures, with 11/12 duty cycle. The bars with different colors correspond to two reference antineutrino spectrum models, Daya Bay (gray) and TAO (black), and corresponding bin-to-bin (b2b) uncertainties.

"right" hypothesis (NO for NO data and IO for IO data), as well as fitting NO and IO data with the "wrong" hypothesis (IO for NO data and NO for IO data).
The corresponding relative uncertainties and best-fit values are provided in Table 5.4. Interestingly, we find that the precision of the solar parameters remains largely unaffected by the mass ordering hypothesis, suggesting its independence from this choice. On the other hand, the atmospheric parameters show a minor dependence on the mass ordering hypothesis. Nevertheless, it is worth noting that the different configurations yield comparable sensitivity.

	Data: NO		**Data: IO**	
Relative precision [%]	Fit: NO	Fit: IO	Fit: NO	Fit: IO
Δm^2_{21}	0.347	0.345	0.347	0.345
Δm^2_{31}	0.185	0.198	0.193	0.179
$\sin^2\theta_{12}$	0.607	0.607	0.613	0.608
$\sin^2\theta_{13}$	12.340	12.799	12.290	11.520
Best fit value	Fit: NO	Fit: IO	Fit: NO	Fit: IO
Δm^2_{21} [eV^2]	7.410×10^{-5}	7.409×10^{-5}	7.411×10^{-5}	7.410×10^{-5}
Δm^2_{31} [eV^2]	2.511×10^{-3}	-2.486×10^{-3}	2.449×10^{-3}	-2.424×10^{-3}
$\sin^2\theta_{12}$	0.303	0.303	0.303	0.303
$\sin^2\theta_{13}$	0.02203	0.02053	0.02140	0.02219

TABLE 5.4: Relative uncertainty and best fit values of the four oscillation parameters after 6 years of data taking at 11/12 duty cycle, for different mass ordering hypotheses. The best fit values for the right mass ordering are given by the input parameters in Table 1.2.

Chapter 6

Conclusions

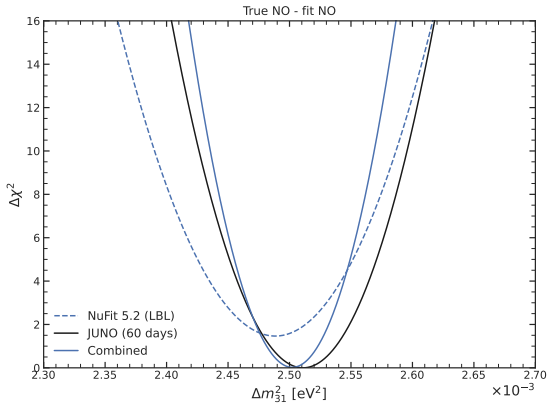
This thesis work provides a thorough investigation of multiple aspects pertaining to reactor antineutrino oscillations. Specifically, I conducted a detailed analysis of their production, interaction, and detection within the JUNO experiment. Furthermore, I examined the event selection process and assessed JUNO's capability to probe oscillation parameters, thereby evaluating its sensitivity in this context. This thesis has provided me with expertise in the entire analysis chain of a neutrino experiment, from handling raw data to deriving physics results, laying a solid foundation for future real-data analysis. Specifically, the starting point was the (simulated) data, with the subsequent event selection process. By combining Monte Carlo simulations and data analysis techniques, I was able to obtain the final outcome, namely the sensitivity to oscillation parameters.

JUNO's outstanding capability to measure the oscillated reactor antineutrino spectrum allows the independent determination of four oscillation parameters: Δm^2_{31}, Δm^2_{21}, $\sin^2\theta_{12}$, and $\sin^2\theta_{13}$. Particular attention was given to Δm^2_{31}, since JUNO measurement is expected to already exceed the current state-of-the-art precision on this parameter in less than 100 days of data-taking. As it was shown in the last section, the sensitivity to Δm^2_{31} relies on complex experimental challenges based on the accurate control of the spectral shape-related systematics arising from energy resolution, energy scale control, and the reactor reference spectrum, to be eventually provided by the TAO experiment. The obtained results are consistently aligned with previous studies [55], demonstrating that the experiment is expected to attain a level of precision surpassing 1% for the determination of the first three parameters. This breakthrough opens up a new era of precise neutrino physics, marking a significant milestone in the field.

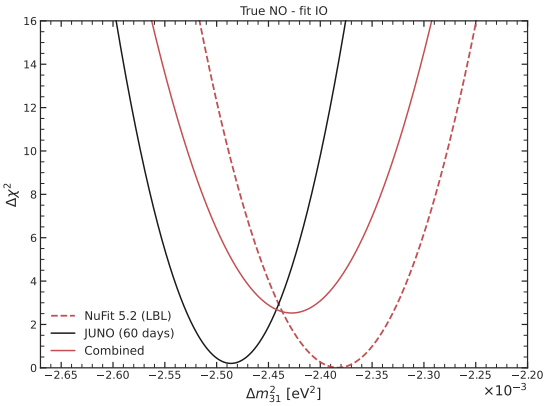
In the current landscape of neutrino experiments, a remarkable synergy exists between reactor and accelerator long baseline (LBL) experiments. Thanks to their complementarity in the determination of the MO [20, 90] and their sensitivity to different detection channels (and thus different parameters), we can expect a bright future of cooperation and interoperability for neutrino physics in the next decade. To date, JUNO stands out as the only experiment sensitive to MO through vacuum-dominated oscillations. Several studies have already explored and demonstrated the synergy between JUNO and LBL and atmospheric experiments, sensitive to MO via matter-dominated effects. Notable examples include the investigation of combined MO sensitivity between JUNO, NOνA, and T2K [20], as well as the study on combined MO sensitivity between JUNO and IceCube-Upgrade/PINGU [24].

The variations observed in the estimation of Δm^2_{31} by different experiments could hint at a favored MO already in the first phase of JUNO data-taking. While JUNO is expected to determine the MO in 6 years, its sub-percent determination of Δm^2_{31} will greatly reduce the allowed parameter space for this parameter, providing a strong handle to test the MO in combined fits [20]. An example of this study is illustrated in Figure 6.1, and details can be found in Appendix F.

In conclusion, the exceptional precision achieved for Δm^2_{31} enables a substantial enhancement in MO sensitivity through combined analyses involving JUNO and LBL experiments [20]. This highlights



(A) $\Delta\chi^2$ profiles for Δm^2_{31} for JUNO (60 days), LBL experiments [27], and combined, i.e., JUNO with external $\Delta m^2_{3\ell}$ information from LBL.



(B) $\Delta\chi^2$ profiles for Δm^2_{31} for JUNO (60 days), LBL experiments [27], and combined, i.e., JUNO with external $\Delta m^2_{3\ell}$ information from LBL.

FIGURE 6.1: 1-dimensional $\Delta\chi^2$ profiles for Δm^2_{31}: (a) for JUNO (60 days) with NO Asimov data fitted with the correct MO hypothesis (NO), LBL with NO hypothesis, and combining the two. (b) for JUNO (60 days) with NO Asimov data fitted with the wrong MO hypothesis (IO), LBL with IO hypothesis, and combining the two. Details on the procedure employed to include the LBL external information on $\Delta m^2_{3\ell}$ (with $\ell = 1$ for NO and $\ell = 2$ for IO hypotheses) in JUNO's analysis is provided in Appendix F.

the pivotal role of JUNO in the MO quest, not only in its unparalleled standalone sensitivity, but also in potential joint investigations.

Bibliography

[1] C. Giunti and C. W. Kim, *Fundamentals of neutrino physics and astrophysics* (Oxford university press, 2007) (cit. on pp. 1, 2, 5, 6, 7, 8, 9, 20).

[2] S. L. Glashow, "Partial-symmetries of weak interactions", Nuclear physics **22**, 579–588 (1961) (cit. on p. 1).

[3] R. L. Workman et al. (Particle Data Group), "Review of Particle Physics", PTEP **2022**, 083C01 (2022) (cit. on pp. 1, 2, 3, 55).

[4] Y. Fukuda et al., "Evidence for oscillation of atmospheric neutrinos", Physical Review Letters **81**, 1562 (1998) (cit. on pp. 2, 9).

[5] B. Pontecorvo, "Neutrino experiments and the problem of conservation of leptonic charge", Sov. Phys. JETP **26**, 165 (1968) (cit. on pp. 2, 3).

[6] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the unified model of elementary particles", Progress of Theoretical Physics **28**, 870–880 (1962) (cit. on p. 2).

[7] Q. R. Ahmad et al., "Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory", Physical Review Letters **89**, 011301 (2002) (cit. on pp. 2, 6, 12).

[8] S. Bilenky, J. Hosek, and S. Petcov, "On oscillations of neutrinos with Dirac and Majorana masses 1980", Physics Letters B **94**, 495 (cit. on p. 3).

[9] N. Agafonova et al., "OPERA tau neutrino charged current interactions", Scientific data **8**, 218 (2021) (cit. on p. 5).

[10] P. Adamson et al., "Measurement of neutrino oscillations with the MINOS detectors in the NuMI beam", Physical Review Letters **101**, 131802 (2008) (cit. on p. 6).

[11] K. Abe et al., "Measurements of neutrino oscillation in appearance and disappearance channels by the T2K experiment with 6.6×10^{20} protons on target", Physical Review D **91**, 072010 (2015) (cit. on pp. 6, 8).

[12] A. Gando et al., "Constraints on θ_{13} from a three-flavor oscillation analysis of reactor antineutrinos at KamLAND", Physical Review D **83**, 052002 (2011) (cit. on pp. 6, 12).

[13] C. Arpesella et al., "Direct Measurement of the ^{7}Be Solar Neutrino Flux with 192 Days of Borexino Data", Physical Review Letters **101**, 091302 (2008) (cit. on pp. 6, 12).

[14] F. An et al., "Spectral measurement of electron antineutrino oscillation amplitude and frequency at Daya Bay", Physical Review Letters **112**, 061801 (2014) (cit. on pp. 6, 9, 21, 30).

[15] G. Bak et al., "Measurement of reactor antineutrino oscillation amplitude and frequency at RENO", Physical Review Letters **121**, 201801 (2018) (cit. on pp. 6, 9, 24).

[16] L. Wolfenstein, "Neutrino oscillations in matter", Physical Review D **17**, 2369 (1978) (cit. on p. 6).

[17] S. Mikheev and A. Y. Smirnov, "Resonance amplification of oscillations in matter and spectroscopy of solar neutrinos", Yadernaya Fizika **42**, 1441–1448 (1985) (cit. on p. 6).

[18] P. F. De Salas, S. Gariazzo, O. Mena, C. A. Ternes, and M. Tórtola, "Neutrino mass ordering from oscillations and beyond: 2018 status and future prospects", Frontiers in Astronomy and Space Sciences **5**, 36 (2018) (cit. on p. 8).

[19] S. Petcov and M. Piai, "The LMA MSW solution of the solar neutrino problem, inverted neutrino mass hierarchy and reactor neutrino experiments", Physics Letters B **533**, 94–106 (2002) (cit. on p. 8).

[20] A. Cabrera et al., "Synergies and prospects for early resolution of the neutrino mass ordering", Scientific Reports **12**, 5393 (2022) (cit. on pp. 8, 75, 95).

[21] A. Abusleme et al., "JUNO physics and detector", Progress in Particle and Nuclear Physics, 103927 (2021) (cit. on pp. 8, 11, 12, 13, 14, 15, 19, 31, 37, 42, 44, 46, 66, 73).

[22] M. Acero et al., "First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NOνA", Physical Review Letters **123**, 151803 (2019) (cit. on p. 8).

[23] R. Acciarri et al., "Long-baseline neutrino facility (LBNF) and deep underground neutrino experiment (DUNE) conceptual design report volume 2: the physics program for DUNE at LBNF", (2015) (cit. on p. 8).

[24] M. Aartsen et al., "Combined sensitivity to the neutrino mass ordering with JUNO, the IceCube Upgrade, and PINGU", Physical Review D **101**, 032006 (2020) (cit. on pp. 8, 75).

[25] Y. Li, Y. Wang, and Z. Xing, "Terrestrial matter effects on reactor antineutrino oscillations at JUNO or RENO-50: how small is small?", Chinese Physics C **40**, 091001 (2016) (cit. on pp. 8, 20, 57).

[26] A. N. Khan, H. Nunokawa, and S. J. Parke, "Why matter effects matter for JUNO", Physics Letters B **803**, 135354 (2020) (cit. on pp. 8, 20, 57, 83, 84).

[27] NuFIT 5.2, *Three-neutrino fit based on data available in November 2022*, (2022) http://www.nu-fit.org/ (cit. on pp. 8, 9, 72, 76, 95).

[28] K. Abe et al., "Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV", Physical Review D **97**, 072001 (2018) (cit. on p. 8).

[29] Y. Abe et al., "Improved measurements of the neutrino mixing angle θ_{13} with the Double Chooz detector", Journal of High Energy Physics **2014**, 1–44 (2014) (cit. on p. 9).

[30] F. An et al., "Neutrino physics with JUNO", Journal of Physics G: Nuclear and Particle Physics **43**, 030401 (2016) (cit. on pp. 11, 13, 14, 19, 21, 23, 44, 45, 46).

[31] A. Abusleme et al., "TAO conceptual design report: a precision measurement of the reactor antineutrino spectrum with sub-percent energy resolution", arXiv preprint arXiv:2005.08745 (2020) (cit. on pp. 11, 66, 73).

[32] A. Abusleme et al., "Radioactivity control strategy for the JUNO detector", Journal of High Energy Physics **2021**, 1–39 (2021) (cit. on pp. 14, 15, 41, 42).

[33] Z. Y. H. Lu and J. Zhao, *^{9}Li/^{8}He background reduction of JUNO*, tech. rep., JUNO-doc-5837 (Institute of High Energy Physics, 2021) (cit. on pp. 14, 38, 44, 45).

[34] A. Abusleme et al., "The JUNO experiment Top Tracker", arXiv preprint arXiv:2303.05172 (2023) (cit. on p. 14).

[35] T. Adam et al., "The OPERA experiment target tracker", Nuclear Instruments and Methods in Physics Research Section A **577**, 523–539 (2007) (cit. on p. 14).

[36] J. Zhao, "Juno status and prospects", 10.5281/zenodo.6683749 (2022) (cit. on p. 14).

[37] F. P. An et al., "Measurement of electron antineutrino oscillation based on 1230 days of operation of the Daya Bay experiment", Physical Review D **95**, 072006 (2017) (cit. on p. 14).

[38] M. Agostini et al., "Spectroscopy of geoneutrinos from 2056 days of Borexino data", Physical Review D **92**, 031101 (2015) (cit. on p. 14).

[39] A. Abusleme et al., "Optimization of the JUNO liquid scintillator composition using a Daya Bay antineutrino detector", Nuclear Instruments and Methods in Physics Research Section A **988**, 164823 (2021) (cit. on p. 15).

[40] P. R. Kampmann, "Energy scale non-linearity and event reconstruction for the neutrino mass ordering measurement of the JUNO experiment", PhD thesis (Universitätsbibliothek der RWTH Aachen, 2020) (cit. on p. 15).

[41] A. Abusleme et al., "The design and sensitivity of JUNO's scintillator radiopurity pre-detector OSIRIS", The European Physical Journal C **81**, 973 (2021) (cit. on p. 15).

[42] A. Abusleme et al., "Calibration strategy of the JUNO experiment", Journal of High Energy Physics **2021**, 1–33 (2021) (cit. on pp. 15, 16, 17, 29, 30, 71).

[43] Y. Han, "Dual Calorimetry for High Precision Neutrino Oscillation Measurement at JUNO Experiment", PhD thesis (Université Paris Cité, 2020) (cit. on p. 16).

[44] Y. Zhang, J. Liu, M. Xiao, F. Zhang, and T. Zhang, "Laser calibration system in JUNO", Journal of Instrumentation **14**, P01009 (2019) (cit. on p. 16).

[45] J. Zou, X. Huang, W. Li, T. Lin, T. Li, K. Zhang, Z. Deng, and G. Cao, "SNiPER: an offline software framework for non-collider physics experiments", in Journal of physics: conference series, Vol. 664, 7 (2015), p. 072053 (cit. on p. 16).

[46] G. Collaboration, S. Agostinelli, et al., "GEANT4–a simulation toolkit", Nuclear Instruments and Methods in Physics Research Section A **506** (2003) (cit. on p. 17).

[47] R. Brun and F. Rademakers, "ROOT-An object oriented data analysis framework", Nuclear Instruments and Methods in Physics Research Section A **389**, 81–86 (1997) (cit. on pp. 17, 35).

[48] A. Abusleme et al., "Mass testing and characterization of 20-inch PMTs for JUNO", The European Physical Journal C **82**, 1168 (2022) (cit. on p. 18).

[49] T. Lin et al., "Simulation software of the JUNO experiment", arXiv preprint arXiv:2212.10741 (2022) (cit. on pp. 18, 35, 38, 41, 51).

[50] G. Cao et al., *CD Electronics Simulation Updates*, tech. rep., JUNO-doc-7211 (2022) (cit. on p. 18).

[51] M. Bellato et al., "Embedded readout electronics R&D for the large PMTs in the JUNO experiment", Nuclear Instruments and Methods in Physics Research Section A **985**, 164600 (2021) (cit. on p. 18).

[52] H. Zhang et al., "Comparison on PMT waveform reconstructions with JUNO prototype", Journal of Instrumentation **14**, T08002 (2019) (cit. on p. 19).

[53] Z. Li et al., "Event vertex and time reconstruction in large-volume liquid scintillator detectors", Nuclear Science and Techniques **32**, 49 (2021) (cit. on p. 19).

[54] G. Huang, W. Jiang, L. Wen, Y. Wang, and W. Luo, "Data-driven simultaneous vertex and energy reconstruction for large liquid scintillator detectors", arXiv preprint arXiv:2211.16768 (2022) (cit. on p. 19).

[55] A. Abusleme et al., "Sub-percent precision measurement of neutrino oscillation parameters with JUNO", Chinese Physics C **46**, 123001 (2022) (cit. on pp. 19, 23, 37, 38, 40, 41, 42, 48, 49, 66, 68, 72, 75).

[56] Y.-F. Li, Z.-z. Xing, and J.-y. Zhu, "Indirect unitarity violation entangled with matter effects in reactor antineutrino oscillations", Physics Letters B **782**, 578–588 (2018) (cit. on p. 19).

[57] S.-F. Ge and W. Rodejohann, "JUNO and neutrinoless double beta decay", Physical Review D **92**, 093006 (2015) (cit. on p. 19).

[58] F. Capozzi, E. Lisi, and A. Marrone, "Neutrino mass hierarchy and electron neutrino oscillation parameters with one hundred thousand reactor events", Physical Review D **89**, 013001 (2014) (cit. on pp. 20, 26, 27, 28, 51, 85).

[59] Y. Huang, V. Chubakov, F. Mantovani, R. L. Rudnick, and W. F. McDonough, "A reference Earth model for the heat-producing elements and associated geoneutrino flux", Geochemistry, Geophysics, Geosystems **14**, 2003–2029 (2013) (cit. on pp. 20, 68).

[60] F. P. An et al., "Improved measurement of the reactor antineutrino flux and spectrum at Daya Bay", Chinese Physics C **41**, 013002 (2017) (cit. on pp. 24, 25, 65, 72).

[61] P. Huber, "Determination of antineutrino spectra from nuclear reactors", Physical Review C **84**, 024617 (2011) (cit. on p. 24).

[62] T. A. Mueller et al., "Improved predictions of reactor antineutrino spectra", Physical Review C **83**, 054615 (2011) (cit. on p. 24).

[63] G. Mention, M. Fechner, T. Lasserre, T. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau, "Reactor antineutrino anomaly", Physical Review D **83**, 073006 (2011) (cit. on p. 24).

[64] F. P. An et al., "Measurement of the reactor antineutrino flux and spectrum at Daya Bay", Physical Review Letters **116**, 061801 (2016) (cit. on pp. 24, 25).

[65] P. Vogel and J. Beacom, "The angular distribution of the reaction $\bar{\nu}_e + p \to e^+ + n$", Physical Review D **60**, 053003 (1999) (cit. on pp. 25, 27, 85).

[66] L. Wei, L. Zhan, J. Cao, and W. Wang, "Improving the energy resolution of the reactor antineutrino energy reconstruction with positron direction", Radiation Detection Technology and Methods **4**, 356–361 (2020) (cit. on p. 26).

[67] A. Strumia and F. Vissani, "Precise quasielastic neutrino/nucleon cross-section", Physics Letters B **564**, 42–54 (2003) (cit. on pp. 27, 85, 86).

[68] F. Capozzi, E. Lisi, and A. Marrone, "Mapping reactor neutrino spectra from TAO to JUNO", Physical Review D **102**, 056001 (2020) (cit. on pp. 27, 28, 71).

[69] J. Birks, *The Theory and Practice of Scintillation Counting, First*, 1964 (cit. on p. 29).

[70] C. Chou, "The nature of the saturation effect of fluorescent scintillators", Physical Review **87**, 904 (1952) (cit. on p. 29).

[71] D. Adey et al., "A high precision calibration of the nonlinear energy response at Daya Bay", Nuclear Instruments and Methods in Physics Research Section A **940**, 230–242 (2019) (cit. on pp. 30, 67).

[72] Y. Wang et al., *JUNO Energy Resolution Prediction Technical Note*, tech. rep., JUNO-doc-8400 (Institute of High Energy Physics, 2023) (cit. on pp. 31, 67).

[73] T. Prohaska et al., "Standard atomic weights of the elements 2021 (IUPAC Technical Report)", Pure and Applied Chemistry **94**, 573–600 (2022) (cit. on p. 31).

[74] L. Zhan et al., *Common inputs for NMO data analysis*, tech. rep., JUNO-doc-5929 (2021) (cit. on pp. 41, 52, 67).

[75] S. Abe et al., "Production of radioactive isotopes through cosmic muon spallation in KamLAND", Physical Review C **81**, 025807 (2010) (cit. on p. 45).

[76] G. Bellini et al., "Cosmogenic backgrounds in Borexino at 3800 m water-equivalent depth", Journal of Cosmology and Astroparticle Physics **2013**, 049 (2013) (cit. on p. 45).

[77] M. Febbraro et al., "New $^{13}C(\alpha, n)^{16}O$ cross section with implications for neutrino mixing and geoneutrino measurements", Physical Review Letters **125**, 062501 (2020) (cit. on p. 49).

[78] M. Agostini et al., "Comprehensive geoneutrino analysis with Borexino", Physical Review D **101**, 012009 (2020) (cit. on p. 49).

[79] J. Cheng, Y. Li, L. Wen, and S. Zhou, "Neutral-current background induced by atmospheric neutrinos at large liquid-scintillator detectors. I. Model predictions", Physical Review D **103**, 053001 (2021) (cit. on p. 49).

[80] A. Gando et al., "Search for extraterrestrial antineutrino sources with the KamLAND detector", The Astrophysical Journal **745**, 193 (2012) (cit. on p. 49).

[81] Y. L. R. Han and L. Zhan, *World reactors' info and contribution to JUNO*, tech. rep., JUNO-doc-7177 (Institute of High Energy Physics, 2021) (cit. on p. 49).

[82] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", The European Physical Journal C **71**, 1–19 (2011) (cit. on pp. 52, 56, 57, 58).

[83] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures.", Analytical chemistry **36**, 1627–1639 (1964) (cit. on p. 52).

[84] P. Virtanen et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python", Nature methods **17**, 261–272 (2020) (cit. on p. 52).

[85] X. Ji, W. Gu, X. Qian, H. Wei, and C. Zhang, "Combined Neyman–Pearson chi-square: an improved approximation to the Poisson-likelihood chi-square", Nuclear Instruments and Methods in Physics Research Section A **961**, 163677 (2020) (cit. on pp. 54, 55, 65).

[86] F. James and M. Roos, "MINUIT: a system for function and analysis of the parameter errors and corrections", Comput. Phys. Commun. **10**, 343–367 (1975) (cit. on pp. 57, 69).

[87] H. Dembinski and P. O. et al., "Scikit-hep/iminuit", 10.5281/zenodo.3949207 (2020) (cit. on pp. 57, 69).

[88] D. Adey et al., "Extraction of the ^{235}U and ^{239}Pu antineutrino spectra at Daya Bay", Physical Review Letters **123**, 111801 (2019) (cit. on pp. 66, 70, 73).

[89] F. An et al., "Antineutrino energy spectrum unfolding based on the Daya Bay measurement and its applications", Chinese Physics C **45**, 073001 (2021) (cit. on pp. 72, 73).

[90] H. Nunokawa, S. Parke, and R. Z. Funchal, "Another possible way to determine the neutrino mass hierarchy", Physical Review D **72**, 013009 (2005) (cit. on p. 75).

[91] P. B. Denton, H. Minakata, and S. J. Parke, "Compact perturbative expressions for neutrino oscillations in matter", Journal of High Energy Physics **2016**, 1–23 (2016) (cit. on p. 83).

[92] Y.-F. Li, J. Cao, Y. Wang, and L. Zhan, "Unambiguous determination of the neutrino mass hierarchy using reactor neutrinos", Physical Review D **88**, 013008 (2013) (cit. on p. 95).

Appendix A

Electron antineutrino survival probability in matter

The explicit expression of the effective matter-induced mixing angles $\widetilde{\theta}_{ij}, (i,j=1,2,3,,i<j)$, and mass squared differences $\Delta\tilde{m}^2_{ij}, (i,j=1,2,3,,i>j)$, used in the survival probability in Eq. (2.4) are taken from Ref. [26]. Firstly, we recall the A parameter in Eq. (2.3),i.e.,

$$A = \pm\, 2\sqrt{2}G_{\rm F}N_eE \simeq \pm\, 1.52\times10^{-4}{\rm eV}^2\cdot Y_e\cdot\frac{\rho}{{\rm g/cm}^3}\cdot\frac{E}{\rm GeV}, \tag{A.1}$$

where the $+(-)$ sign refers to the neutrino (antineutrino) channel. We then introduce an effective mass difference:

$$\begin{aligned}\Delta m^2_{ee} &\equiv \cos^2\theta_{12}\Delta m^2_{31}+\sin^2\theta_{12}\Delta m^2_{32}\\ &= \cos^2\theta_{12}\Delta m^2_{31}+\sin^2\theta_{12}(\Delta m^2_{31}-\Delta m^2_{21})\\ &= \Delta m^2_{31}-\sin^2\theta_{12}\Delta m^2_{21}\end{aligned} \tag{A.2}$$

From the approximate parameterization derived by Denton, Minakata and Parke in [91], $\widetilde{\theta}_{13}$ has the following expression:

$$\cos 2\widetilde{\theta}_{13} \simeq \frac{\Delta m^2_{ee}\cos 2\theta_{13}-A}{\Delta\widetilde{m}^2_{ee}}, \tag{A.3}$$

where

$$\Delta\widetilde{m}^2_{ee} \equiv \Delta m^2_{ee}\sqrt{\left(\cos 2\theta_{13}-\frac{A}{\Delta m^2_{ee}}\right)^2+\sin^2 2\theta_{13}}. \tag{A.4}$$

Then $\widetilde{\theta}_{12}$ is such that

$$\cos 2\widetilde{\theta}_{12} \simeq \frac{\Delta m^2_{21}\cos 2\theta_{12}-A'}{\Delta\widetilde{m}^2_{21}}, \tag{A.5}$$

where

$$\Delta\tilde{m}^2_{21} \simeq \Delta m^2_{21}\cdot\sqrt{\left(\cos 2\theta_{12}-\frac{A'}{\Delta m^2_{21}}\right)^2+\cos^2\left(\widetilde{\theta}_{13}-\theta_{13}\right)\sin^2 2\theta_{12}}. \tag{A.6}$$

An effective matter potential for solar (1-2) sector, dubbed A', can be defined as:

$$A' \equiv \frac{A+\Delta m^2_{ee}-\Delta\tilde{m}^2_{ee}}{2}, \tag{A.7}$$

as well as

$$\cos^2\left(\tilde{\theta}_{13}-\theta_{13}\right) = \frac{\Delta\tilde{m}^2_{ee}+\Delta m^2_{ee}-A\cos 2\theta_{13}}{2\Delta\tilde{m}^2_{ee}} \tag{A.8}$$

and finally $\Delta\tilde{m}^2_{31}$ and $\Delta\widetilde{m}^2_{32}$;

$$\Delta\tilde{m}^2_{31} = \Delta\tilde{m}^2_{ee} + \sin^2\tilde{\theta}_{12}\Delta\tilde{m}^2_{21}, \tag{A.9}$$

$$\Delta\tilde{m}^2_{32} = \Delta\tilde{m}^2_{31} - \Delta\tilde{m}^2_{21}. \tag{A.10}$$

In [26], they performed a Taylor series expansion in the parameters $A/\Delta m^2_{ee}$ and $A/\Delta m^2_{21}$ on equations (A.3)–(A.6), (A.9) and (A.10) to calculate all the oscillation parameters up to the required accuracy level for JUNO. The following expression are obtained and used in the calculation of the $\overline{\nu}_e$ survival probability in Eq. (2.4):

$$\sin^2\widetilde{\theta}_{12} \simeq s^2_{12}\cdot\left[1 + 2c^2_{12}\left(\frac{c^2_{13}A}{\Delta m^2_{21}}\right) + 3c^2_{12}\cos 2\theta_{12}\left(\frac{c^2_{13}A}{\Delta m^2_{21}}\right)^2\right], \tag{A.11}$$

$$\Delta\tilde{m}^2_{21} \simeq \Delta m^2_{21}\cdot\left[1 - \cos 2\theta_{12}\left(\frac{c^2_{13}A}{\Delta m^2_{21}}\right) + 2s^2_{12}c^2_{12}\left(\frac{c^2_{13}A}{\Delta m^2_{21}}\right)^2\right], \tag{A.12}$$

$$\sin^2\widetilde{\theta}_{13} \simeq s^2_{13}\cdot\left[1 + 2c^2_{13}\left(\frac{A}{\Delta m^2_{ee}}\right)\right], \tag{A.13}$$

$$\Delta\tilde{m}^2_{31} \simeq \Delta m^2_{31}\cdot\left[1 - \left(\frac{A}{\Delta m^2_{31}}\right)\left(\left(c^2_{12}c^2_{13} - s^2_{13}\right) - s^2_{12}c^2_{12}c^2_{13}\left(\frac{c^2_{13}A}{\Delta m^2_{21}}\right)\right)\right], \tag{A.14}$$

$$\Delta\tilde{m}^2_{32} \simeq \Delta m^2_{32}\cdot\left[1 - \left(\frac{A}{\Delta m^2_{32}}\right)\left(\left(s^2_{12}c^2_{13} - s^2_{13}\right) + s^2_{12}c^2_{12}c^2_{13}\left(\frac{c^2_{13}A}{\Delta m^2_{21}}\right)\right)\right], \tag{A.15}$$

$$\Delta\tilde{m}^2_{ee} \simeq \Delta m^2_{ee}\cdot\left[1 - \cos 2\theta_{13}\left(\frac{A}{\Delta m^2_{ee}}\right)\right]. \tag{A.16}$$

In [26] it is shown using this approximate matter oscillation probability has a fractional precision better than 10^{-6} for antineutrino energies below 10 MeV.

Appendix B

Detailed IBD kinematics

The IBD reaction has a threshold depending on the nucleus the target proton is bound to. The lighter the target nucleus is, the lower the IBD threshold is. Considering an IBD reaction of a $\overline{\nu}_e$, whose mass can be neglected, on a hydrogen nucleus (i.e., a free proton) in the center of momentum frame, the threshold is defined by the positron and neutron being produced at rest [65]

$$E_\nu^{\text{thr,CM}} = \frac{(m_n + m_e)^2 - m_p^2}{2\,(m_n + m_e)} = 1.803\,\text{MeV} \tag{B.1}$$

with m_p, m_n and m_e respectively the proton, neutron and electron masses. In the laboratory frame, where the proton is at rest

$$E_\nu^{\text{thr,lab}} = \frac{(m_n + m_e)^2 - m_p^2}{2m_p} = 1.806\,\text{MeV} \tag{B.2}$$

Above this threshold, the relation between the neutrino and positron energies at zeroth order is given by [58]:

$$E_{e^+}^{(0)} = E_{\bar{\nu}_e} - \Delta_{np}, \tag{B.3}$$

with $\Delta_{np} = m_p - m_n$. However, this rough approximation leads to a non-negligible bias in the cross section estimation, especially at energies above 10 MeV.
The recoil of the neutron induces an angle-dependent deficit in the positron energy E_{e^+}, which falls inside a kinematical range, namely

$$E_{e^+} \in [E_1, E_2]. \tag{B.4}$$

Labeling the 4-momenta as

$$\overline{\nu}_e(p_\nu) + p(p_p) \to e^+(p_e) + n(p_n),$$

we can define the Mandelstam variables in the proton rest frame:

$$\begin{aligned} s &= (p_\nu + p_p)^2 = m_p^2 + 2m_p E_\nu \\ t &= (p_\nu - p_e)^2 = m_n^2 - m_p^2 - 2m_p\,(E_\nu - E_e) \\ u &= (p_\nu - p_n)^2 = m_p^2 + m_e^2 - 2m_p E_e, \end{aligned} \tag{B.5}$$

The allowed values of E_{e^+} in Eq. (B.4), correspond to the possible scattering angles θ_{CM} in the center of mass (CM) frame [67]:

$$E_{1,2} = E_\nu - \delta - \frac{1}{m_p} E_\nu^{\text{CM}} \left(E_e^{\text{CM}} \pm p_e^{\text{CM}}\right), \qquad \text{with} \qquad \delta \equiv \frac{m_n^2 - m_p^2 - m_e^2}{2m_p} \tag{B.6}$$

where the energy and momenta in the CM have the following expressions:

$$E_\nu^{\mathrm{CM}} = \frac{s - m_p^2}{2\sqrt{s}}, \quad E_e^{\mathrm{CM}} = \frac{s - m_n^2 + m_e^2}{2\sqrt{s}}, \quad p_e^{\mathrm{CM}} = \frac{\sqrt{\left[s - (m_n - m_e)^2\right]\left[s - (m_n + m_e)^2\right]}}{2\sqrt{s}} \tag{B.7}$$

As mentioned in Section 3.2, in the "mid-recoil" approximation, the average positron energy is given by:

$$\langle E_e \rangle \equiv E_{e^+}^{\mathrm{mid}} = \frac{E_1 + E_2}{2} = E_\nu - \delta - \frac{E_\nu^{\mathrm{CM}} E_e^{\mathrm{CM}}}{m_p}. \tag{B.8}$$

This yields a much better approximation than Eq. (B.3), incorporating a large part of the effect due to the recoil of the nucleon.

B.1 Cross section

The differential cross section at tree level in the weak interactions, averaged (summed) over initial (final) polarizations is given by [67]:

$$\frac{d\sigma}{dt} = \frac{G_F^2 \cos^2\theta_C}{2\pi \left(s - m_p^2\right)^2} |\mathcal{M}|^2, \tag{B.9}$$

where G_F is the Fermi constant, $\cos\theta_C$ is the Cabbibo angle, $\mathcal{M}$ is the invariant amplitude for the IBD reaction. The complete expression of $|\mathcal{M}|^2$ can be found in Ref. [67].
The differential cross section in Eq. (B.9), can be expressed as a function of the neutrino and positron energies:

$$\frac{d\sigma}{dE_e}(E_\nu, E_e) = 2m_p \frac{d\sigma}{dt}. \tag{B.10}$$

The total cross section, shown in Figure 3.3a, can be expressed including the following one-loop radiative corrections [67]:

$$d\sigma(E_\nu, E_e) \to d\sigma(E_\nu, E_e)\left[1 + \frac{\alpha}{\pi}\left(6 + \frac{3}{2}\log\frac{m_p}{2E_e} + 1.2\left(\frac{m_e}{E_e}\right)^{1.5}\right)\right], \tag{B.11}$$

where α is the fine-structure constant. This correction is valid at $E_\nu \ll m_p$. The total cross section is still given by Eq. (B.11), if all final state leptons are detected.
Moreover, the authors of Ref. [67] computed an approximation of the total cross section, given by:

$$\sigma(\bar{\nu}_e, p) \simeq 10^{-43} p_e E_e E_\nu^{-0.07056 + 0.02018 \ln E_\nu - 0.001953 \ln^3 E_\nu} \mathrm{cm}^2, \tag{B.12}$$

which is proven to be valid for positron energies given by Eq. (B.3), and is in agreement within a few per-mille for $E_\nu \leq 300$ MeV with the complete expression.

Appendix C

Bayesian inference

The Bayes' theorem is a fundamental principle in statistics that relates conditional probabilities. It provides a framework for updating our beliefs about a set of parameters, denoted as $\boldsymbol{\theta}$, given observed data and a selected model. The theorem can be stated as:

$$P(\boldsymbol{\theta} \mid data, model) = \frac{P(data \mid \boldsymbol{\theta}, model) \cdot P(\boldsymbol{\theta} \mid model)}{P(data \mid model)} \tag{C.1}$$

The probability distribution $P(\boldsymbol{\theta} \mid data, model)$ is commonly referred to as the *posterior*, which expresses the probability to obtain a particular set of parameters $\boldsymbol{\theta}$ given the observed data and the selected model. The distribution $P(data \mid \boldsymbol{\theta}, model)$ is known as the *likelihood*, which quantifies the degree of agreement between the model and the observed data. The *prior*, denoted as $P(\boldsymbol{\theta} \mid model)$, encapsulates any external constraints, prior knowledge, or limitations imposed on the fitting of the data. Finally, the *evidence* distribution, denoted as $P(data \mid model)$, serves as a normalizing factor and captures the overall probability of observing the data under the chosen model.
The probability $P(data \mid \theta)$ can be interpreted as a probability density function (PDF) when viewed as a function of data with fixed θ. On the other hand, when viewed as a function of θ with fixed data, it becomes a likelihood function. To simplify our calculations, we will work with a fixed model and drop the model dependence. Additionally, we will refer to the observed number of events in each bin as *data*, denoted by $\mathbf{M}$ as in Eq. (5.1).
Bayesian parameter estimation involves using Bayes' theorem to update our knowledge of the model parameters $\boldsymbol{\theta}$ given the observed data $\mathbf{M}$. The posterior distribution $P(\boldsymbol{\theta} \mid \mathbf{M})$ is proportional to the product of the likelihood function $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M})$ and the prior distribution $P(\boldsymbol{\theta})$:

$$P(\boldsymbol{\theta} \mid \mathbf{M}) \propto \mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M})\mathbf{P}(\boldsymbol{\theta}). \tag{C.2}$$

Here, the likelihood function $\mathcal{L}(\boldsymbol{\theta}|\mathbf{M})$ is the probability of observing the data $\mathbf{M}$ given the model parameters $\boldsymbol{\theta}$, and the prior distribution $P(\boldsymbol{\theta})$ expresses our knowledge or beliefs about the model parameters before observing the data. The mode of the posterior distribution corresponds to the maximum a posteriori (MAP) estimate of the parameters, which can be obtained by maximizing the logarithm of the posterior distribution, i.e., finding the values of $\boldsymbol{\theta}$ that maximize

$$\ln P(\boldsymbol{\theta} \mid \mathbf{M}) = \ln \mathcal{L}(\boldsymbol{\theta} \mid \mathbf{M}) + \ln \mathbf{P}(\boldsymbol{\theta}) + \text{const.} \tag{C.3}$$

If we assume a flat prior distribution (i.e., $P(\boldsymbol{\theta})$ is constant), then the mode of the posterior distribution coincides with the ML estimate. This means that maximizing the likelihood function is equivalent to finding the mode of the posterior distribution under a flat prior.

Appendix D

Likelihood profile scans results

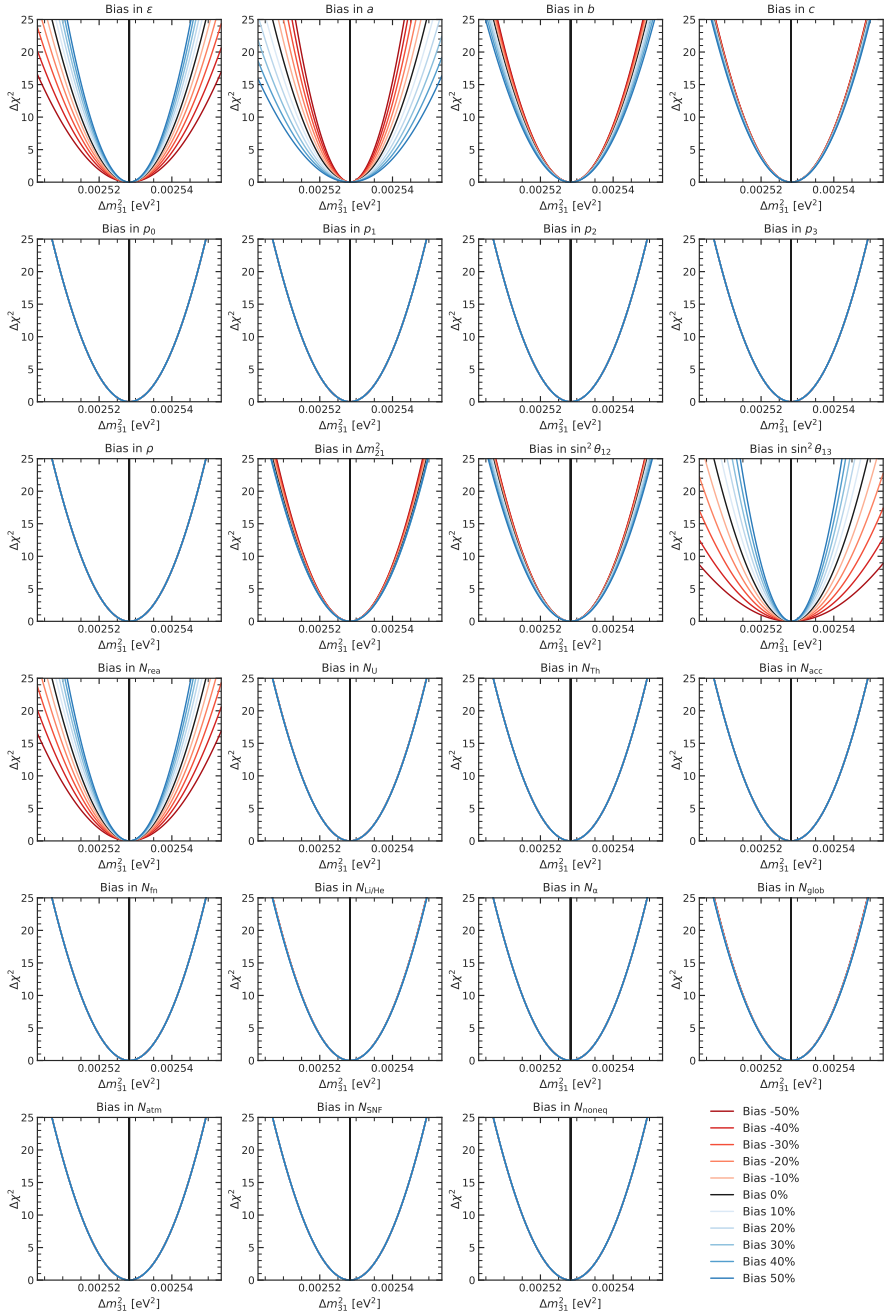


FIGURE D.1: Conditional likelihood scans for $\theta = \Delta m^2_{31}$ and all nuisance parameters. The profiles are obtained using Eq. (5.14).

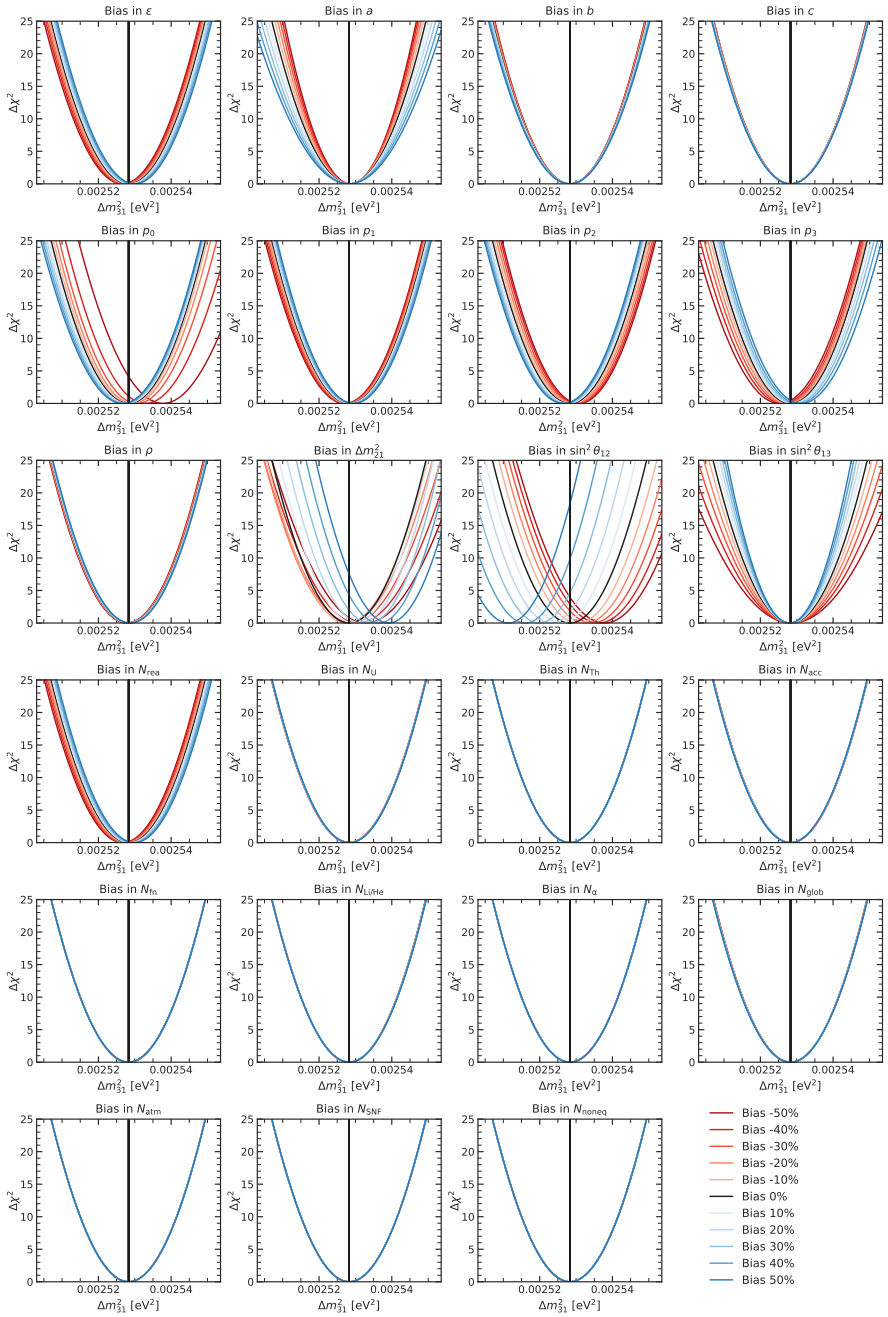
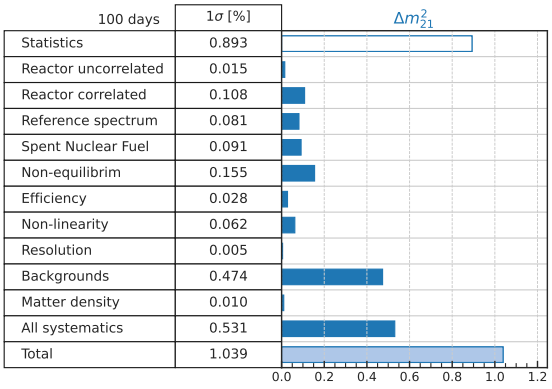


FIGURE D.2: Conditional likelihood scans for $\theta = \Delta m^2_{31}$ and all nuisance parameters. The profiles are obtained using Eq. (5.15).

Appendix E

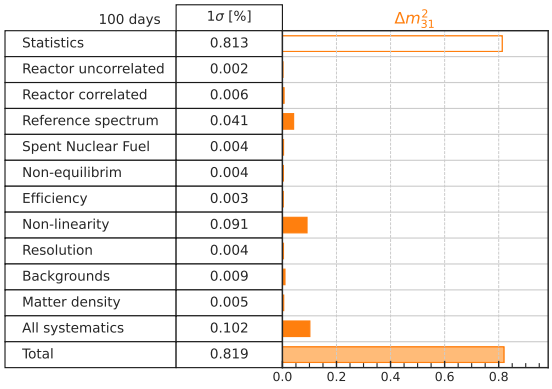
Sensitivity results

100 days	1σ [%]
Statistics	0.893
Reactor uncorrelated	0.015
Reactor correlated	0.108
Reference spectrum	0.081
Spent Nuclear Fuel	0.091
Non-equilibrim	0.155
Efficiency	0.028
Non-linearity	0.062
Resolution	0.005
Backgrounds	0.474
Matter density	0.010
All systematics	0.531
Total	1.039



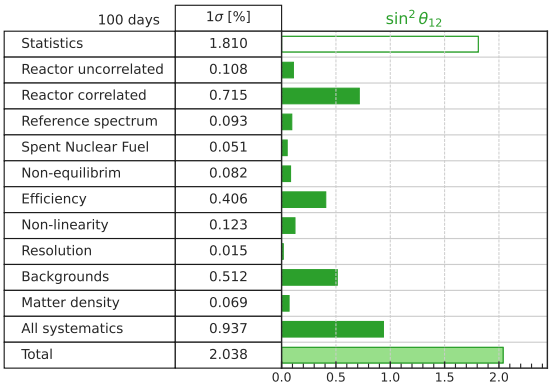
(A) Relative uncertainty contributions to the precision of Δm^2_{21}.

100 days	1σ [%]
Statistics	0.813
Reactor uncorrelated	0.002
Reactor correlated	0.006
Reference spectrum	0.041
Spent Nuclear Fuel	0.004
Non-equilibrim	0.004
Efficiency	0.003
Non-linearity	0.091
Resolution	0.004
Backgrounds	0.009
Matter density	0.005
All systematics	0.102
Total	0.819



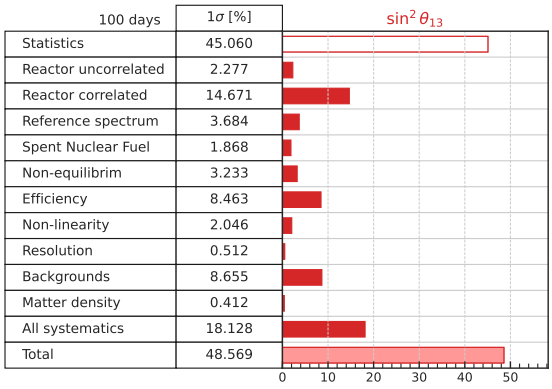
(B) Relative uncertainty contributions to the precision of Δm^2_{31}.

100 days	1σ [%]
Statistics	1.810
Reactor uncorrelated	0.108
Reactor correlated	0.715
Reference spectrum	0.093
Spent Nuclear Fuel	0.051
Non-equilibrim	0.082
Efficiency	0.406
Non-linearity	0.123
Resolution	0.015
Backgrounds	0.512
Matter density	0.069
All systematics	0.937
Total	2.038



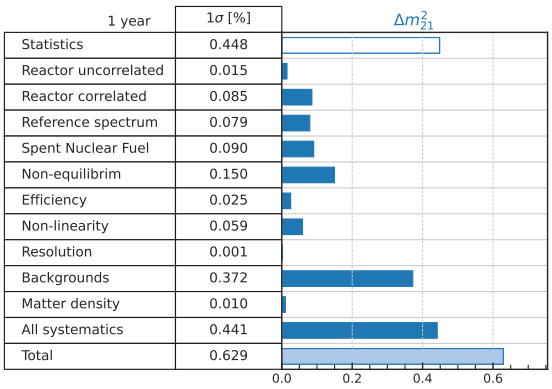
(C) Relative uncertainty contributions to the precision of $\sin^2\theta_{12}$.

100 days	1σ [%]
Statistics	45.060
Reactor uncorrelated	2.277
Reactor correlated	14.671
Reference spectrum	3.684
Spent Nuclear Fuel	1.868
Non-equilibrim	3.233
Efficiency	8.463
Non-linearity	2.046
Resolution	0.512
Backgrounds	8.655
Matter density	0.412
All systematics	18.128
Total	48.569

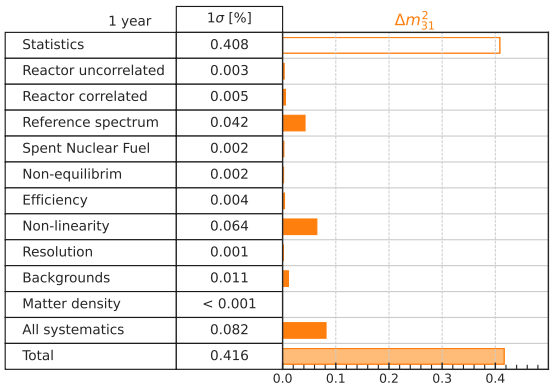


(D) Relative uncertainty contributions to the precision of $\sin^2\theta_{13}$.

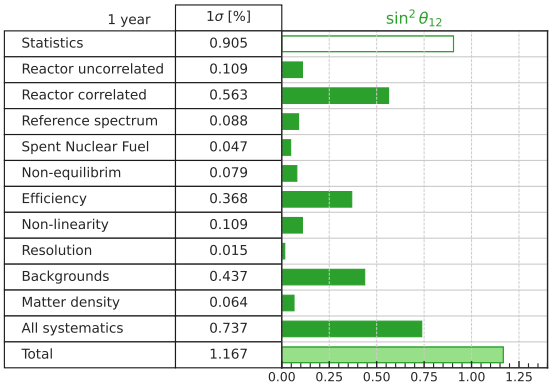
FIGURE E.1: Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). All uncertainties correspond to 100 days of JUNO data with 11/12 duty cycle.



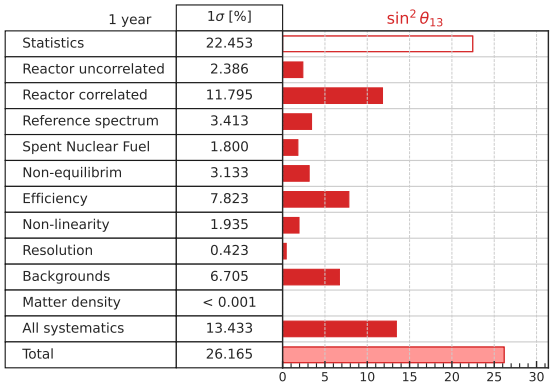
(A) Relative uncertainty contributions to the precision of Δm^2_{21}.



(B) Relative uncertainty contributions to the precision of Δm^2_{31}.

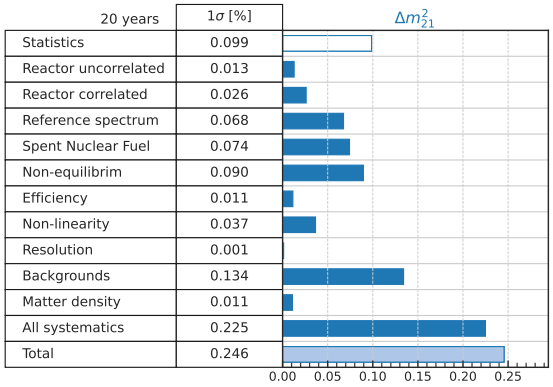


(C) Relative uncertainty contributions to the precision of $\sin^2\theta_{12}$.

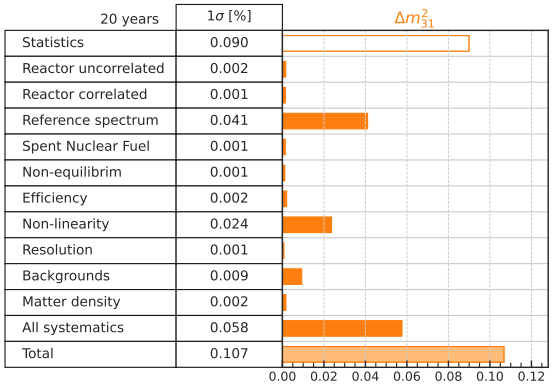


(D) Relative uncertainty contributions to the precision of $\sin^2\theta_{13}$.

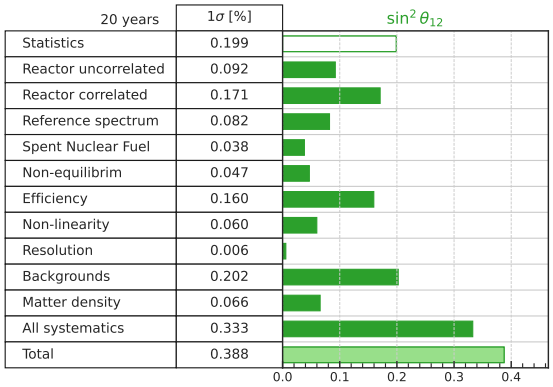
FIGURE E.2: Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). All uncertainties correspond to 1 year of JUNO data with 11/12 duty cycle.



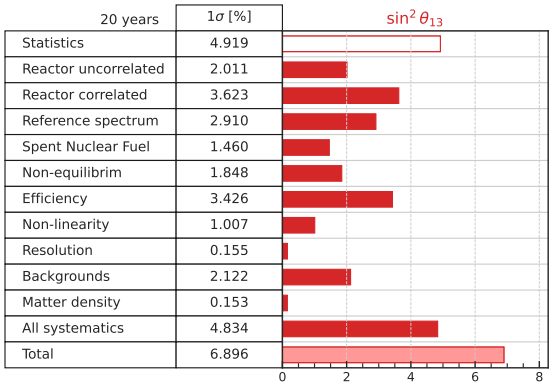
(A) Relative uncertainty contributions to the precision of Δm^2_{21}.



(B) Relative uncertainty contributions to the precision of Δm^2_{31}.



(C) Relative uncertainty contributions to the precision of $\sin^2\theta_{12}$.



(D) Relative uncertainty contributions to the precision of $\sin^2\theta_{13}$.

FIGURE E.3: Relative impact of individual sources of uncertainty on the total precision of Δm^2_{21} (a), Δm^2_{31} (b), $\sin^2\theta_{12}$ (c), and $\sin^2\theta_{13}$ (d). All uncertainties correspond to 20 years of JUNO data with 11/12 duty cycle.

Appendix F

Neutrino MO synergy between JUNO and LBL experiments

The synergy for MO determination sensitivity, achieved by combining JUNO and LBL experiments, arises from the observed differences in the fitted values of $\Delta m^2_{3\ell}$ corresponding to incorrect MO solutions [20]. Indeed, differences in the estimation of $\Delta m^2_{3\ell}$ obtained by various experiments could potentially indicate a favored MO even during the initial phase of JUNO's data taking. While JUNO is anticipated to determine the MO within six years, its highly precise determination of Δm^2_{31} at a sub-percent level will significantly narrow down the parameter space, offering a robust means to test the MO through combined analyses [20].

Following the procedure presented in [20, 92], we can include in JUNO's analysis the external information on $\Delta m^2_{3\ell}$, provided by the combined fit of LBL experiments (NuFIT 5.2 [27]), with an additional pull term in the chi-square:

$$\chi^2_{\text{comb}} = \chi^2_{\text{JUNO}} + \left(\frac{\Delta m^2_{3\ell} - \Delta m^2_{3\ell,\text{LBL}}}{\sigma(\Delta m^2_{3\ell,\text{LBL}})} \right)^2 , \tag{F.1}$$

where χ^2_{JUNO} is the final objective function in Eq. (5.21), $\Delta m^2_{3\ell,\text{LBL}}$ is the measurement obtained by LBL experiments (with $\ell = 1$ for NO and $\ell = 2$ for IO hypotheses), and $\sigma(\Delta m^2_{3\ell,\text{LBL}})$ is the corresponding uncertainty. The results of this procedure are reported in Figure 6.1, where an exposure of 60 days is considered (such that the precision on $\Delta m^2_{3\ell}$ is of the order of 1%). Under the assumption of true NO data, the discriminator of the neutrino MO can be defined as

$$\Delta\chi^2_{\text{MO}} = |\chi^2_{\text{min}}(\text{NO}) - \chi^2_{\text{min}}(\text{IO})|. \tag{F.2}$$

Note that in our this approach based on Asimov datasets, by construction, $\chi^2_{\text{min}}(\text{NO}) = 0$ for NO data. Figure 6.1b shows the 1-dimensional $\Delta\chi^2$ profiles for Δm^2_{31} for the global LBL results (assuming IO hypothesis) and fitting JUNO Asimov NO data with the IO hypothesis, and the combined profile obtained by introducing the external Δm^2_{32} information in JUNO chi-square function. The combination of JUNO with LBL experiments results in an enhanced intrinsic sensitivity for JUNO, as illustrated in Figure 6.1b. This enhancement is evident from the increase in $\Delta\chi^2_{\text{MO}}$, which rises from approximately 0.2 to around 2.5.