Analysis of Neutrino Oscillations at Current and Future Accelerator Neutrino Experiments

Henry Wallace

Department of Physics and Astrophysics, University of Sheffield

Supervisor: Dr Susan Cartwright

A Dissertation Submitted to the University of Sheffield for the Degree of

Doctor of Philosophy

October 17, 2024

Abstract

The T2K and Hyper-Kamiokande experiments are current and future long baseline off-axis neutrino oscillation experiments located in Japan. They monitor a primarily $\overset{(-)}{\nu}_{\mu}$ at two points, one 280 m from the beam source and the other 295 km from the source allowing for the measurement of the flavour content of the neutrino beam over a long distance. The Bayesian results from the T2K experiment using the data set from runs 1–11 are presented in this thesis. In addition this thesis presents future sensitivities studies from the Hyper-Kamiokande experiment and improvements to the Bayesian fitter.

When fitting with only T2K data for run 1–11 the best-fit oscillation parameters are $\sin^2\theta_{23} = 0.48^{+0.014}_{-0.077}$, $\Delta m^2_{32} = 2.51 \times 10^{-3}$ eV^2 $[2.44, 2.57] \cup [-2.59, -2.51]$, $\sin^2\theta_{13} = 0.024^{+0.006}_{-0.003}$, $\delta_{CP} = -1.84^{+1.15}_{-0.92}$. The 95.45% credible interval for δ_{CP} excludes $[0.25, 2.51]$ and the 90% credible interval excludes $[-0.13, 2.89]$.

When fitting with T2K data with the PDG 2020 reactor constraint [1] on $\sin^2\theta_{13}$ for run 1–11 the best-fit oscillation parameters are $\sin^2\theta_{23} = 0.552^{+0.022}_{-0.053}$, $\Delta m^2_{32} = 2.51 \times 10^{-3}$ eV^2 $[2.42, 2.58] \cup [-2.56, -2.54]$, $\sin^2\theta_{13} = 0.0220 \pm 0.0007$, $\delta_{CP} = -1.59^{+0.65}_{-0.98}$. The 95.45% credible interval for δ_{CP} excludes $[-0.25, 3.02]$ and the 90% credible interval excludes $[-3.14, -3.02] \cup [-0.44, 3.14]$.

Acknowledgements

My thanks first go to my supervisor Susan Cartwright, who's seemingly endless knowledge and support has made thesis both possible and removed my constant comma, splices. I would also like to thank the entire MaCh3 group particularly Patrick Dunne, Ed Atkin and Kamil Skwarczyński and Dan Barrow whose constant patience with my constant "wacky" new MCMC algorithm suggestions and support throughout the PhD have spurred a fascination with all things Bayesian.

Starting this PhD was made substantially easier by the Sheffield group's constant support and regular games of Among Us and EniGame puzzle solving! My thanks go to Jordan McElwee, Sam Jenkins, Celeste Pidcott, Andrew Scarff and Dom Barker for welcoming me into their extended friendship group and providing support/gentle mocking throughout the PhD. Current members of the group have also been a massive help, particularly for making sure I don't go too deep down the Bayesian rabbit hole while writing and I am grateful for Steve Wilson, Jack Fannon, James Greer, Rob Foster, Matt Thiesse and Liz Kneale for keeping me sane!

Travelling around the world to Japan as part of my LTA is an experience I will always be grateful for. I'd like to thank Menai Lamers James, Ewan Miller, Katharina Lachner, Adam Speers, Nick Latham, Paul Morrison, Matt Lawe, Jake McKean and Andres Lopez Moréno, particular highlights include the "radiation" beach bonfire, trips to the Ōnsen, Hanami, and the many many adventures attempting to find vegetarian food! Being a part of the T2K collaboration has been an absolute joy and I couldn't have wished for a more welcoming and supportive collaboration. I would also like to

acknowledge the Alliance Canada computing cluster which provided the computing resources required to run the analyses in this thesis.

I'd like to say bijî to Fran, Adam, Becky, Dec and the many other former Warwick folks[1] for constantly reminding me that there is more to life than academia! The multiple attempted DnD campaigns, constant online board games and many pub trips have been a continually source of joy throughout the PhD.

Finally I'd like to thank my parents and brother for supporting in everything I do. Hopefully my explanation of neutrino oscillations via ice-cream flavours has made this thesis make a bit more sense!

[1] and Marshall who wanted to be somewhere funny in the acknowledgements

The story so far: In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move

Douglas Adams - The Restaurant at the End of the Universe

Contents

List of Figures

List of Tables

List of Algorithms

List of Acronyms

The following is a brief translation of acronyms used within this thesis

- AC: Autocorrelation
- ACF: Autocorrelation function
- CP: Charge-Parity
- BANFF: Beam Analysis and Near detector Fitting Framework
- ESS: Effective Sample Size
- FGD: Fine-Grain Detector
- FHC: Forward Horn Current (ν-mode)
- HK/Hyper-K: Hyper-Kamiokande
- JAMS: Jumping Adaptive Metropolis Sampler
- MaCh3: Markov Chains for 3-flavour oscillation
- MC: Monte Carlo
- MCMC: Markov Chain Monte Carlo
- MSW: Mikheyev, Smirnov, Wolfenstein
- PMT: Photomultipler Tube

- PMNS: Pontecorvo, Maki, Nakagawa, Sakata
- RHC: Reverse Horn Current ($\bar{\nu}$-mode)
- SGFD: Super Fine-Grain Detector
- SK/Super-K: Super-Kamiokande
- WLS: Wavelength Shifting (fibre)

Declaration

This thesis presents work from the author and any external contributions are either listed here or clearly cited.

- Chapter 4 contains the results of the P-Theta fit performed by Lucile Mellet and Denis Carabadjac which has been used for comparison with the MaCh3 fit performed by the author. Further comparisons were also performed with the MaCh3 fit from a previous analysis which was performed by Thomas Holvey.
- Chapter 5 contains work from the P-Theta, VALOR and Osc 3++ groups which have been used for the comparison with the MaCh3 fits produced by the author. In addition work by Andrés Lopez Moreno has been included to demonstrate the difficulty performing fits which include both mass hierarchies.

Chapter 1

Introduction and Neutrino Physics Overview

1.1 Proposal, Discovery, and The Solar Neutrino Problem

1.1.1 Early Neutrino History

The existence of neutrinos was first proposed by Wolfgang Pauli in his now famous 1930 letter [3] as a response to anomalous seeming electron energy measurements seen in beta-decay experiments [4]. Beta-decay electrons were found to be emitted with continuous energy spectra [Fig. 1.1] rather than the expected delta-like peak required for 4-momentum conservation. Neutrinos were thus proposed as a "ghost-like" spin-1/2 fermion that can simultaneously carry away excess momentum and be completely invisible to the standard equipment of the time. This hypothesis was given further credibility by Fermi's theory of the weak interaction [5] which requires neutrinos.

Pauli's hypothesis was finally confirmed 26 years later by Cowan and Reines [6] using a liquid scintillator-based detector. By capturing neutrinos produced from the

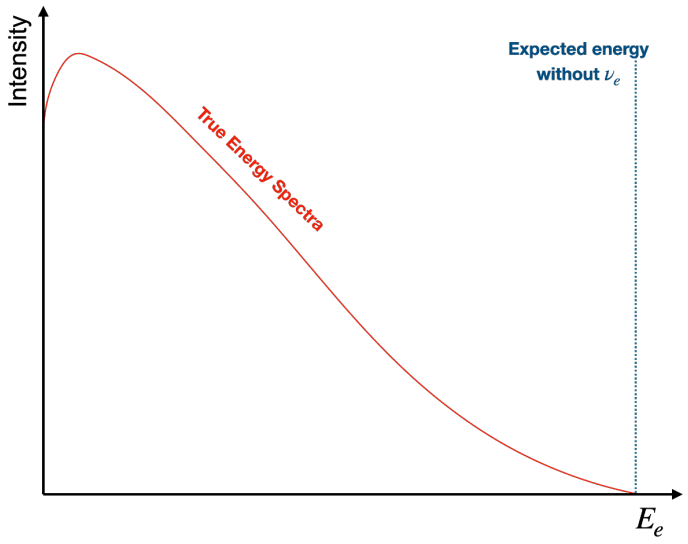

Figure 1.1: Plot showing the expected distribution of beta-decay electron energies in a universe without ν_e (Blue) and the measured distribution of energies. Notably the true energy spectrum is strictly less than the expected spectrum due to conservation of energy. The expected electron energy spectrum is dependent on the nucleus atomic nucleus with a E_e $\mathcal{O}(1$ MeV) [2]

Savannah River nuclear reactor[1] through $\bar{\nu}_e p \rightarrow e^+ n$, it was possible to generate a unique neutrino signal as can be seen in figure 1.2. Firstly, the emitted e^+ annihilates with electrons in the water tank producing two photons; these can be immediately observed in the liquid scintillator. Secondly, dissolved $CdCl_2$ in the tank was used to absorb the neutron. This resulted in the Cd nucleus entering a meta-stable excited state. The de-excitation out of this meta-stable state resulted in a photon cascade which could be detected in the liquid scintillator a short time after the initial photons produced from the aforementioned annihilation event. Combined these two signals create a unique neutrino signature thus allowing for the first observation of these elusive particles. A second experiment located at BNL used non-detection of neutrino capture on Chlorine, $Cl^{37} + \nu \rightarrow e^- + Ar^{37}$ to determine that neutrinos and anti-neutrinos are different particles [8]. Following the discovery of $\bar{\nu}_e$, a team located at Brookhaven National Laboratory [9] used the fact that pion decays do not produce $e^{\pm}$ to deduce the existence of muon neutrinos [Fig. 1.3].

[1]The quite frankly insane first proposal was to use nuclear bombs as the neutrino source [7]

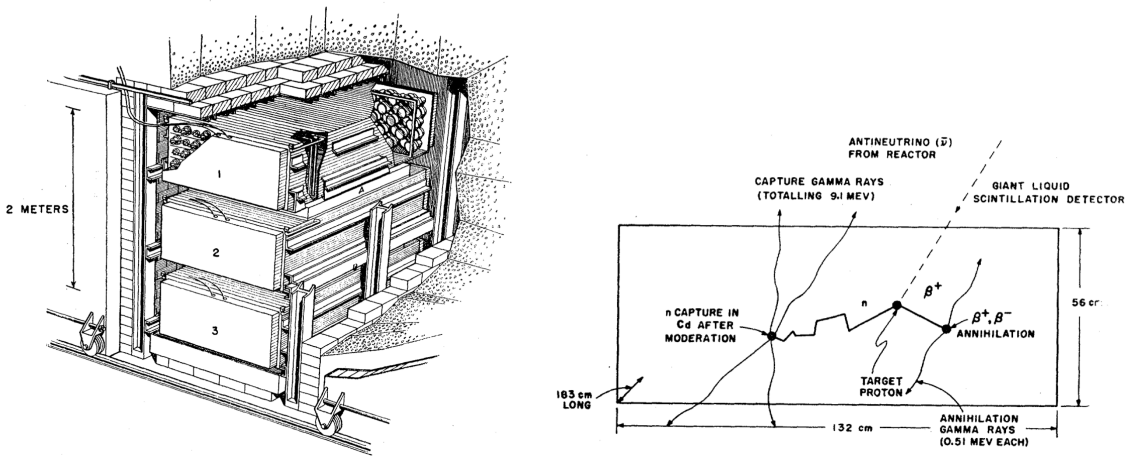

Figure 1.2: (Left) "Club sandwich" structure of Reines and Cowan's seminal experiment. The tanks (labelled 1,2,3) are filled with liquid scintillator which is observed by a set of photomultiplier tubes (PMTs). The CdCl tanks are arranged in layers and a sketch of a neutrino event is depicted (right). This diagram shows the annihilation event and metastable cadmium de-excitation. Figures adopted from [8] and [10]

.

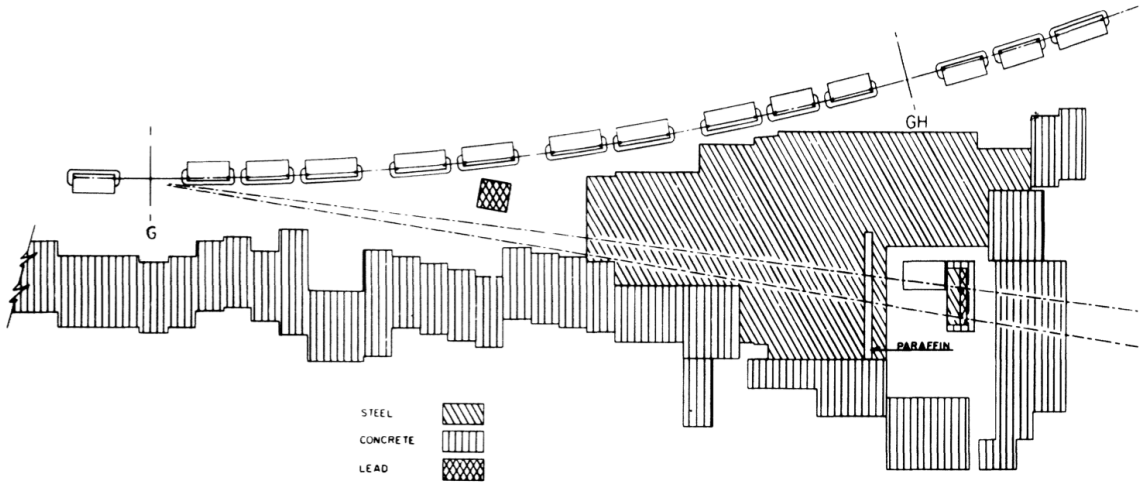

Figure 1.3: Sketch of the experimental setup used by Lederman, Steinberger, and Schwartz to detect the existence of multiple flavours of neutrino. Figure adopted from [9]

.

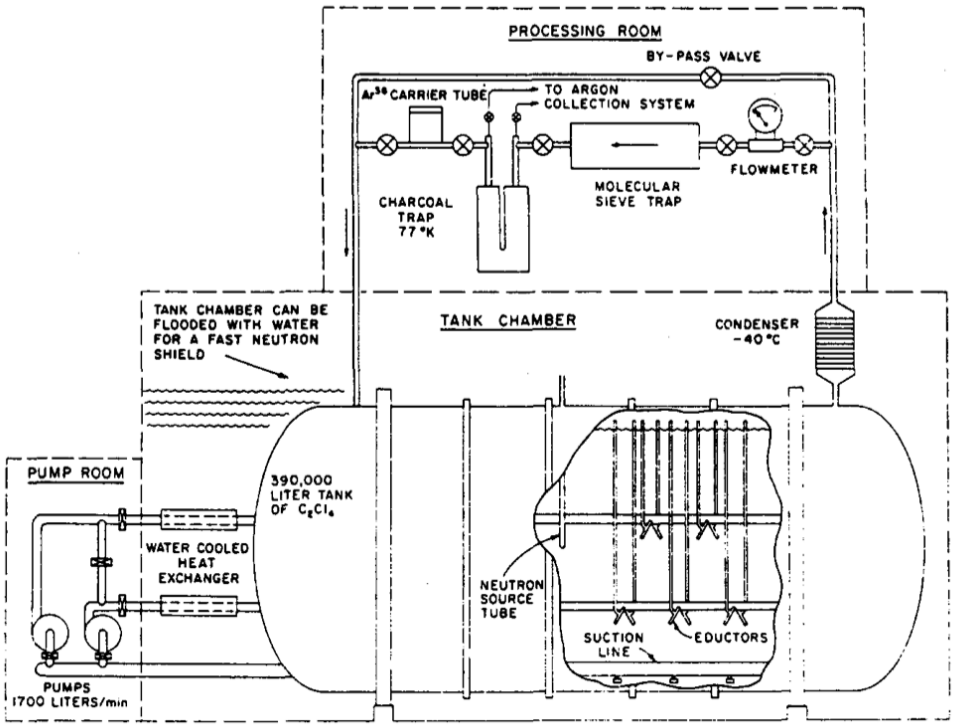

Figure 1.4: Diagram of the Homestake experimental setup taken from [11]. The main detector mass is the 390,000 L tank containing petrochloroethylene (C_2Cl_4). The argon created from neutrino interactions is then extracted by circulating helium through the tank every few minutes. Argon is then separated from the petrochloroethylene via a combination of condensation and molecular sieving resulting in ≈95% of the argon produced in the experiment being collected.

1.1.2 Homestake and the Solar Neutrino Problem

Commencing operation in 1970, the Homestake experiment [11] produced a result that has come to define the field of neutrino physics, the ν_e solar deficit. Homestake aimed to measure the solar neutrino flux through $\nu_e + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar}^+ + e^-$. The argon atoms produced by this interaction could then be counted and the number of argon argon atoms was then used to deduce the total solar ν_e flux. After several years of careful experimentation, the experiment found ≈60% of the expected ν_e flux was missing. The solution lay in a proposal by Pontecorvo [12] and later by Maki, Nakagawa, and Sakata [13], neutrino oscillations.

1.2 Neutrino Oscillations

1.2.1 A Mathematical basis for Neutrino Oscillations

Neutrino oscillations are phenomena whereby the measured weak flavour states of neutrinos change as a function of neutrino energy and distance propagated through space. Since only ν_e and $\bar{\nu}_e$ were known, Pontecorvo proposed oscillations of the form $\nu_e \leftrightarrow \bar{\nu}_e$ [12] similar to neutral meson mixing described in [14]. This was later modified by Maki, Nakagawa, and Sakata [13] after the discovery of ν_μ to instead describe weak flavour changing neutrino oscillations. This only predicted a deficit of 1/2 and could not fully explain the solar neutrino deficit, but, after the discovery of the τ-lepton [15] it was determined that a 3rd flavour of neutrino existed (this was directly detected by DONUT in 2000 [16]).

In order to understand 3-flavour neutrino oscillation it helps to consider the case for 2 neutrino flavours first [17]. Neutrino flavour states, ν_e, ν_μ and mass states ν_1, ν_2 are related in the following way:

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \end{pmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{pmatrix} |\nu_1\rangle \\ |\nu_1\rangle \end{pmatrix}. \tag{1.1}$$

Where the matrix is required to be unitary to conserve transition probability, hence the choice of representation by real-valued angles using a parameter θ. In order to see exactly why this results in energy/propagation distance dependent oscillations, consider the following scenario. Firstly, explicitly write the flavour states in terms of mass states. For $|\nu_e\rangle$ this is,

$$|\nu_e\rangle = \cos(\theta)\,|\nu_1\rangle + \sin(\theta)\,|\nu_2\rangle \tag{1.2}$$

Similarly, the expression for ν_μ in terms of flavour states is

$$|\nu_\mu\rangle = -\sin(\theta)\,|\nu_1\rangle + \cos(\theta)\,|\nu_2\rangle \tag{1.3}$$

Now, consider a neutrino with space-time dependent wavefunction $|x,t\rangle$ with initial state $|0,0\rangle = |\nu_e\rangle$. If we consider the mass-states to be solutions to the Schrödinger equation, this yields

$$|\psi(x,t)\rangle = \cos(\theta)\,|\nu_1\rangle\, e^{i\phi_1} + \sin(\theta)\,|\nu_2\rangle\, e^{i\phi_2} \tag{1.4}$$

where $\phi_i = p_{\mu,i}x_i^{\mu}$ is a real-valued 4-momentum and 4-position dependent phase. Now, converting this back to the flavour basis yields

$$\begin{aligned} |\psi(x,t)\rangle = &[\cos^2(\theta)e^{i\phi_1} + \sin^2(\theta)e^{-i\phi_2}]\,|\nu_e\rangle \\ &+\sin(\theta)\cos(\theta)[-e^{i\phi_1} + e^{i\phi_2}]\,|\nu_\mu\rangle \end{aligned} \tag{1.5}$$

Two assumptions can now be made:

- Neutrinos move linearly along the x-axis
- Neutrino masses are small hence the 3-momenta can be shown to satisfy $\|\vec{p}_1\| = \|\vec{p}_2\| = p \ll 1$. Furthermore, the small mass assumption ensures that neutrino speeds are close to c hence, for a neutrino that's travelled a distance L in time t, $L = t$.

Hence, $\phi_i = E_i t - px$. These simplifications ensure that $\Delta\phi_{12} = (\phi_1 - \phi_2) = \Delta E_{12} t$. In general for variables x_1, x_2, the notation $\Delta x_{12} = x_1 - x_2$ will be used. Using Einstein's famous energy/momentum relationship it can be shown that

$$\Delta E_{12} = \left(\sqrt{p^2 + m_1^2} - \sqrt{p^2 + m_2^2}\right) \tag{1.6}$$

Under the assumption of small p,

$$E_i = \sqrt{p^2 + m_i^2} = p\sqrt{1 + \frac{m_i^2}{p^2}} \approx p\left(1 + \frac{m_i^2}{2p^2}\right) \tag{1.7}$$

From this approximation, it follows that

$$\Delta\phi_{12} = \Delta E_{12} t = \frac{\Delta m^2_{12}}{2p} t. \tag{1.8}$$

Now using the 3 assumptions above, this allows the phase difference to be reformulated as (for a neutrino with energy E and propagation distance L)

$$\Delta\phi_{12} = \frac{\Delta m^2_{12}}{2E} L. \tag{1.9}$$

Consider a neutrino starting in state $|\psi(0,0)\rangle = |\nu_e\rangle$ at time T having propagated a distance L. The probability of the neutrino having transitioned to a state $|\nu_\mu\rangle$ is given by

$$P(\nu_e \to \nu_\mu) = |\, \langle \nu_\mu | \psi(L,T) \rangle \,|^2 \tag{1.10}$$

Expanding into the full form from equation 1.5 yields [2]

$$P(\nu_e \to \nu_\mu) = |\, \langle \nu_\mu | \left(\sin(\theta)\cos(\theta)\,|\nu_\mu\rangle\,[-e^{i\phi_1} + e^{i\phi_2}]\right) |\nu_e\rangle \,|^2 \tag{1.11}$$

Simplifying again yields,

$$P(\nu_e \to \nu_\mu) = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2_{12}}{4E} L\right) \tag{1.12}$$

crucially, the existence of neutrino oscillations requires that there is at least one massive neutrino.

Extending this to the 3 neutrino case, the PMNS representation becomes $U(3)$. The most commonly used parametrisation of this has three real mixing angles (θ_{12}, θ_{13}, θ_{23})

[2] note that the neutrino flavour basis is orthonormal so $\langle \nu_\alpha | \nu_\beta \rangle = \delta_{\alpha\beta}$

and a complex phase (δ_{CP}). This is then represented in the following way [18]:

$$U_{\text{PMNS}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.13}$$

where $c_{ij} = \cos(\theta_{ij})$, $s_{ij} = \sin(\theta_{ij})$. Oscillation probabilities can then be calculated in the same way as the two-neutrino case. The key behavioural difference to the 2 neutrino case is the addition of a complex phase as it allows for CP violation in neutrino oscillations. If neutrinos are Majorana particles i.e. neutrino flavour and anti-flavour states are the same, additional complex phases can also be introduced into the oscillation calculation [19].

1.2.2 CP Violation and 3-Flavour Oscillations

It can be shown that, for 3-neutrino oscillations [20],

$$p(\nu_\mu/\bar{\nu}_\mu \to \nu_e/\bar{\nu}_e) = \sin^2(2\theta_{13})\sin^2(\theta_{23})\sin^2\left(\frac{\Delta m^2_{32}L}{2E}\right) \tag{1.14}$$

$$\mp 8J_{CP}\sin^2\left(\frac{\Delta m^2_{31}L}{4E}\right)\sin^2\left(\frac{\Delta m^2_{21}L}{4E}\right) \tag{1.15}$$

where J_{CP} is a constant known as the Jarlskog invariant,

$$J_{CP} = \frac{1}{8}\sin 2\theta_{12}\sin 2\theta_{23}\sin 2\theta_{13}\cos\theta_{13}\sin\delta_{CP} \tag{1.16}$$

As a result, provided $J_{CP} \neq 0$, neutrino and anti-neutrino oscillation probabilities can differ. This difference violates CP symmetry, the transformation required to change a particle state into an antiparticle state[3]. In order for CP violation to occur the complex component of the PMNS matrix is required to be non-zero.

[3]This will be discussed in greater detail in section 1.4

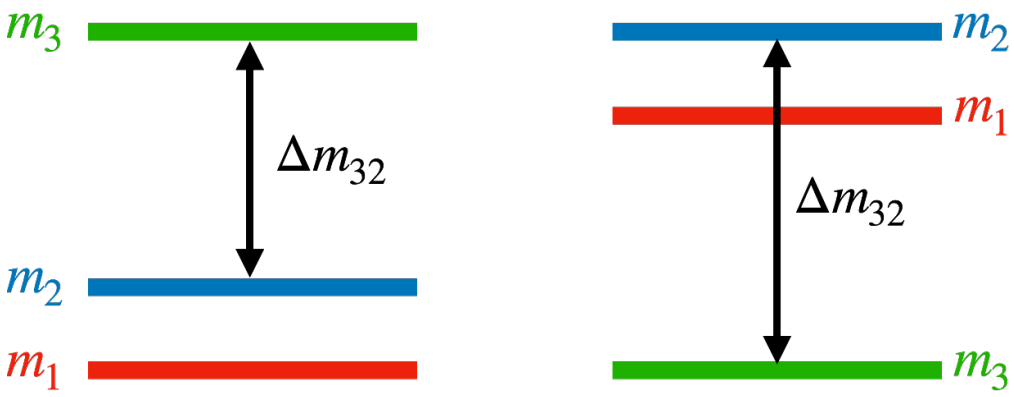

Figure 1.5: Diagrammatic View of mass ordering showing normal (left) and inverted (right) ordering. The mass states m_i are arranged vertically in order of increasing mass (bottom to top) and the mass difference Δm_{32} is shown.

1.2.3 Degeneracies in Neutrino Oscillations

A key feature of the neutrino oscillation calculation is that many of the terms are "masked" by $\sin^2$ or $\cos^2$ which hides information about the sign of parameters and can lead to multiple "correct" values due to the cyclical nature of trigonometric functions. As a result vacuum oscillations lead to several degeneracies when attempting to measure θ_{ij} and resolve the ordering of neutrino masses. Two degeneracies, $\Delta m^2_{21} > 0$ and $\sin^2(\theta_{12}) < 0.5$ have been lifted through solar neutrino measurements. In addition, $\sin^2(\theta_{13})$ has been well constrained by both reactor and accelerator neutrino experiments. Two remaining degeneracies are yet to be resolved, the mass ordering and the octant of θ_{23}. The mass ordering problem arises from the fact that all terms including Δm^2_{32} are of the form $f(\Delta m^2_{32})$ where f is some symmetric function about 0. This means that, in a vacuum, it is impossible from oscillations alone to determine if $m_2 < m_3$ (normal ordering) or $m_3 < m_2$ (inverted ordering) [Fig. 1.5]. Similarly, the neutrino octant problem merges from terms of the form $f(\theta_{32})$; typically this is posed as $\sin^2(\theta_{32}) > 0.5$ (upper octant), $\sin^2(\theta_{32}) < 0.5$ (lower octant) or $\sin^2(\theta_{32}) = 0.5$ (so-called "maximal mixing"). Thankfully for experimentalists, neutrino interactions with matter lift these degeneracies. This allows for the determination of mass ordering and octant in future experiments.

1.2.4 Oscillations in Matter

Vacuum neutrino oscillations alone cannot fully explain the solar neutrino problem. Nuclear fusion can only produce sufficient energy per reaction to produce electrons and hence the initial solar neutrino flux is entirely ν_e [21]. Since vacuum oscillations are dependent on L/E and the distance between the Earth and the Sun is variable, the pure vacuum oscillation prediction would result in seasonal changes in neutrino oscillation probability. This result is incompatible with the observed constant 30% survival rate seen in solar neutrino experiments such as Super-Kamiokande [22]. The solution comes from interactions between neutrinos and matter. By considering matter interactions as a perturbation to the vacuum Hamiltonian i.e. $H = H_{\text{vacuum}} + \delta H_{\text{matter}}$, the total Hamiltonian becomes non-diagonal. Hence mass states evolve in the following way [23],

$$\dot{\nu}_i = -i \sum_{ij} H_{ij} \nu_j \tag{1.17}$$

This results in off-diagonal terms in the Hamiltonian mass states that are allowed to mix generating mass-state oscillations; this is called the MSW effect. This is dependent on the electron number density of a material[4]. As a result, neutrinos travelling through the sun oscillate rapidly into pure-mass states and thus have very little oscillation when travelling from the surface of the sun to the earth.

In accelerator neutrino experiments, MSW can allow for additional degeneracies to be lifted [24]. Unlike solar neutrinos, the electron number density in these experiments constant to a very good approximation. This means that there is no resonance effect and the mass states don't "oscillate" as seen in the previous example. In the vacuum case, discerning the mass ordering from oscillation probabilities is impossible due to the symmetries in the calculation. MSW lifts this and thus it becomes possible to observe the true hierarchy using long baseline oscillation experiments. Figure 1.6 shows the expected bi-probability plot on the Hyper-Kamiokande (HK) experiment. In particular,

[4]This is an approximation given that there is no large block of matter containing heavier leptons.

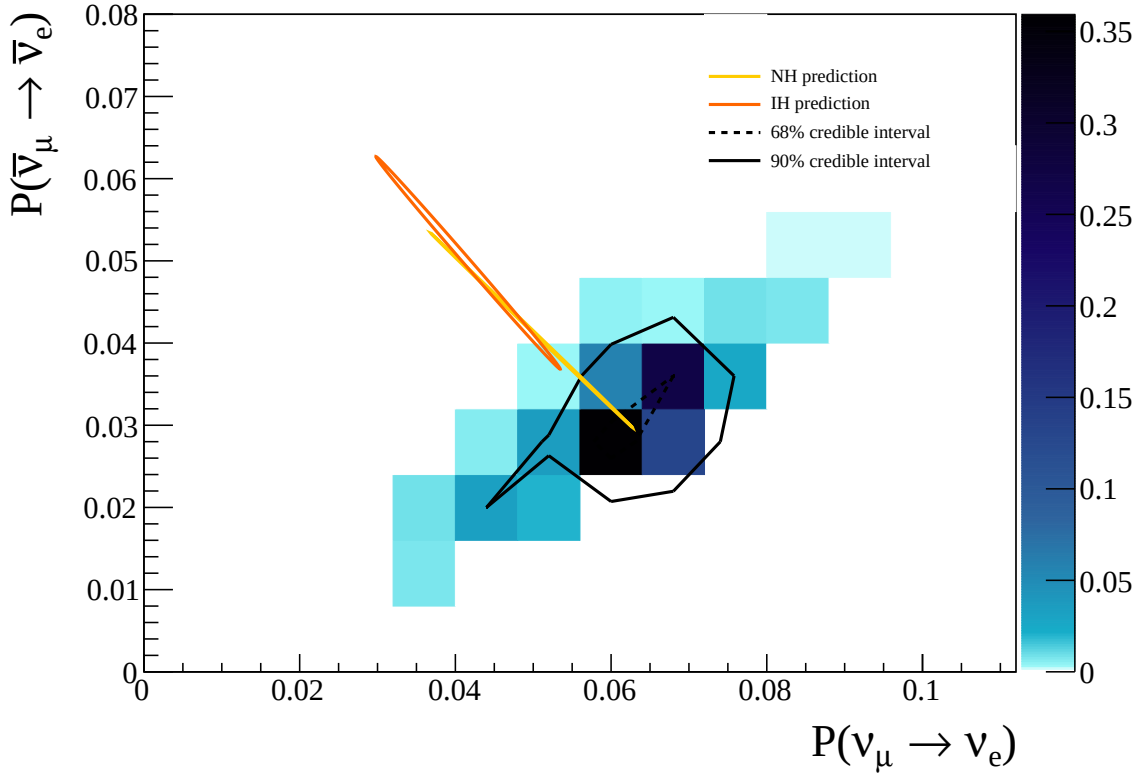

Figure 1.6: Bi-probability plot for the Hyper-Kamiokande experiment Best fit contours are drawn assuming true normal ordering and maximal CP violation. The plot shows the asymmetry in oscillation probability for neutrinos vs anti-neutrinos.

when the true ordering is normal ($\Delta m^2_{32} > 0$) and CP-violation is maximal ($\delta_{cp} = -\frac{\pi}{2}$), the degeneracy lifts almost completely at the HK baseline and thus the true ordering can be determined in this case.

1.3 Observations of Oscillation

1.3.1 Atmospheric Oscillations and Super-Kamiokande

The first direct observation of neutrino oscillations came from the Super-Kamiokande (SK) experiment in Japan [25]. SK is a 50 kt water Cherenkov detector that observes particles using the Cherenkov radiation emitted when they travel faster than light through water [26]. The photons emitted form the light cone that can then be recorded

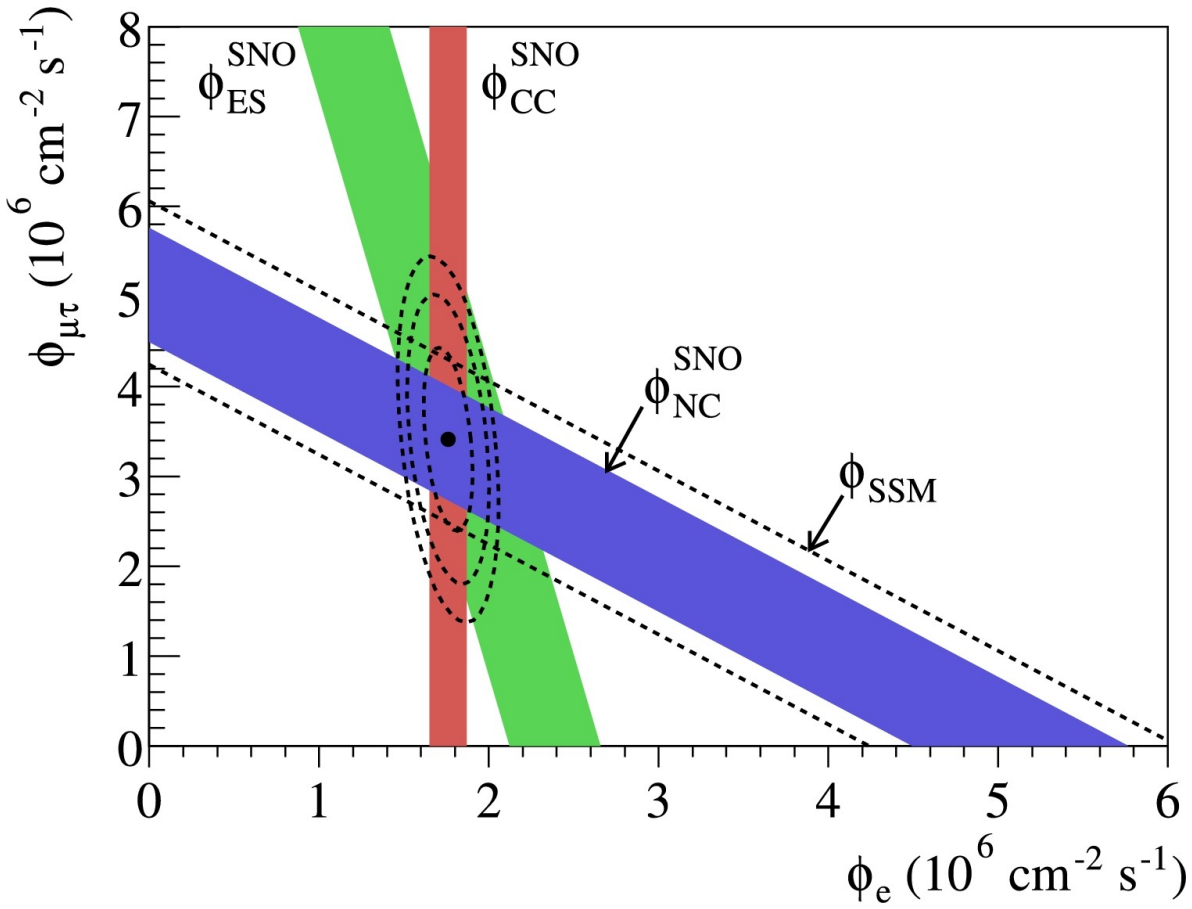

Figure 1.7: Flux of ν_μ and ν_τ vs ν_e flux observed at SNO with 1σ error bands for each interaction channel. Dashed lines are from the standard solar model prediction. Figure from [27].

using photo-multiplier tubes that line the interior and exterior of the detector with the interior PMTs facing inwards towards the center of the tank and exterior facing outwards. Neutrinos interacting in the tank are then identifiable since they will not produce final state particles that generate Cherenkov rings in the outward-facing PMTs first. Furthermore, ν_μ-like and ν_e-like signals can be further distinguished using the "fuzziness" of the resultant Cherenkov ring. Further details about this experiment can be found in section 2.1.9.

The 1997 Atmospheric oscillation result [25] used the deficit in upward going ν_μ-like events relative to downward going events and the zenith-angle distribution to show that neutrinos exhibited flavour changing behaviour consistent with PMNS-like oscillations.

Since the number of events without oscillation was expected to be roughly the same for up and down-going neutrinos, the observed ratio of $\nu_{\text{up}}/\nu_{\text{down}} = 0.5$ showed there was an oscillation-like mechanism present for neutrinos. Whilst this result was vital in proving neutrino oscillations are occurring, the final proof for neutrino oscillations came from the SNO experiment.

1.3.2 Solar Oscillations and SNO

SNO, like SK, used Cherenkov radiation to detect neutrinos; unlike SK it was filled with heavy water rather than simply using ultrapure water [27]. Whilst individual flavour could not be trivially identified through Cherenkov rings in this medium, interactions could be discriminated by one of three interaction channels:

- The neutral current channel, $\nu + d \rightarrow n + p + \nu$
- The charged current channel $\nu_e + d \rightarrow p + p + e^-$
- The electron scattering channel $\nu + e^- \rightarrow \nu + e^-$

Since only ν_e can interact through the CC at the solar neutrino energy scale, the ν_e flux could be measured simultaneously with the total neutrino flux. In 2001 SNO measured a total neutrino flux of $5.44 \times 10^6\,\text{cm}^2\,\text{s}^{-1}$ and a ν_e flux of $1.75 \pm 0.07(\text{stat.})^{0.12}_{-0.11}(\text{sys.}) \pm 0.05(\text{theor.}) \times 10^6\,\text{cm}^2\,\text{s}^{-1}$. This agreed with both the prediction from the solar model [Fig. 1.7] and results from the Homestake experiment, and provided concrete evidence for neutrino oscillations.

1.4 Neutrino Mass

1.4.1 The Weak Interaction

A key consequence of neutrino oscillations is that they require non-zero neutrino mass. To understand why this contradicts the Standard Model (SM), it is important to first

look at the structure of the weak interaction. The weak interaction current was shown by Marshak and Sudarshan to have a $V-A$ structure [28]. This means it is the linear composition of a vector current $V=\gamma^{\mu}$ and an axial current $A=\gamma^{\mu}\gamma^{5}$, where γ^{μ} can be represented by the Dirac gamma matrices and $\gamma^{5}=i\gamma^{0}\gamma^{1}\gamma^{2}\gamma^{3}$. Fermion spinors ψ are commonly described in terms of left and right-chiral states [29]

$$\psi_L = \frac{1}{2}(1-\gamma^5)\psi \tag{1.18}$$

$$\psi_R = \frac{1}{2}(1+\gamma^5)\psi \tag{1.19}$$

$$\psi = \psi_L + \psi_R \tag{1.20}$$

Consider the weak interaction in terms of these chiral states,

$$\gamma^{\mu}(1-\gamma^5)\psi = \gamma^{\mu}(1-\gamma^5)(\psi_L+\psi_R) \tag{1.21}$$

$$\tag{1.22}$$

Since $(1-\gamma^5)(1+\gamma^5)=0$,

$$\gamma^{\mu}(1-\gamma^5)\psi = \gamma^{\mu}(1-\gamma^5)\psi_L \tag{1.23}$$

For antiparticles, consider $\bar{\psi}$ which, through a similar process, yields

$$\bar{\psi}\gamma^{\mu}(1-\gamma^5) = \bar{\psi}_R\gamma^{\mu}(1-\gamma^5) \tag{1.24}$$

As a result, the weak interaction can only "see" left-chiral particle and right-chiral antiparticle states. Since SM neutrinos can only be produced from weak interactions, it immediately follows that only left-chiral neutrinos (and right-chiral anti-neutrinos) are permitted within the SM. From this, it follows that in order for the weak force to transform particle and antiparticle states a CP transformation is required [30].

1.4.2 Chirality and the Higgs Mechanism

The Higgs mechanism generates mass within the SM. The mass Lagrangian can be considered to be a mixing of left and right-chiral states [31]

$$-\mathcal{L}_{mass} = m_D \bar{\psi}_R \psi_L + h.c \tag{1.25}$$

where m_D is the Dirac mass and can be obtained directly from the Higgs mechanism [32]

$$m_D = G_\nu \frac{\langle h \rangle}{\sqrt{2}} \tag{1.26}$$

where $\langle h \rangle$ is the Higgs vacuum expectation value and G_ν the neutrino-Higgs coupling strength. The immediate consequence of this is that, due to their uni-chiral nature, neutrinos cannot have mass within the SM.

1.4.3 Alternative Mechanisms for Neutrino Mass Generation

The simplest way to generate neutrino mass is by introducing right-chiral particle and left-chiral antiparticle states for neutrinos. In order to agree with current observations of the universe, these additional neutrinos must be 'sterile', that is to say they can only couple with the Higgs field.

Another method of introducing neutrino mass is by allowing them to be Majorana particles. A Majorana fermion is unique in that its particle state is conserved by $\hat{C}$ symmetry i.e. $\hat{C}\psi \mapsto \psi$. As a result, it is possible to construct a mass term using only left-handed particle states:

$$-\mathcal{L}_{\text{mass},Maj} = -\frac{1}{2} m_M \bar{\psi}^c_L \psi_L + h.c. \tag{1.27}$$

where m_M is the Majorana mass and $\psi^c = \hat{C}\bar{\psi}^T$ the particle spinor after charge conjugation. This has two notable features. Firstly there is an additional contribution to

the PMNS matrix [19] which can be parametrised in terms of two phases, α_1, α_2

$$U_{\text{PMNS,Maj}} = U_{\text{PMNS}} \begin{pmatrix} e^{i\alpha_1} & 0 & 0 \\ 0 & e^{i\alpha_2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.28}$$

As this is an additional complex component it leads to further terms similar to the Jarlskog invariant and hence allows for additional CP-symmetry violating properties in neutrinos. The second, slightly more subtle, feature is that Majorana and Dirac masses are not mutually exclusive. This results in true neutrino mass states being a linear combination of Dirac and Majorana mass states [33].

1.5 Modern Neutrino Physics

1.5.1 What's Left to Discover?

In previous sections, mechanisms for oscillation and neutrino mass generation have been discussed/ However, many of the relevant properties have either not been measured to enough precision to remove degeneracies or have not been observed at all. The following non-exhaustive list details the remaining known unknowns in the field:

- Does δ_{CP} exclude CP-conserving values?
- Which octant does θ_{23} lie in?
- Is the mass ordering inverted or normal?
- What are the individual neutrino masses?
- Are neutrinos Majorana?
- Do sterile (right-handed) neutrinos exist at low mass scales?
- Are there additional left-handed states beyond the 3 observed weak flavour states?

There are currently several experiments both running and under construction that aim to answer these questions.

1.5.2 Reactor Experiments

Since SK and SNO's initial results, neutrino physics has rapidly progressed. Modern neutrino oscillation experiments can be broadly classified into four groups: reactor, low and high energy astrophysical, and accelerator. Reactor experiments follow the same principles as Reines and Cowan's original detector, using the large antineutrino flux generated by nuclear fission to observe short (<1 km) and medium (<100 km) baseline neutrino oscillations. Recent reactor experiments include Daya Bay, RENO, and Double Chooz [34]. The survival probability for $\bar{\nu}_e$ is given by

$$P(\bar{\nu}_e \to \bar{\nu}_e) \approx 1 - \cos^4(\theta_{13}) \sin^2(2\theta_{12}) \sin^2(\Delta_{21}) - \sin^2(2\theta_{13}) \sin^2 \Delta_{3\ell} \tag{1.29}$$

Where $\Delta_{ij} = 1.27\Delta m^2_{ij} L/E$ and $\Delta_{3\ell} = \Delta_{32} \approx \Delta_{31}$ is used to simplify this expression. It is clear from this that reactor experiments are sensitive to θ_{12}, θ_{12}, Δm^2_{31} and Δm^2_{21}. The eventual exclusion of $\sin^2(\theta_{13}) = 0$ was given in 2012 by Daya Bay which presented a value of $\sin^2(\theta_{13}) = 0.092 \pm 0.016(\text{stat}) \pm 0.005(\text{syst})$ [35]. This is particularly vital for neutrino physics as the Jarlskog invariant [Eqn. 1.16], which enables neutrino CP violation, requires all mixing angles to be non-zero. Future experiments in this field include JUNO, based in Jiangmen, China. In addition to improving the precision of previous reactor results, this experiment will use a much larger ($\sim$ 50 km) baseline to measure Δm^2_{32} and determine the mass ordering [36].

1.5.3 Astrophysical Experiments

Astrophysical neutrinos have a wide energy range from sub-eV relic neutrinos [37] to ultra-high energy PeV and above objects produced in violent cosmological events [38] allowing for observations of a large range of neutrino energy spectra. In order to observe

the full spectrum of astrophysical neutrino energies many techniques are required. The largest[5] currently running neutrino observatory is IceCube, located in Antarctica [39]. By using an array of photo-multiplier tubes (PMTs) inside of digital optical modules over a $\approx$ 1 km^3 region, the detector can observe Cherenkov radiation produced by particles travelling through the ice sheet.

On a smaller scale, experiments such as Super-Kamiokande (SK) and its upgrade Hyper-Kamiokande (HK) also use Cherenkov radiation produced by the products of neutrino interactions. In this case, this is done through liquid water rather than ice and will be discussed in more detail in 2.1 and 2.2.

Combined together the array of neutrino detectors become part of the Supernova Early Warning System (SNEWS) [40]. 99% of the energy produced in a core-collapse supernova is carried away by neutrinos. Since neutrinos have very small cross-sections relative to photons and travel very close to the speed of light they can escape the supernova faster than photons and hence, the neutrino pulse can arrive at Earth first [41]. This means that, by passively observing the neutrino background, astronomers can be alerted to the position of a supernova before light reaches Earth. This positional information can be obtained both from the incoming neutrino direction via elastically scattered ν_e and from triangulation using the full set of SNEWS detectors.

Alongside astrophysics, these experiments can be used to measure atmospheric neutrino properties. As was discussed earlier, Super-Kamiokande provided the first evidence of oscillations using atmospheric neutrinos. This sector is particularly sensitive to neutrino mass ordering and future experiments such as Hyper-Kamiokande and DUNE aim to combine atmospheric and accelerator neutrino data to exclude the incorrect ordering [42].

[5] and coolest...badum tst

1.5.4 Accelerator Experiments

Accelerator experiments are the primary method for measuring neutrino CP violation and are the main focus of this thesis. Generally speaking, accelerator experiments produce neutrinos using the method described by Figure 1.8. This allows for control of both the neutrino energy distribution and whether the beam produces primarily neutrinos or anti-neutrinos. This class of experiments can be further separated into on-axis and off-axis. On-axis experiments place all detectors on the central axis of the beam which results in a maximised neutrino flux and a large range of neutrino energies. Off-axis experiments exploit conservation of momentum in to produce a narrower range of neutrino energies. This results in an increased flux for the desired peak oscillation energy but will have a reduced overall flux relative to an on-axis experiment with the same neutrino source. This helps to increase the number of neutrinos with L/E close to the value for maximal oscillation.

Detector distance also plays a factor in these experiments; nearby detectors, such as ND280 [43], MiniBOONE [44], and LSNd [45] can be used to search for short baseline oscillations since the value of L/E picked has minimal oscillation. In order to efficiently measure oscillation properties it is common to place further detectors at oscillation maxima [where $P(\nu_\mu \rightarrow \nu_\mu)$ is minimised]. The experiments allow for precisions measurements of $\sin^2(\theta_{23})$, $\sin^2(\theta_{13})$, Δm^2_{32} and, since they can use both neutrino and antineutrino beams, can also measure δ_{cp} to a very high precision.

The next generation of accelerator neutrino experiments are being led by DUNE [46] and Hyper-Kamiokande (HK) [47], which aim to provide ultra-high precision measurements of PMNS parameters. Both experiments are aiming to remove the mass-hierarchy degeneracy[6] and constrain δ_{CP}. DUNE uses the MSW effect to "automatically" lift the degeneracy and will only observe a single mass ordering. For HK the resolution is dependent on the value of δ_{cp} which will be discussed in greater detail in chapter 5.

[6]This may hav been achieved by JUNO by the time HK and DUNE switch on [48]

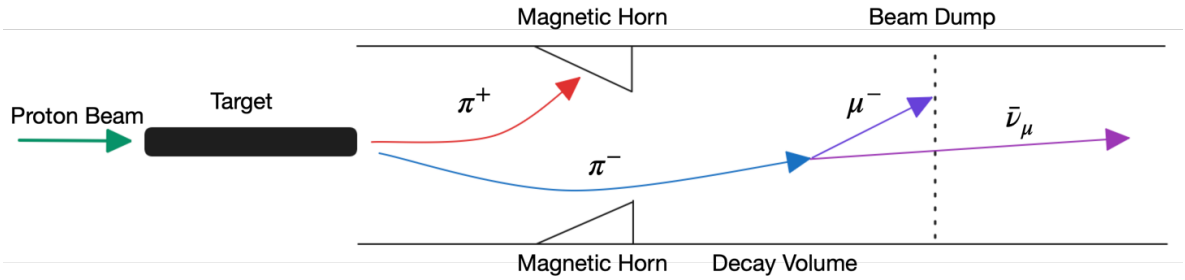

Figure 1.8: Basic Diagram of an accelerator experiment setup. A target is bombarded by an accelerated proton source. Magnetic horns are used to select either positively or negatively charged particles and the selected particles are allowed to decay. Since the most common particles produced are pions and pions primarily decay through $\pi^{\pm} \rightarrow \mu^{\pm} \overset{(-)}{\nu}_{\mu}$, this results in a large number of neutrinos in the final state. Finally, the non-neutrino particles are absorbed by the beam dump resulting in a neutrino beam.

1.5.5 Non-Oscillation Experiments

Along with experiments measuring neutrino oscillation, there is also a rich ecosystem of non-oscillation experiments aiming to determine further properties of neutrinos. These include KATRIN, which uses β-decay of tritium to determine the mass of $\bar{\nu}_e$ [49]. In addition, several experiments such as LEGEND [50] and KamLAND-Zen [51] are investigating neutrino mass generation mechanisms and, in particular whether or not neutrinos are Majorana. This is primarily done by looking for neutrino-less double β-decays [52]. This is a process whereby neutrons can decay into two electrons without requiring a outgoing neutrino.

Neutrino cross-sections are another area of the field under rigorous investigation, detectors such as MIνERVa [53], NINJA [54], and ND280 are used to measure neutrino interactions with minimal oscillation. Understanding this area is vital for all neutrino experiments as understanding systematics related to neutrino/matter interaction improves sensitivity in all detectors.

1.5.6 Current Best Fit Values

The current global PDG best-fit values for neutrino oscillation parameters are as listed in table 1.1. Note that the convention for δ_{cp} in the PDG is $[0, 2\pi]$, T2K uses $[-\pi, \pi]$ which will be the convention for most plots in this thesis.

Parameter	True Inverted Ordering	True Normal Ordering
$\sin^2(\theta_{12})$	0.307 ± 0.013	0.307 ± 0.013
$\sin^2(\theta_{23})$	$0.534^{+0.021}_{-0.024}$	$0.537^{+0.018}_{0.024}$
Δm^2_{21}	$(7.53 \pm 0.18) \times 10^{-5}$ eV^2	$(7.53 \pm 0.18) \times 10^{-5}$ eV^2
Δm^2_{32}	$(-2.519 \pm 0.033) \times 10^{-3}$ eV^2	$(2.437 \pm 0.033) \times 10^{-3}$ eV^2
$\sin^2(\theta_{13})$	$(2.20 \pm 0.07) \times 10^{-2}$	$(2.20 \pm 0.07) \times 10^{-2}$
δ_{cp}	$(1.23 \pm 0.21)\pi$ rad	$(1.23 \pm 0.21)\pi$ rad

Table 1.1: PDG best-fit values for PMNS parameters [55]

1.6 Description of Thesis Structure

Now that the history of neutrino physics has been explored and the theory detailed, the rest of the thesis will describe the specifics of performing a Bayesian oscillation analysis. Chapter 2 will detail the T2K and Hyper-Kamiokande detectors. Chapter 3 discusses Bayesian statistics with particular focus on the Markov chain Monte Carlo (MCMC) methods used in this thesis. Chapter 4 will present the result of the 2023 T2K oscillation analysis which were produced in collaboration with the P-Theta fitting group. Chapter 5 is an overview of the early stages of the Hyper-Kamiokande Bayesian fit and includes validations produced with the VALOR and P-Theta groups as well as work from Andrés Lopez Moreno. Penultimately, chapter 6 will detail the introduction of novel MCMC techniques to the MaCh3 fitter used in all analyses in this thesis. Finally, chapter 7 will contain some concluding remarks.

Chapter 2

T2K and Hyper-Kamiokande

2.1 The T2K Experiment

2.1.1 Overview of T2K

T2K is a long-baseline off-axis accelerator neutrino experiment that uses two detector complexes, separated by 295 km, to probe neutrino oscillations [Fig. 2.1]. Whilst the method used to generate the neutrino beam is shown schematically in figure 1.8 it is important to understand the specifics. The J-PARC synchrotron [56] is used to generate a 30 GeV proton beam; this then bombards a graphite target to produce a large number of mesons, most of which are charged pions [57] [58].

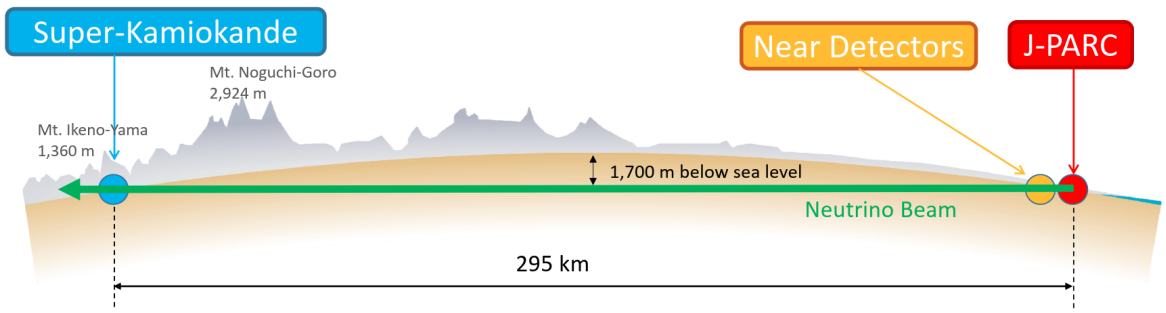

Figure 2.1: Cross-sectional view of the T2K experiment layout. Figure acquired from [43].

Pions can then be selected by charge using magnetic focusing horns [59] to remove either π^+ or π^-; the remaining pions then primarily decay through either

$$\pi^+ \rightarrow \nu_\mu + \mu^+ \tag{2.1}$$

or

$$\pi^- \rightarrow \bar{\nu}_\mu + \mu^- \tag{2.2}$$

This process occurs in the 96m long decay pipe before the remaining particles collide with the beam dump (a large wall of graphite and sandstone) [60] removing most non-neutrino particles, thus resulting in a neutrino beam. The near detector complex [61], located 280 m from the beam target, consists of four detectors, WAGASCI/babyMIND, INGRID, and ND280. These use the high neutrino flux to constrain beam properties and measure neutrino cross-sections. The oscillation group currently only uses data from ND280.

The final detector, located 295 km away is, Super-Kamiokande [26], a 40 kt water Cherenkov detector located under Mount Ikena. This is located at the first oscillation maximum for $L/E = 295\,\text{km}/0.625\,\text{GeV}$ and hence sees the maximal number of ν_e from oscillation.

This setup allows for precision measurements of neutrino oscillation parameters [62]; T2K is primarily sensitive to δ_{CP}, Δm^2_{32}, $\sin(\theta_{13})$ and $\sin(\theta_{23})$. In addition, the physics program also includes cross-section measurements and searches for exotic physics at the near detector.

2.1.2 The J-PARC Beamline

J-PARC's proton accelerator uses a 3-stage process to produce the 30 GeV proton beam required by T2K [63]. Firstly, a linear accelerator [64] boosts a proton beam to 400 MeV. This beam then enters the Rapid Cycling Synchrotron (RCS) [65] in which electrons are stripped from the ion beam which results in a pure proton beam. The RCS boosts

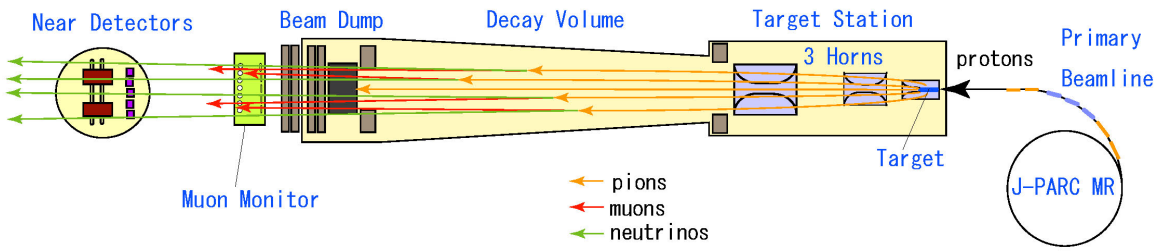

Figure 2.2: Diagram showing the secondary T2K beamline. Protons from the J-PARC main ring (MR) are propelled through the primary beamline towards the graphite target. The resulting shower of particles is focused by the 3 beam horns and allowed to decay in the beam dump. The surviving particles then travel through the decay volume and non-neutrino particles are absorbed by the beam dump. Figure taken from an internally produced diagram.

the beam energy up to 3 GeV at a periodicity of 25 Hz. Four beam pulses or bunches are injected into the RCS per cycle; 5% of these are then injected into the main ring (MR) [66] every 40 ms resulting in 8 bunches of protons in the MR at a time. The MR accelerates these bunches up to 30 GeV at a cycling rate of 0.5 Hz. protons are then extracted from the MR using a series of kicker magnets. These are the sent into the T2K beamline.

2.1.3 T2K Beamline, Target, Decay Pipe and Magnets

The T2K beamline [Fig. 2.2] [43] aims the proton beam at the target in two stages. Firstly, the beam is bent by ≈80°; this points it such that Super-Kamiokande is 2.5° away from the central beam axis. It then curves downwards in order to be at the correct angle to impinge onto the beam target producing primarily pions. The target itself is a graphite rod surrounded by a layer of titanium. Helium gas cooling is used to dissipate the heat produced from high-energy proton collisions.

The particles are then focused using 320 kA magnetic focussing horns. The horn current direction will collimate either negatively (reverse horn current) or positively (forward horn current) charged pions into decay volume resulting in a primarily $\bar{\nu}$ or ν neutrino beam [67]. As mentioned above, the pions then almost all entirely decay to

produce $\nu_\mu/\bar{\nu}_\mu$ and the corresponding (anti-)muon [68]. As can be seen in figure 2.3, the neutrino beam is significantly purer than the anti-neutrino beam The decay volume then terminates at the beam dump, a large wall of sandstone and graphite that absorbs most particles produced in the decay pipe aside from the neutrinos and muons with energies $> 5\,\text{GeV}$. These high-energy muons are then monitored by MUMON which gives constraints on the beam direction and intensity.

2.1.4 Why Off-Axis?

Before discussing the neutrino detectors, it is useful to understand the reasoning behind the off-axis configuration. By positioning the detectors off of the main beamline axis, T2K can observe a much narrower spectrum of neutrino energies; this can be seen in figure 2.4. In particular, by using an off-axis angle of 2.5°, the neutrino energy spectrum will sharply peak at $0.625\,\text{GeV}$. To understand this phenomenon, consider the decay process $\pi \to \nu + \mu$. By following the method in [71], it can be shown that the neutrino energy, E_ν, is related to off-axis beam angle (θ) way. 4-momentum conservation requires

$$p^\alpha_\mu = p^\alpha_\pi - p^\alpha_\nu \tag{2.3}$$

Where $p^\alpha_{\mu,\nu,\pi}$ are the 4-momenta of the muon, pion and neutrino involved in the decay. If this is then "squared",

$$p^\alpha_\mu p_{\mu,\alpha} = p^\alpha_\pi p_{\pi,\alpha} - 2p^\alpha_\pi p_{\nu,\alpha} + p^\alpha_\nu p_{\nu,\alpha} \tag{2.4}$$

Since $p^\alpha_i p_{\alpha,i} = m^2_i$ where m_i is the mass of the particle, this expression can be further simplified to

$$m^2_\mu = m^2_\pi - 2p^\alpha_\pi p_{\nu,\alpha} \tag{2.5}$$

Where m^2_ν has been removed as it negligible relative to the other terms in the equation. In addition, the "mixed" particle term can be written in terms of energy (E_i) and

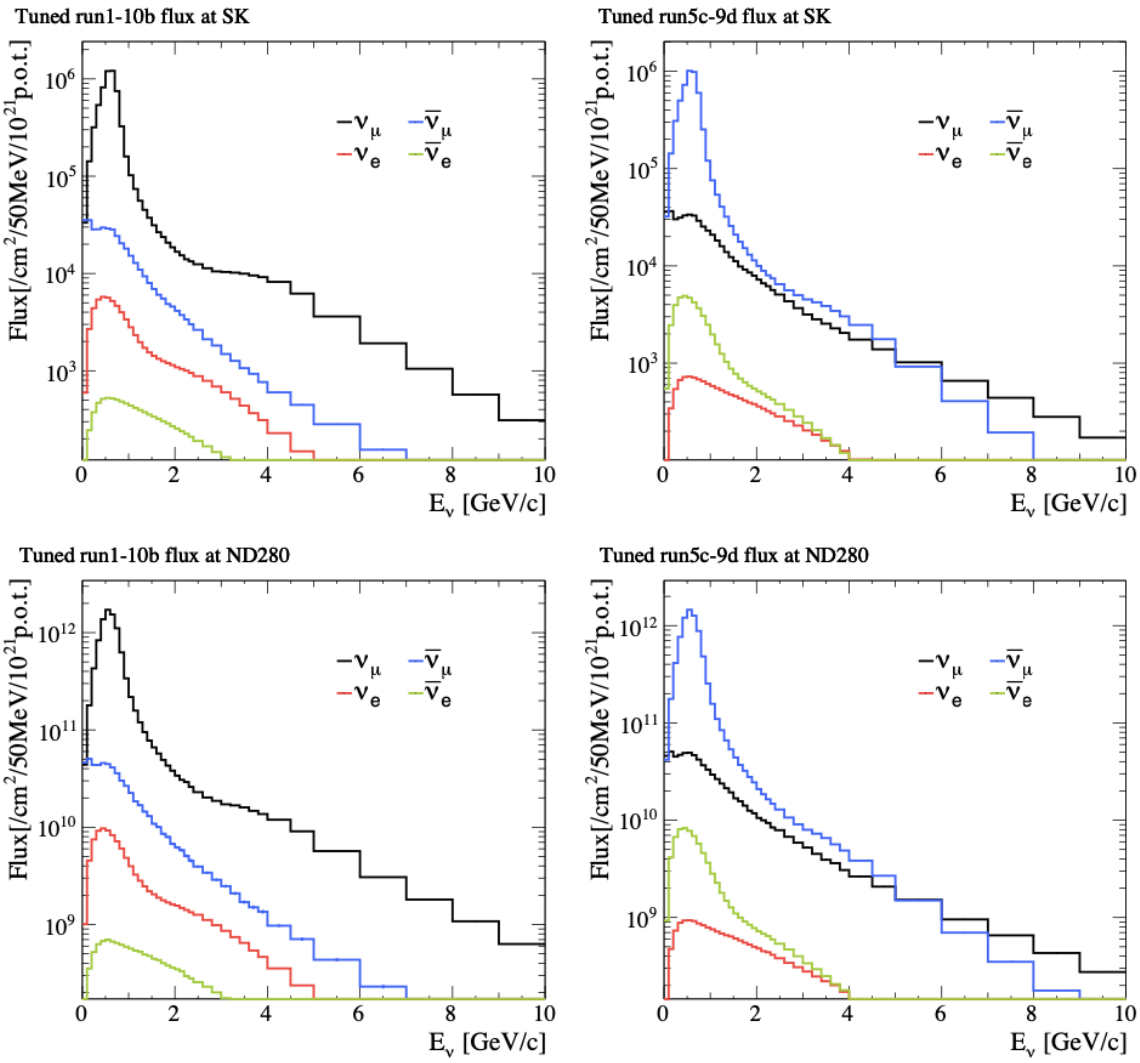

Figure 2.3: The T2K beam flux prediction at Super-Kamiokande (SK) (top) and ND280 (bottom) including additional tuning from the NA61-SHINE experiment [69]. The ν beam mode prediction (left) is notable purer than the $\bar{\nu}$-mode (right) prediction, particularly at large beam energies. Taken from [70].

3-momentum ($\vec{p}_i$) components,

$$p^{\alpha}_{\nu} p_{\pi,\alpha} = E_\nu E_\pi - \vec{p}_\nu \circ \vec{p}_\pi \tag{2.6}$$

Using the small neutrino mass approximation, $|\vec{p}_\nu| \approx E_\nu$ and the simple dot product formula, $\vec{p}_\pi \circ \vec{p}_\nu = |\vec{p}_\pi||\vec{p}_\nu|\cos(\theta)$ where θ is the angle between incoming pion and the outgoing ν directions it can be shown that

$$p^{\alpha}_{\pi} p_{\nu,\alpha} = E_\nu E_\pi - |\vec{p}_\pi| E_\nu \cos(\theta) \tag{2.7}$$

Hence, the following expression can be obtained

$$m^2_\mu = m^2_\pi - 2E_\nu \left[E_\pi - |\vec{p}_\pi|\cos(\theta)\right] \tag{2.8}$$

This can then be rearranged to

$$E_\nu = \frac{\Delta m^2_{\pi\mu}}{2[E_\pi - |\vec{p}_\pi|\cos(\theta)]} \tag{2.9}$$

where $\Delta m^2_{\pi\mu}$ is the difference between squares of pion and muon masses. The maximum neutrino energy can then be calculated in the following way

$$\frac{\partial E_\nu}{\partial E_\pi} = -\frac{\Delta m^2_{\pi\mu}}{2}\left(\frac{2}{(E_\pi - |\vec{p}_\pi|\cos(\theta))^2}\right)\left(1 - \frac{E_\pi}{|\vec{p}_\pi|}\cos(\theta)\right) \tag{2.10}$$

Setting $\partial E_\nu / \partial E_\pi = 0$ yields that E_ν is maximised when $|\vec{p}_\pi| = E_\pi \cos(\theta)$. Using the mass-energy-momentum relationship, it can then be shown that, at this maximal value of E_ν,

$$m^2_\pi = E^2_\pi - |\vec{p}_\pi|^2 \Rightarrow m^2_\pi = E^2_\pi - E^2_\pi \cos^2(\theta) \tag{2.11}$$

And hence $m_\pi = E_\pi \sin(\theta)$. As a result, the maximal value of E_ν can be written in purely terms of particles masses and the outgoing pion angle in the form,

$$E_{\nu,\max} = \frac{\Delta m^2_{\pi\mu}}{2m_\pi \sin(\theta)} \tag{2.12}$$

The result is a narrower band of neutrino energies with a distinct peak at the maximal energy. In T2K θ is tuned so E_ν peaks at the maximal oscillation probability. This maximises the number of observed ν_e at the far detector.

2.1.5 Muon Monitor

Located downstream of the beam dump, the T2K Muon Monitor (MUMON) monitors the neutrino beam direction and intensity using the muons produced in the decay volume and beam dump [Fig. 2.5] [73]. It is capable of measuring the position of the beam central axis to a precision of 3 cm (0.25 mrad) and can use the muon data to constrain the neutrino beam stability to a precision of 3% or better.

2.1.6 Near Detectors

Whilst the main body of this thesis is primarily concerned with neutrino oscillations using data from the far detector, the near detectors are vital for constraining beam flux properties and neutrino cross-section parameters. All detectors are located $\approx 280\,\text{m}$ from the beam target with WAGASCI/babyMIND and ND280 located off-axis and INGRID located on the main beam axis.

2.1.6.1 WAGASCI/babyMIND

WAGASCI (WAter Grid And SCIntillator) and babyMIND (baby Magnetised Iron Neutrino Detector) form an experiment aiming to constrain the differences between neutrino cross-sections on water and liquid scintillator (LS). This is particularly useful since ND280 and INGRID are primarily made of iron and carbon-based detectors which

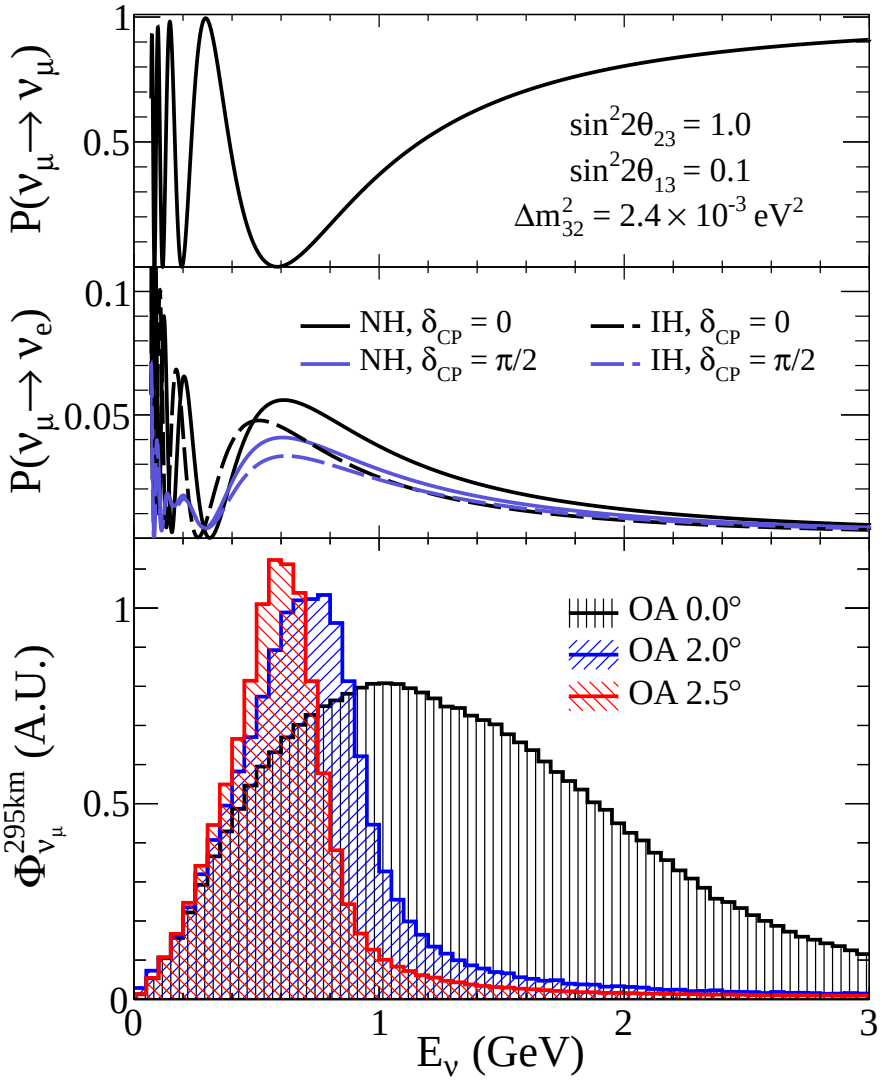

Figure 2.4: (Top, middle) Neutrino oscillation probability plotted against neutrino flux for a detector located at 295 km away with a range of neutrino energies. (Bottom) Neutrino energy spectra at SK at 3 angles relative to the main beamline. Figure adopted from [72]

makes scaling cross-section for the water-based far detector somewhat non-trivial.

WAGASCI is comprises two water tanks sandwiched between two hydrocarbon-filled (HC) tanks. The structure containers a grid structure made of scintillator bars that pass through these tanks. Scintillator light is then detected using Multi-Pixel Photon Counters (MPPCs). In addition, the detector uses exterior muon range detectors (MRDs) on either side to measure outgoing muon kinematics to aid with internal event

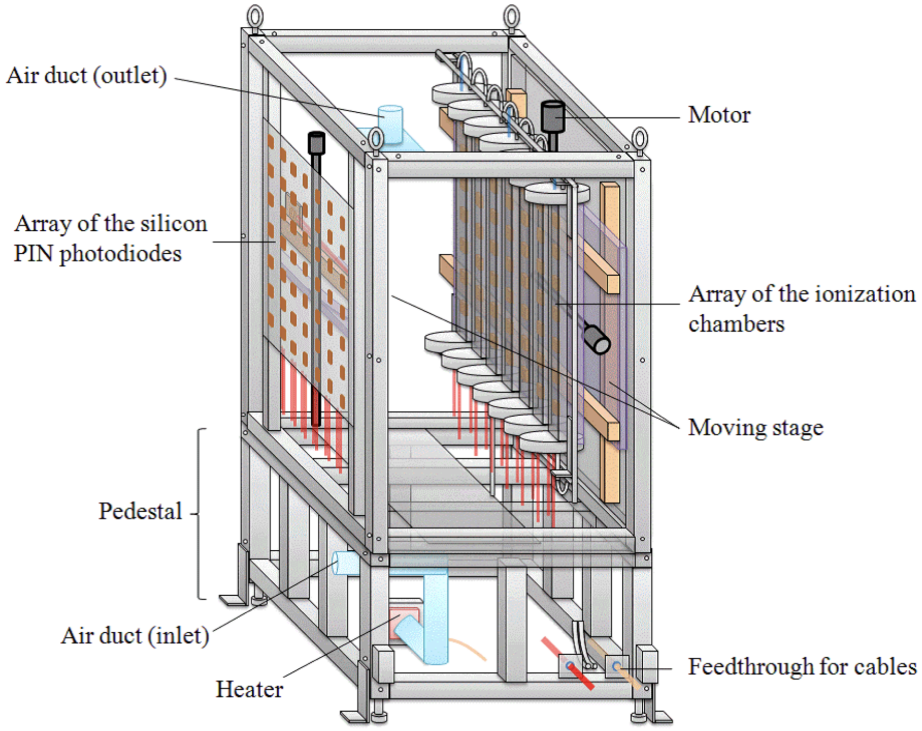

Figure 2.5: Diagram showing a side on view of MUMON. Figure taken from [73].

reconstruction. The babyMIND detector is used to improve the charge and momentum measurements for muons produced within WAGASCI. This detector consists of magnetised iron plates interleaved between scintillator planes. Currently neither of these detectors is used within the oscillation group; for further information see [74] and [75].

2.1.7 INGRID

INGRID (Interactive Neutrino GRID) [Fig. 2.6] is the only on-axis detector used by T2K to measure beam direction and intensity. It consists of modules [Fig. 2.6] which are made up of iron plates interleaved with plastic scintillator planes [78]. The modules are arranged in a cross-shape which allows the detector to measure beam direction and intensity along an x-y plane. Scintillator light is then readout by MPPCs which allows for event reconstruction. INGRID is primarily sensitive to muons and detects charged-current neutrino events via outgoing muons. In addition to the modules in the cross,

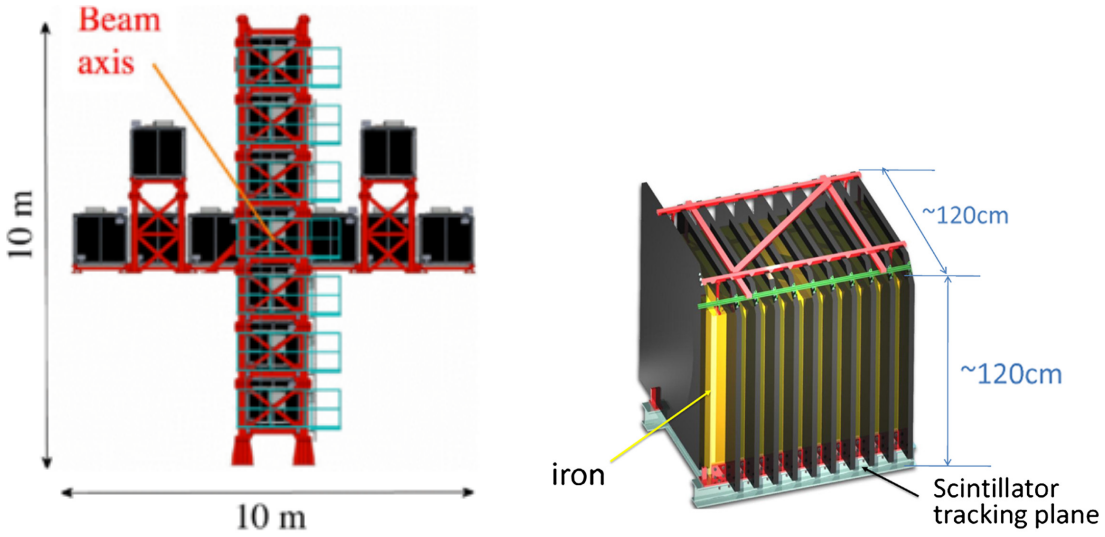

Figure 2.6: (Left) Schematic view of the INGRID detector showing the arrangement of INGRID modules, the detector cross-shaped and embedded into the side of the ND pit. (Right) An INGRID module, with 9 iron planes and 11 scintillator planes can be seen. (Figures obtained from [76] and [77], resp.)

INGRID contains a proton module [79] located at the central point of the detector, upstream of a standard module. This aims to observe additional interactions, primarily CC0π, however it has been used in addition cross-section analyses for example, CC1π^+ events with a proton in the final state topology which are not well constrained by the regular INGRID modules.

2.1.8 ND280

ND280 is a large off-axis detector composed of multiple sub-detectors. Figure 2.7 shows two detector configurations. The configuration used in chapter 4 contains the PØD in the upstream section whilst all future analyses use the ND280-upgrade configuration containing the time of flight (TOF) detector, super fine grain detector (SFGD) and the high-angle TPCs (HA-TPCs). As with all the near detectors, the primary aims of ND280 are to constrain beam properties such as the neutrino energy spectrum and to measure neutrino interaction cross-sections. The entire detector is encased in a 0.2 T magnet to enable measurements of the sign of the charge and momentum of parti-

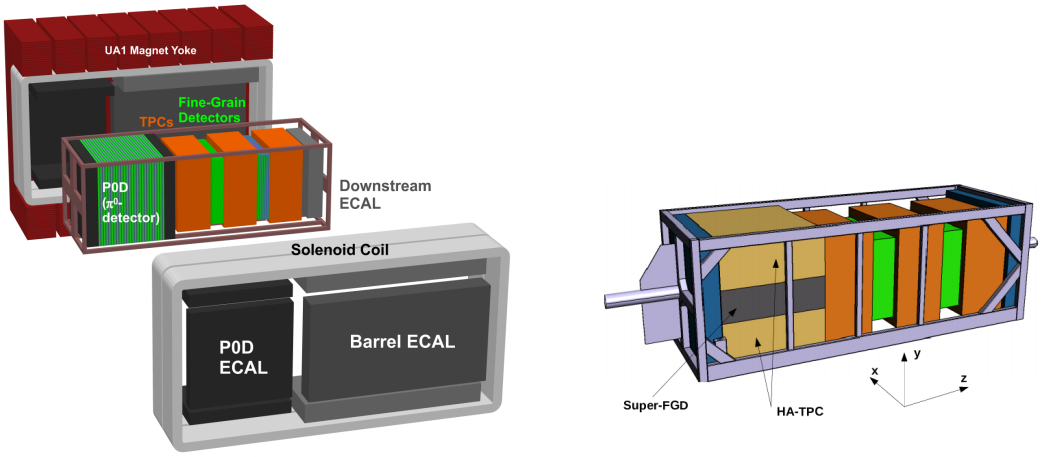

Figure 2.7: (Left) An exploded diagram of the configuration of ND280 with the PØD configuration. (Right) Diagram of the ND280 upgrade, the P0D has been replaced with the Super-FGD and high-angle TPCs. The magnet and time of flight (TOF) detector are not shown. [Figures obtained from [43] and [80]]

cles. Currently, ND280 is the only near detector from which data is used in oscillation analysis.

2.1.8.1 Electromagnetic Calorimiters

ND280 is surrounded by three large electromagnetic calorimeters (ECALs) [81]: the P0D ECAL surrounds the region formerly occupied by the PØD, the Barrel ECAL surrounds the TPCs and FGDs, and the downstream ECAL caps the end of the detector. The ECALs have three purposes: firstly the barrel and downstream ECALs are designed to complement the TPCs by measuring the energy of neutral particles produced within the downstream end of the detector. Since the PØD performed shower reconstruction internally, the PØD ECal is designed to just monitor events that were not fully contained in the PØD. Each ECAL is made up of a total of 13 modules, with each module consisting of several layers of lead plate along with a grid of scintillator bars.

2.1.8.2 Fine Grained Detectors

ND280 contains two Fine Grain Detectors (FGDs) [82] located in between the time projection chambers (TPCs) which act as the target mass for the downstream TPCs. These detectors are also used to measure tracks from charged-current (CC) interactions, particularly from backwards-going and short-lived particles. The detector itself is made of 192 scintillator bars which are arranged in a grid shape. These bars then have a wavelength-shifting fibre passed through them which passes scintillator light from the detector to an MPPC. FGD2 consists of 7 double layers of scintillator and 6 layers of water.

2.1.8.3 Time Projection Chambers

ND280 contains three TPCs which are used to provide particle path reconstruction [83]. The T2K TPCs are argon-based gas-filled detectors with a constant electric potential applied through them by a large cathode plate. Particle interactions generate electrons which are repelled from the cathode towards 12 MicroMegas (Micro-Mesh Gaseous Structure) detectors [84]. The MicroMegas modules use a wire mesh to locate the y-z position of the emitted electrons whilst the constant electrical current allows for an exact reconstruction of the particle's x-position. This x-position reconstruction is possible from timing information obtained by the FGDs. The electrons produced from these events will travel at a constant velocity towards the MicroMegas and, since the interaction time and incoming direction is already known, the positional information can be reconstructed from this information. In addition, since the entire ND280 is magnetised, the particle charge can be calculated using the curvature of particle tracks.

2.1.8.4 Side Muon Ranging Detector

As its name suggests, the side muon ranging detector (SMRD) [85] is used to measure the momenta of muons penetrating the sides of ND280. This serves two purposes:

- Measurement of muons produced by events within the detector that manage to

penetrate the exterior.

- Detection of muons created by cosmic events within the magnet and exterior of ND280.

The SMRD consists of plastic scintillator layers sandwiched between the iron plates around the magnet and readout by MPPCs. This detector is used to veto events generated by muons entering the detector from the sides as well as providing measurements of muons that escape the internal detectors.

The Pi-Zero Detector (PØD) was a large detector that filled most of the upstream region of ND280 [86]. The primary goal of the PØD was to measure pions produced by

$$\nu_\mu + N \rightarrow \nu_\mu + \pi^0 + X + N' \tag{2.13}$$

This forms a particularly important background since NCπ^0 events can be mischaracterised as electron events in SK if only one outgoing photon is detected. The detector consisted of 40 modules (PØDules) consisting of 134 horizontal and 126 vertical scintillator bars. As with many of the other detectors, scintillator light is passed to an MPPC via and WLS fibre. The central region of the detector consists of a repeating water, brass, scintillator structure as shown in figure 2.8. The end caps were used as ECALs which used lead sheets rather than brass and did not contained water. This allowed for the measurement of events in water in order to better constrain errors when the same interactions are measured in the far detector target.

2.1.8.5 PØD

2.1.8.6 ND280 Upgrade

In 2023 ND280 was reconfigured; this new configuration aims to improve the angular coverage of the detector to 4π. In order to do this, the PØD was removed and replaced with three new detectors [Fig. 2.7]. The first of these is the super fine-grain detector (SFGD) [87]. Unlike the older FGDs, the SGFD consists of 2 million $1\times1\times1\,\mathrm{cm}^3$ plastic

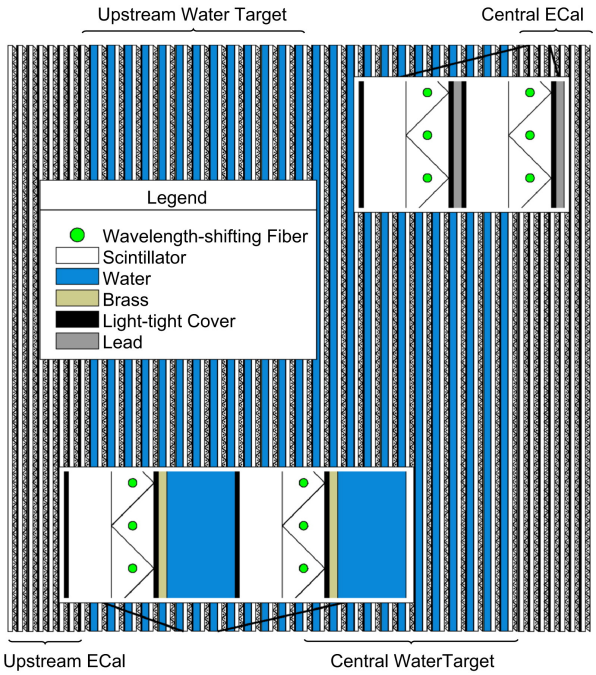

Figure 2.8: Diagram of the P0D, The water/scintillator layering and surrounding ECALs can be seen. Figure obtained from [86]

scintillator cubes arranged in a 192×192×56 grid. Each cube is coated with a thin layer of polystyrene to insulate it and has two holes through which WLS fibres are threaded. The SFGD aims to improve upon the existing FGDs by allowing for the acceptance of backward-going and high angle particle tracks as well as providing a much higher granularity.

The SFGD is sandwiched between two high-angle time projection chambers (HA-TPCs) [88]. The design of these TPCs is similar to those used in the older configuration with the largest difference being the use of resistive MicroMegas [89] and a thinner field cage. The thin cage allows for a larger fiducial volume within the TPCs which increases the capabilities of the detector. The HA-TPCs will allow for the measurements of high-angle charged particles, which were not measured in the previous configuration.

The final piece of the upgrade is the time of flight detector (TOF) [90]. The TOF consists of a series of large plastic scintillator bars surrounding the SFGD, HA-TPCs, and upstream ECal and aims to provide timing information for charged particles leaving

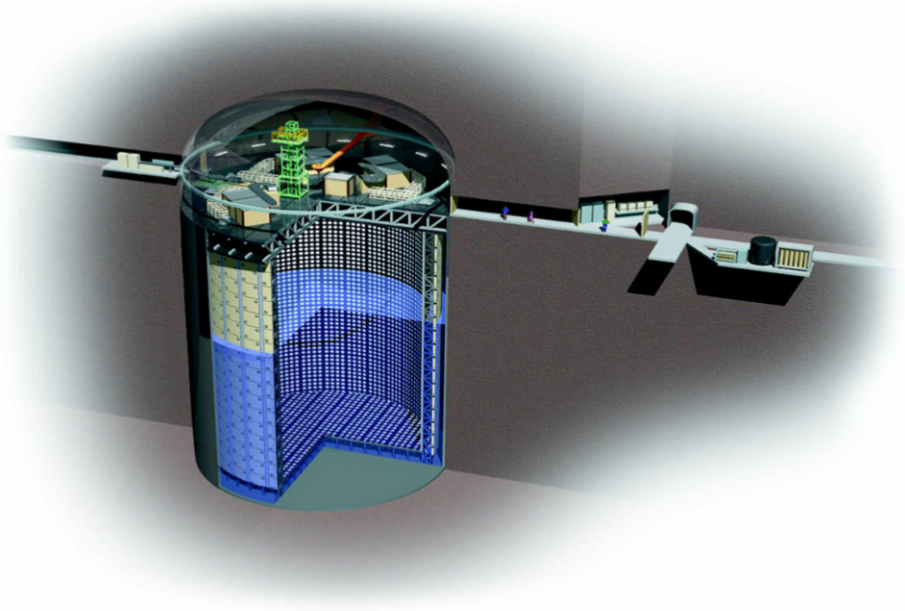

Figure 2.9: Diagrammatic view of the Super-Kamiokande detector. The ID/OD structure can be seen with the Tyvek® separator laying between the scaffolding and the PMTs

and entering the detector. The scintillator signal is readout by a series of MPPCs.

2.1.9 Super-Kamiokande

Super-Kamiokande (SK) is the far detector for the T2K experiment [Fig. 2.9]. As mentioned previously, SK is a 50 kt water Cherenkov detector with 1 km rock overburden to minimise cosmic background interactions. The detector has two distinct regions separated by a wall of Tyvek® [91]. The inner detector (ID) is monitored by 11,146 50 cm PMTs and is used for neutrino event reconstruction. The outer detector (OD), containing 1885 20 cm PMTs [92] is used as a veto region to detect charged particles produced outside the water tank. This improves background rejection since particles produced through neutrino interactions within the ID will not have a preceding OD signal. Detected particles can be identified based on the "fuzziness" of Cherenkov rings seen on the PMTs [93]. Electrons produce "fuzzier" rings as a result of multiple scattering which may induce electromagnetic showers.

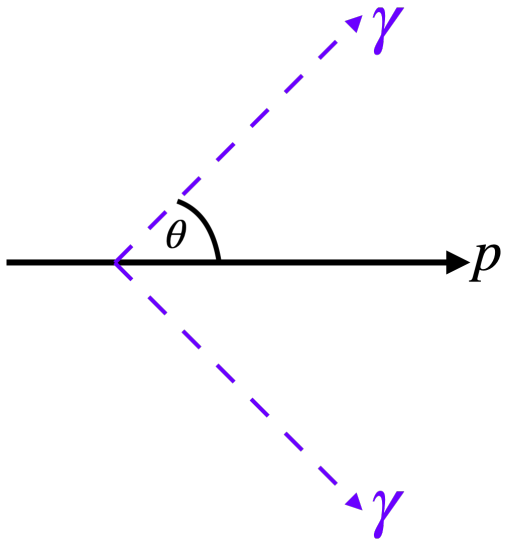

Figure 2.10: Diagrammatic view of the production of Cherenkov radiation. For a particle p, the Cherenkov photons are emitted at characteristic angle given by equation 2.14.

2.1.9.1 Cherenkov Radiation

Cherenkov radiation is produced when particles travel through a medium faster than c/n where n is the refractive index of the material and c the speed of light in a vacuum [Fig. 2.10] [94]. Particles travelling at speed β emit a cone of light analogous to a "sonic boom" which travels in a cone with characteristic angle

$$\theta = \cos^{-1}\left(\frac{1}{n\beta}\right) \tag{2.14}$$

The geometry of Cherenkov cones can be used to both identify particles and measure their velocity with lighter particles producing fuzzier cones. As can be seen in figure 2.11, this leads to clear visual differences between muons and electrons produced within the detector. The SK event discrimination for single-ring events is shown on figure 2.12. Since the addition of gadolinium via $Gd_2(SO_4)_3$ in 2022 [95], SK also has the ability to discriminate between final state particles by charge through neutron capture. This can be used to discriminate incoming neutrino flavour in simple interactions with few final state particles, for example the CCQE interaction $\bar{\nu}p \rightarrow \bar{l}n$ results in a final state neutron whilst $\nu n \rightarrow lp$ does not. This process uses large neutron capture

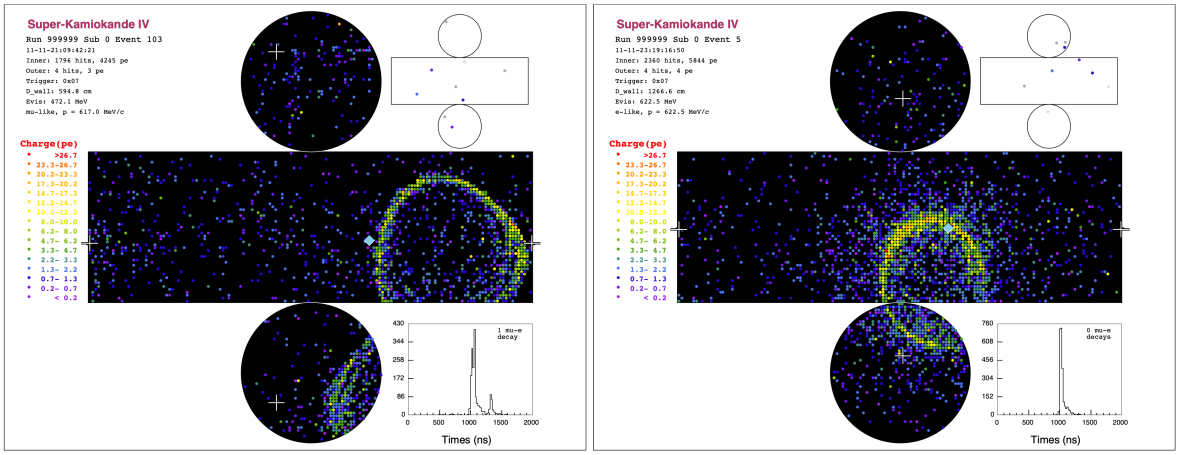

Figure 2.11: Two SK event displays. Both are identifiable as neutrino events by a lack of events in the outer detector (top right) observed to have occurred before the interactions in the inner detector. Flavour can be distinguished from the "fuzziness" of the rings. The left plot shows a μ-like event with a relatively sharp ring whilst the right shows an e-like event with a notably fuzzier image.

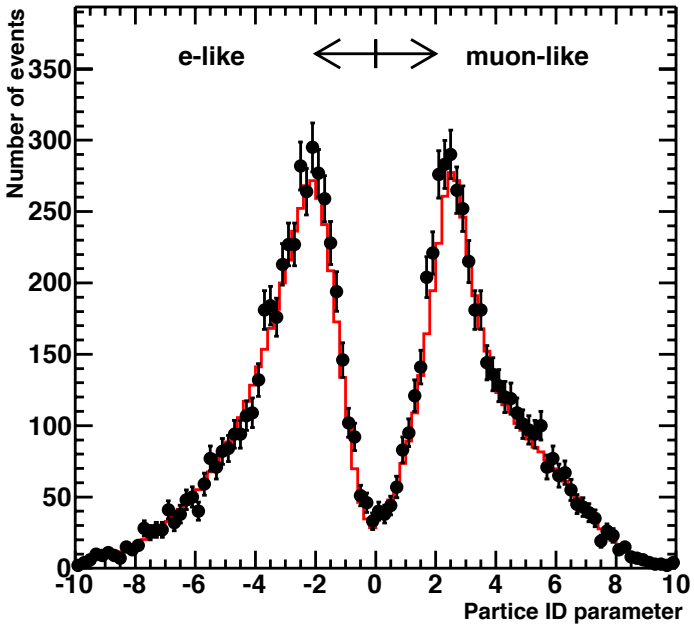

Figure 2.12: Event discrimination for single-ring μ and e like events in SK. The bi-modal distribution from event reconstruction demonstrates the ability of the SK event reconstruction algorithm to separate neutrino flavours.

cross-section of Gd to absorb neutrons produced from neutrino interactions. The Gd is excited into a meta-stable excited state which produces a characteristic 8 MeV photon cascade when it de-excites. This allows the outgoing nucleon, present in many neutrino interactions, to be identified and thus for discrimination of outgoing lepton charge. Primarily this technology is being used for low energy inverse-beta decay ($\bar{\nu}p \rightarrow l^{+}n$) events but there are plans to use it for T2K beam energy events in the future.

2.2 Hyper-Kamiokande

Hyper-Kamiokande (HK) is the successor experiment to T2K and SK [47]. HK will use the same near detector complex and beamline as T2K. The experiment is expected to observe roughly 100× more neutrinos at its far detector than its predecessors [96]. In order to do this the experiment has several notable upgrades. Firstly the J-PARC beam power is being increased to 1.3 MW [63]; this will result in an order of magnitude more neutrinos being observed throughout both the near and far detectors. Work on this has already commenced and, as of March 2024, the beam power has been successfully increased to 750 kW.

Currently, the near detector complex is not expected to differ from that used by T2K. The largest difference is the addition of the IWCD [97]. Located approximately 1 km from the beam target, the Intermediate Water Cherenkov Detector (IWCD) is a planned 10 m×10 m cylindrical Water Cherenkov detector instrumented with 3000 20 cm PMTs filled with 300 t of water. In addition to being a large water target, the detector is allowed to move vertically which allows it to change off-axis angle from 1° to 4°. This will allow it to measure a range of neutrino energy distributions and, as a result, provide a far greater constraint on cross-section parameters than is currently possible with only the T2K ND complex and the HK far detector.

Finally, the far detector (HK-FD)[1] [98]. The detector is 258 kt with a fiducial volume

[1]The names “Hyper-Kamiokande” and “Hyper-Kamiokande far detector” are often used interchangeably, this thesis will use Hyper-Kamiokande to refer to the full long-baseline program and Hyper-Kamiokande far detector when referring to the far detector itself.

8.31 times that of SK. The HK-FD is expected to contain ≈20,000 PMTs and 800 multi-PMTs in the inner detector and 3000 20 cm PMTs in the OD. Construction is expected to finish in 2027.

Chapter 3

Bayesian MCMC

3.1 What is Probability?

The nature of probability has been a source of contention in mathematics for centuries [99]. Formally, probability is defined as a measure P on a space $\mathcal{E}$ that satisfies the Kolmogorov axioms [100]. These state the following, for all sets A, $B \subseteq \mathcal{E}$:

- $P(\mathcal{E}) = 1$: *The space ($\mathcal{E}$) defines the set of all events that could ever occur, hence something must always happen.*
- $A \subseteq B \Rightarrow P(A) \leq P(B)$: *If a set of events contains all the events in another set, it is at least as likely to happen.*
- $P(\mathcal{E} \backslash A) = 1 - P(A)$: *The probability of an event not occurring is equal to 1−the probability of it occurring. Note that since* $P(\mathcal{E}) = 1 \Rightarrow P(\emptyset) = P(\mathcal{E} \backslash \mathcal{E}) = 0$. [1]

This description, however, does not provide an interpretation of probability. Broadly speaking, statisticians have decided on two mutually exclusive interpretations of probability: frequentist and Bayesian [101]. Frequentists interpret probability as a fixed property of a system given by the ratio of the number of times a particular event occurs to the total number of events in a given system [102]. Bayesians, on the other hand,

[1] To clarify notation, $A \backslash B$ is the set of elements in A which are not in B

treat probability as mutable [103]. In the Bayesian interpretation, probability is the "degree of belief" that a particular event will occur and can be updated to account for new evidence. This thesis will primarily focus on the Bayesian approach.

3.2 Bayes Theorem and Bayesian Inference

In Bayesian statistics, probability is a quantity that can be constantly updated. To do this systematically, Bayes' Theorem [104] is utilised in a technique called Bayesian inference [105]. For a data set D and hypothesis θ, Bayes' theorem states that

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{3.1}$$

where $P(\theta|D)$ is known as the posterior, $P(D|\theta)$ the likelihood $P(\theta)$ the prior, and $P(D)$ the marginal likelihood. More intuitively, the prior contains the set of assumptions about the hypothesis, the likelihood describes how well the model suits the data, and the posterior is the updated prior with new evidence. Typically the marginal likelihood is treated as a normalisation constant and often is not computed. Even though this formula appears simple, in reality, the calculation can be complex [106]. In particular, the evaluation of the likelihood is often non-trivial and numerical techniques are required to integrate over the full likelihood space.

3.3 Practical Probability

3.3.1 Introduction

Since the explanation for probability thus far has been very theoretical, this next section will detail the practical differences between Bayesian and frequentist approaches. The canonical example for this is to consider an experiment to determine whether a coin is biased; in this experiment, a coin is flipped 100 times and lands on heads (H) 61 times. The approaches and questions asked by frequentists and Bayesians differ but will end

up reaching similar conclusions.

3.3.2 Frequentist Approach

Let's first consider the frequentist approach to this problem. In order to perform a frequentist hypothesis test we need to propose some hypotheses. For simplicity these will be:

- **Null Hypothesis** (H_0): The coin is unbiased
- **Alternative Hypothesis** (H_1): The coin is biased towards landing on heads

A p-value is then assigned. This gives a maximum likelihood under, below which we reject the null hypothesis; for simplicity we'll set this to be $p_C = 0.05$. In order to evaluate the likelihood of the coin being biased, the frequentist approach looks at the likelihood of a coin being flipped at least 61 times. Since coin flips are distributed binomially, we can find that the probability of coin landing on heads at least as often as observed is:

$$P(h \geq 61 \mid p_0 = 0.5, N = 100) = \sum_{k=61}^{100} \binom{k}{100} 0.5^k (1 - 0.5)^{100-k} = 0.018 \tag{3.2}$$

where h is the number of flipped heads, p_0 the probability of flipping heads under the null hypotheses and N the total number of coin flips. Since $0.018<0.05$, we deem the null hypotheses to be unlikely and hence we reject the null hypotheses in favour of the alternative. More philosophically, the frequentist interpretation of this is that if this test was repeated, only 1.8% of experiments using an unbiased coin would see the coin land on heads at least 61 times.

3.3.3 Bayesian Approach

In Bayesian stats, rather than focusing on a single hypothesis, the focus instead is on the distribution of probabilities of the coin landing on heads. As shown already in

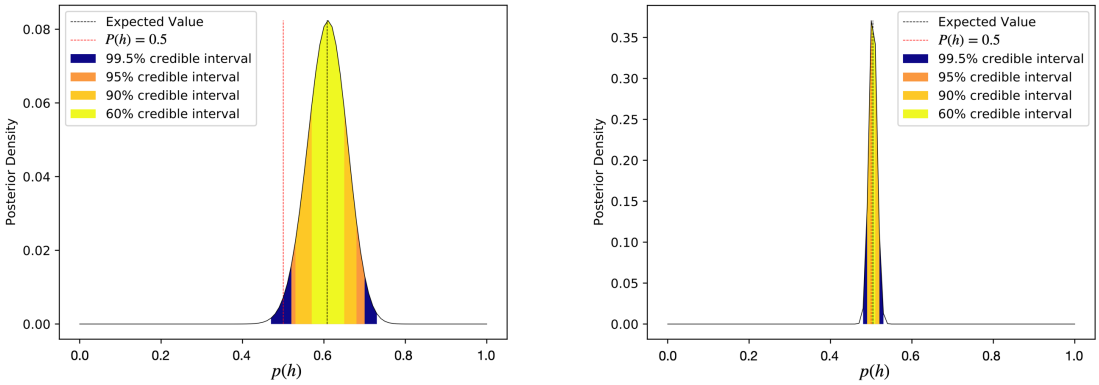

Figure 3.1: Two posteriors are shown for the probability of a coin having a particular bias in the probability of flipping heads ($P(h)$) given 61 coins out of 100 have landed on heads already. On the left is the posterior with a flat prior whilst the right has the posterior with a Gaussian prior centred on $p(h) = 0.5$ with a width of 0.01. The former case excludes the coin being unbiased ($P(h) = 0.5$) at the 95% credible interval whilst coin being unbiased lies within the 60% credible interval of the latter. The highly constraining Gaussian prior is a good example of how priors can potentially bias results since it heavily biases the posterior.

equation 3.1, this distribution can be calculated using Bayes' theorem. In the case of biased coin flips, this can be written as:

$$P(h \mid H = 61, N = 100) \propto P(H = 61, N = 100 \mid h)P(h) \tag{3.3}$$

Where h is a free variable corresponding to the probability of a given coin landing on heads. In order to understand this, we can first consider the likelihood term, $P(H = 61, N = 100 \mid h)$. This is given by

$$P(H = 61, N = 100 \mid h) = \binom{61}{100} h^{61}(1-h)^{100-61} \tag{3.4}$$

Secondly the prior term is designed to encapsulate all reasonably trusted knowledge about biased coins. In this example we consider two priors. Firstly, a relatively uninformed prior, $P(h) = 1$ which just assumes no previous knowledge about coins. Secondly, a Gaussian prior with deviation of 0.01 is picked i.e. $P(h) \sim \mathcal{G}(0.5, 0.01)$. This might correspond to a case where this coin has been randomly picked out of circulation

and there is information from the national mint about coin bias.

The posteriors, which have been normalised such that

$$\int P(h \mid H = 61, N = 100)dh = 1 \tag{3.5}$$

are displayed in figure 3.1. Credible intervals have been drawn in both figures, these correspond to regions where there is a probability of $X\%$ that $P(h)$ lies within them. This means that, in the case of the flat prior, there is at least a 95% chance that the coin is biased whereas, in the more constrained case the coin has a less than 60% chance of being biased.

3.3.4 Comparison of Methods

At this point it is worth considering a comparison of these two techniques. In the case of flat priors both analyses come to essentially identical conclusions i.e. that there is a less than 5% chance that the coin is unbiased. The key difference lies in the questions asked by both schools of thought. Frequentists ask the question "Given a number of experiments, how likely am I to see this result?" whilst the Bayesian approach is to ask "Given I've seen this result,how much do I believe it?". The rest of this thesis will be primarily concerned with Bayesian analyses and the numerical techniques used to perform them.

3.4 Monte Carlo Methods

Unlike the above example, most instances of Bayesian inference do not have "nice" likelihoods which are simple to calculate analytically. For example many physics simulations may require substantial computational time per likelihood calculation. Instead the distribution must be sampled numerically. Many techniques can be used for this, the most popular class of which are Monte Carlo methods [107]. These use pseudo-random number generation to sample the phase space of a distribution and then apply an

accept/reject condition to perform integration over the space. Standard accept/reject Monte Carlo samples a sequence of independent and identically distributed (IID) points across the phase space. The ratio of sampled points within the distribution to the total number of sampled points is then calculated and this will approach the true integral as $N_{\text{points}} \to \infty$.

This method is very efficient when sampling low dimensional spaces, but becomes increasingly inefficient as the number of dimensions gets larger [108]. In order to demonstrate this, consider an N-sphere of unit radius. The volume of the N-sphere decreases rapidly as the spatial dimension gets large [fig. 3.2] [109]. This directly correlates to a decreased number of accepted samples when doing accept/reject Monte Carlo and thus greatly increased run times when attempting to perform integration over high-dimensional spaces. To combat this, more alternative methods need to be applied. One such class of methods is known as Markov Chain Monte Carlo (MCMC) [110].

3.5 Markov Chains

3.5.1 Overview of Convergent Markov Chains

Unlike standard Monte Carlo, where samples are IID, Markov Chain Monte Carlo (MCMC) requires samples to be Markovian [111]. A sequence of samples $\{X_0, \ldots, X_N\}$ is said to be Markovian if, for all steps t,

$$P(X_t|X_{t-1}) = P(X_t|X_{t-1}, X_{t-2}, \ldots, X_0) \tag{3.6}$$

That is, each sample is dependent on the previous sample in the sequence but, crucially, not on any other samples. Such a sequence is also called a Markov chain. If a Markov chain also satisfies the following conditions:

1. Aperiodicity – the chain has no inescapable loops

2. Recurrence – any possible state can be accessed in finite time,

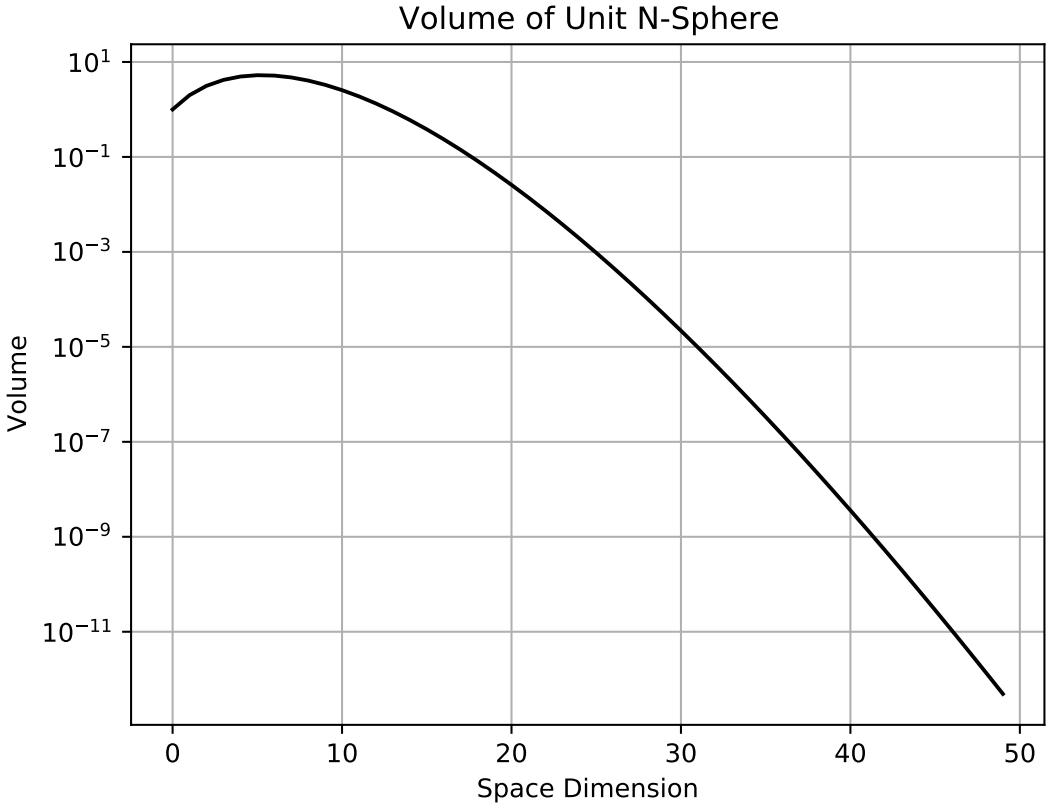

Figure 3.2: Plot showing the volume of an N-dimensional sphere of radius 1 as a function of spatial dimension (N). The volume of the sphere decreases rapidly as N becomes large. As a result, when trying to integrate this object numerically with independent and identically distributed Monte-Carlo throws, the number of accepted throws will decrease correspondingly. This slows the convergence of these methods and is known as the curse of dimensionality.

then the chain is said to be ergodic and the Markov chain process will eventually converge on some stationary distribution [112]. Bayesian inference exploits this property by performing MCMC with the posterior as the final stationary distribution.

3.5.2 Stationary Distributions

Before looking at the practical uses of MCMC, it is informative to look briefly at the properties of Markov chains in general. As stated in the previous section, ergodic Markov chains will converge to a stationary distribution. That is to say that for a Markov process $M(X_t) = X_{t+1}$, there is some probability distribution π such that $X_t \sim \pi$ produces an equivalent distribution of samples (under the law of large numbers) to the Markov process and vice versa. π is known as the stationary distribution of the chain.

3.5.3 Markov Chain Central Limit Theorem

As with IID processes, convergent (ergodic) Markov chains have a central limit theorem (CLT). The proofs for the following statements can be found in [113]. For a Markov chain $\{X_k\}_{k=0}^{\infty}$ such that $X_i \in \mathbb{X}$ and has stationary distribution π the following holds, $\forall\ f : \mathbb{X} \to (-\infty, \infty)$:

$$\bar{f}_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i) \tag{3.7}$$

Since the chain is ergodic, $X_i \sim \pi$ and hence,

$$\lim_{n\to\infty} \bar{f}_n = \mathbb{E}_\pi[f] = \int_{\mathbb{X}} f(x)\pi(dx) \tag{3.8}$$

i.e. the expected value of f integrated over the full parameter space "weighted" by the stationary distribution[2]. The CLT then states the following

$$\lim_{n\to\infty}\left[\sqrt{n}(\bar{f}_n - \mathbb{E}_\pi)\right] \to \mathcal{G}(0, \sigma_f^2) \tag{3.9}$$

where

$$\sigma_f^2 = \mathrm{var}_\pi f(X_0) + 2\sum_{k=0}^{\infty} \mathrm{cov}_\pi[f(X_0), f(X_k)] \tag{3.10}$$

Where $\mathrm{var}_\pi(\circ)$ and $\mathrm{cov}_\pi(\circ,\circ)$ are the variance and covariance of f "weighted" by π and $\mathcal{G}$ is the normal distribution. Concretely this means that, eventually, Markov Chains can be shown to display normally distributed mixing behaviour which is incredibly useful for evaluating errors as a result of MCMC non-convergence.

3.6 Markov Chain Monte Carlo

3.6.1 Metropolis Hastings

Algorithm 1 Metropolis Hastings Algorithm

Input: Initial point x_0, proposal function $Q(\circ,\circ)$, total number of step to run for N_{steps} and likelihood function $\mathcal{L}(\circ)$. Let t=0 for the first step.
while $t < N_{steps}$ **do**
  Propose $y \sim Q(x_t, .)$
  Evaluate $\alpha = \min(1, \frac{\mathcal{L}(y)Q(y|x_t)}{\mathcal{L}(x_t)Q(x_t|y)})$
  Generate $u \sim U(0,1)$
  if $\alpha > u$ **then**
    $x_{t+1} = y$
  else
    $x_{t+1} = x_t$
  $t = t + 1$

One of the simplest implementations of MCMC is a technique known as the Metropolis-Hastings (MH) algorithm [Alg. 1] [115]. Conceptually, this method can be thought

[2]More precisely this is the integral under the probability measure of π. An understanding of measure theory is unnecessary for this thesis, for a comprehensive introduction to this field please see [114]

of as a random walk across the likelihood space where new steps, y, are proposed using some probability distribution (dependent on the previous step), $Q(x_t, .)$, and are accepted/rejected with a probability

$$\alpha = \min\left(1, \frac{\mathcal{L}(y)Q(y|x_t)}{\mathcal{L}(x_t)Q(x_t|y)}\right) \tag{3.11}$$

where x_t is the current step, $Q(.,.)$ the proposal function, and $\mathcal{L}(.)$ the likelihood. There is a considerable amount of freedom allowed when choosing the proposal function, but, the typical choice is to pick a Gaussian centred on x_t. Not only is this simple to evaluate, but it has the additional property that $Q(x_t|y) = Q(y|x_t)$. The acceptance ratio can thus be reduced to

$$\alpha = \min\left(1, \frac{\mathcal{L}(y)}{\mathcal{L}(x_t)}\right) \tag{3.12}$$

MH with this condition is commonly referred to as Metropolis rather than Metropolis-Hastings. In addition to choosing the class of proposal function, there are often hyper-parameters associated with the class which can affect the convergence speed of the MCMC algorithm. In the case of Gaussian proposal functions, the covariance matrix of the proposal Gaussian is a free variable and is referred to as the step size [116].

3.6.2 Adaptive Metropolis Hastings

The step size chosen in Metropolis-Hastings can have a drastic effect on the rate of convergence, with poorly tuned hyper-parameters potentially generating incorrect results due to the large convergence times. Though step sizes can be chosen manually through trial and error, for most "well-behaved" posteriors the optimal proposal function covariance is proportional to the covariance of the posterior [117]. Whilst this cannot be calculated exactly, the covariance of the MCMC will eventually converge on this matrix; as a result, the proposal function can be updated iteratively to be proportional to this [Alg. 2] [118]. The scaling factor of $\frac{2.38^2}{d}$ is chosen to optimise this algorithm for Gaussian proposals targetting Gaussian posterior distributions [117]. It turns out that

this optimisation is sufficient to ensure convergence for most well-behaved posteriors.

Algorithm 2 Adaptive Metropolis-Hastings

Input: Initial point x_0, proposal function $Q(\circ, \circ)$, total number of step to run for N_{steps} and likelihood function $\mathcal{L}(\circ)$. Let $t = 0$ for the first step.
while $t < N_{steps}$ **do**
  Propose $y \sim \mathcal{G}(x_t, \frac{2.38^2}{d}\Sigma_t)$ where $\Sigma_t = \mathrm{cov}(x_0, \ldots, x_t)$ and $d = \dim \Sigma_t$
  Evaluate $\alpha = \min\left(1, \frac{\mathcal{L}(y)Q(y|x_t)}{\mathcal{L}(x_t)Q(x_t|y)}\right)$
  Generate $u \sim U(0, 1)$
  if $\alpha > u$ **then**
    $x_{t+1} = y$
  else
    $x_{t+1} = x_t$
  Calculate Σ_{t+1}
  $t = t + 1$

This implementation is effective in many cases but does not guarantee convergence since the updated covariance can remove the ergodic property from the fit [117]. It is possible to mitigate this with a few modifications to the adaption process. The first and simplest is to not adapt every step. Instead, while the covariance is calculated every step, the matrix used for step proposals is only updated every N_{update} steps [Alg. 3] which results in the chain "remembering" fewer previous steps and thus being allowed more freedom to explore the space.

Additionally, the following modification can be made to the proposal function; instead of proposing purely from a single Gaussian, the proposal function can be changed to

$$(1 - \beta)\mathcal{G}\left(x_t, \frac{2.38^2}{d}\right) + \beta\Sigma_n\mathcal{G}(0, \sigma^2) \tag{3.13}$$

where σ and β are some small user inputted real numbers [117]. The primary motivation for this modification is to account for the error caused by the current matrix not being fully converged to the true posterior covariance. Alternatively, as proposed by Haario,

Algorithm 3 Adaptive Metropolis-Hastings with Delayed Covariance Updates

Input: Initial point x_0, proposal function $Q(.,.)$, total number of step to run for N_{steps}, likelihood function $\mathcal{L}(.)$, number of steps between updates N_{update}. Let $t = n = 0$ for the first step.

while $t < N_{steps}$ **do**

- Propose $y \sim \mathcal{G}(x_t, \frac{2.38^2}{d}\Sigma_t)$ where $\Sigma_t = \mathrm{cov}(x_0, \ldots, x_t)$ and $d = \dim \Sigma_t$
- Evaluate $\alpha = \min\left(1, \frac{\mathcal{L}(y)Q(y|x_t)}{\mathcal{L}(x_t)Q(x_t|y)}\right)$
- Generate $u \sim U(0,1)$
- **if** $\alpha > u$ **then**
  - $x_{t+1} = y$
- **else**
  - $x_{t+1} = x_t$
- Calculate Σ_{t+1}
- **if** $t|N_{update}$ **then**
  - $\Sigma_{n+1} = \Sigma_{t+1}$ Update the proposal covariance matrix
  - $n = n + 1$
- $t = t + 1$

Saksman, and Tamminen [118], the matrix can be updated in such a way that

$$y \sim \mathcal{G}\left(x_t, \frac{2.38^2}{d}\Sigma_t + \epsilon I_d\right) \tag{3.14}$$

where ϵ is some small positive number and I_d the $d \times d$ identity matrix. This constant contribution, whilst not as good an error approximation, still prevents the matrix from becoming singular and hence is often sufficient to ensure convergence. In either case, there are still several potential issues that may lead to this technique producing erroneous results [119]. The most common issue is a posterior distribution having multiple disconnected modes. In this instance the proposal function may adapt to one mode and prevent the chain exploring the others. These issues may lead to the adaptive chain being non-ergodic and thus not converging to the expected stationary state i.e. the posterior.

3.6.3 Jumping Adaptive Metropolis Sampler

The Jumping Adaptive Metropolis Sampler (JAMS) [120] is an extension to the standard adaptive Metropolis-Hastings algorithm which aims to handle convergence issues encountered as a result of multi-modal posterior distributions. The full algorithm is complex but the basic implementation can be summarised in the following way (we assume Gaussian proposal functions for local steps)

- Modes are identified using a simple maximum likelihood fit.
- For each mode, a burn-in chain is run with that mode fixed. This chain uses the adaptive Metropolis-Hastings algorithm [Alg. 2].
- A new chain is run using the step proposal in equation 3.15 for jump steps and adaptive Metropolis Hastings for local steps.
- The probability of a chain performing a jump step is a fixed number p_J. This means that each step has a probability p_J of proposing a move into another mode.

In order to propose steps in different modes (jump steps) to regions of equivalent likelihood, the following proposal function is used:

$$y = \mu_k + \Lambda_k \Lambda_i^{-1}(x_t - \mu_i) \tag{3.15}$$

where y is in mode k which has central value μ_i and x_t in mode i. The matrix Λ_k is the (non-unique) matrix that satisfies the following relationship

$$\Sigma_k = \Lambda_k \Lambda_k^T \tag{3.16}$$

where Σ_i is the covariance matrix used for throws in the Markov Chain. In all instances where this algorithm is used, Λ_k is the Cholesky decomposition of Σ_k; since all covariance matrices are trivially real valued this satisfies the required condition. Since this

is non-random a different step acceptance function is also required. In this case, jump steps from x_t in mode i to y in mode k are accepted with probability

$$P(\text{accept}) = \min\left[1, \frac{\mathcal{L}(y)\mathcal{G}(\mu_i, \Sigma_i)P(k \to i)\sqrt{\det \Sigma_k}}{\mathcal{L}(x_t)\mathcal{G}(\mu_k, \Sigma_k)P(k \to i)\sqrt{\det \Sigma_i}}\right] \tag{3.17}$$

where $\mathcal{G}(\mu_i, \Sigma_i)$ is a normal distribution with variance given by the covariance of the chain in mode i which has central value μ_i. Note that this is simply the proposal function for a given state. Secondly the transition probability of $p(i \to k)$ is just the probability transitioning between two states.

The initial paper that proposes JAMS [120] also suggests updating distributions to prevent proposing steps in the incorrect mode. This has been omitted since the main use case for JAMS within MaCh3 is for highly disjoint bi-modal distributions, for example Δm^2_{32}.

Algorithm 4 JAMS Algorithm

Input Parameters: Current step position and mode (X_t, m), list of modes, probabilities to jump into each mode and covariance matrices for each mode $(\mu_i, p(\mu_i), \Sigma_i)$, probability of proposing a jump step, p_J and list of posteriors covariance matrices for each mode
Generate a random number $u_J \sim U[0, 1]$
if $u_J > P_J$ **then**
    Propose Local Step:
    Propose and accept standard adaptive Metropolis Hastings step [Alg. 2]
else
    Propose Jump Step:
    Propose a new mode where a mode μ_i has probability $p(\mu_i)$ of being selected
    Propose a new step using equation 3.15
    Accept step with probability calculated from 3.17

This algorithm is considerably more computationally expensive than other adaptive regimes but has the advantage of improved efficiency on multi-modal spaces which standard adaptive MCMC may struggle with.

3.7 Markov Chain Convergence Metrics

3.7.1 Trace and Acceptance Rate

MCMC convergence is highly dependent on the choice of hyper-parameters (step size, number of steps, etc.). This can lead to slow convergence or, in extreme cases, a sufficiently poorly tuned chain will lose its ergodicity entirely and hence not have guaranteed convergence to the correct target distribution. As a result, it is insufficient to claim convergence without closely monitoring the fit. The simplest metric available is to look at the step acceptance rate. For high-dimensional MCMC, it can be shown that the most efficient sampling rate is 23.4% [121]. In practice, this is hard to achieve but acceptance rates that deviate significantly from this are indicative of a poorly tuned fit for example a rate of 1% would indicate that the chain is not exploring the space effectively.

The Markov chain's exploration of parameter space can be examined directly using a trace plot [122]. The trace is a plot of a parameter value at each step of the chain. Figure 3.3 shows the traces of two Markov chains, one of which is well-tuned and the other poorly tuned; as a result, the differences in parameter space exploration between these two fits are stark. The poorly tuned chain has very low local variance and explores the space very slowly, whereas the well-tuned chain can be seen rapidly exploring the entire space. Although trace is a useful indicator of convergence, a well-behaved trace plot is not a guarantee that the chain is properly exploring the space: for example, parameters stuck in local minima can generate well-behaved trace plots. In order to minimise the impact of this, additional metrics are also required. The most commonly used metric is autocorrelation, which measures how correlated steps in the chain are with each other.

3.7.2 Autocorrelation

Recall from section 3.5.1 that MCMC convergence is guaranteed if the fit is an ergodic Markov chain [123]. Since the dependence of each step on the previous steps in the chain is partly dependent on the tuning of hyper-parameters in the fitter, poorly tuned fits are often not fully Markovian. Correlations between steps in the same chain can be measured using autocorrelation. This can calculated through

$$f_A(l) = \frac{1}{N}\sum_{k=l}^{N} A(k, \bar{k} - l) \tag{3.18}$$

Where $A(k, k-l)$ is simply the average correlation between steps k and $k-l$ across all steps. An example of the autocorrelation function for both a well behaved and poorly chain can be seen in figure 3.3.

3.7.3 Burn-in

MCMC fits require some time to converge on the optimal starting location; this requires the chain to traverse many low-likelihood regions before converging on the true stationary distribution. This stage of the fit is commonly referred to as burn-in. The characteristic burn-in length can be measured [116] by looking at the stability of contours generated by the fit and proposed likelihoods around steps. Once these variations are deemed sufficiently small, all earlier steps are discarded. Whilst this is standard practice for many MCMC fitters, there exist methods to avoid it, for example by calculating regeneration times [124] or similar metrics.

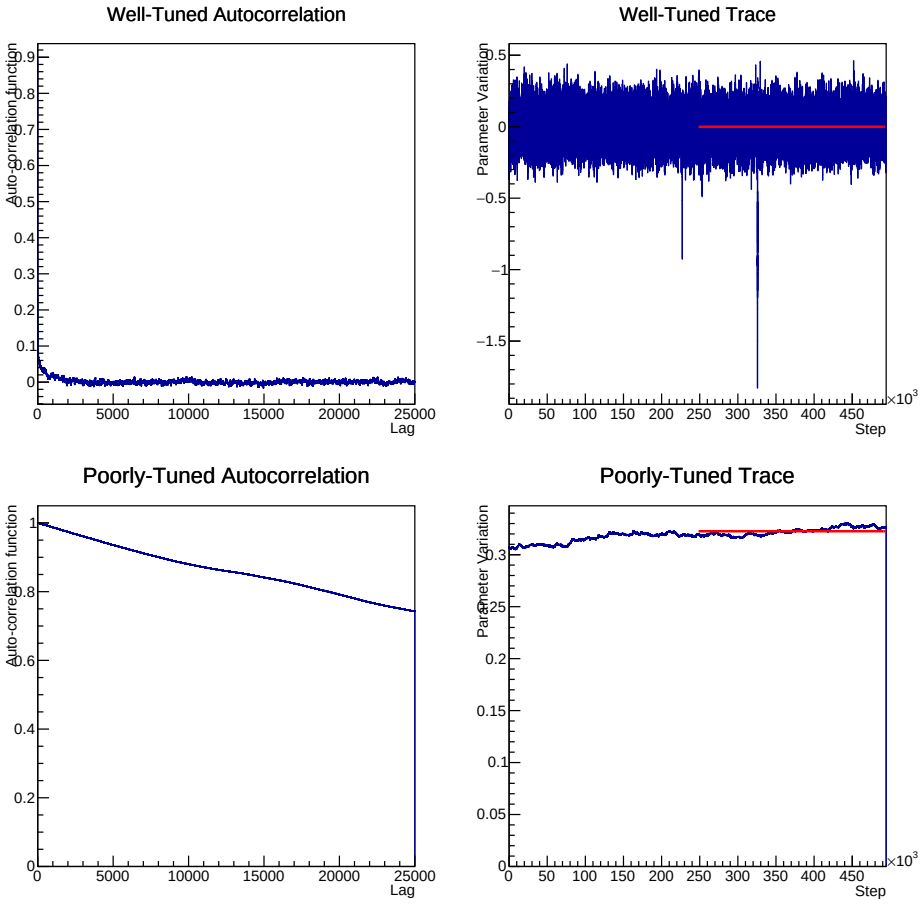

Figure 3.3: Trace (left) and autocorrelation (right) plots for well-tuned (top) and poorly-tuned (bottom) parameters in an MCMC fit. The well-tuned autocorrelation rapidly tends towards 0 autocorrelation while the poorly-tuned parameter maintains a large autocorrelation even at large values of lag. The trace for the well-tuned parameter varies rapidly but statically about the best-fit value (red line) while the poorly-tuned trace has very little variation and obvious trends are visible.

3.8 Analysis of MCMC Fits

3.8.1 Point Estimates and Posteriors

Bayesian results are usually presented in terms of credible intervals [125], the region of the posterior containing a specified percentage of the MCMC steps. Calculation of this for MCMC simply requires counting the number of MCMC steps in a given region of

space. Of course, picking a random region of space with 90% of the MCMC will not correspond to any useful interpretation of this result. Instead, posterior calculation has to be done systematically. For an N-dimensional posterior, the mean (central) value can be found by integrating over the full posterior. $X\%$ credible intervals are then constructed by selecting bins with the largest the number entries which contain $X\%$ of the data. The Bayesian interpretation of this is that events within the credible interval have an $X\%$ chance of occurring. The central value is known as the point estimate.

Although evaluating posterior credible intervals over the full posterior can be useful, it is often more practical to reduce the dimension of the space in order to both visualise posteriors graphically and remove nuisance parameters, i.e. parameters that are important to the fit but not to the final result.[3] This is done by integrating over the posterior to "average" away nuisance parameters in a process called marginalisation, i.e.

$$P(\theta_{\text{marg}} \mid D) = \int P(\theta_{\text{marg}}; \theta_{\text{nuis}} \mid D) d\theta_{\text{nuis}} \tag{3.19}$$

These marginalised posteriors can then be treated identically to the full posterior and used to obtain point estimates and credible intervals.

3.8.2 Bayes Factors

Unlike frequentist statistics, which requires a complicated hypothesis testing apparatus [102], the basic Bayesian hypothesis test is relatively simple. For two hypotheses θ_1 and θ_2 with equal priors, the Bayes factor $\mathcal{B}$ is just the ratio of the posteriors [126]

$$\mathcal{B} = \frac{P(\theta_1|D)}{P(\theta_2|D)} \tag{3.20}$$

Whilst an exact mapping from Bayes factor to particle physics-like σ value does not exist, table 3.1 shows the mapping used within T2K and T2K+NOvA. In statistics it is more common to use the Jeffrey's scale [127], but since the 5σ discovery threshold

[3] Typically all systematic parameters and any physics parameters that are not being directly analysed

Particle Physics Value	Bayes Factor Ratio
2σ	20.74
3σ	369.4
4σ	15800
5σ	1745000

Table 3.1: Table showing the equivalent value of equation 3.20 required for a given hypothesis to be preferred to the same degree as a standard particle physics σ value. This will be referred to as Kaboth-Dunne Scale to follow T2K convention. Table taken from internal technical note 435 [128].

corresponds to a Bayes factor of 1745000 and the maximum evidence in the Jeffrey's scale is given by a Bayes factor greater than 100 ($\approx 3\sigma$) this more standard scale has limited use.

Bayes factors can be obtained from a Markov chain by counting the number of steps that correspond to a given hypothesis and two hypotheses can be compared by looking at the ratio of steps of two distinct hypotheses. For example, the Bayes factor for mass ordering can be obtained through

$$\mathcal{B}_{MO} = \frac{\#\text{steps in NO}}{\#\text{steps in IO}} \tag{3.21}$$

3.8.3 Uncertainties

Since a Markov Chain is a stochastic process, it is important to understand the uncertainties generated by the algorithm itself. The simplest measure of error is the Markov Chain Standard Error (MCSE) [110],

$$\sigma^2_{\text{MCSE}} = \frac{1}{N}\sum \text{tr}\left(\Sigma\right)^2 \tag{3.22}$$

where Σ is the covariance of the chain and N is the number of steps. This is done under the assumption that the Markov Chain central limit theorem holds [113], which it does for almost all instances of MCMC. This provides a useful metric for the convergence of the chain as a whole, but it is often the case that the chain incorrectly fits individual

parameters. In order to assess this, in addition to using the metrics mentioned above (trace, autocorrelation etc.), the uncertainty can be quantified using effective sample size (ESS) [129]. ESS is a measure of sample size given that there is uncertainty within the chain generated by autocorrelations. It can be calculated through

$$N_{\text{eff}} = \frac{N}{\sum_{l=0}^{N} f_A(l)} \tag{3.23}$$

i.e. the sum over the autocorrelation function $f_A(l)$. The error can then be expressed in terms of the effective sample size instead of the true sample size

$$\sigma^2_{\text{eff}} = \frac{1}{N_{\text{eff}}} \sum \text{tr}\,(\Sigma)^2. \tag{3.24}$$

3.8.3.1 Goodness of Fit

Once a fit has been performed, it is essential to understand how well the Markov chain fits the data. Several techniques exist to evaluate this; MaCh3 uses posterior-predictive p-values and compares the Monte Carlo-generated model to the true data with the null hypotheses corresponding to good model/data agreement. In order to generate a Bayesian p-value, N steps are chosen at random from the chain and the model is evaluated at these points. The p-value is then calculated using the technique from [116]

$$p \approx \frac{1}{N} \sum_{i=1}^{N} I\Big[T(\text{model}) > T(\text{Data})\Big] \tag{3.25}$$

where T is a test statistic and I the indicator function, which is 1 if the condition is true and 0 if it is false. If the model and data match perfectly, it is expected that $p = 0.5$ since it would be expected that $T(\text{model}) > T(\text{data})$ half the time on average. To understand this more concretely, it is helpful to consider the MaCh3 implementation of this technique. As with the general technique, N steps are selected from the chain. The model predictions are then the MC spectrum for each sample re-weighted to the value of the chain at each of these randomly generated points. Initially, a sub-sample of events

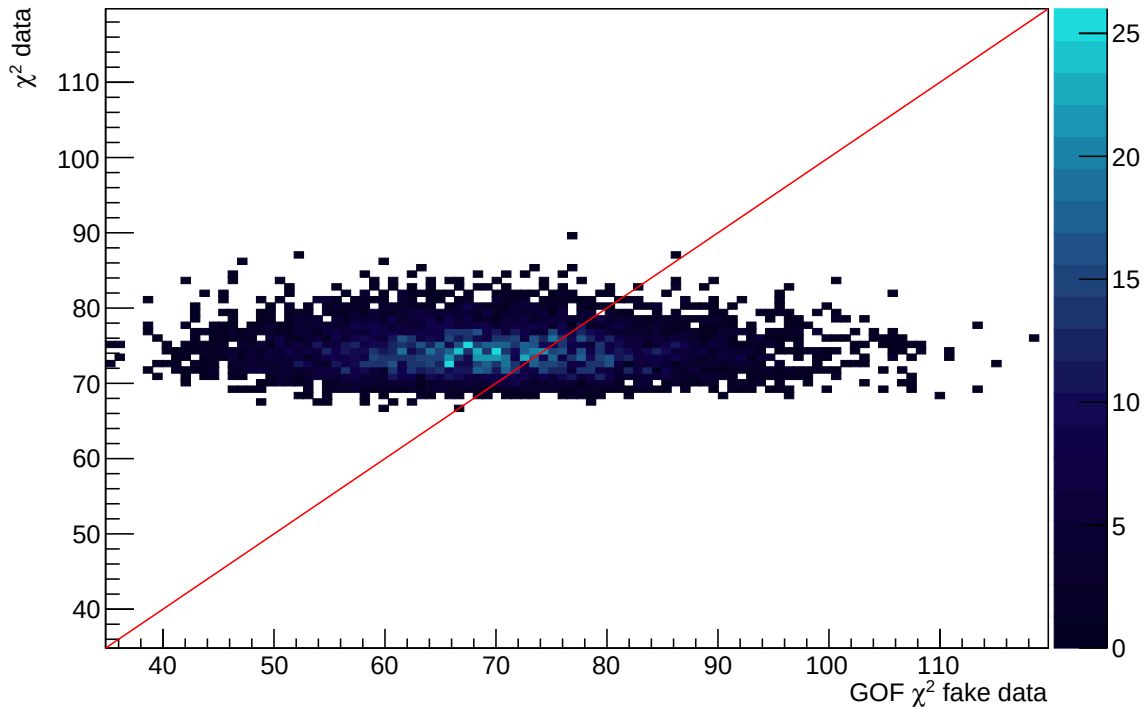

Figure 3.4: Rate+Shape goodness of fit plot for a T2K far detector sample. The distribution of $\Delta\chi^2$ for data and fake data are expected to be evenly distributed about $y = x$ (red) for a well behaved model.

is drawn from the model spectra; the total number of events is given by the Poisson distribution $N_{\text{draw}} \sim P(\int \text{spectrum})$. This then forms a fake test spectrum. The test is given by a $\Delta\chi^2$ comparison between the model spectrum and data spectrum. $\Delta\chi^2$ can be calculated in two ways: firstly, by the difference between the total number of events in each spectrum can be calculated; this is known as the rate information. Secondly, the data, model and fake MC can be binned and the total $\Delta\chi^2$ difference for each bin is used instead; this is known as rate+shape. In either case, the total p-value is then given by

$$p \approx \frac{1}{N}\sum_{i=1}^{N} I\Big[\Delta\chi^2(\text{model} - \text{fake}) > \Delta\chi^2(\text{data} - \text{fake})\Big] \tag{3.26}$$

In addition to the p-value, it is also useful to plot $\Delta\chi^2$(data) against $\Delta\chi^2$(model) as seen in figure 3.4. This gives a clear distribution of points which, for a good fit, should be distributed evenly about $y = x$.

Chapter 4

Oscillation Analyses at T2K

4.1 Overview of the T2K Oscillation Program

The T2K oscillation analysis program is led by two key analyses:

- **Near Detector (ND) only**: Focuses on constraining systematics related to fluxes, cross-sections and the near detector within the fit
- **Far + Near Detector (FD+ND)**: Focuses on constraining neutrino oscillation parameters using far+near detector constraints

The rest of this chapter will focus on the latter. The FD+ND fit consists of two fitters, PTheta, a frequentist fitter which performs fits using far detector data and has an ND constraint provided from the ND-only fit, and MaCh3 which fits FD+ND data simultaneously using Markov chain Monte Carlo (MCMC). This group structure allows for two semi-independent analyses to be performed and the results compared to ensure consistency. In addition, since the groups use radically different approaches, an agreement between frequentist likelihoods and Bayesian posteriors enhances the robustness of the result.

This chapter will detail the full FD+ND fit pipeline from implementation validations to the final T2K oscillation result.

4.2 Data Inputs

4.2.1 Far Detector Inputs

4.2.1.1 Overview

The far detector fits 6 'samples' based on Cherenkov rings seen in Super-Kamiokande (SK). These are split into μ-like samples and e-like samples. The μ-like configuration has two samples correspond to events with a single muon-like Cherenkov ring in both the ν-mode and $\bar{\nu}$-mode beam as well as a sample corresponding to events with multiple Cherenkov ring. This final sample corresponds to two possible event topologies as can be seen in figure 4.1 and is used to quantify CC1π events. The e-like also samples consist of two samples which represent events with 1 electron-like Cherenkov ring in the final state and a third 1Re1 decay e sample which represents events with a two electron-like rings produced through CC-pion production. All single ring samples are assumed to produce their final state lepton through charged-current quasi-elastic (CCQE) scattering.

Whilst the data is simply split into one of these 6 samples, the Monte Carlo-generated model samples are also split into 6 oscillation channels (expected start and end states for each neutrino): $\nu_\mu \rightarrow \nu_\mu$, $\bar{\nu}_\mu \rightarrow \bar{\nu}_\mu$, $\nu_e \rightarrow \nu_e$, $\bar{\nu}_e \rightarrow \bar{\nu}_e$, $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$. These represent neutrinos that start in an initial state ν_i and oscillate to a new state ν_j which is observed in the far detector. Each μ-like sample is binned in terms of true energy (E_{true}) and reconstructed energy (E_{rec}) whilst the e-like samples are also binned in terms of outgoing lepton angle. Binning varies depending on both the sample flavour and its oscillation channel. The reconstructed energy calculation varies depending on sample; for the 1Re/μ, the interaction is assumed to be CCQE, hence,

$$E_{\text{rec}} = \frac{m_p^2 - (m_n - E_b)^2 + m_l^2 + 2(m_n - E_b)^2 E_l}{2\Big[m_n - E_b - E_l + p_l \cos(\theta_l)\Big]} \tag{4.1}$$

where m_p, m_n and m_l correspond to the proton, neutron, and outgoing lepton masses, E_b the nuclear binding energy, and p_l and θ_l the outgoing lepton momentum and angle.

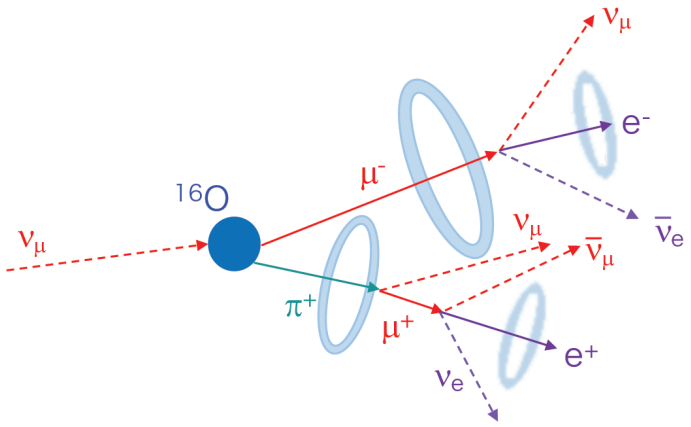

Figure 4.1: Diagram showing a ν_μCC1π event. The incoming ν_μ generates a Δ resonance (not in diagram) through CC deep inelastic scattering. The produced μ^- will decay to produce two rings, one from μ^- and a secondary ring from the decay electron. In addition the π^+ may produce between 1 and 2 rings dependent on its energy, with the intermediate μ^+ assumed to be below the Cherenkov threshold. This naturally creates two samples separated by the number of decay electrons. Plots taken from an internal technical note [130]

The 1Re1d.e sample assumes the electron is generated through charged current pion production via a Δ^{++} resonance and thus can contain final state pions. The reconstructed energy for this sample can be given by

$$E_{\text{rec}} = \frac{2m_n E_e + m_{\Delta^{++}}^2 - m_n^2 - m_e^2}{2[m_n - E_e + p_e \cos(\theta_e)]} \tag{4.2}$$

Unlike the other samples, the ν_μCC1π selection consists of several sub-samples split by the number of decay electrons in the final state. The diagram in figure 4.1 shows the possible topologies considered for these events. As with the 1Re1d.e. sample, fiTQun assumes that these events are generated through CC interaction with a Δ-baryon resonance [131]. As this hypothesis is the same is the 1Re1d.e. case, the reconstructed energy calculation is nearly identical,

$$E_{\text{rec}} = \frac{2m_n E_\mu + m_{\Delta^{++}}^2 - m_n^2 - m_\pi^2}{2[m_n - E_\mu + p_\mu \cos(\theta_\mu)]} \tag{4.3}$$

The two major differences are the primary outgoing lepton which is a μ^- and the

Parameter	Nominal
$\sin^2(\theta_{12})$	0.307
$\sin^2(\theta_{23})$	0.561
$\sin^2(\theta_{12})$	0.0220
Δm^2_{21}	7.53×10^{-5} eV^2
Δm^2_{32}	2.494×10^{-3} eV^2
δ_{cp}	-1.601 rad

Table 4.1: The Asimov A22 parameter set.

inclusion of the π^+ mass in the calculation.

4.2.1.2 Aside on Jargon

In order to make explanations more compact the following jargon will be used. Firstly, when discussing the nominal values of cross-section parameters, three configurations are used:

- **Generated Values**: Systematic nominal set to best fit value given by the event generator, this results in no changes to the input MC model.
- **pre-BANFF Values**: Systematic Nominal set to a suggested value given by the Neutrino Interaction Working Group (NIWG) rather than the generator. This can differ from the generated value and will result in systematic re-weighting for some parameters.
- **post-BANFF Values:** Nominal value is now determined by the ND fit best-fit value. This will change the value of all cross-section systematic parameters.

Where BANFF here refers to the T2K ND-only fitter which provides the best-fit value. Secondly, the configuration of oscillation parameters will be referred to as Asimov A22 which corresponds to the global PDG best-fit values [55] assuming normal ordering. These are listed in table 4.1.

4.2.1.3 Event Selection

Event reconstruction is performed by the T2K+SK group using the fiTQun reconstruction toolkit. All samples have a fiducial volume cut that ensures the event occurred within the fiducial volume of the SK inner detector. In addition, the 1Re, 1Rμ and 1Re1d.e. samples have cuts ensuring that the Cherenkov ring is the correct flavour. The e-like samples are further selected in the following way

- The visible energy (total energy of reconstructed final state particles) is larger than 100 MeV.
- The reconstructed neutrino energy is less than 1.25 GeV.
- Events do not contain a π^0.

The 1Re and 1Re1de samples are then distinguished via the reconstruction of a single decay electron caused by π^+ decay such that the 1Re samples do not contain this electron. The single ring μ-like samples have a similar set of selection cuts to the 1Re samples with the following differences:

- Reconstructed μ momentum >200 MeV/c.
- Fewer than two decay electrons.
- The reconstructed energy cut is replaced by a cut at 30 GeV.
- Rather than not containing π^0 events are selected to not include π^+ decays.

The 30 GeV E_{rec} cut exists effectively as an upper bound at $+\infty$ and is there to prevent the model re-weighting large numbers of events far outside the expected data energy range for T2K beam data. The ν_μCC1π sample has cuts dependent on the final state topology of events contained in the sample:

- Events are fully contained in the fiducial volume of the detector.
- The reconstructed energy cut is the same as the 1Rμ samples.

- Total energy of final state particles (visible energy) is less than 30 MeV.
- If the the sample only has a single ring there is an additional cut based on the missing energy of 300 MeV.

The energy of final state particles varies by event and full details can be found in [132]. Since events can migrate between energy bins, events with energies slightly outside these boundaries are included in the processed MC and excluded at run time during the fit.

The full data spectra for these samples are shown in figure 4.2 and the total number of events for each sample summarised in table 4.2. The MC spectra split by interaction mode are shown in figures 4.3, 4.4 and 4.5 and have been re-weighted assuming oscillation parameters have true values given by the Asimov A22 parameter set and all cross-section systematics are at post-BANFF values. These definitions are summarised in section 4.5.2.1.

Selection	Run 1-11 POT	Events in Data
ν-mode 1Rμ	21.428×10^{20}	357
ν-mode ν_μCC1π		140
ν-mode 1Re		102
ν-mode 1Re1d.e		15
$\bar{\nu}$-mode 1Rμ	16.34556×10^{20}	137
$\bar{\nu}$-mode 1Re		16

Table 4.2: The total number of data events for each SK selection for runs 1-11 split by T2K horn current direction.

4.2.2 Near Detector Inputs

As in the far detector, the near detector data is split into samples based on the expected final state topology. The samples used in this analysis are briefly summarised in table 4.3. Each ND sample is binned in terms of outgoing muon momentum and outgoing muon angle.

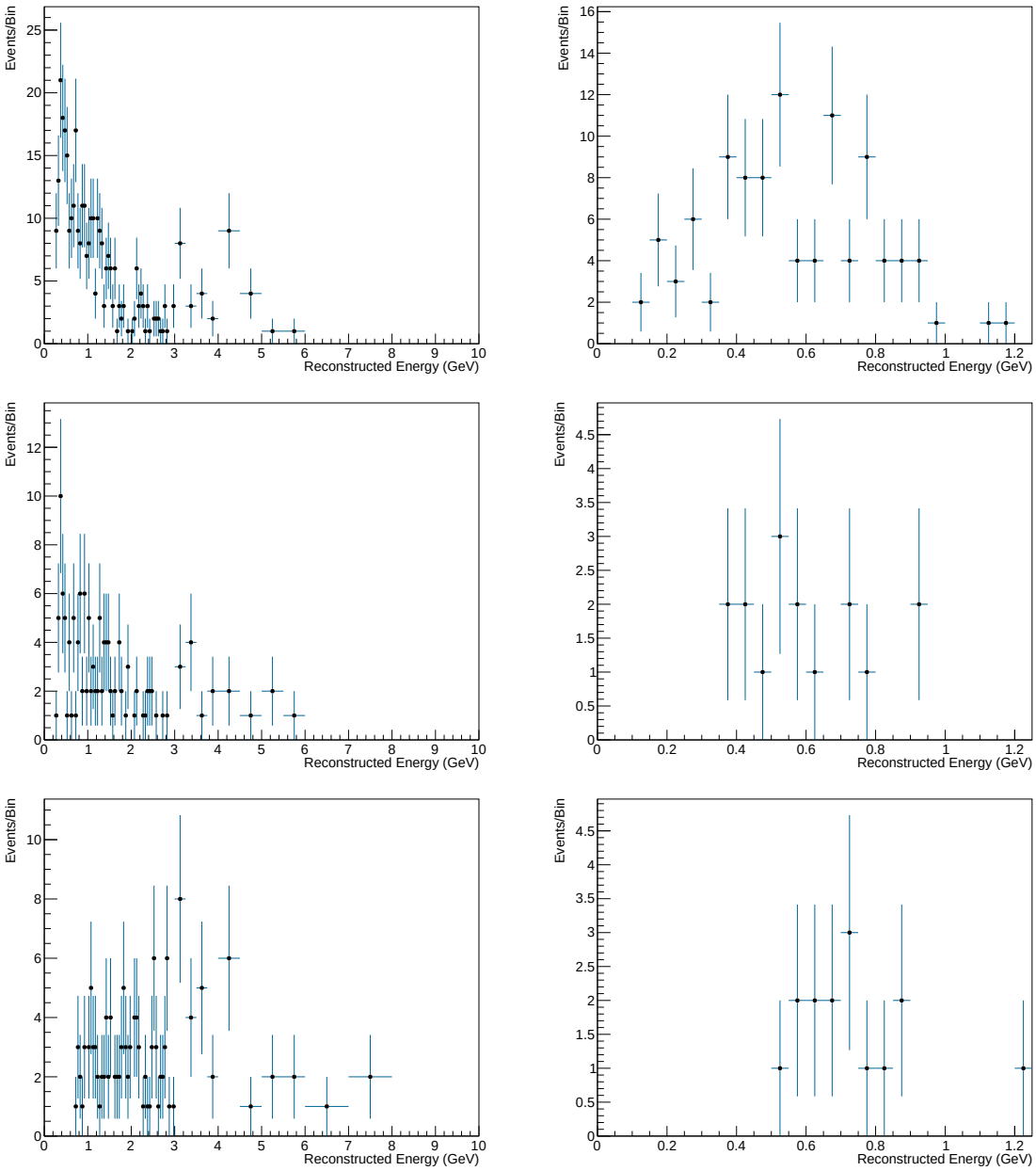

Figure 4.2: SK run 1-11 Data used in this analysis shown for SK event selection. The binning used is the same as in the MC selections. Data corresponds to ν-mode 1Rμ (top left), $\bar{\nu}$-mode 1Rμ (top right), ν-mode 1Re (middle left), $\bar{\nu}$-mode 1Re (middle right), ν-mode ν_μCC1π (bottom left), ν-mode 1Re1d.e. (bottom right).

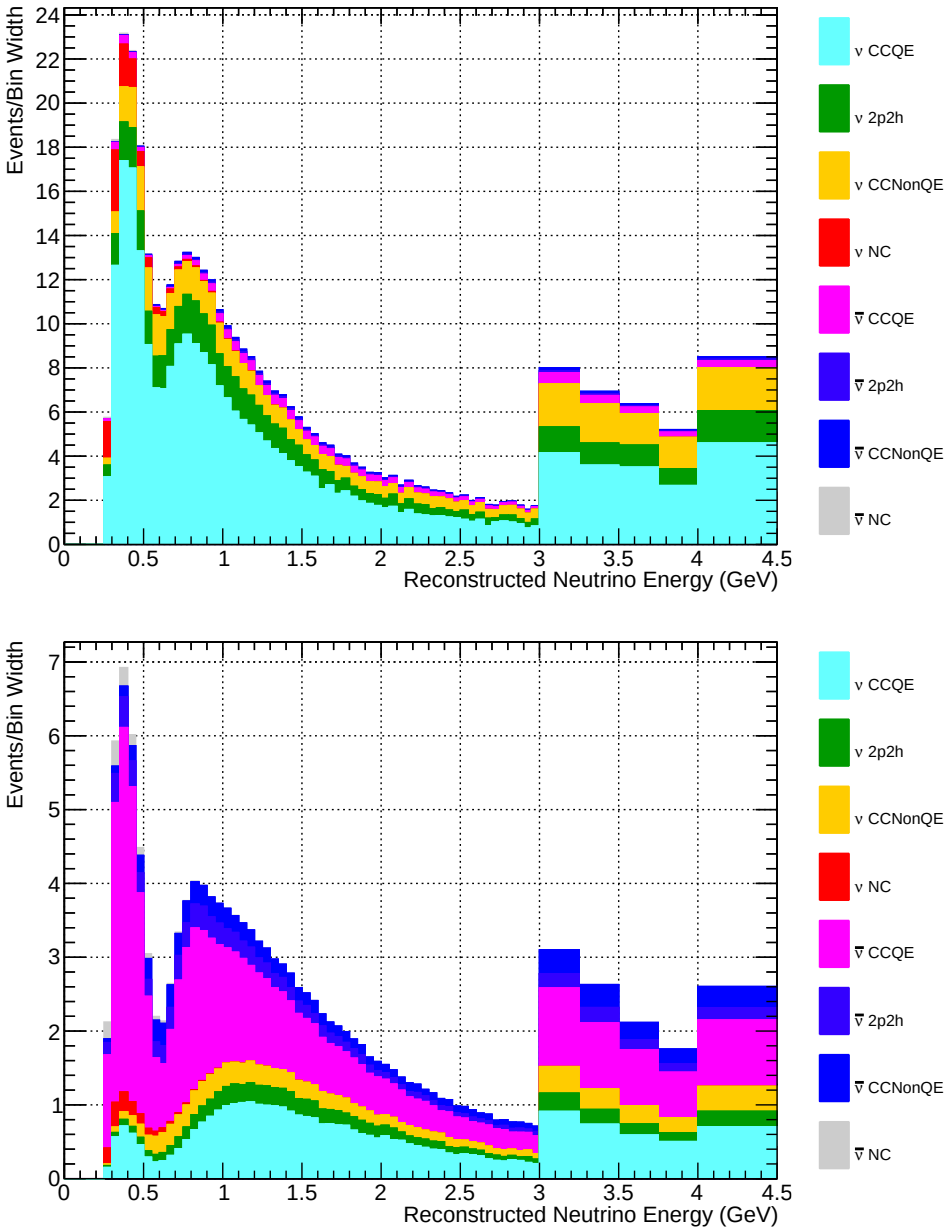

Figure 4.3: Oscillated event rate spectra in reconstructed energy for single Cherenkov ring μ-like far detector samples for ν-mode (top) and $\bar{\nu}$-mode (bottom). Plots are broken down by true interaction mode Plots use AsimovA22 oscillation parameter set and post-BANFF configuration.

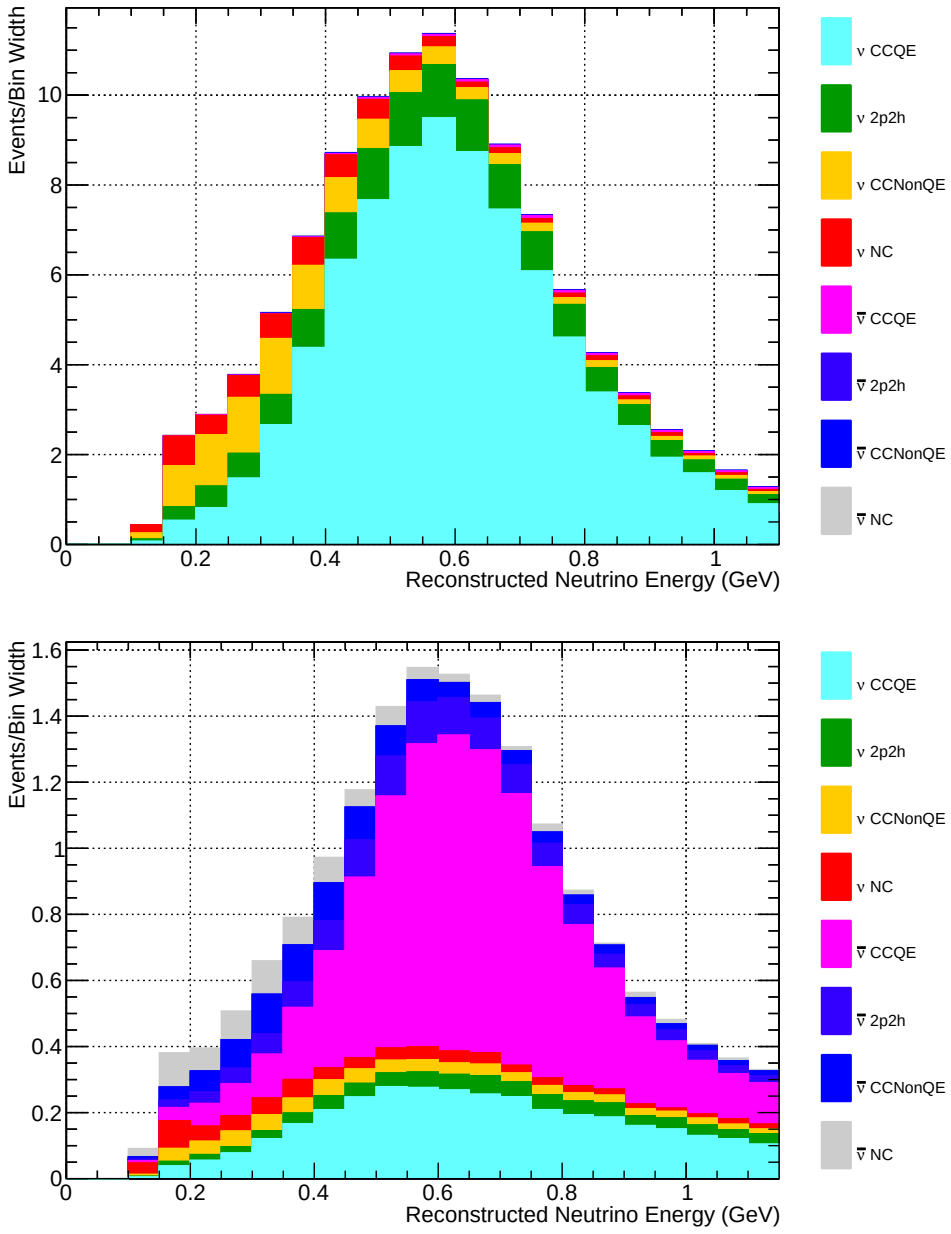

Figure 4.4: Oscillated event rate spectra in reconstructed energy for single Cherenkov ring e-like far detector samples broken down by true interaction mode for ν-mode (top) and $\bar{\nu}$-mode (bottom). Plots use AsimovA22 oscillation parameter set and post-BANFF configuration.

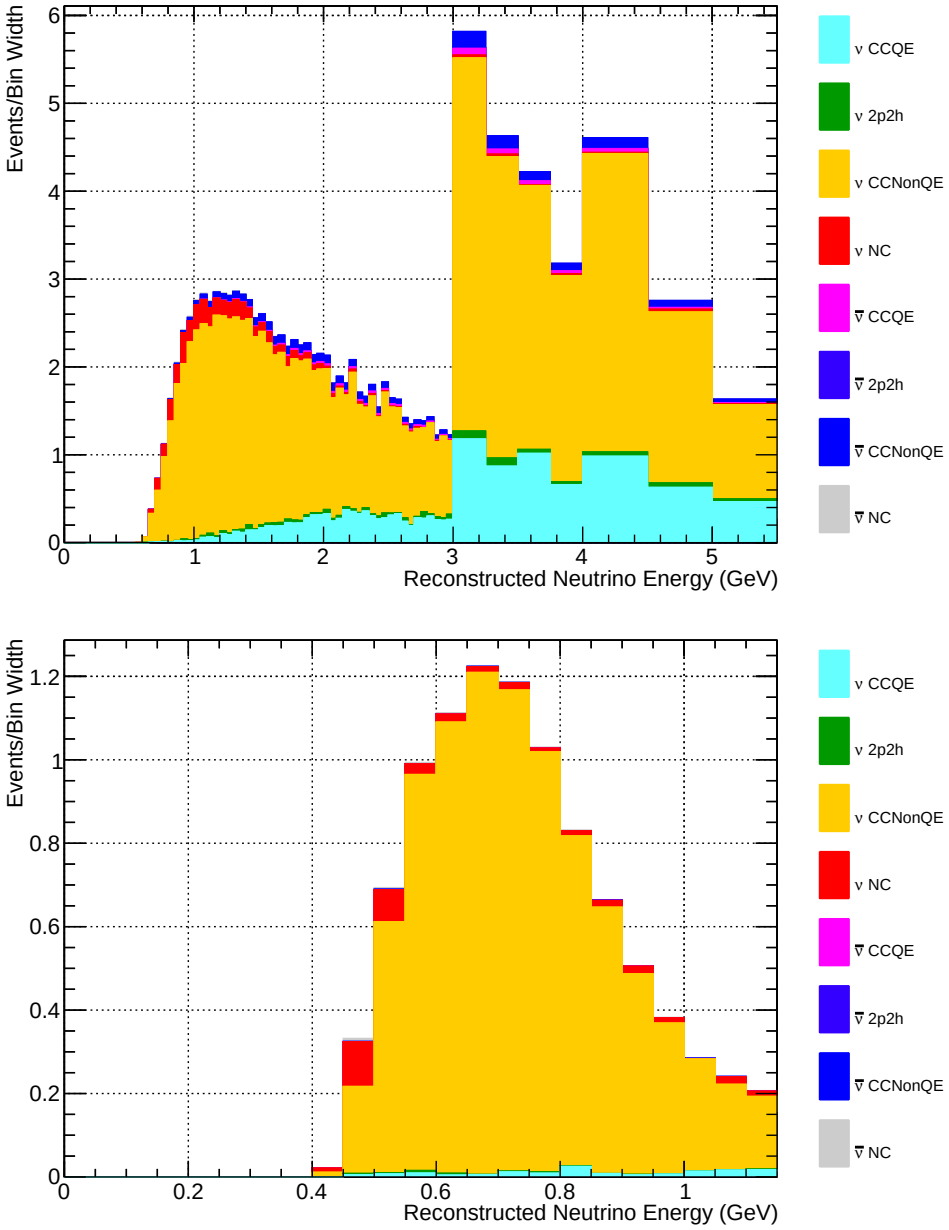

Figure 4.5: Oscillated event rate spectra in reconstructed energy for multi-ring ν-mode ν_μ CC1π (top) and ν-mode 1Re1d.e (bottom) samples broken down by true interaction mode. Plots use AsimovA22 oscillation parameter set and post-BANFF configuration.

Sample Topology	Neutrino Flavour	Description
CC0π0p	ν_μ, $\bar{\nu}_\mu$	CC interaction with only muon tracks in the final state
CC1π^+	ν_μ, $\bar{\nu}_\mu$	CC interaction with single final state pion track and single muon track
CC0$\pi N p$	ν_μ, $\bar{\nu}_\mu$	CC interaction with at least one final state proton
CCγ	ν_μ	Charged current interaction with final state photons
CC other	ν_μ, $\bar{\nu}_\mu$	Used to classify events that don't fit into the above categories, typically with many final state pions

Table 4.3: Table containing description for near detector samples. Each sample is split by detector (FGD1 and FGD2) and by incoming neutrino flavour. $\bar{\nu}_\mu$ samples also include an additional background sample to account for $\bar{\nu}_\mu$ produced outside of ND280, for example in the sand surrounding the detector.

The bin edges for each sample are non-uniform and implemented in such a way that, for each bin, there is a single data event and ≈20 MC events. In addition, bin width is also limited by detector resolution so it is often not possible to only contain a single data event per bin. Event selection in the near detector is omitted from this thesis as the analyses are concerned with implementing the far detector data and systematics, with the ND implementation being performed separately. More details on this process as it pertains to the T2K oscillation analysis can be found in [133] and the full details of the selections used are detailed in an internal T2K technical note [134].

4.3 Systematics

4.3.1 Cross-Section Model

A full summary of the cross-section systematics can be found in Appendix A. This set of systematics has the effect of re-weighting the cross-section for each MC event used in the T2K model. For a far detector sample, an event with cross-section Φ has a

re-weighted cross-section of

$$\Phi_{\text{re-weight}} = \Phi \prod_{i=1}^{N_{\text{syst}}} w_i(b) \tag{4.4}$$

where $w_i(b)$ are the weights for each systematic and b the bin the event lies in. Calculation of the weight is non-trivial but can be achieved through several mechanisms. The simplest set of parameters use normalisation-based re-weighting. Here the weight is just linear based on the value of the systematic variable in the fit. As this is applied uniformly to all events in the model it is generally used for systematics that change the total expect number of events for a given sample.

Cubic splines approximate functions by evaluating the function at set points (knots) and interpolating between each knot with a cubic polynomial and are commonly used when the systematic response varies depending on the kinematics of each event. For a systematic with full response $S(x)$ and N knots at $\{x_i\}_{i=1}^{N}$, the spline consists of $N-1$ bounded cubic polynomials $\{P_i\}_{i=1}^{N-1}$ with $P_i(x) = 0$ if $x \notin [x_i, x_{i+1}]$ which satisfy the following conditions:

- for each knot, i $P_i(x_i) = S(x_i)$, $P_i(x_{i+1}) = S(x_{i+1})$;
- ignoring the first and final knots, $P_i'(x_{i+1}) = P_{i+1}'(x_{i+1})$;
- ignoring the first and final knots, $P_i''(x_{i+1}) = P_{i+1}''(x_{i+1})$.

For far detector cross-section systematics, the spline response for each bin in the analysis is evaluated by varying the value of the systematic to pre-defined knot locations, typically intervals of the prior error. This can be seen in figure 4.6. The near detector uses a slightly different spline treatment and instead assigns each event in the sample a spline and can thus apply weights to each event rather than to bins.

The final systematic implementation is functional systematic. Unlike splines, the exact functional form of the systematic is known and, as a result the modification to the cross-section can be calculated without needing interpolative methods. All functional

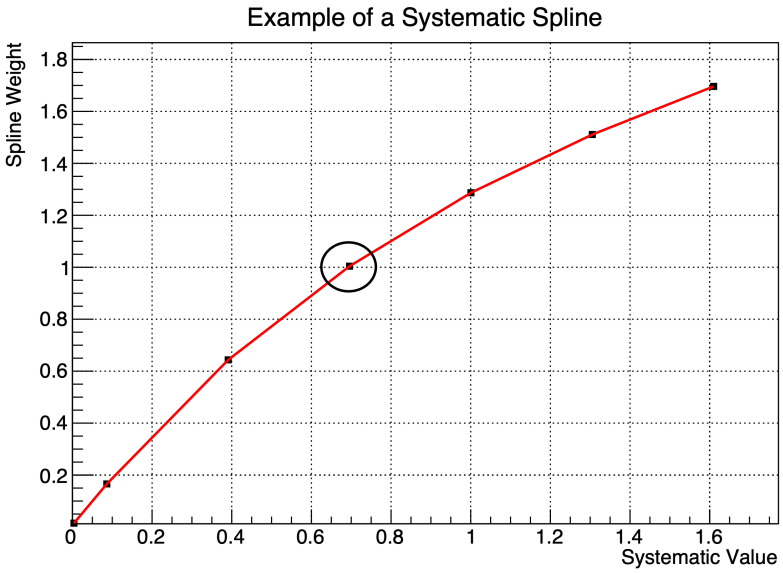

Figure 4.6: An example of a cross-section systematic spline with 7 knots. The nominal knot has been circled and represents the generator expected value for a given systematic. Whilst most knots are at integer multiples of the prior error (i.e. nominal+$1k\sigma$), the knot position at 0 is enforced by a physical bound on this parameter.

systematics are calculated on an event-by-event basis rather than by bin. Currently, MaCh3 uses this for calculating the Coulomb correction and the binding energy (E_b) for each event.

4.3.2 Improving the Cross-Section Spline Implementation

One of the major bottlenecks within MaCh3 prior to this analysis was caused by an inefficient implementation of the spline re-weighting algorithm for far detector cross-section splines. Previously splines were split by oscillation channel and stored in a struct accessible through a large nested vector. This implementation was both memory inefficient since the stored spline objects required far more space than was really necessary for the fitter and slow both due to the oscillation channel split limiting parallelisation and increasing function calls, and through a non-contiguous data-structure which increased lookup times.

In order to improve this a "monolithic" approach was used instead. All spline

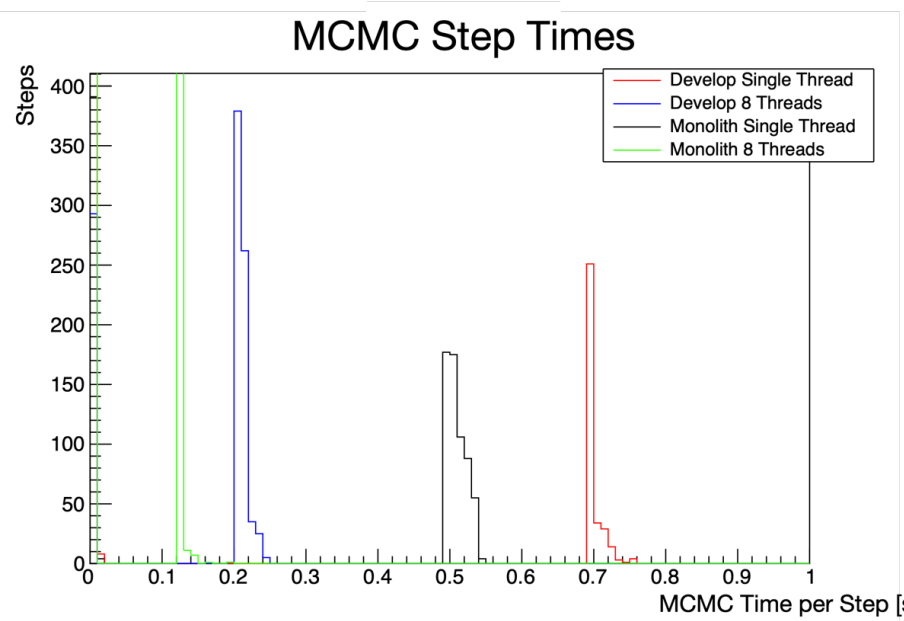

Figure 4.7: Improvement is MCMC step times as a result of the change in spline storage configuration using both single threaded MaCh3 and multi-threading. The previous structure has been labelled as "develop" due to this being the stable branch of MaCh3 during the development process.

coefficients[1] and knot values (x) are stored in a single 1 dimensional array, of the form $[x_{11}, a_{11}, b_{11}, c_{11}, ...]$, where x_{ij} is the j^{th} knot of the i^{th} spline parameter. This allows for far more efficient parallelisation when looking up spline coefficients and, since it is stored in a single 1D vector, all events are contiguous in memory. As a result of these efficiency improvements, there is roughly a factor 2 improvement in overall step time for the MaCh3 MCMC algorithm [Fig. 4.7].

4.3.3 Flux Systematics

The neutrino flux model consists of 100 parameters which are provided based on data from the T2K beam group and simulation using a replica target at NA61-SHINE [135]. These parameters quantify the uncertainty on the flux of each neutrino at the near and far detectors split by neutrino energy. This flux model MC is generated using several event generators due to the large range of energies required for the reconstruction.

[1] Coefficients for a cubic spline are the values a, b, c, d for each polynomial $y = ax^3 + bx^2 + cx + d$.

Uncertainties for each neutrino energy, for example on the beam direction or effects from hadron FSI, are then quantified into a total uncertainty for each bin. MaCh3 implements these systematics as part of the cross-section model and treats them all as normalisation parameters.

4.3.4 Detector Uncertainties

4.3.4.1 Near Detector Systematic Errors

The ND280 systematics model assigns an uncertainty to each ND sample bin. These uncertainties include both detector systematics and errors inherent to event reconstruction. This is the largest component of the T2K analysis model, accounting for over 500 of the ≈760 systematic dials present in the full analysis[2]. All parameters are roughly Gaussian and thus this is a relatively good prior approximation of the posterior covariance. As will be seen later, this makes step size tuning relatively manageable even for this large parameter set.

In addition to the uncertainty model, due to the large number of events observed in ND280, the full ND280 data+error likelihood calculation also includes a Barlow-Beeston likelihood [136]. This extra addition to the (statistical) likelihood is given by

$$\mathcal{L}_b = \frac{(\beta_i^2 - 1)^2}{2\sigma_{\beta_i}^2} \tag{4.5}$$

where $\sigma_{\beta_i}^2$ is the relative error on the bin calculated using a simple error calculation based on the number weight (w_i) given to each MC event

$$\sigma_{\beta_i}^2 = \frac{\sqrt{\sum w_i^2}}{N_{\text{events,MC}}} \tag{4.6}$$

and β_i is a scaling parameter. If each bin is assumed to be Gaussian, β_i can be calculated

[2]This lowered bin number occurs due to "bin merging" which reduces the number of energy bins actually used for each sample. To do this bins with similar systematic responses are combined cutting the total number of bins down by around 90%.

iteratively through

$$\beta_i^2 + (\mu\sigma_{\beta_i}^2 - 1)\beta_i - n\sigma_{\beta_i}^2 = 0 \tag{4.7}$$

where n is the number of data events in each bin and μ the number of simulated events in the bin. This additional condition quantifies how likely data events are to be a statistical fluctuation of the MC and how much MC events are likely to be fluctuations of MC if the simulation was allowed to run indefinitely. Full details of this can be found in the original paper by Barlow and Beeston [136].

4.3.4.2 Far Detector Systematic Errors

The final source of error present in the T2K systematics model comes from Super-Kamiokande (SK). SK errors are used to account for mischaracterisation of Cherenkov rings seen in the far detector. In this analysis SK systematic errors are treated as normalisation parameters. Systematic errors are applied for each sample which, as with ND280, is split into energy bins. The initial matrix is generated using a fit from the T2K+SK group using data from Super-Kamiokande atmospherics analyses. This fit does not account for FSI information, secondary interactions (SI) of pions and photo-nuclear (PN) effects. FSI information is well constrained by the near detector and so not incorporated, however, a secondary matrix is created to include SI+PN effects. The final matrix is then generated through

$$C_{\text{SK,full}} = C_{\text{SK,fit}} \oplus C_{\text{SK,SI+PN}} \tag{4.8}$$

where $\oplus$ indicates the matrices have been summed in quadrature.

The analysis presented in this chapter includes a new treatment for the SK detector matrix. In this matrix correlations between single and multi-ring samples are correlated, whereas in previous analyses the correlation was assumed to be 0 and a combined matrix was used instead with the assumption that these samples have 0 prior correlation. In addition this new matrix was generated using atmospheric events as these have far

higher statistics and a larger energy range than beam events in T2K. The impact of this matrix on the bin-by-bin uncertainties for each sample are shown in figures 4.8, 4.9 and 4.10.

4.4 MaCh3 MCMC Implementation

MaCh3 uses a modified version of the Metropolis-Hastings algorithm [Alg. 1] using the binned log-likelihood calculation in algorithm 5 to evaluate the likelihood of each MCMC step. In the full T2K model, near and far detector data are fitted simultaneously; due to the high statistics, a non-zero Barlow-Beeston correction is applied to each ND bin when calculating model likelihood. The systematics model consists of 4 parameter sets: Super-Kamiokande detector uncertainties, ND280 uncertainties, cross-section model, and the oscillation parameters. MaCh3 uses a re-weighting scheme to calculate the value of each event in the Monte-Carlo model at any given set of parameter values where the total weight contribution from each is the product of the weights obtained from each systematic.

4.5 2023 T2K Statistical Update

4.5.1 Overview

The rest of this chapter will detail the T2K "statistical" update from its 2023 oscillation analysis (OA2023). This analysis differs from previous oscillation analyses in the following ways:

- Approximately 10% increase in ν-mode protons on target (POT) from 19.664×10^{20} POT to 21.428×10^{20} POT from the 11th T2K run.
- Improved flux prediction using this new data.

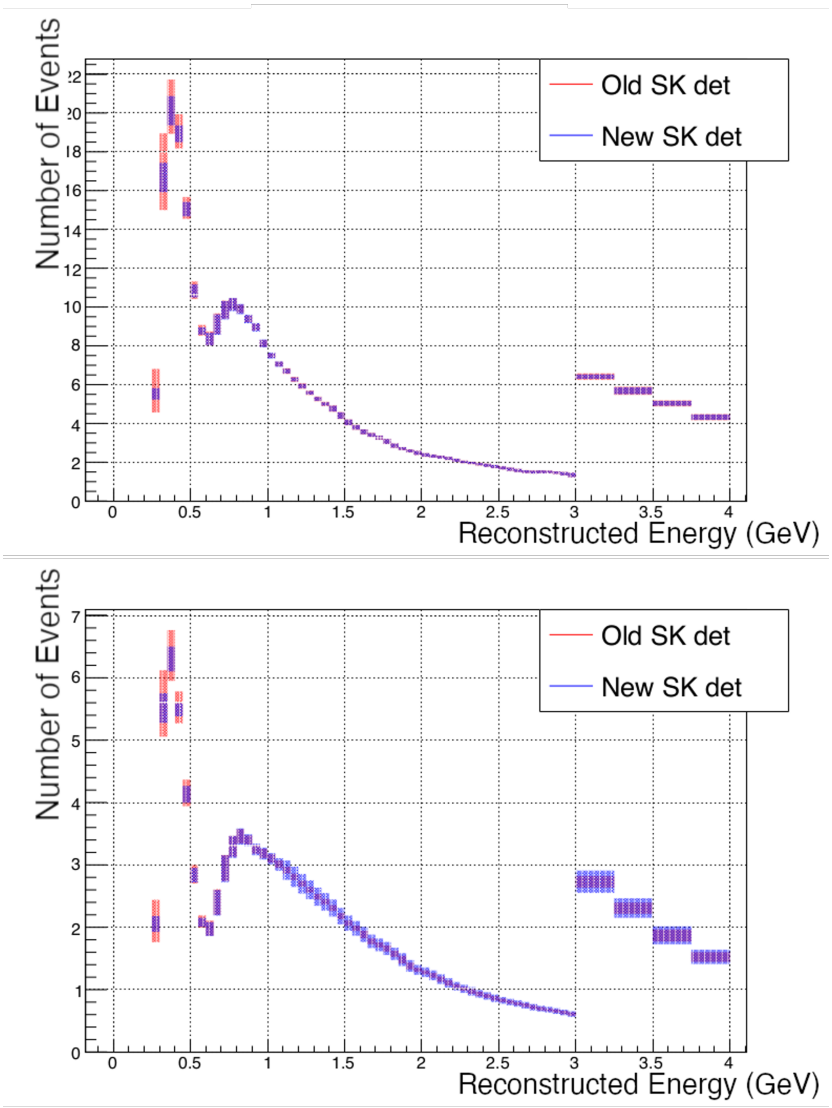

Figure 4.8: The change on uncertainties between the SK detector matrix on MC data between the previous analysis (red) and the current analysis (blue) for single ring μ-like samples split by ν-mode (top) and $\bar{\nu}$-mode (bottom). There is a slight tightening of errors at low reconstructed energy values.

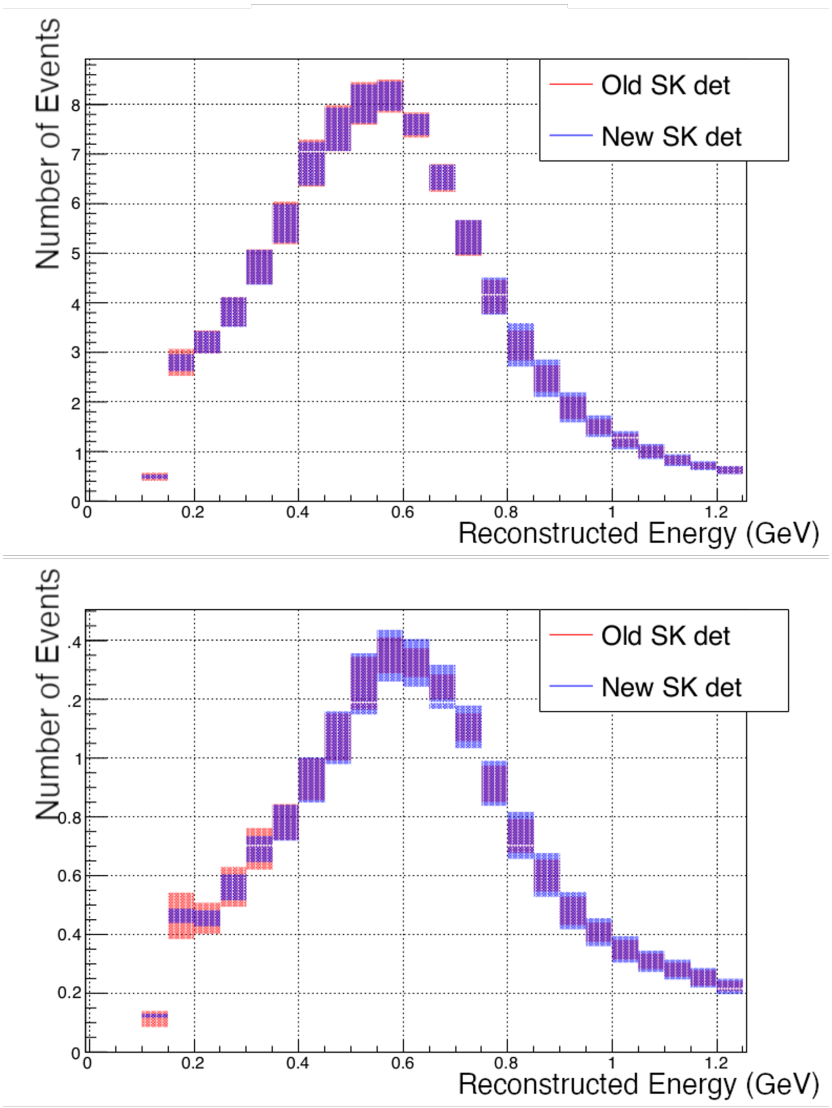

Figure 4.9: The change on uncertainties between the SK detector matrix on MC data between the previous analysis (red) and the current analysis (blue) for single ring e-like samples split by ν-mode (top) and $\bar{\nu}$-mode (bottom). There is a slight tightening of errors at low reconstructed energy values.

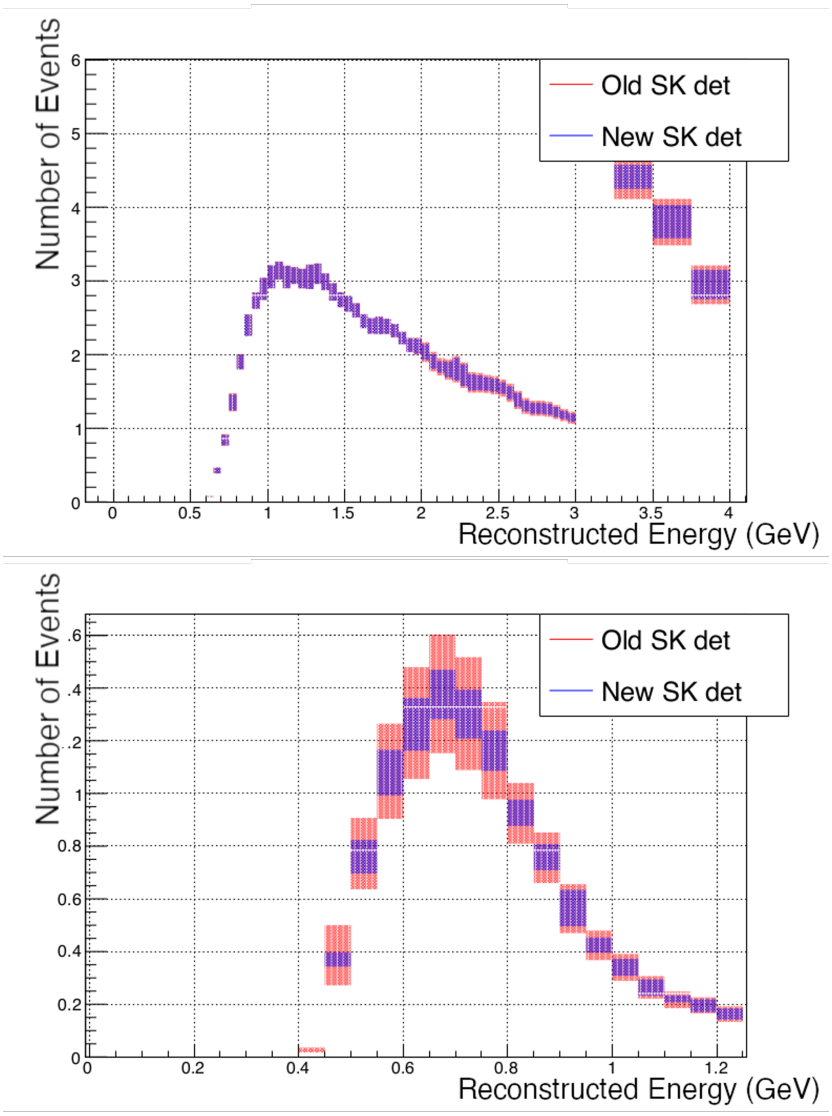

Figure 4.10: The change on uncertainties between the SK detector matrix on MC data between the previous analysis (red) and the current analysis (blue) for multi-ring samples split by ν_μCC1π (top) and 1Re 1d.e. (bottom). There is a slight tightening of errors across the full energy spectrum.

Algorithm 5 MaCh3 Log-Likelihood Calculation

Inputs: Prior Covariance Matrices for parameter sets (C_{par}), current and prior-nominal value for each parameter (X_{par}, $X_{\text{par,nom}}$), input data (D), and a set of model-generated samples ($\mathcal{S}$). Data and model samples should have the same binning. High statistics detectors require Barlow-Beeston corrections on each bin parametrised by (β, σ_β) (At low statistics, $\beta = \sigma_\beta = 1$ for all bins).

Set initial (negative) log-likelihood: $\mathcal{L} = 0$
Evaluate Parameter Penalty Terms:
for p **in** parameters **do**
    if Gaussian prior **then**
        Calculate Penalty: $\mathcal{L}_p = \frac{1}{2}(X_p - X_{p,\text{nom}})^T C_p^{-1}(X_p - X_{p,\text{nom}})$
    else
        Assume flat prior $\mathcal{L}_p = 0$
    $\mathcal{L} = \mathcal{L} + \mathcal{L}_p$
if $\mathcal{L} > 1 \times 10^9$ **then**
    End calculation and output $-\mathcal{L}$ since a subset of parameters are outside their physical boundaries
Calculate Model Likelihood
for s, d in $\mathcal{S}, D$ **do**
    Get the model prediction $s(X_{pars})$ for all parameter sets
    $\mathcal{L}_s = 0$
    Loop over sample bins
    for i **in** s_{bins} **do**
        Calculate Poisson Difference: $\mathcal{L}_s = \mathcal{L}_s + s_i - \left[\log\left(\frac{d_i}{s_i}\right) - 1\right] d_i$
        Apply Barlow-Beeston Correction: $\mathcal{L}_s = \mathcal{L}_s + \frac{(\beta_i - 1)^2}{2\sigma^2_{\beta_i}}$
    $\mathcal{L} = \mathcal{L} + \mathcal{L}_s$
Output: $-\mathcal{L}$

- Inclusion of far detector data that uses Gadolinium with a new momentum-based decay electron cut in this data.
- The SK detector matrix now combines single-ring and multi-ring uncertainties.

The analysis data selection and systematics have been described above [Sec. 4.2] so the remaining chapter will detail the analysis validation procedure and the results of fits.

4.5.2 Validation

4.5.2.1 Event Rate Validation

The bulk of time in the analysis is spent validating that systematics models have been implemented in the same way in both the frequentist fitter (PTheta) and MaCh3. The first step is to validate that both analyses produce the correct event rates. That is to say, when the systematics model is set to the same values for both fitters, is the integral of model spectra consistent? This is primarily a test of cross-section model implementation and, as a result, the groups only change the values of cross-section parameters. There are three configurations that are currently tested are summarised in section sec:jargon. All validations and Asimov Fits used in this analysis used the AsimovA22 oscillation parameter set which is discussed in the same section. Event rates are expected to have a difference of at most $\mathcal{O}(0.1\%)$ in all interaction modes within each sample. For this analysis it was found that generated tune provided good agreement, but, both pre and post-BANFF there were significant differences present in the ν_μCC1π sample in several modes, particularly NCπ^0 [Sec 4.5.2.3].

4.5.2.2 Sigma Variations

In addition to examining the event rates for samples, both fitters also perform so-called 'σ-variations'. This method of validation varies each cross-section and far detector systematic by multiples of the prior error and looks at the relative change in each bin of the MC model spectrum. These changes are compared between fitters to further determine if the interaction model produces near identical responses. Figure 4.11 shows an example of σ-variation for M_{res}^{A} where there is clear agreement between both fitters. As will be seen below, validation that this sample has the same response for both fitters is vital in understanding the major differences seen between expected event rates for this analysis.

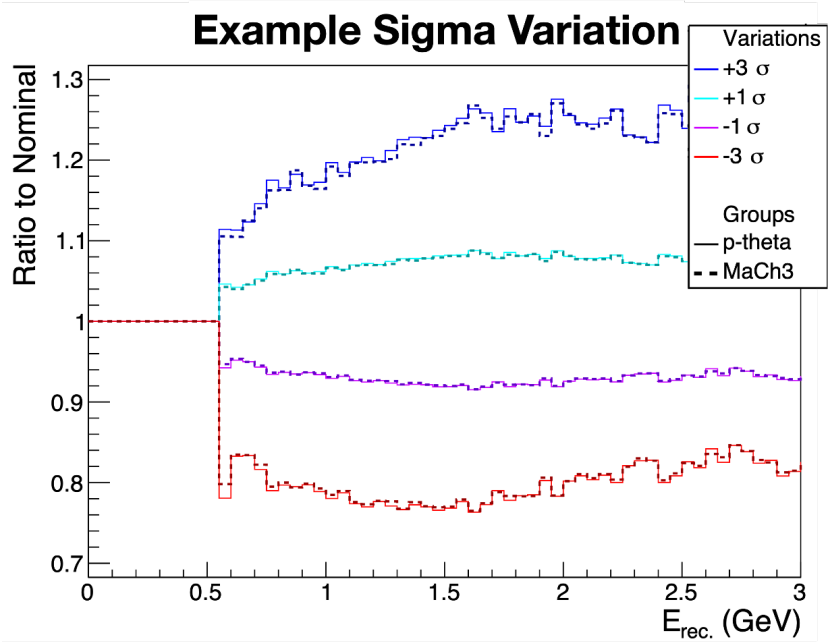

Figure 4.11: Figure showing the MaCh3 and PTheta σ-variation comparisons for the resonant axial mass cross-section parameter with the nominal set to its generated value in the ν-mode ν_μCC1π FD sample. By varying the systematic by $\pm1\sigma$ and $\pm3\sigma$ the spline response can be measured between both fitters. The small level of disagreement seen at the 3σ variation is expected and has been seen on all previous analyses using this parameter

4.5.2.3 Assessing the ν_μCC1π Event Rate Bug

When preparing for this analysis it was found that there were significant differences between MaCh3 and PTheta's expected event rates for several modes in the ν_μCC1π sample at pre and post-BANFF tunes. The differences in event rates at post-BANFF tune are detailed in table 4.4. As explained above, the key feature of pre/post-BANFF tuning is that the nominal values for systematics are moved away from the generator predicted nominal value. In order to assess the cause of this in the fitters, event rates were generated with all systematics weights disabled. Initial testing was done to ensure that it was purely a fault within the systematics re-weighting by disabling all weightings within the fitters beyond even the systematics weights, for example POT weighting and oscillation weighting. The first is a multiplicative weight which multiplies the expected

Interaction Mode	MaCh3	PTheta	Relative Difference
CCQE	16.87561	16.88758	0.07%
CC1π^0	3.75661	3.71598	-1.08%
CC1$\pi^\pm$	59.76295	59.56966	-0.32%
CCcoh	1.93179	1.93359	0.09%
2p2h	1.90862	1.93452	1.36%
CCDIS	8.7856	8.79105	0.06%
CCMpi	24.36831	24.36797	0.00%
CCMisc	1.70302	1.69308	-0.58%
NCπ^0	0.07492	0.05988	-20.07%
NC1$\pi^\pm$	0.80476	0.76772	-4.60%
NCcoh	0	0	0.00%
NCoth	4.08743	4.06912	-0.45%
NC1γ	0	0	0.00%
Total	124.05961	123.79	-0.22%

Table 4.4: Comparison of MaCh3 and PTheta post-BANFF event rates for the ν-mode ν_μCC1π sample. The NCπ^0 row has been highlighted as it is the largest source of difference between the two fitters. The interaction column links to the various interaction modes used in T2K. More information on the interaction modes listed can be found in [137].

number of events in the spectrum by POT/10^{20} and the second modifies the spectrum to have the distribution expected from the inputted set of oscillation parameter values (AsimovA22). As the differences were only seen after all systematics were enabled + pre/post-BANFF fits were run it was concluded that this was purely a cross-section systematics issue.

The next stage of validation was to disable re-weighting for all sets of cross-section parameters and then systematically enable them based on re-weighting type i.e. normalisation, splined and functional. It was found that the largest differences occurred when spline parameters were re-enabled and hence the largest effort was spent investigating these. Initial investigation looked at directly comparing MaCh3 and PTheta spline implementations. As can be seen in figure 4.12, the differences in splines were negligible and it was found that knot positions were identical between fitters. In addition both fitters were found to be able to reproduce the inputs used in the previous oscillation analysis using the current implementation of the software. It was found

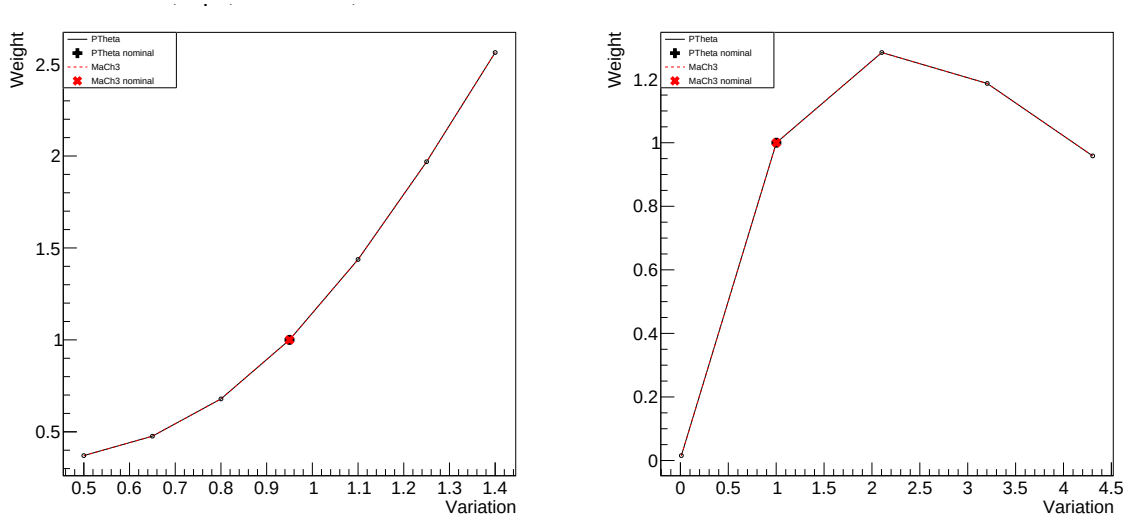

Figure 4.12: Small selection of cross-section spline comparisons for MaCh3 and PTheta splines form the ν_μCC1π mode. As can be seen the fitters agree at the weights for each spline knot. The spline interpolation between is knots is not shown, however, since both fitters use the ROOT TSpline3 implementation, this was not thought to be important.

that the largest impact on event rates came from splined final state interaction (FSI) parameters. Table 4.5 illustrates that this was driven almost entirely by resonant pion systematics.

The investigation into pion production systematics looked at two areas. Firstly sigma variations were produced and compared between fitters. Due to the low number of expected events in the problematic mode, these plots were susceptible to statistical noise and, as a result, were inconclusive. The second method was to repeat the above process but, rather than enabling systematics in groups, to instead enable each pion production systematic individually. The results of this study are detailed in table 4.6.

Whilst these studies demonstrated there were substantial differences between fitter implementations, both the spline treatment analysis and the fitters' abilities to reproduce the previous years analysis made finding the source of these differences tricky. It was decided that, due to the time constraints of the analysis, the log-likelihood scans for both fitters should be compared. These evaluate the impact of individual systematics on the total log-likelihood by varying each individual systematic over a range of values

Mode	No Spline Weights	Non-Pion Systematics	Resonant Pions	CCMπ (All)
CCQE	0.16%	0.19%	0.18%	0.18%
CC1π^0	0.03%	-0.39%	-0.67%	-0.67%
CC1$\pi^\pm$	-0.01%	-0.02%	-0.16%	-0.16%
CCcoh	0.13%	0.13%	0.13%	0.13%
2p2h	-0.09%	0.10%	1.35%	1.35%
CCDIS	0.06%	0.03%	0.03%	0.15%
CCMπ	-0.04%	0.00%	0.00%	-0.02%
CCMisc	-0.01%	-0.19%	-0.19%	-0.19%
NC1π^0	-0.10%	2.97%	-19.74%	-19.74%
NC1$\pi^\pm$	-0.03%	-0.05%	-3.96%	-3.96%
NCcoh	0.00%	0.00%	0.00%	0.00%
NCoth	-0.05%	-0.22%	-0.22%	-0.22%
NC1γ	0.00%	0.00%	0.00%	0.00%
Total	0.01%	0.00%	-0.10%	-0.09%

Table 4.5: Table showing the relative differences between PTheta and MaCh3 post-BANFF events rates when enabling splined systematic weights by expected interaction mode. These weights are enabled sequentially so each column also has the weights in the previous column enabled. Somewhat expectedly the largest different comes from systematics that directly affect pion production. More information on the interaction modes used can be found in [137]

Mode	CA5	M_{res}^A	Non-Res $I\frac{1}{2}$	Non-Res $I\frac{1}{2}$, low	Res E_b, ν	Res E_b, $\bar{\nu}$	Pion Ejection
CCQE	0.18%	0.18%	0.18%	0.18%	0.18%	0.18%	0.18%
CC1π^0	-0.57%	-0.97%	-0.97%	-0.89%	-0.91%	-0.91%	-0.91%
CC1$\pi^\pm$	-0.02%	-0.10%	-0.10%	-0.11%	-0.23%	-0.23%	-0.23%
CCcoh	0.13%	0.13%	0.13%	0.13%	0.13%	0.13%	0.13%
2p2h	1.35%	1.35%	1.35%	1.35%	1.35%	1.35%	1.35%
CCDIS	0.15%	0.15%	0.15%	0.15%	0.15%	0.15%	0.15%
CCMπ	-0.02%	-0.02%	-0.02%	-0.02%	-0.02%	-0.02%	-0.02%
CCMisc	-0.19%	-0.19%	-0.19%	-0.19%	-0.19%	-0.19%	-0.19%
NC1π^0	-8.18%	-21.96%	-21.96%	-20.12%	-20.12%	-20.12%	-20.12%
NC1$\pi^\pm$	-2.58%	-5.18%	-5.18%	-4.22%	-4.22%	-4.22%	-4.22%
NCcoh	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NCoth	-0.22%	-0.22%	-0.22%	-0.22%	-0.22%	-0.22%	-0.22%
NC1γ	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total	-0.01%	-0.08%	-0.08%	-0.08%	-0.13%	-0.13%	-0.13%

Table 4.6: Table showing the relative difference between post-BANFF PTheta and MaCh3 event rates for the ν_μCC1π sample when (resonant) pion production systematics are enabled sequentially. Each column shows the difference in event rates when both the systematic in the header of the column and all preceding columns are enabled. The largest difference comes for the resonant axial mass (M_{res}^A). More information on the neutrino interaction modes listed can be found in [137].

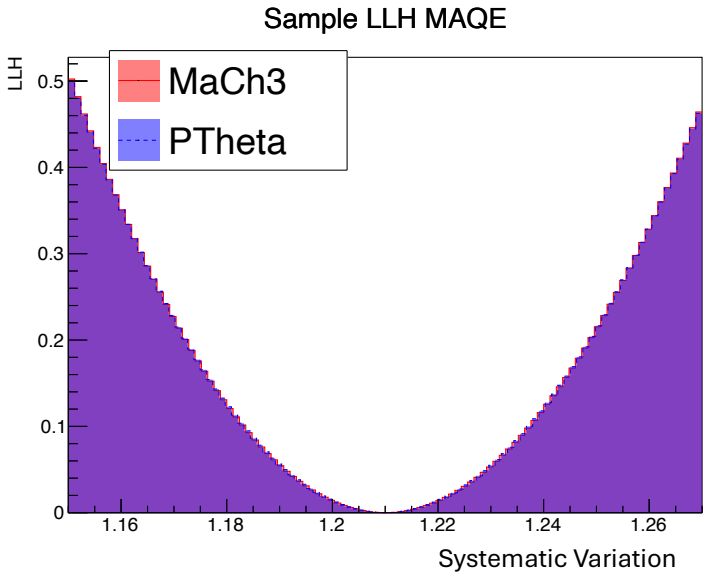

Figure 4.13: Log-Likelihood scans for M_{res}^{A} for PTheta and MaCh3 are compared across all samples. It can be seen that, when the full systematics model is accounted for, the MaCh3 and PTheta sample likelihood calculations produce the almost same result.

[Fig. 4.13]. As the LLH scans were in good agreement and due to the relatively small contribution of NC1π^0 to the total ν_μCC1π sample (0.0604% of expected events), it was decided that the impact of this bug on the analysis would be negligible and thus this error was left unresolved.

4.6 Asimov Fits

4.6.1 Introduction

The first set of fits run in this analysis are Asimov fits. These fits use Asimov "data" generated by re-weighting the MC prediction to some fixed set of systematic values. This tests the fitter robustness in several ways: firstly, the results of these fits can be compared between groups and internally and used to examine cross-fitter agreement and self-consistency between analyses. The second, MCMC specific, use is to tune the Markov Chain step sizes to suit the systematics model for a particular analysis. All Asimov fits reported in this section use real ND280 data. The Asimov data presented

in this section uses the post-BANFF cross-section parameter configuration and the Asimov A22 [Tab.4.1] oscillation parameter set. The ND and FD nominals are set to their priors which can be found in appendix A. All fits present in this chapter have two configurations determined by the "reactor constraint" (RC). This is a prior constraint on $\sin^2(\theta_{13})$ which constrains the space using a Gaussian centred on 2.2×10^{-2} with width 7×10^{-4} [138]. This is a prior based on data from reactor experiments and is applied to the posteriors post-fit. The results are also split into three categories based on mass ordering (NO, IO and both). The fit is run such that steps are allowed to enter both hierarchies and ordering is determined post-fit based on the values of Δm^2_{32} at each step. The following posteriors are produced by running an MCMC chain for 124×10^6 steps using the systematics configuration detailed above. There are some small statistical fluctuations at larger credible intervals; this is due to the low number of steps in this region and the effect can be reduced by running longer chains. For Asimov fits the requirement is only to have smooth contours up to the 90% credible interval since running MCMC is both computationally expensive and time-consuming.

4.6.2 Asimov Fit Posteriors

There are three sets of posteriors shown (with and without RC), figure 4.14 shows the 1D δ_{CP} posterior whilst figures 4.16 and 4.15 show 2D contours. These are produced by marginalising over all parameters in the chain not present in the final plot. Furthermore, two sets of credible intervals are shown. The motivation for this is one of convention with percentage-based credible intervals being a more traditional Bayesian metric for creating credible intervals than particle physics-like σ values.

In all 2D plots the best fit value shown is the 2D marginalised best fit value. As discussed in section 3.8.1, marginalisation can modify the best fit value and hence this may not be identical to the Asimov point. In addition, in figure 4.15 the best fit point without RC does not lie in the same octant as the Asimov point in both the fit with both orderings and NO only fit. This is a feature present in previous T2K analyses and

is due to the geometry of the $\sin^2(\theta_{23})$ space.

4.6.3 Comparisons with PTheta

Figures 4.17, 4.14, 4.19 and 4.20 show comparisons between the PTheta and MaCh3 Asimov fits. Figures 4.17 and 4.14 show comparisons without RC, and 4.19 and 4.20 with RC. Unsurprisingly, by constraining θ_{13} with the RC, the appearance contours ($\sin^2(\theta_{13})$ vs δ_{CP}) agree far better between fitters. The differences in contour shape are a well-understood feature [139] and a result of small analysis differences between fitters. In particular, PTheta and MaCh3 use slightly different far-detector sample binning and a different near-detector constraint, with MaCh3 fitting ND+FD simultaneously and PTheta requiring an external input. As will be seen in the data fits (below), when PTheta do use a "MaCh3-like" setup, these differences vanish. It was decided that this was unnecessary for the Asimov fits due to both the time required to perform extra analyses and the fact that this effect is well understood.

4.7 Data Fits

This section details the results of simultaneously fitting ND280 run 2–9 and SK run 1–11 data. This corresponds to 21.428×10^{20} protons on target (POT) in ν-mode mode and 16.346×10^{20} POT in $\bar{\nu}$-mode at SK, and for the ND 11.531×10^{20} POT in ν-mode and 8.336×10^{20} POT in $\bar{\nu}$ beam mode. Model validation diagnostics will be shown first before a discussion of the results.

4.7.1 Model Validation

4.7.1.1 Posterior Predictive

The posterior predictive can be thought of as the "mean" prediction from the fit. Analysing the value of the model at each step of the chain is inefficient. Instead, the finished chain is sampled randomly at a 5000 points and the MC spectra is evaluated

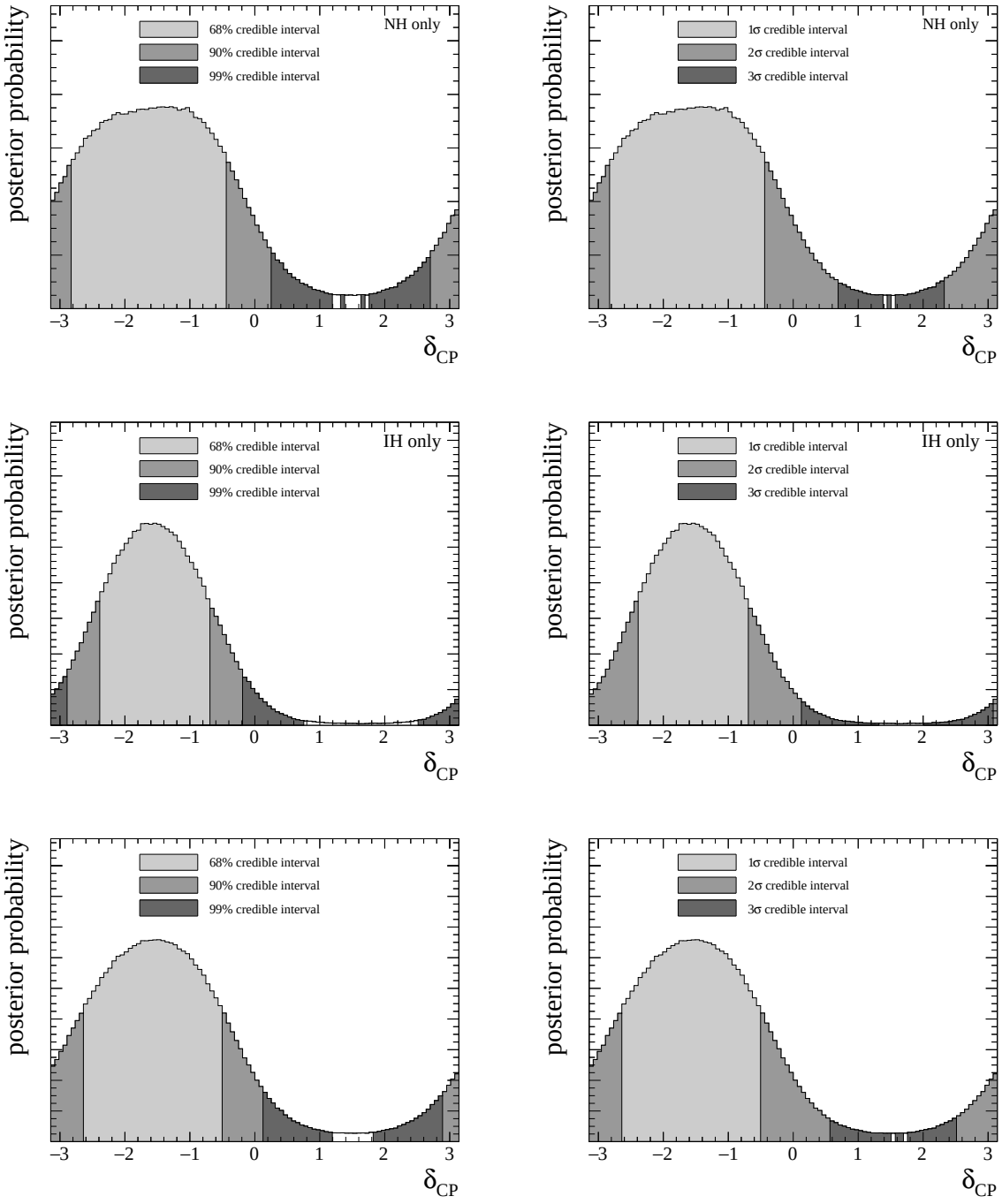

Figure 4.14: 1D credible interval for δ_{CP} from the Asimov A22 parameter fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals. Fits are displayed marginalised over normal ordering (top), inverted ordering (middle), without ordering marginalisation (bottom).

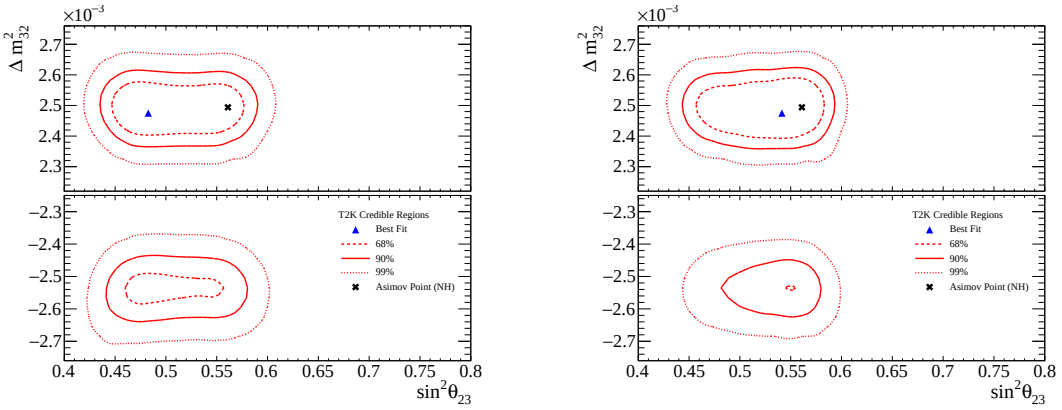

Figure 4.15: 2D $\sin^2\theta_{23}$-Δm^2_{32} contours from the Asimov A22 fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals.

at each of those points. Figures 4.21, 4.22 and 4.23 show the post-fit distribution and the 1σ error band for the number of events in each sample bin. The data used for the analysis are overlaid on top. Whilst higher energy data events exist, the ranges chosen in the plot correspond to the oscillation 'dip' which corresponds to the region of the reconstructed energy spectra most affected by neutrino oscillation as this is the area the has the largest impact on the posterior distributions.

4.7.1.2 Goodness of Fit

As discussed in section 3.8.3.1, MaCh3 performs a goodness of fit test by comparing MC and data to fake spectra. The results of this test are shown in table 4.7. The critical P-Value used by MaCh3 for a single sample is $\alpha = 0.05$, but, when multiple tests are compared the probability of rejecting the null hypothesis increases. This is known as the family-wise error rate (FWER). For n statistical tests, the FWER is given by

$$\text{FWER} = 1 - (1-\alpha)^n = 1 - (1-0.05)^6 = 26.5\% \tag{4.9}$$

where the factor of 6 comes from the 6 samples used in the MaCh3 far detector. In addition to this extra error, the size of the parameter space explored can increase the likelihood of apparently statistically significant events occurring. This is known

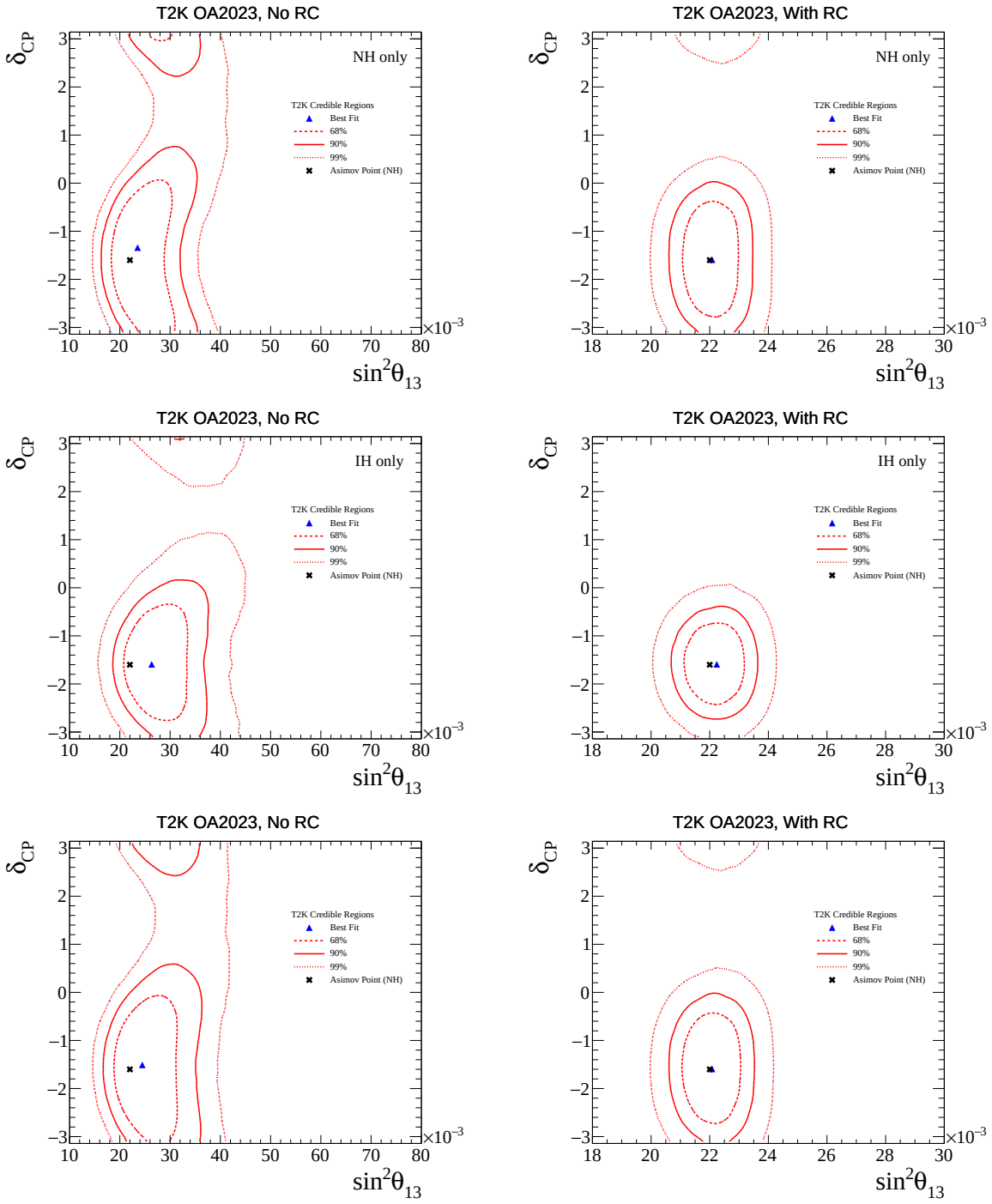

Figure 4.16: 2D $\sin^2\theta_{13}$-δ_{CP} contours from the Asimov A22 fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals. Fits are displayed marginalised over normal ordering (top), inverted ordering (middle), without ordering marginalisation (bottom).

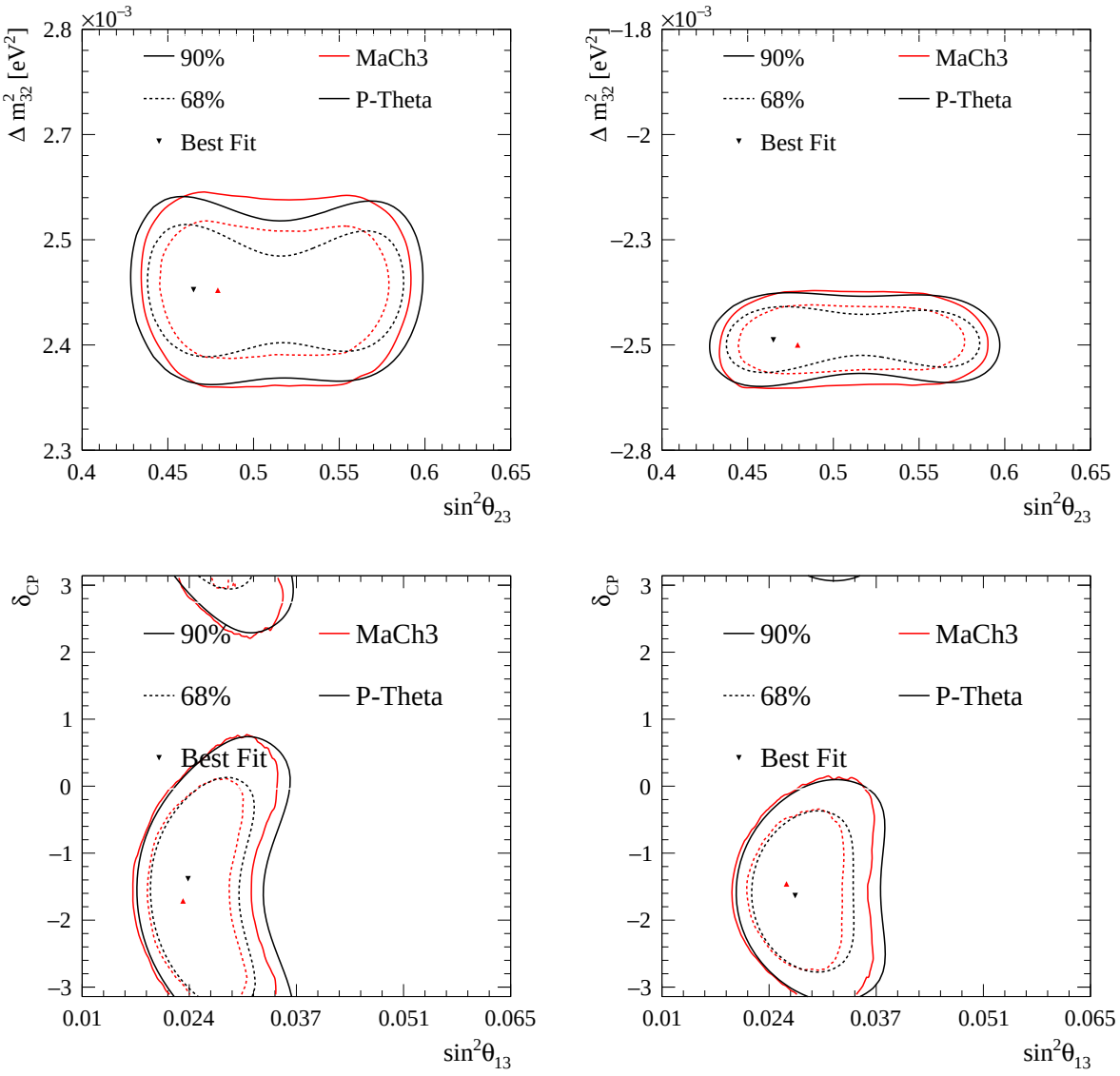

Figure 4.17: Comparison of oscillation parameter contours from the Asimov A22 fits of MaCh3 and PTheta without applying the reactor constraint. (Top) disappearance contours (bottom) appearance contours (left) normal hierarchy (right) inverted hierarchy.

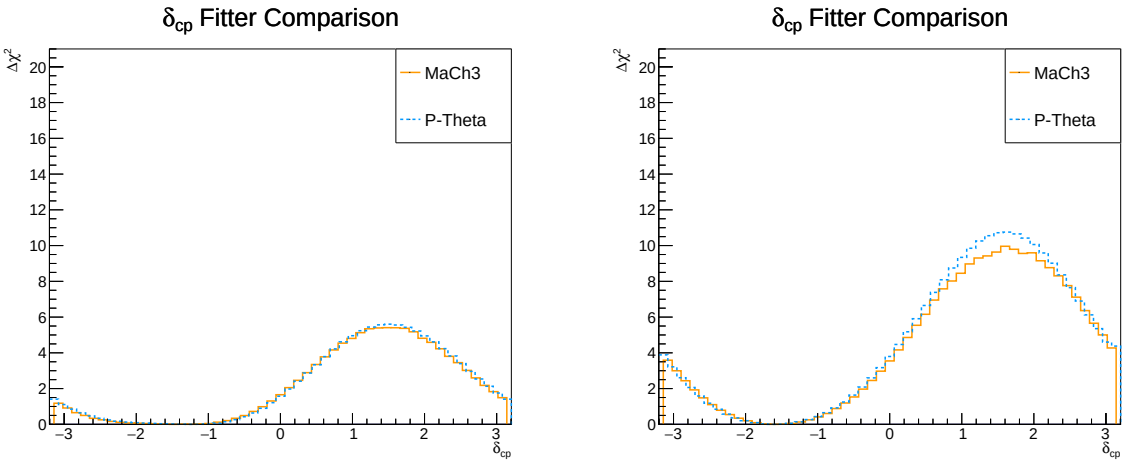

Figure 4.18: Comparison of 1D δ_{CP} $\Delta\chi^2$ between the MaCh3 and PTheta Asimov A22 fits without the reactor constraint. (Left) normal hierarchy (right) inverted hierarchy.

Sample / p-value	**Shape-based**	**Total Rate-based**
ν-mode 1Rμ	0.3254	0.0206
$\bar{\nu}$-mode 1Rμ	0.8426	0.5782
ν-mode ν_μCC1π	0.4168	0.2526
ν-mode 1Re	0.1210	0.4826
$\bar{\nu}$-mode 1Re	0.6378	0.4774
ν-mode 1Re1d.e.	0.7272	0.1142
Total	0.5064	0.0626

Table 4.7: Breakdown of goodness-of-fit p-values, quoted separately for bin-by-bin (Shape-based) and total rate (Total Rate-based) based χ^2 calculation, used as a test for the compatibility between the best-fit model and the data, using T2K-only fit.

as the look elsewhere effect (LEE). In order to account for this, for each sample, the Bonferroni corrected significance level [140], α_B, is used instead of the critical p-value. This is simple to calculate,

$$\alpha_B = \frac{\alpha}{n} = \frac{0.05}{6} = 0.008 \tag{4.10}$$

As can be seen in table 4.7, all samples have p-values above this significance level.

In order to ensure the robustness of MaCh3 fits, it was deemed necessary to investigate the lower rate-based p-values (difference in the total predicted number of events

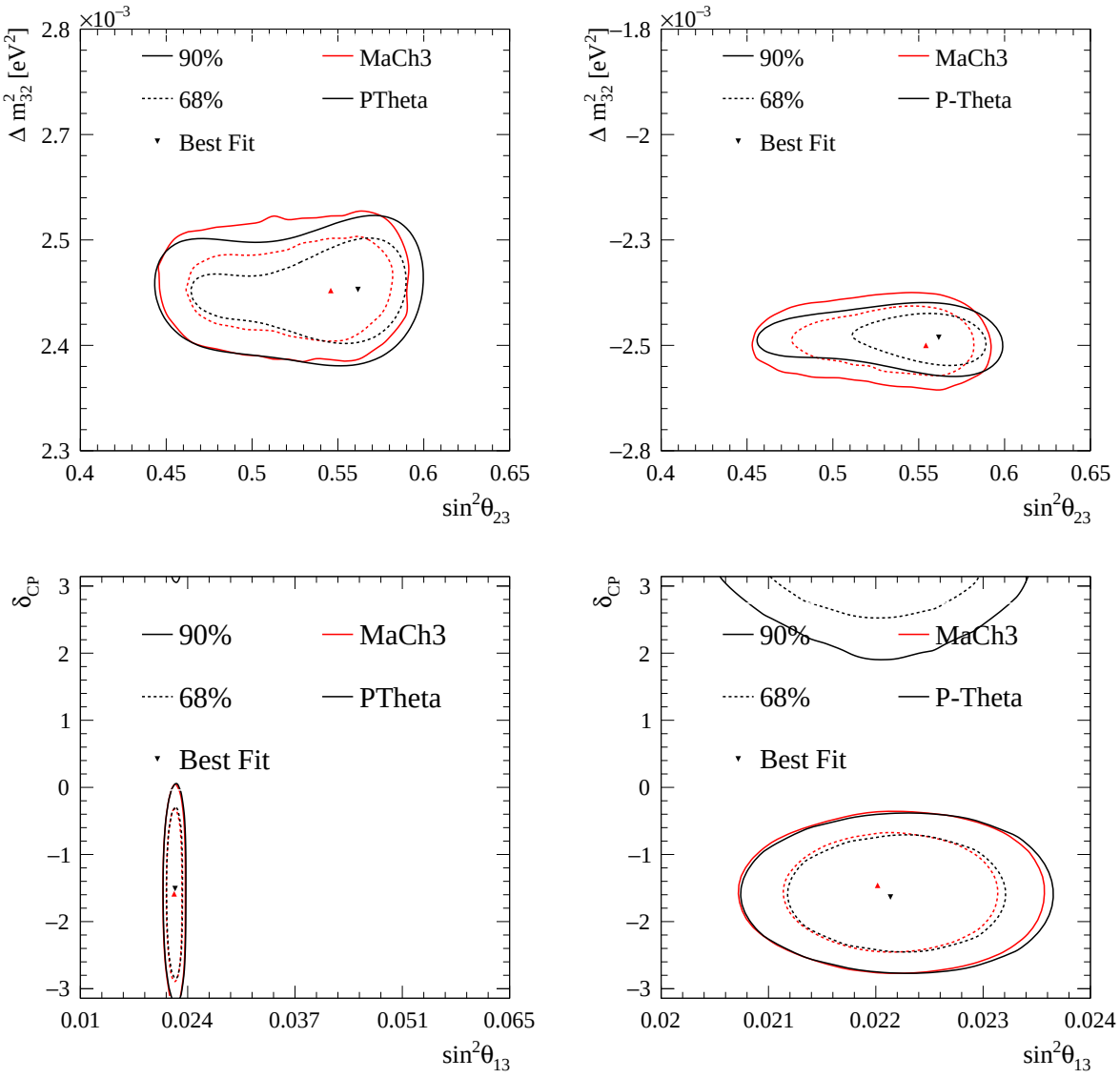

Figure 4.19: Comparison of oscillation parameter contours from the Asimov A22 fits of MaCh3 and PTheta with the reactor constraint. (Top) disappearance contours (bottom) appearance contours (left) normal hierarchy (right) inverted hierarchy. The roughness seen in the disappearance contours is a result of the relatively lower number of steps after applying the reactor contraint.

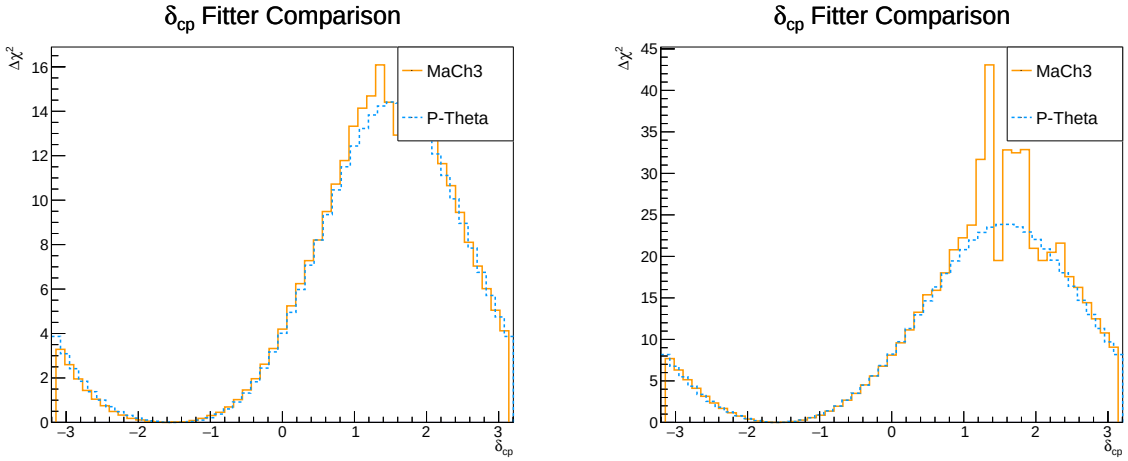

Figure 4.20: Comparison of 1D δ_{CP} $\Delta\chi^2$ between the MaCh3 and PTheta Asimov A22 fits with the reactor constraint; (Left) normal hierarchy (right) inverted hierarchy. The instability in the MaCh3 fit at larger values of $\Delta\chi^2$ is due to the reactor constraint moving steps outside of this region. Note that in figure 4.18, the maximum value of $\Delta\chi^2$ is much lower than that seen after application of the RC. This is due to the reactor constraint heavily restricting the range of appearance parameters. The large statistical noise at the higher values of $\Delta\chi^2$ seen in the inverted hierarchy plot is simply due to the low number of steps in this region made more evident by the RC weighting down many of the steps seen in figure 4.18.

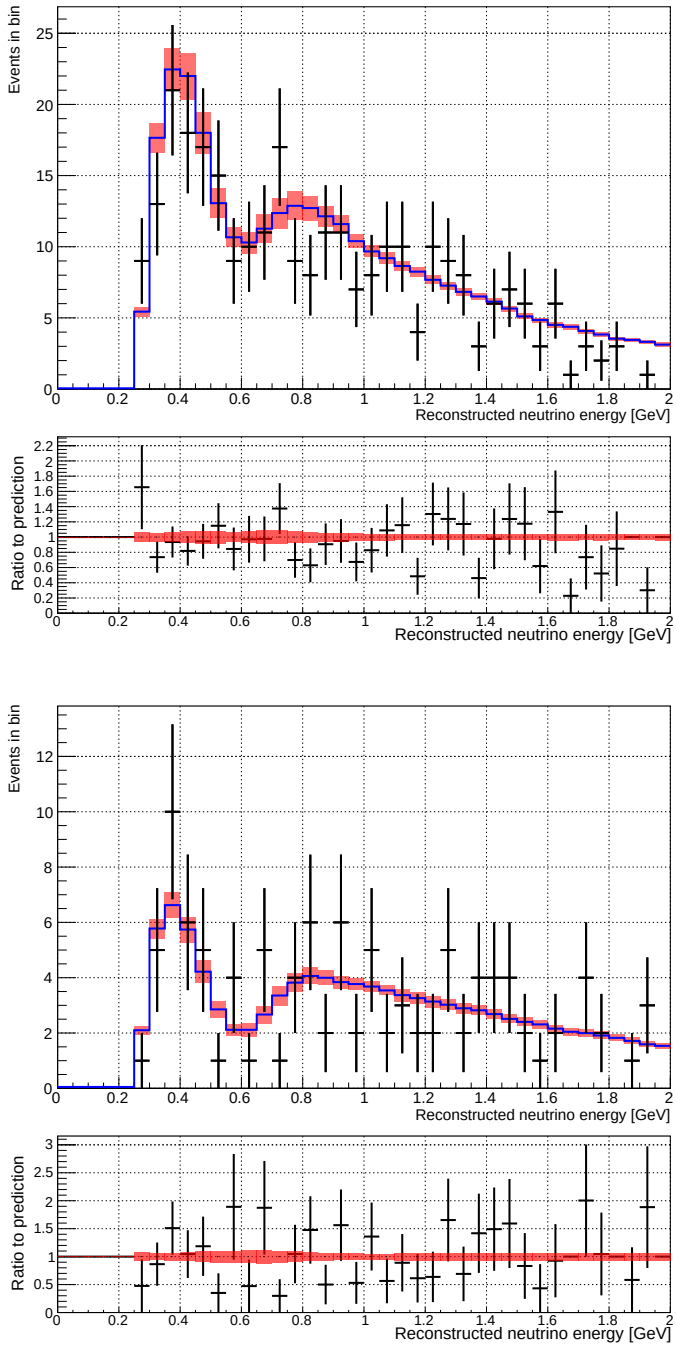

Figure 4.21: Plot shows the post-fit posterior predictive distribution (blue) generated by tuning the MC spectra to 5000 random throws from the FD+ND Markov Chain for the 1Rμ SK samples. The ν-mode sample is displayed at the top and $\bar{\nu}$-mode at the bottom. The 1 sigma error (red band) and data (black) used for these selections are overlaid.

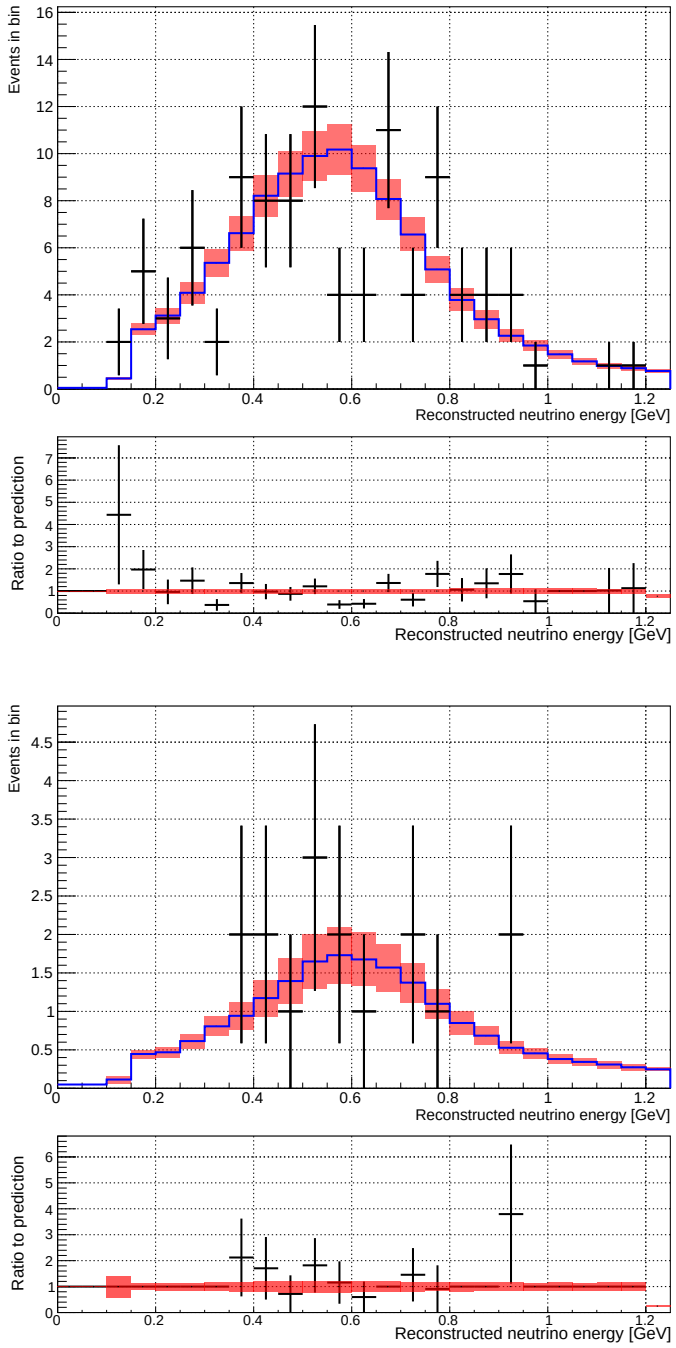

Figure 4.22: Plot shows the post-fit posterior predictive distribution (blue) generated by tuning the MC spectra to 5000 random throws from the FD+ND Markov Chain for the 1Re SK samples. The ν-mode sample is displayed at the top and $\bar{\nu}$-mode at the bottom. The 1 sigma error (red band) and data (black) used for these selections are overlaid.

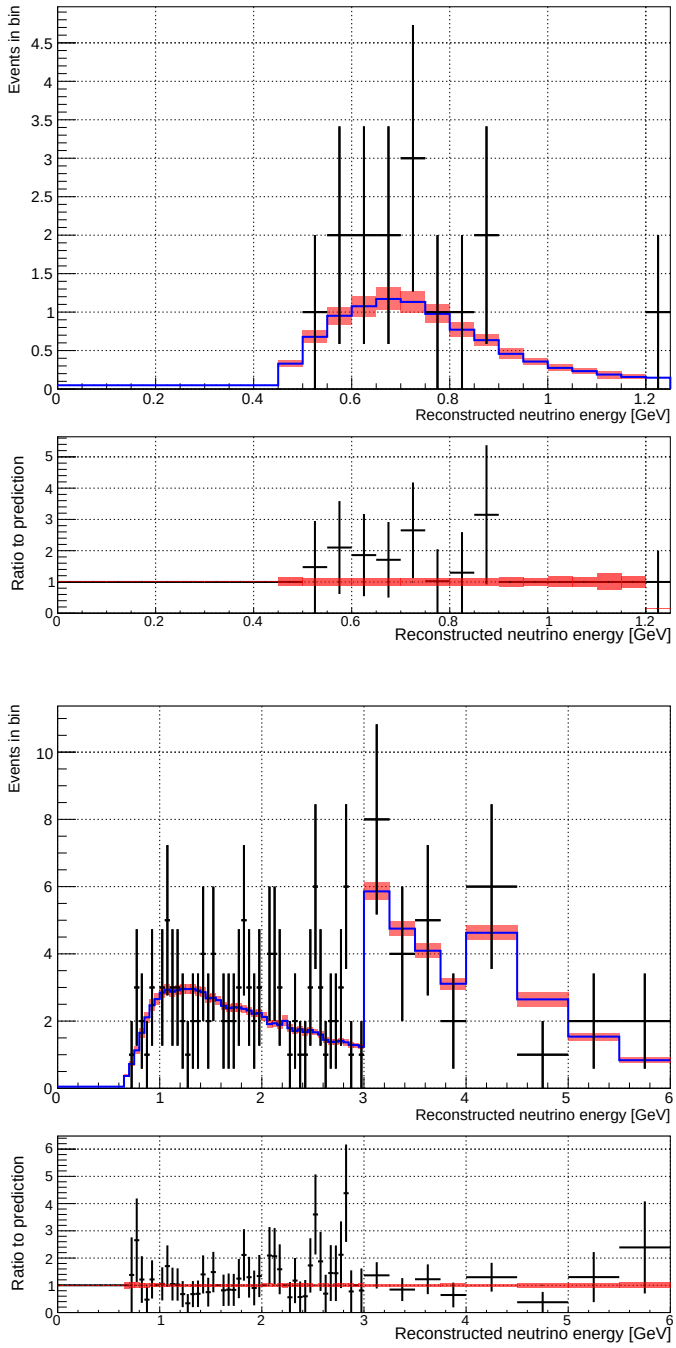

Figure 4.23: Plot shows the post-fit posterior predictive distribution (blue) generated by tuning the MC spectra to 5000 random throws from the FD+ND Markov Chain for the multi-ring SK samples. The ν-mode 1Re1d.e. sample is displayed at the top and ν-mode ν_μCC1π at the bottom. The 1 sigma error (red band) and data (black) used for these selections are overlaid.

between samples). It was found that this low value was primarily driven by the 1Rμ sample. Figure 4.24 shows the distribution of the number of MC events for several intervals of reconstructed energy. In particular, whilst the model and data are in relatively good agreement for the oscillation dip range [0.0, 0.9] GeV, the two higher energy ranges, [0.9, 2.0) GeV and [2.0, 10] GeV predict very different results to the spectra. Since the bulk of the data for this sample lie in this oscillation dip region and the fit is not a rate-only fit (the likelihood assessment is bin based and hence shape+rate information is used), it was decided that this low rate-based p-value was acceptable for this analysis.

4.7.2 MaCh3 Posteriors

As in section 4.6.2, the posteriors are presented with and without reactor constraint (RC). All posteriors are shown with marginalisation over normal ordering, inverted ordering and without ordering-based marginalisation.

4.7.2.1 1D Posteriors

Figures 4.27, 4.26, 4.28, and 4.28 show the 1D posterior distributions for the oscillation parameters. The first notable conclusion of this analysis is that CP conserving values of δ_{CP} ($\delta_{CP} \in \{k\pi : k \in \mathbb{Z}\}$) are excluded at the 90% confidence interval when the reactor constraint is applied; this can be seen more clearly in figure 4.29. In addition, the reactor constraint also changes the octant of the best fit point [Fig. 4.25] and increases the preference for normal ordering [Fig. 4.26]. The impact of this constraint on $\sin^2(\theta_{13})$ can be seen in figure 4.28 with phase space being limited to a far smaller range of values. The proportion of MCMC steps in each octant and hierarchy are displayed in tables 4.8 (without RC) and 4.9 (with RC).

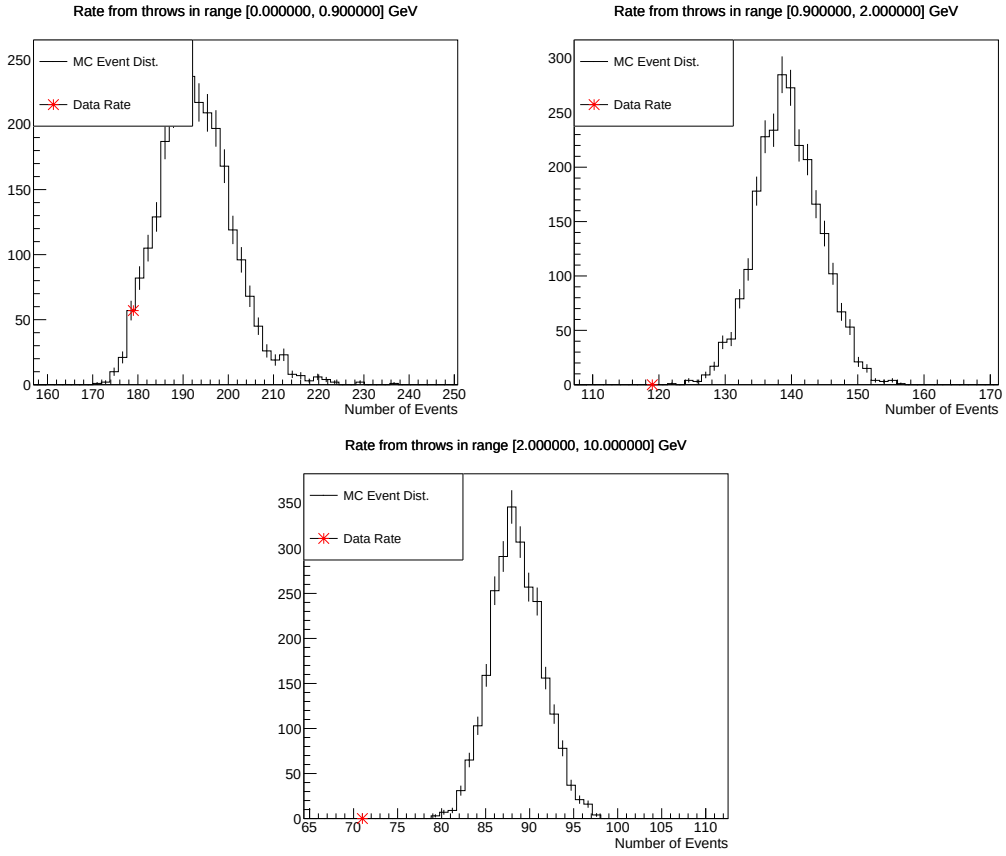

Figure 4.24: Distribution of the expected number of events for 2500 model generated spectra at parameter values sampled from the 2023 data fit (black) within fixed energy bins is compared to to observed number of data events (red) in those bins in the ν-mode 1Rμ sample. Poor agreement is seen for energy ranges outside the "oscillation dip" ([0.9, 2.0] GeV and [2.0, 10.0] GeV) but the MC re-weight agrees relatively well within the oscillation dip region [0, 0.9] GeV.

	$\sin^2\theta_{23} < 0.5$	$\sin^2\theta_{23} > 0.5$	Sum
NH ($\Delta m^2_{32} > 0$)	0.27	0.37	0.63
IH ($\Delta m^2_{32} < 0$)	0.17	0.20	0.37
Sum	0.43	0.57	1.000

Table 4.8: Proportion of steps in the normal and inverted hierarchies, as well as upper and lower octants, from the posterior without RC.

	$\sin^2\theta_{23} < 0.5$	$\sin^2\theta_{23} > 0.5$	Sum
NH ($\Delta m^2_{32} > 0$)	0.23	0.54	0.77
IH ($\Delta m^2_{32} < 0$)	0.05	0.18	0.23
Sum	0.28	0.72	1.00

Table 4.9: Proportion of steps in the normal and inverted hierarchies, as well as upper and lower octants, from the posterior with RC.

4.7.2.2 2D Posteriors

2D disappearance and appearance contours are shown in figures 4.30 and 4.31 respectively. As is expected from the Asimov fit the application of the reactor constraint results in a change of octant preference in the 2D marginalised posterior [Fig 4.30]. Application of the reactor constraint [Fig. 4.32] significantly limits the phase space of the appearance contours resulting in far more values of δ_{CP} being excluded.

4.7.3 The Jarlskog Invariant

A key feature of MCMC is that individual parameters from the chain can be combined to examine composite parameters. The additional key parameter interest in oscillation analysis which is not directly fitted by MaCh3 is the Jarlskog invariant [Eqn. 1.16]. $J_{CP} \neq 0$ is the full condition required for PMNS CP violation, thus, it is convenient to consider it as well as non-CP conserving values of δ_{CP}. The plots in figure 4.33 show the

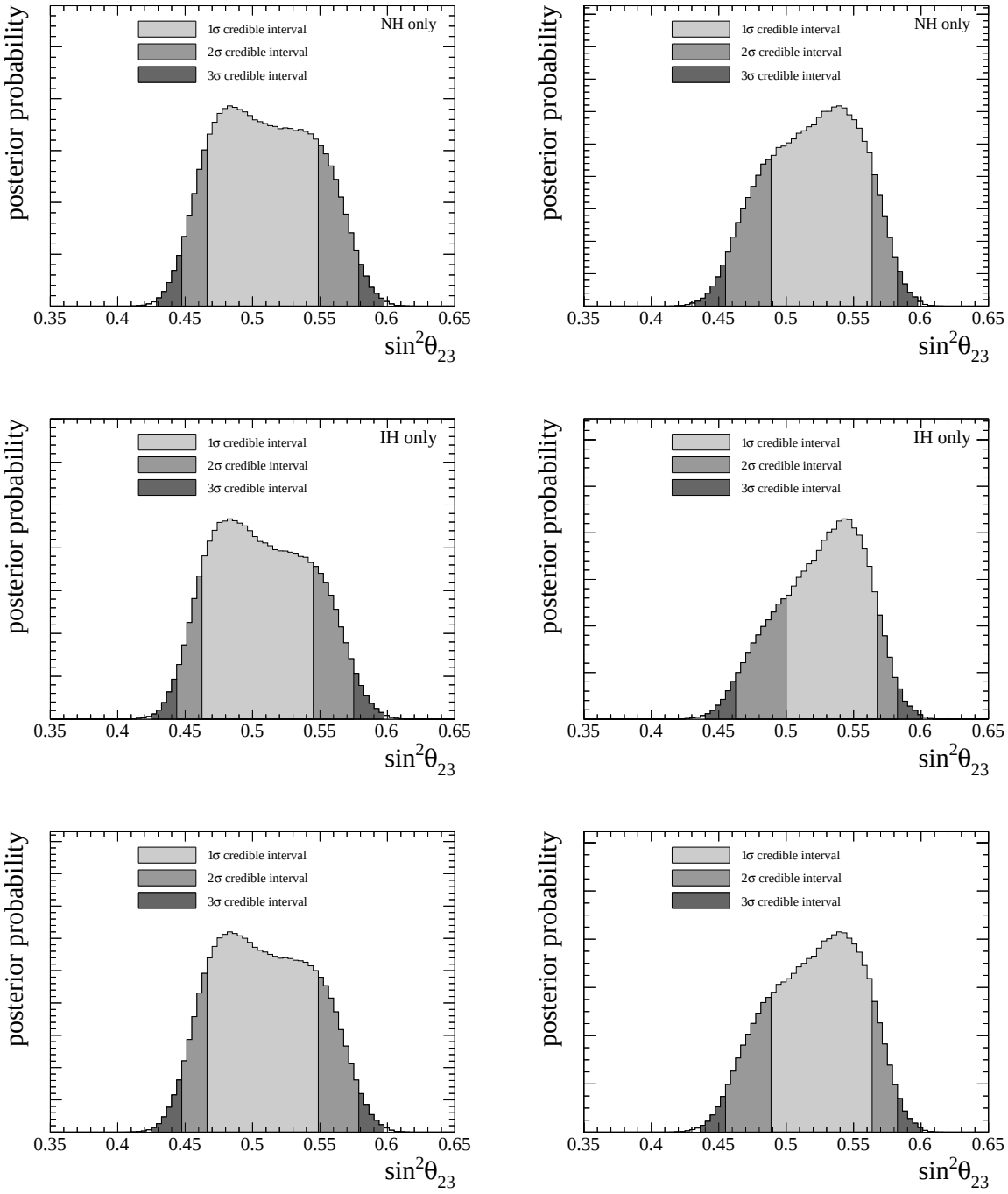

Figure 4.25: 1D 60%, 90%, and 99% credible intervals for $\sin^2\theta_{23}$ from the run 1-11 data fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals. Fits are displayed marginalised over normal ordering (top), inverted ordering (middle), without ordering marginalisation (bottom).

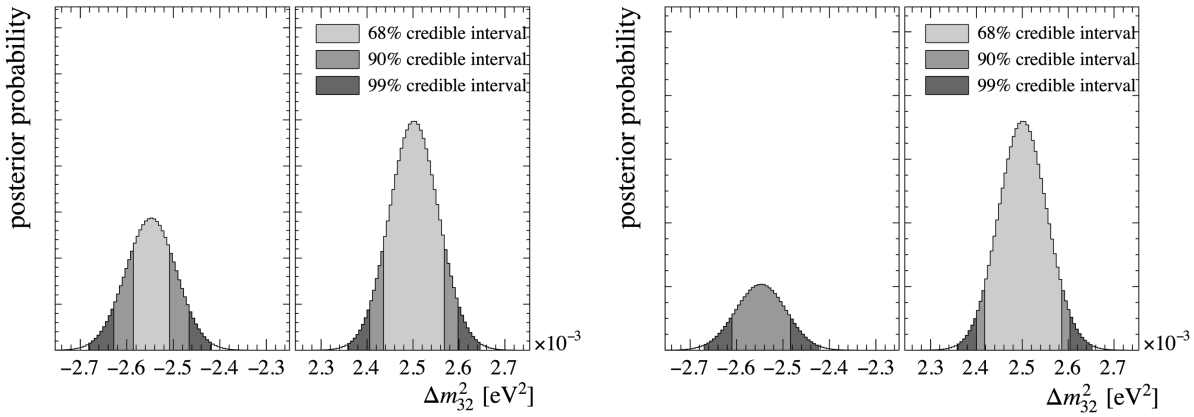

Figure 4.26: 1D 60%, 90%, and 99% credible intervals for Δm^2_{32} from the run 1-11 data fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals. Fits are displayed marginalised over both orderings

posteriors of the Jarlskog invariant under two different choices of prior. The flat $\sin\delta_{CP}$ prior was achieved by re-weighting a chain post-fit. In the normal ordering only and non-marginalised posteriors choice of prior also has an effect on the physics conclusions. In particular under the flat δ_{CP} prior, $J_{CP} = 0$ is excluded at the 2σ credible interval in both the normal ordering and the non-marginalised posteriors. By changing prior to one that is flat in $\sin(\delta_{CP})$, the credible intervals shift to include $J_{CP} = 0$ within this credible interval. This illustrates the importance of good prior justification alluded to in chapter 3 as it can and does alter physics conclusions. This effect is also shown in the 2D posteriors in figure 4.34 which shows J_{CP} plotted against $\sin^2(\theta_{23})$ for the same choice of priors. The choice of $\sin^2(\theta_{23})$ is driven by T2K sensitivities since, alongside δ_{CP}, it is the parameter present in the Jarlskog invariant that T2K is most sensitive to. As in the 1D case changing priors also moves the credible intervals to include/exclude CP conservation at the 2σ credible interval.

4.7.4 Comparisons to Previous Analyses

This section compares the data fits from this analysis to the previous T2K analysis. The major differences between these analyses are an increase in POT and the new SK

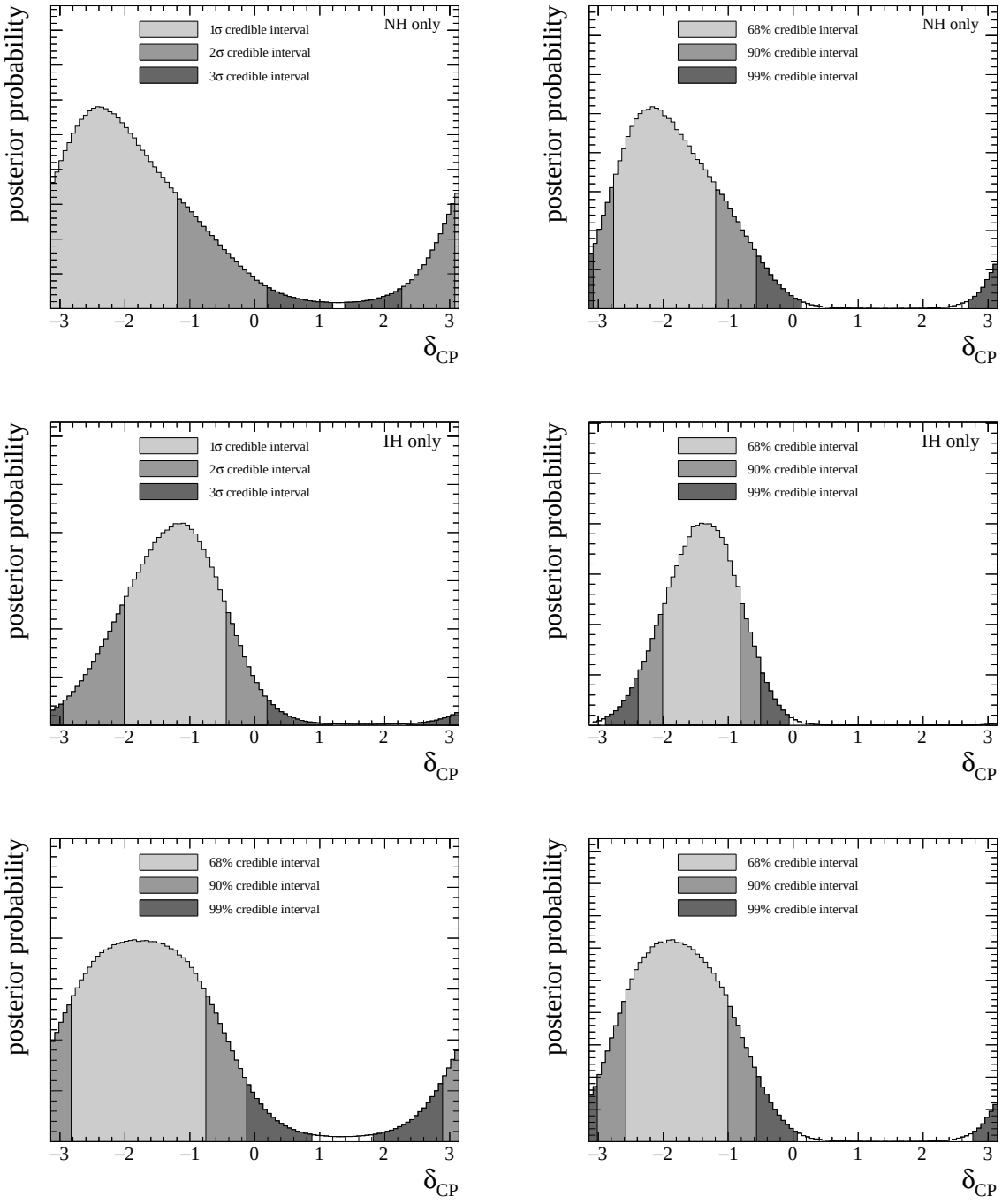

Figure 4.27: 1D 60%, 90%, and 99% credible intervals for δ_{CP} from the run 1-11 data fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals. Fits are displayed marginalised over normal ordering (top), inverted ordering (middle), without ordering marginalisation (bottom).

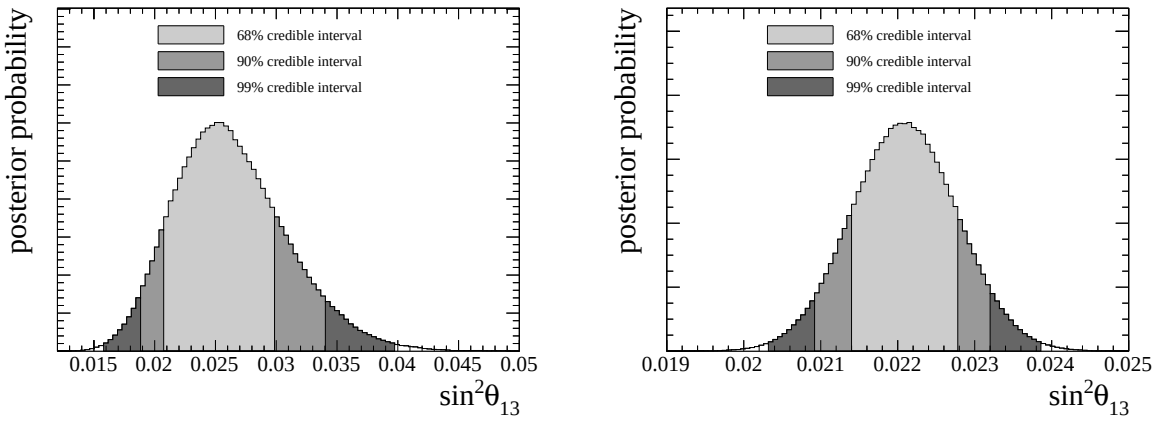

Figure 4.28: 1D 60%, 90%, and 99% credible intervals for $\sin^2\theta_{13}$ from the run 1-11 data fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals. Fits are displayed marginalised over normal ordering (top), inverted ordering (middle), without ordering marginalisation (bottom).

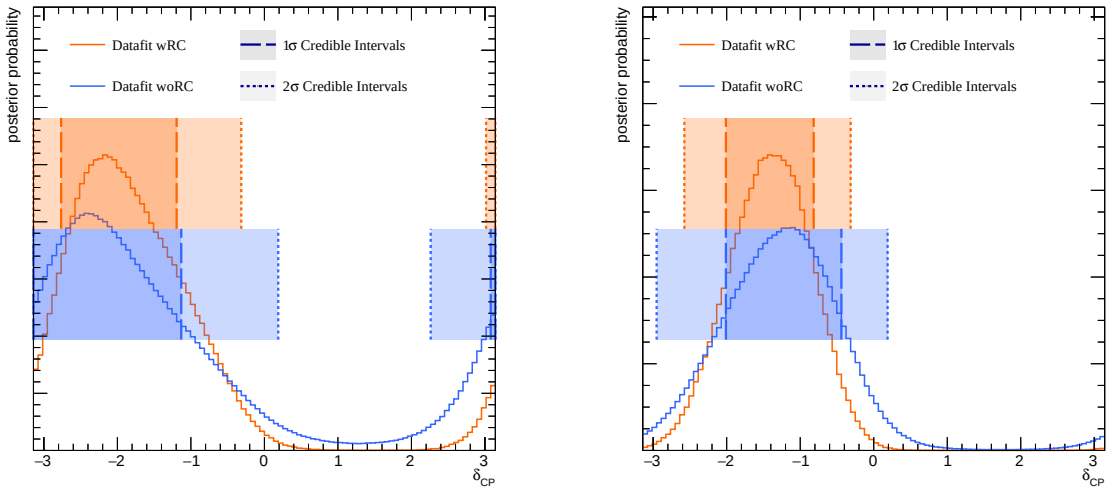

Figure 4.29: Comparison of 1D δ_{CP} posterior probability projections and 1 and 2σ credible intervals between data fits with and without the reactor constraint. (Left) normal hierarchy (right) inverted hierarchy.

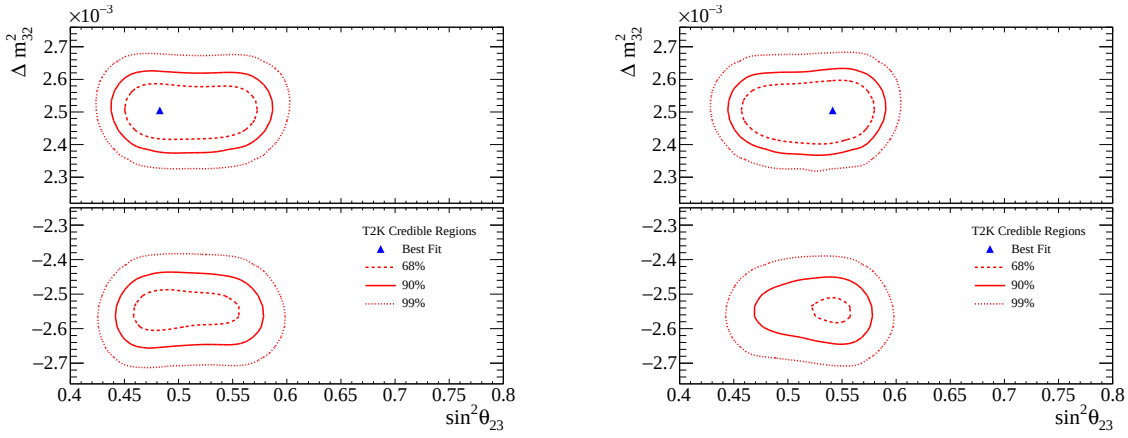

Figure 4.30: 2D $\sin^2\theta_{23}$-Δm^2_{32} contours from the run 1-11 data fit with (right) and without (left) reactor constraint showing the 68%, 90%, and 99% credible intervals.

detector uncertainty model. The near detector and cross-section systematics models have remained unchanged. As would be expected from this small POT increase and minor systematics change the differences are relatively minor with no changes to physics conclusions. More positively, a small tightening in constraints for all parameters can be seen and the δ_{CP} constraint is moving towards rejecting CP conservation at the 2σ CI ($\approx$95%).

4.7.5 Comparisons to PTheta

Figure 4.37 shows 2D contours from reactor constrained MaCh3 data fit compared to those from PTheta. Figure 4.38 shows a corresponding comparison of the 1D marginalised posterior distribution for δ_{CP} and $\sin^2\theta_{23}$. Comparisons without the reactor constraint are omitted for brevity.

Here we show two PTheta contours: the black contours show the results from a standard PTheta fit, and the blue contours, labelled 'PTheta, MaCh3-like' are results from a PTheta fit using the MaCh3 SK sample binning constraints from a MaCh3 ND fit.

Only small differences in the 1D $\Delta\chi^2$ of δ_{CP} can be seen in Figure 4.38. Further, χ^2 values above 9 (3σ) are not meant to be included in the report of this analysis. This is

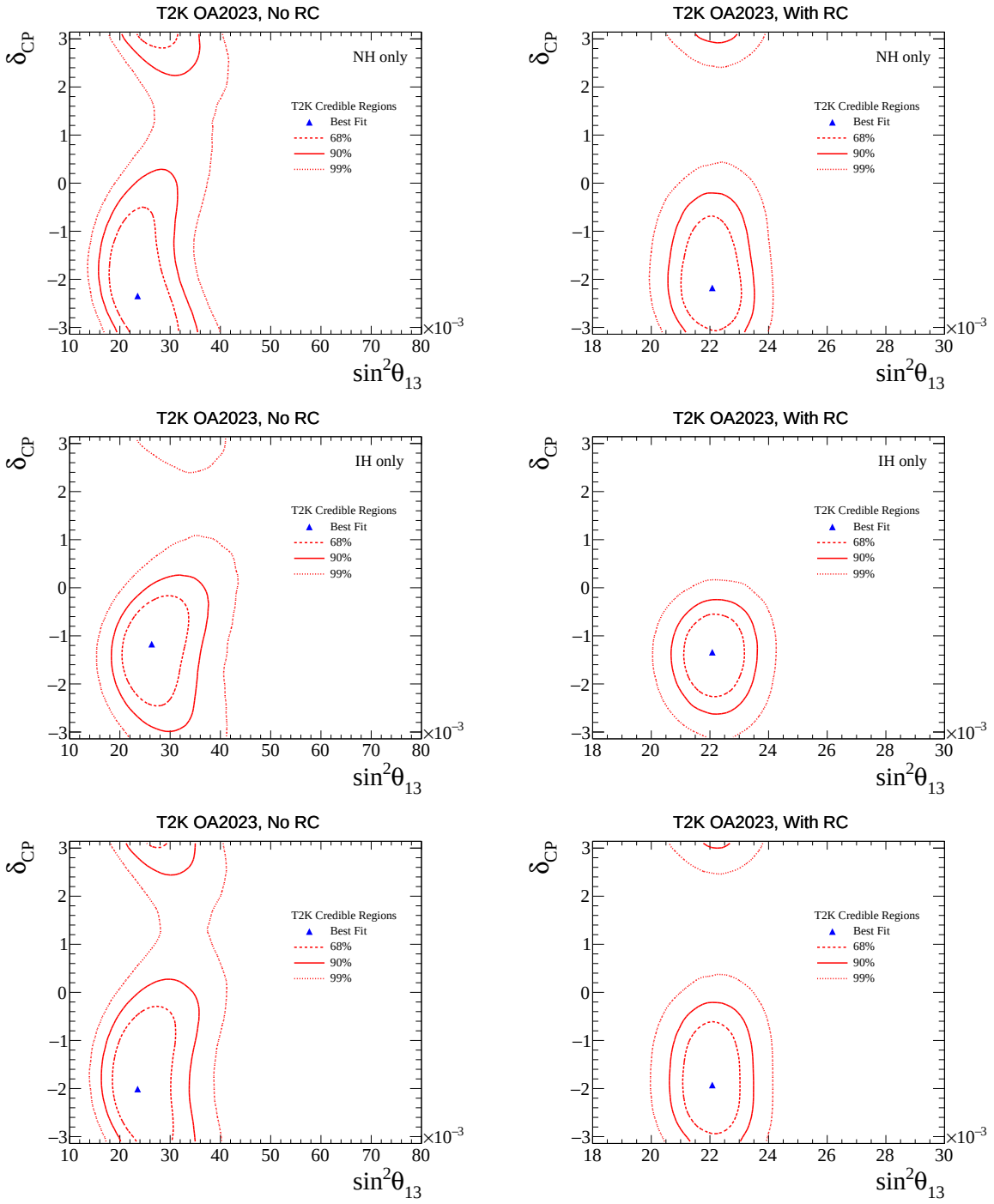

Figure 4.31: 2D $\sin^2\theta_{13}$-δ_{CP} contours from the run 1-11 data fit without reactor constraint. (Left) 68/90/99% credible intervals, (right) $1/2/3\sigma$ credible intervals. (Top) normal hierarchy (middle) inverted hierarchy (bottom) marginalized over both hierarchies.

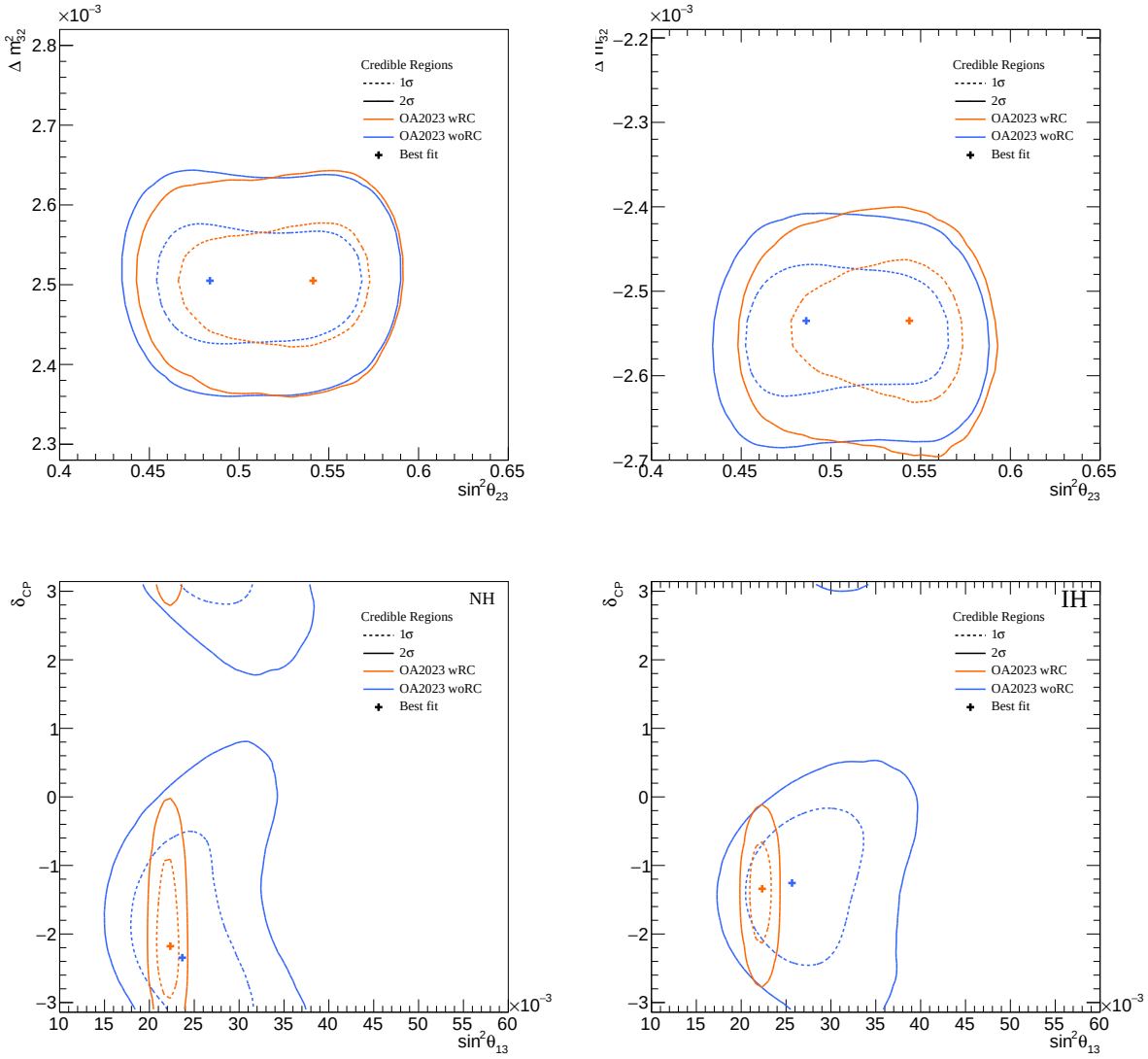

Figure 4.32: Comparison of oscillation parameter contours from the data fits of MaCh3 with and without the reactor constraint. (Top) disappearance contours (bottom) appearance contours; (left) normal hierarchy; (right) inverted hierarchy.

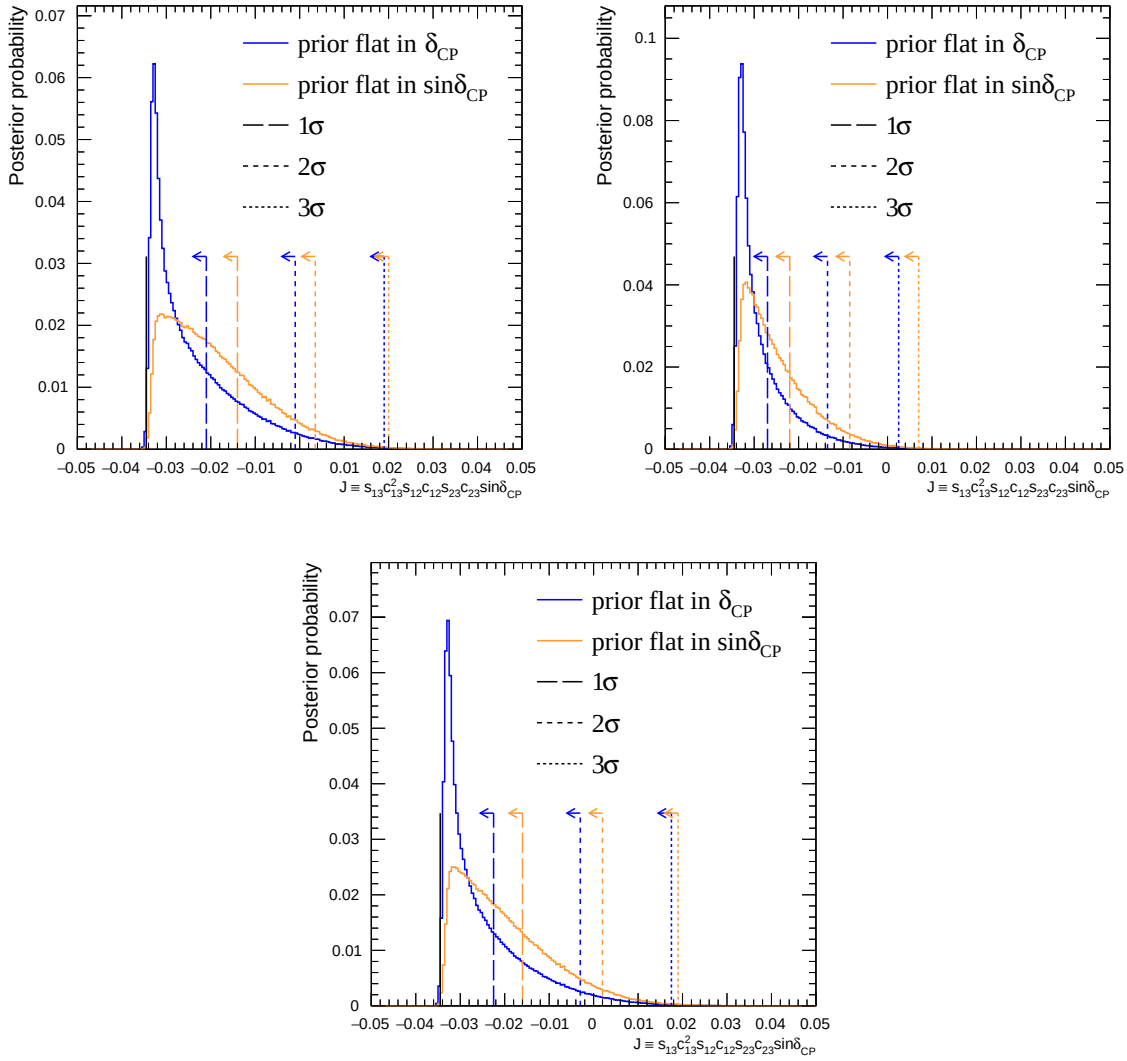

Figure 4.33: Comparison of 1D Jarlskog posterior distributions from priors that are flat in either (i) δ_{CP} or (ii) $\sin\delta_{CP}$. Plots broken up by hierarchy [NH top left, IH top right, both bottom].

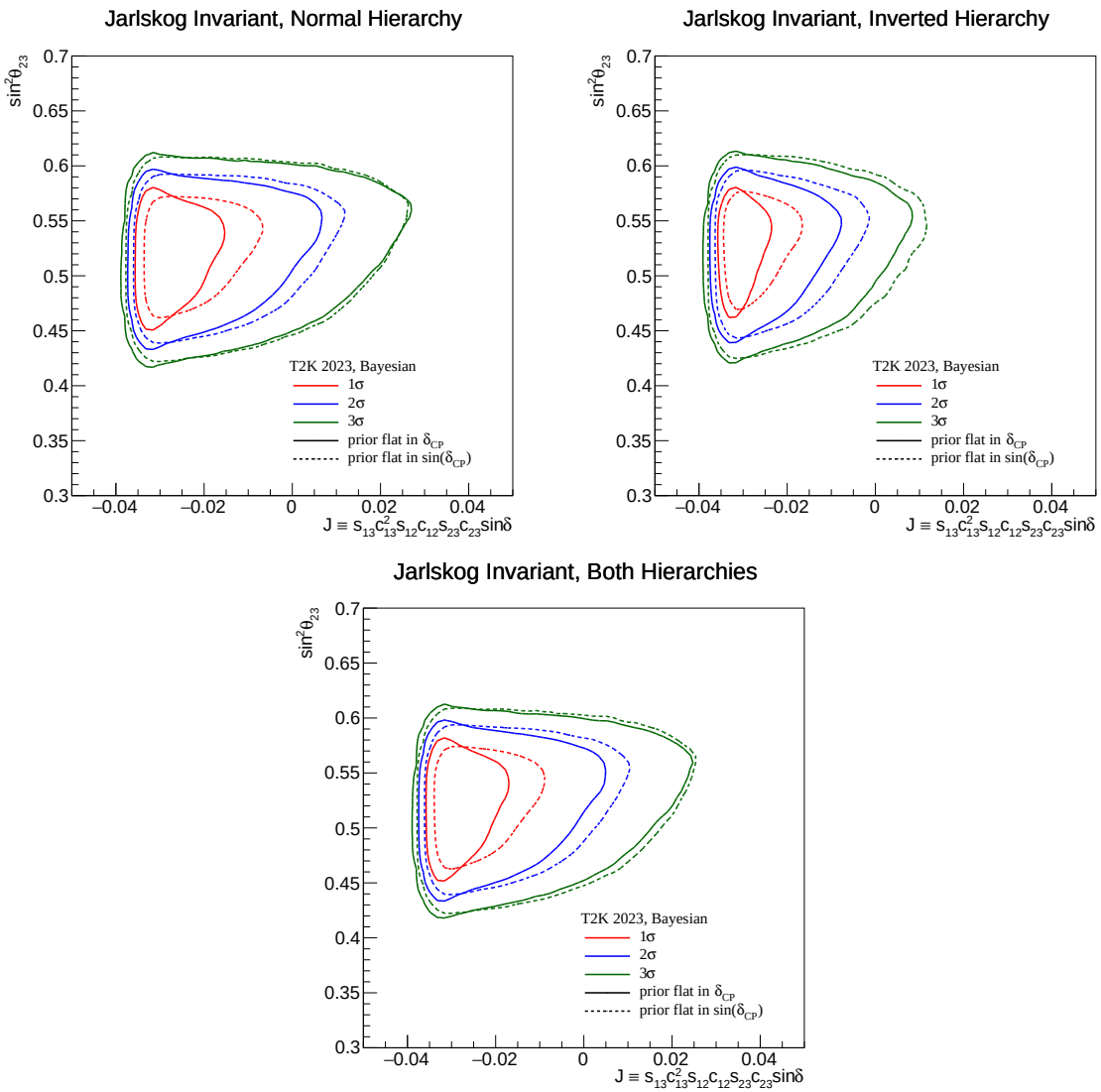

Figure 4.34: Comparison of 2D Jarlskog posterior distributions against $sin^2(\theta_{23})$ from priors that are flat in either (i) δ_{CP} or (ii) $\sin\delta_{CP}$ marginalised over the normal hierarchy (top left), inverted hierarchy (top right) or without hierarchy marginalisation (bottom).

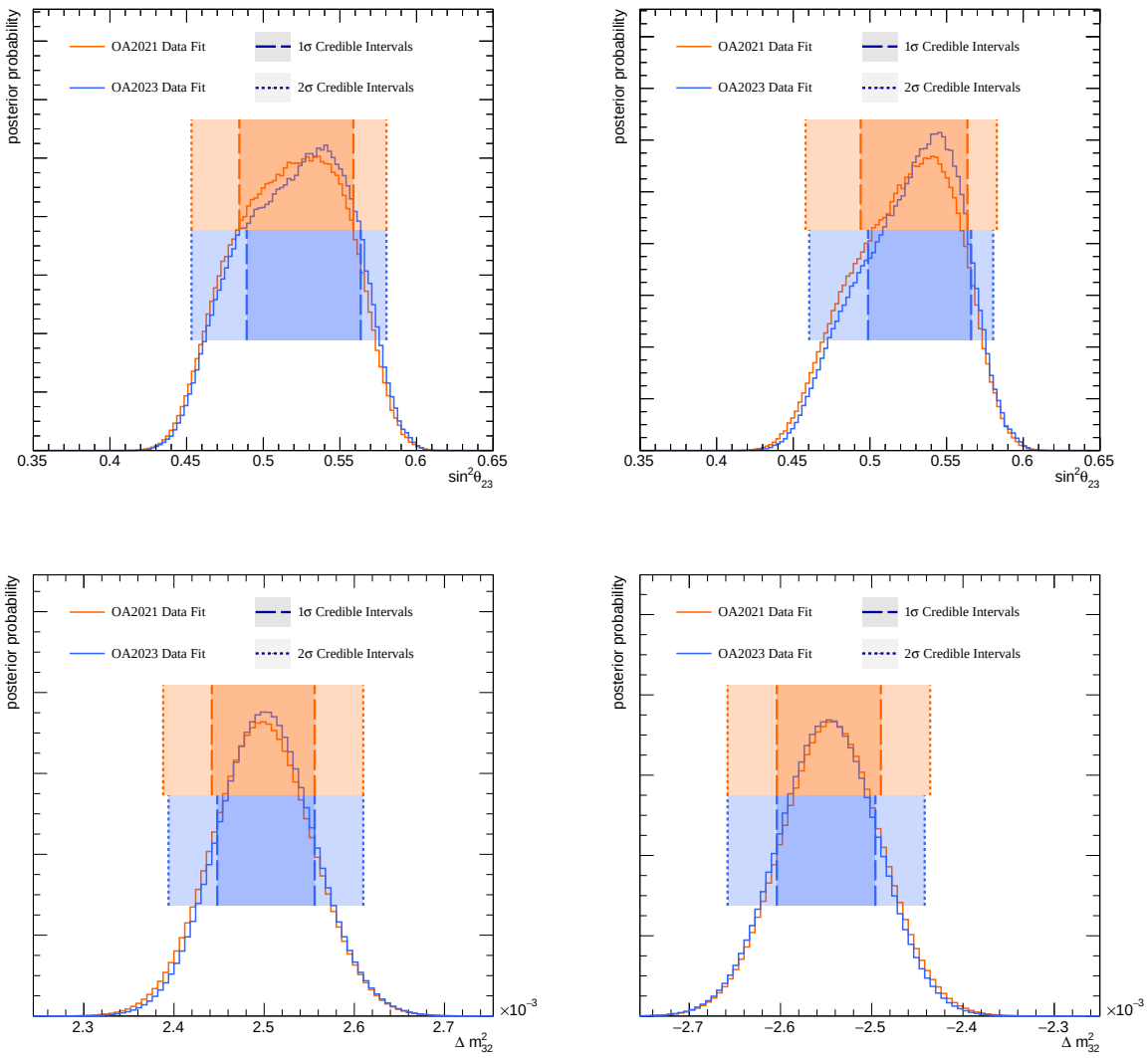

Figure 4.35: Comparison of 1D $\sin^2\theta_{23}$ and Δm^2_{32} posterior probability projections between the data fits, with the reactor constraint, of MaCh3 from the previous analysis (orange) and current analyses (blue). (Left) normal hierarchy (right) inverted hierarchy.

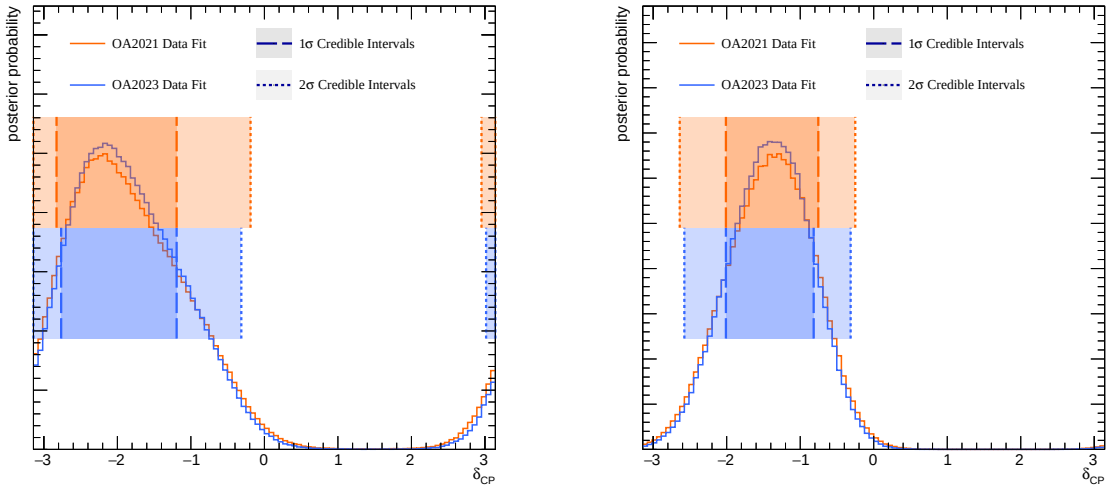

Figure 4.36: Comparison of 1D δ_{CP} posterior probability projections between MaCh3 data fits with the reactor constraint between the previous analysis [139] (orange) and current analyses (blue). (Left) normal hierarchy (right) inverted hierarchy.

a heavily disfavoured region of phase space, and to properly sample this region MaCh3 would require additional steps in the Markov chain. Further, there are small disagreements in the lower octant in the 2D $\sin^2\theta_{23}$ $\Delta\chi^2$, shown in Figure 4.37. Crucially, the choice of PTheta analysis does not meaningfully change the final physics conclusions and, as a result, it was concluded that PTheta and MaCh3 provide a valid analysis of T2K data.

4.8 Summary of Data Fits

The T2K data fit pipeline is a rigorous statistical examination of the experiment's oscillation program requiring very different approaches to statistical analysis to produce near-identical results. The final best-fit results with reactor constraint are displayed in table 4.10 and show the best fit point and 1σ credible interval marginalised over the disappearance and appearance parameters respectively. This result represents the first to use SK with Gd and paves the way for future analyses using this data. Further results from this analysis are displayed in appendix B.

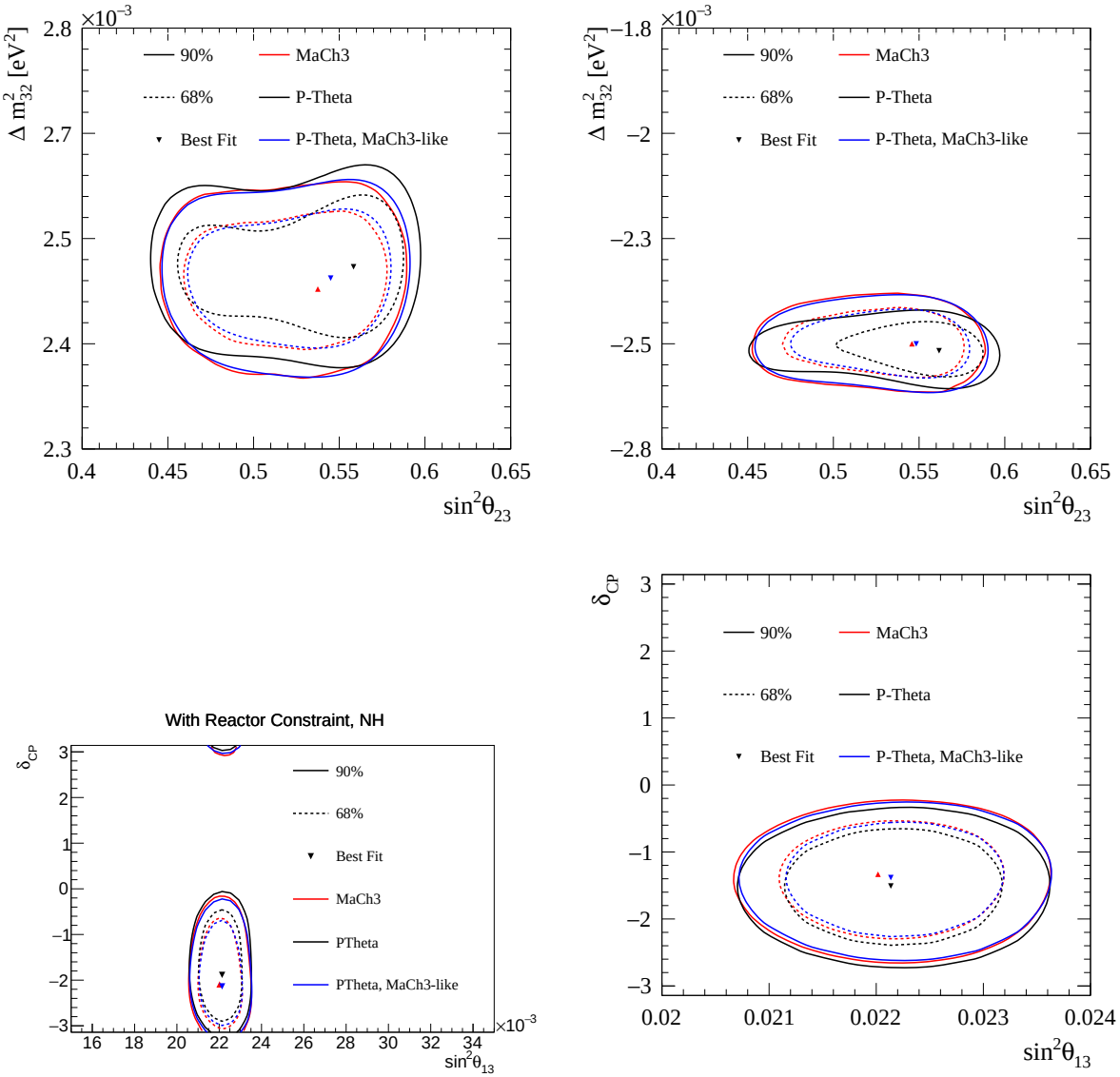

Figure 4.37: Comparison of oscillation parameter contours from the data fits of MaCh3 and PTheta applying the reactor constraint. (Top) disappearance contours (bottom) appearance contours (left) normal hierarchy (right) inverted hierarchy.

	$\sin^2\theta_{23}$	$\Delta m^2_{32}(\times 10^{-3})\text{eV}^2$	$\sin^2\theta_{13}$	δ_{CP}
2D best fit	0.540	2.50	0.022	-1.85
68% C.I. (1σ) range	0.49 – 0.56	-2.63 – -2.47 $\cup$ 2.38 – 2.62	0.021 – 0.023	-2.60 – -1.01

Table 4.10: Best-fit values for disappearance oscillation parameters, with reactor constraint applied. The 2D best-fit values are taken from the mode of the 2D marginal posterior distributions in $\sin^2\theta_{23} - \Delta m^2_{32}$ space for disappearance parameters and $\sin^2\theta_{13} - \delta_{CP}$ space for appearance parameters. The 1D 68% credible intervals correspond to the 1σ central area of the marginalised posterior distributions.

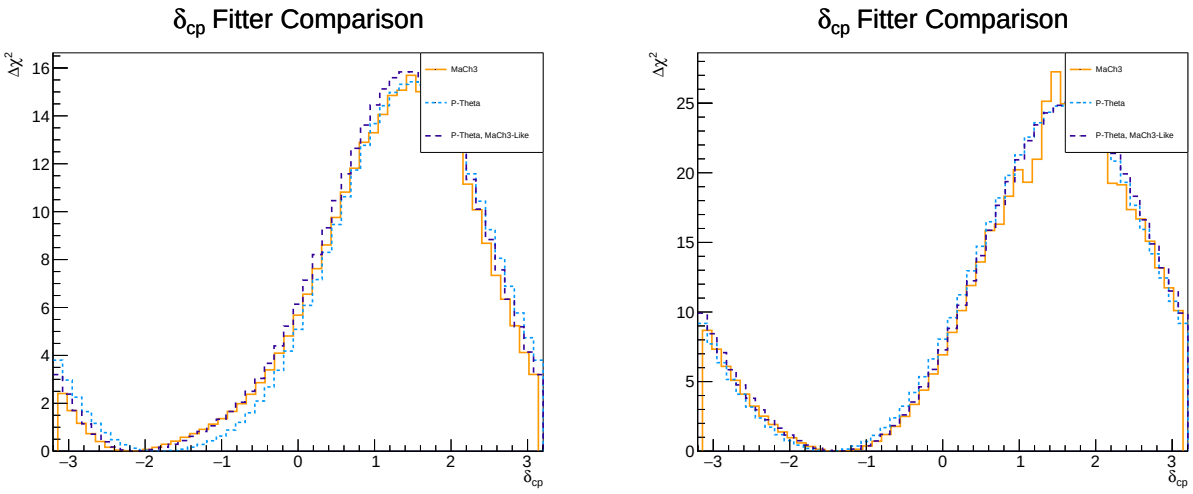

Figure 4.38: Comparison of 1D δ_{CP} posterior probability projections between the MaCh3 and PTheta data fits with the reactor constraint. (Left) normal hierarchy (right) inverted hierarchy.

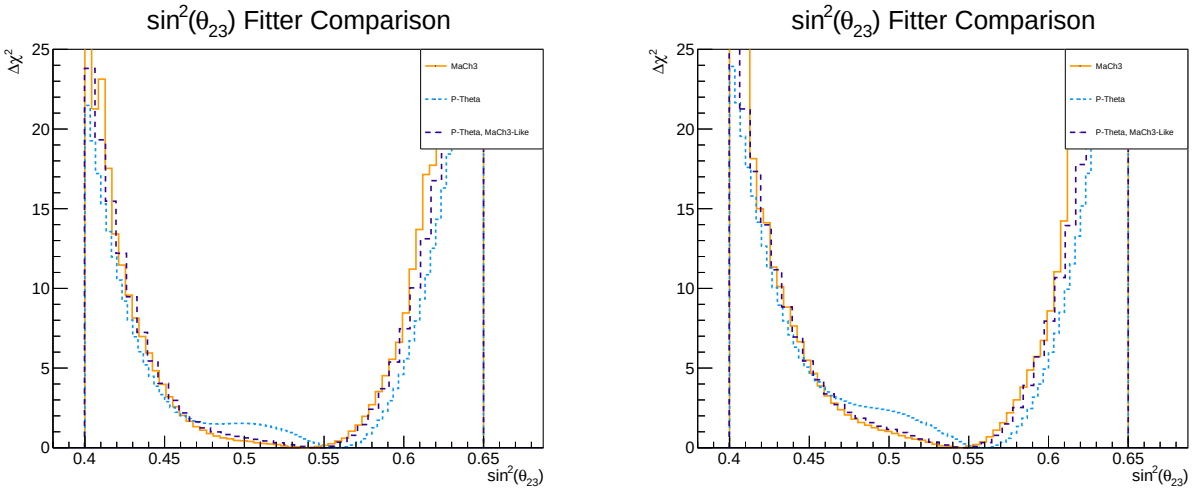

Figure 4.39: Comparison of 1D $\sin^2\theta_{23}$ posterior probability projections between the MaCh3 and PTheta data fits with the reactor constraint. (Left) normal hierarchy (right) inverted hierarchy.

Chapter 5

Sensitivity Studies for Hyper-Kamiokande

5.1 Overview

The Hyper-Kamiokande (HK) experiment is the next step in long-baseline Water-Cherenkov based accelerator neutrino experiments. Although it is due to be constructed in 2027, systematics studies are being performed now in order to both analyse expected physics results and ensure that analyses can correctly converge given the large increase in statistics present at HK.

The HK systematics group currently bases analyses on previous T2K fits. In order to make these sensitivity studies more "HK-like", the following changes are made to MaCh3:

- 10 year run time is assumed with a total of 2.7×10^{22} protons on target
- POT is split into a 3:1 $\bar{\nu}$:ν-beam ratio
- The detector fiducial volume is scaled by a factor of 8.31 and the surface area by 3.56

The analysis inputs are then simply those used in previous T2K analyses. The two fits

Parameter	Nominal
$\sin^2(\theta_{12})$	0.307
$\sin^2(\theta_{23})$	0.528
$\sin^2(\theta_{12})$	0.0218
Δm^2_{21}	7.53×10^{-5} eV^2
Δm^2_{32}	2.509×10^{-3} eV^2

Table 5.1: The AsimovA20 parameter set, note that δ_{CP} is omitted as it is varied on a fit-by-fit basis.

presented in this chapter use inputs from the 2018 and 2020 T2K oscillation analyses with the change in analysis occurring due to technical difficulties maintaining and running substantially older versions of the code. Regardless of analysis version all fits use the AsimovA20 oscillation parameter set [Tab. 5.1] with varying values of δ_{CP}.

This section will detail the early stages of developing a fitter as well as the convergence issues encountered when attempting to run these fits.

5.2 Fits using the 2018 T2K Model

5.2.1 Overview of the 2018 Model

The T2K 2018 model is, in many respects, very similar to the analysis presented in the previous chapter. The most notable differences for these fits are the lack of the ν_μCC1π sample and a substantially smaller number of cross-section parameters. In addition there are several major code differences within MaCh3 which lead to far slower step times in this version of the code. These include separated class objects for each far detector sample and a different spline treatment.

5.2.2 Validation

Validation was performed with the frequentist VALOR [141] fitter using the same event rate comparison seen in T2K analyses. In order to check self consistency, MaCh3 first applied each scaling step sequentially. For 2018 fits additional flux tuning was available

Sample	Valor	MaCh3	%Difference
ν-mode 1Rμ	9349.3	9345.52886	0.0403442
ν-mode 1Re	2739.76	2739.07106	0.02514916
ν-mode 1Re1d.e	257.63	257.43359	0.07626631
$\bar{\nu}$ 1Rμ	12375.02	12365.37756	0.07794895
$\bar{\nu}$ 1Re	1623.97	1622.66617	0.08031882

Table 5.2: Table showing final OA2018 validations between MaCh3 and VALOR using the AsimovA20 parameter set at post-BANFF tune.

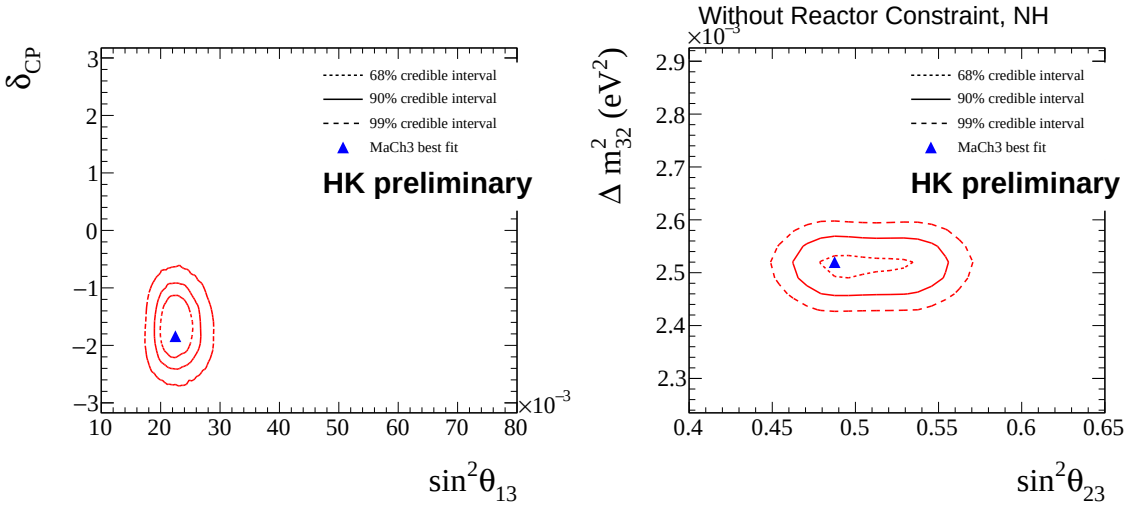

Figure 5.1: Appearance (left) and disappearance (right) contours for a HK Asimov fit using the 2018 systematics model. Oscillation parameters are set to the AsimovA20 nominal values [Tab. 5.1] with $\delta_{CP} = -\pi/2$. The disappearance contour has been marginalised over the normal hierarchy as there are almost no steps in the inverted hierarchy.

to account for the differences in expected neutrino flux at SK vs HK and allows for the fit to account for the extra beam power planned for HK. Practically speaking this is the scaled MC spectra by the ratio of the the expected HK flux to the ratio of the expected SK flux for each energy bin. The event rate comparison with VALOR after flux tuning has been applied is shown in table 5.2. In addition sigma variation comparisons showed good agreement.

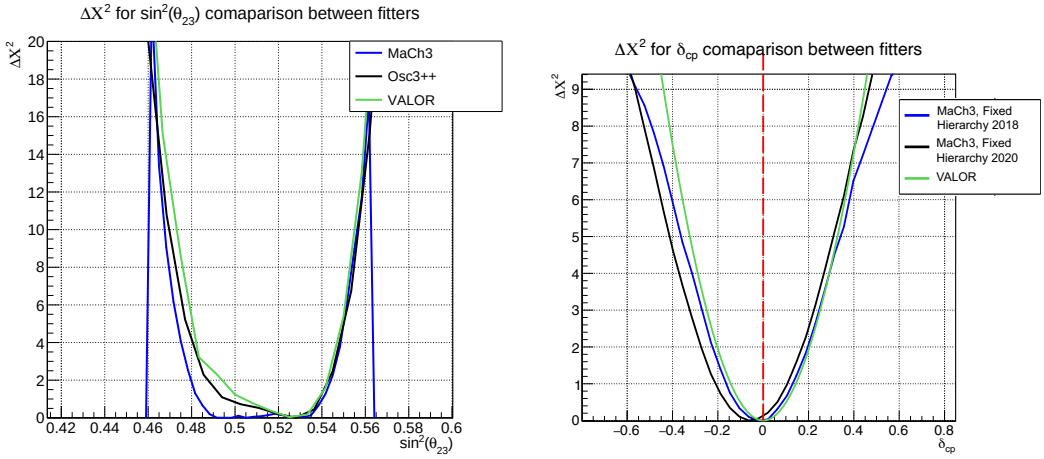

Figure 5.2: Comparison of fit results for MaCh3, VALOR and Osc3++ for multiple fits. (Left) Shows a comparison between all 3 fitters using the 2018 T2K Systematics model; notably MaCh3 predicts a very different shape to the other two fitters. This was later thought to be caused by the incorrect application of the reactor constraint within MaCh3. (Right) Shows a comparison between the 2018 and 2020 MaCh3 models with VALOR for a fixed normal hierarchy fit assuming $\delta_{CP} = 0$. Whilst VALOR and 2018 MaCh3 are in relatively good agreement, it is clear there are some substantial differences between these two fits and 2020 MaCh3.

5.2.3 Initial Fits

Early sensitivity studies were performed with two sets of oscillation parameters, these used the Asimov A20 set detailed table 5.1 but with $\delta_{CP} = 0$ and $\delta_{CP} = -\frac{\pi}{2}$ as these are values that show minimal/maximal CP violation. In addition, these fits were performed without varying the values systematics in order to test the "statistics-only" capabilities of the model. Since the oscillation parameter likelihood space has a different geometry to the T2K space, significant effort was spent tuning the step sizes of the oscillation parameters. This primarily involved running several fits with a variety of step scales and examining traces, autocorrelations and acceptance rates [for more information on this see section 3.5.1]. The first full fits were run and resulted in the contours seen in figure 5.1. Somewhat confusingly, this resulted in almost no steps being accepted in the inverted hierarchy when for the parameter set with $\delta_{CP} = -\frac{\pi}{2}$. In order to understand this properly it helps to consider the HK bi-probability plot.

5.2.4 Bi-probability Plot in HK

The bi-probability plot used in neutrino physics plots the probability of ν_e appearance against $\bar{\nu}_e$ appearance, with $P(\nu_\mu \rightarrow \nu_e) = P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ indicating no CP violation is present. This plot has already been shown in figure 1.6 and was discussed in some detail in section 1.2.4. In particular, it can be seen that the region with Asimov A22 + $\delta_{CP} = -\frac{\pi}{2}$ (the maximal CP violating region with normal ordering), there is no degeneracy. It was, however, expected that fits with $\delta_{CP} = 0$ would explore this region.

5.2.5 Initial HK Convergence Issues

In order to assess the capabilities of the MaCh3 HK fitter to produce posteriors that are consistent with other HK fitters (VALOR and Osc3++), comparisons were made using the 2018 T2K model as a baseline with systematic variations disabled. It was found that, whilst the frequentist HK fitters agreed relatively well, there were significant discrepancies between them and MaCh3. This is most evident in the $\sin^2(\theta_{23})$ plot where the frequentist fitters obtain a very different final result to MaCh3. Whilst differences in analysis choice (similar to the PTheta "MaCh3-like" binning in the previous chapter) and the use of marginalisation in MaCh3 vs likelihood profiling in the other fitters were discussed, due to the presence of the several known analysis bugs and the lack of maintainability for this version of MaCh3, it was decided that MaCh3 should move to using a newer version of the codebase.

5.3 Fits using the 2020 T2K Model

5.3.1 Initial Validation

As with the 2018 T2K model, initial validations were performed to ensure self-consistency within MaCh3, the fitter was scaled in steps and event rates were compared to a numerically scaled set of T2K event rates. Unlike the previous analysis, flux scaling

Sample	MaCh3	PTheta	Relative Difference
ν-mode 1Rμ	8850.40216	8826.46	0.27%
$\bar{\nu}$ 1Rμ	12019.2137	12016.4	0.02%
ν-mode 1Re	2486.6921	2482.81	0.16%
$\bar{\nu}$ 1Re	1546.5219	1545.58	0.06%
ν-mode 1Re1d.e	313.64378	311.424	0.71%

Table 5.3: Pre-BANFF event rate comparison for MaCh3 and PTheta. Event rates for each sample and the relative differences are shown. The agreement condition was set to be slightly looser (at most 1% difference) than that used for T2K analyses due to the code development required for HK analyses.

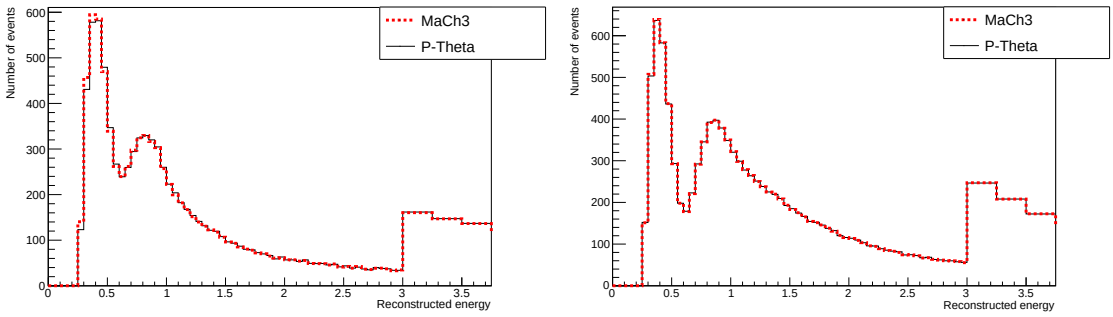

Figure 5.3: Predicted HK event rate spectra for the ν-mode (left) and $\bar{\nu}$-mode (right) 1Rμ samples at HK statistics compared between MaCh3 and PTheta. Small differences are seen in the ν-mode event rate spectra but these were deemed to be acceptable for the purposes of sensitivities and studies.

for T2K-OA2020 does not exist, and due to binning differences in between the 2018 and 2020 analyses it was deemed to be an unnecessary addition, particularly since the initial aim was to validate that MaCh3 could both converge and explore both hierarchies effectively. A summary of pre-BANFF event rate differences for each sample is shown in table 5.3 and shows good initial agreement between MaCh3 and PTheta. In addition the MaCh3 and PTheta MC event rate spectra were compared and found to be in very good agreement [Fig 5.3]

5.3.2 Fits with Assumed CP Conservation

The most interesting fits from this analysis at present are those with $\delta_{CP} = 0$. At first glance figure 5.5 seems to be fairly reasonable based on the bi-probability plot

[Fig. 1.6] with steps in both mass orderings. However, upon closer inspection of the trace plot 5.4 it can be seen that the fit has not truthfully explored both mass orderings simultaneously and instead each chain has become stuck in a single ordering. As the Asimov point and true value of δ_{CP} are 0 in the normal ordering, it is expected that the posterior would be maximal here, but clearly this has not happened. The most likely explanation for this is that it is the result of a poorly formed "ordering flip" within MaCh3. As can be seen in figure 1.5, Δm^2_{32} in one mass ordering can be calculated from the its value in the alternative ordering through

$$\Delta m^2_{32,\text{NO}} = -\Delta m^2_{32,\text{IO}} - \Delta m^2_{21} \tag{5.1}$$

In the T2K version of MaCh3, $-\Delta m^2_{21}$ is much smaller than the width of the Δm^2_{32} posterior and, as a result, is an unnecessary correction. This lets T2K-MaCh3 propose steps in the other hierarchy as described in algorithm 6. This allows for both sides of this very bimodal space to be explored; however, it makes the assumption that there are no other parameters that have strong correlations with the mass hierarchy. This is not true for HK fits: instead δ_{CP} will in general become a disjoint bimodal distribution. This will in general mean that a second δ_{CP} flip is required in tandem when the mass hierarchy changes. The following plots created by Andrés Lopez Moreno using the pyExotics library [142] demonstrate that this flip depends on the true value of δ_{CP} and thus implementation of this feature is non-trivial [Fig. 5.6].

Algorithm 6 T2K-MaCh3 Mass Ordering Proposal Function

Propose step using step proposal in algorithm 1

Generate random number $s \sim U(0, 1)$

if $s > 0.5$ **then**

    step(Δm^2_{32}) = -step(Δm^2_{32})

Continue MCMC procedure

The suggested implementation of this additional flipping is shown in algorithm 7,

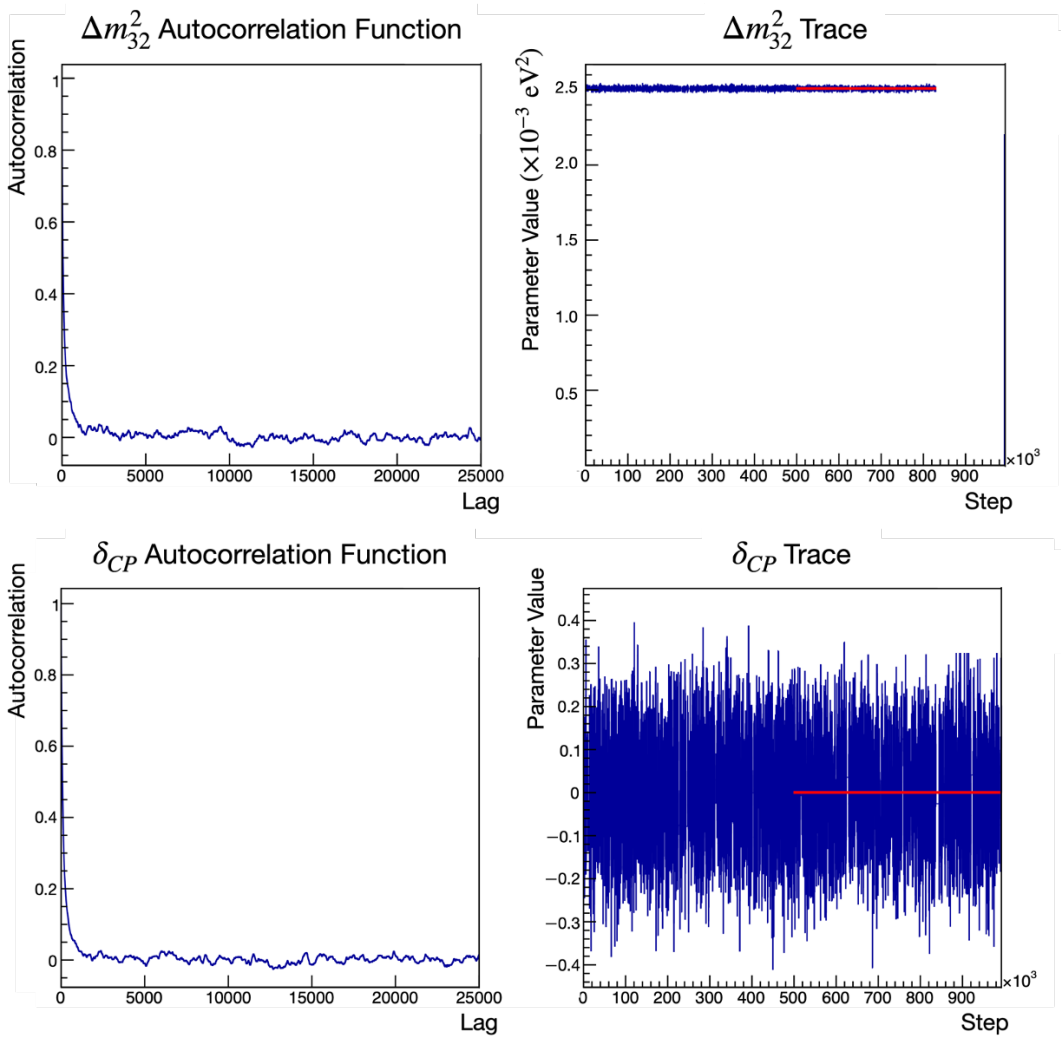

Figure 5.4: Diagnostics for an OA2020 HK fit with using the AsimovA20 parameter set with $\delta_{CP} = 0$ without an explicit δ_{CP} flip. The diagnostics are for Δm^2_{32} (top) and δ_{CP} (bottom) and show the chain has become 'stuck' with normal ordering. Combined result for multiple chains with this characteristic result in the posterior in figure 5.5.

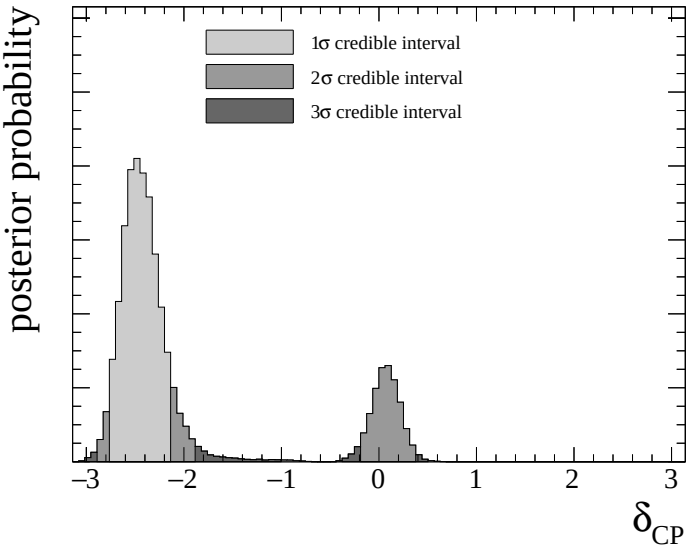

Figure 5.5: δ_{CP} posterior from an OA2020 fit with using the AsimovA20 parameter set with $\delta_{CP} = 0$ without an explicit δ_{CP} flip. Despite the Asimov point lying in the normal ordering, the best fit point is stuck in the inverted ordering. This is due to a high proportion of chains used to create it not entering the normal ordering.

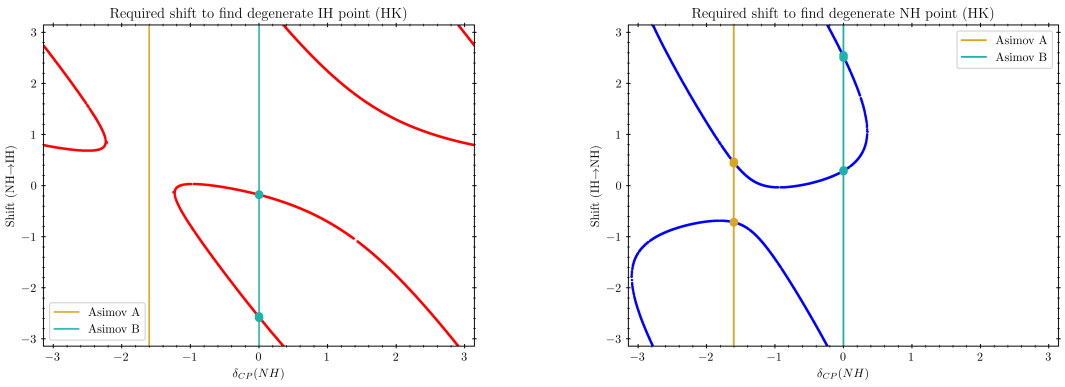

Figure 5.6: Plot showing the shift in the value of δ_{CP} required to find the mode in the opposite mass ordering dependent on the true value of δ_{CP} at HK, assuming 10 years of run time. For regions with no flip the degeneracy in mass ordering is lifted Plots produced by Andres Lopez Moreno using pyExotics [142].

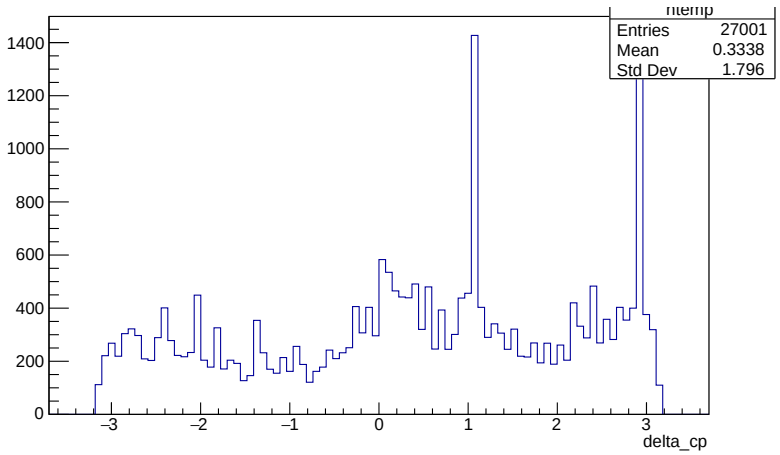

Figure 5.7: Early results from a very short JAMS chain with $\delta_{CP} = 0$ + the AsimovA20 parameter set. Whilst the chain shown is very short, it is not expected to explore the entire parameter space which can be seen here. Steps to the majority of points in this region are expected to be rejected at a far higher rate than is seen and not accepted at a roughly equal rate. This is likely the result of an incorrectly implemented jump step acceptance method.

Algorithm 7 HK-MaCh3 Mass Ordering Proposal Function

Propose step using step proposal in algorithm 1

Generate random number $s \sim U(0, 1)$

if $s > 0.5$ **then**

Perform mass flip in full

$\text{step}(\Delta m^2_{32}) = -\text{step}(\Delta m^2_{32}) - \Delta m^2_{21,\text{nominal}}$

$\text{step}(\delta_{CP}) = \text{step}(\delta_{CP})$ + flip from lookup table

Continue MCMC procedure

This case is fine if the true ordering and value of δ_{CP} are known, but, the implementation for fits where both of these are unknown, i.e. data fits, is less trivial. It was suggested that an alternative algorithm could be utilised, for example JAMS [Alg. 4], but implementation of this is still in the exploratory phase. The results of early fits using this algorithm clearly demonstrate that this requires substantial validation, as can be seen in figure 5.7.

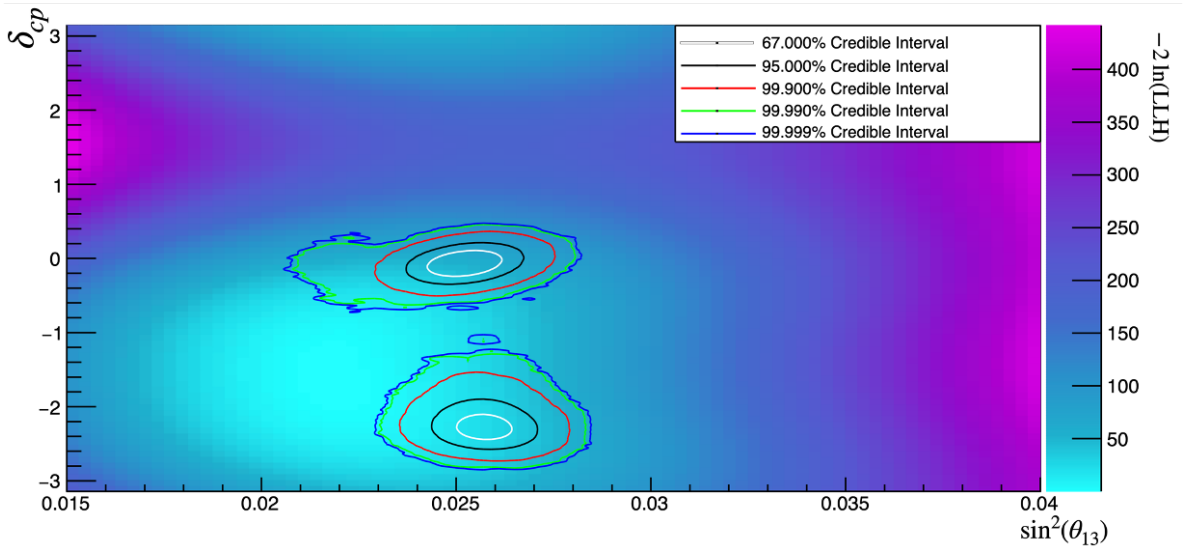

Figure 5.8: Figure overlaying the marginalised appearance parameter posterior from an OA2020 AsimovA20, $\delta_{CP} = 0$ chain onto a 2D log-likelihood scan of the two parameters. It can clearly be see than both parameters lie slightly outside the expected maximal log-likelihood region. Whilst this is expected in the case of δ_{CP} due to its ordering dependent bi-modality, $\sin^2(\theta_{13})$ has minimal ordering dependence and as a result, the fit would be expected to be centred closer to the Asimov point at $\sin^2(\theta_{13}) = 0.022$.

5.3.3 Alternative Hypotheses for Non-Convergence

Before examining a converged fit, it is important to assess other potential reasons for non-convergence, in part because the flipping behaviour has not been confirmed to be a true cause and because a substantial amount of effort has been expended attempting to rule out these causes. The first is simply poor step size tuning. In order to remove this issue statistics only fits were used, which meant that only the oscillation parameters required tuning. As can be seen in figure 5.4, these parameters are very well tuned and hence unlikely to be the cause of the convergence issues.

The second cause is that the chain is potentially stuck in a "deep" local minimum. By overlaying the 2D MCMC posterior on top of a likelihood scan for the non-marginalised parameters, it is clear that it is not converging to the region of maximum likelihood in the 2D space. This is indicative of a local minimum but investigating these issues with MCMC is somewhat tricky.

Finally, this issue could be caused by a previously unknown bug within the MaCh3 fitter. The HK software was based on a frozen version of the 2020 MaCh3 analysis toolkit which had been modified to include several quality of life features, for example the using a single handler object for all samples rather than each sample requiring separate objects, and speed improvements. This version was not rigorously validated and the next fully validated version changed many components of the analysis and fixed several known bugs in this version. It is unclear currently if a bug was causing these non-convergence issues and it was hoped that the HK analysis would move to a new refactored version of MaCh3 within the timeline of this thesis. Further evidence for this potential bug can be seen in figure 5.2 whereby the OA2018 and VALOR fits agree well whilst the 2020 fit seems to converge to a slightly different value of δ_{CP}.

5.3.4 Fits with Fixed Mass Hierarchy

The following fits were run using MaCh3 with fixed inverted/normal ordering. These are oscillation parameter only fits and have no systematic variation. As a result the constraints are somewhat tighter than would be expected from a full HK fit. This was done in order to minimise the amount of step size tuning required and to improve the convergence rate for the fit. Fits were run with $\delta_{CP} = 0$ and to demonstrate the fitter is capable of converging if there are no "jumps". The fit contains 60M steps, which results in stable contours up to 3σ. Running MaCh3 fits with stable contours up to 5σ is prohibitively slow and would take several years of continuous computing time.

The 1D posteriors (with no external constraints) are shown in figure 5.9 and demonstrate the capabilities of HK relative to T2K. For all parameters shown in figure 5.10, the HK contours are unsurprisingly tighter than those seen in chapter 4. This is most evident in the δ_{CP} contour which now excludes a substantially larger region of the available phase space to the 99% CI.

In addition, if the external constraint from reactor experiments (Gaussian prior with central value 0.022 and error 0.0007) is applied, HK gains preference for the upper

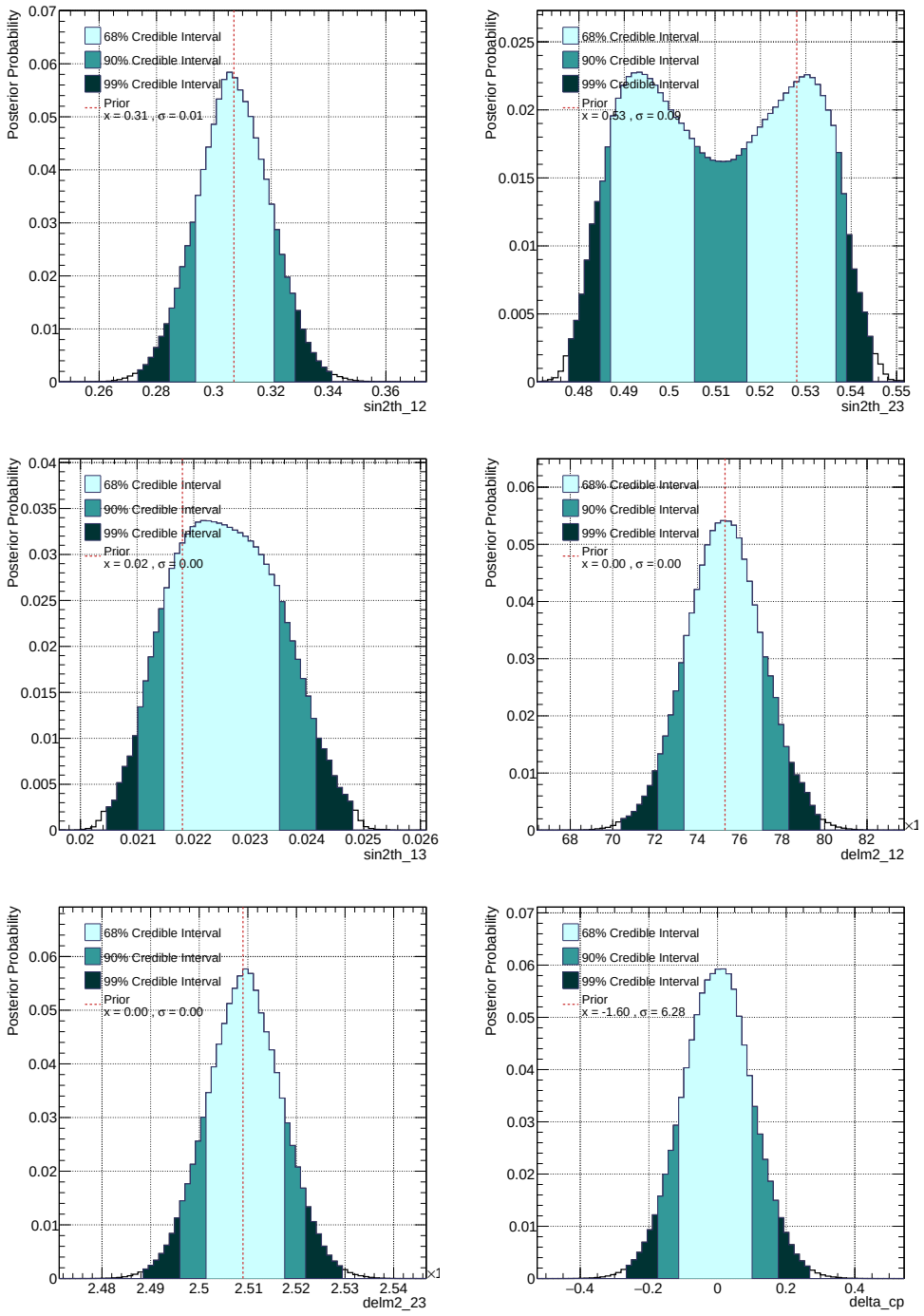

Figure 5.9: Posteriors for a fixed normal ordering MaCh3 HK chain including solar parameters with $\delta_{CP} = 0$. Credible intervals are stable up to the 99% credible interval. The central point for the prior used for each parameter is shown on each plot

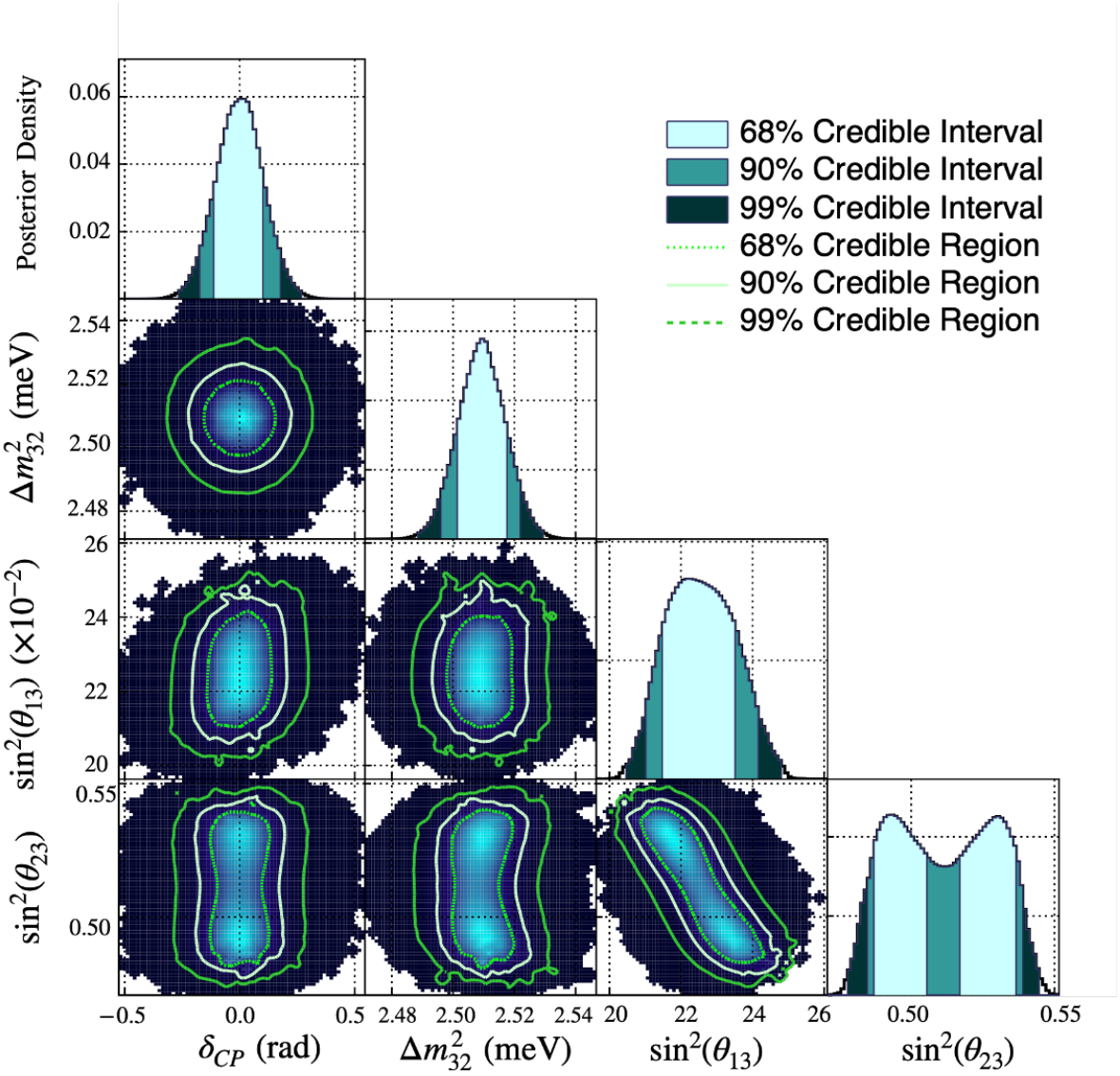

Figure 5.10: Triangle plot showing 2D contours for a MaCh3 HK chain in fixed normal ordering with $\delta_{CP} = 0$. The solar parameters, $\sin^2(\theta_{12})$ and Δm^2_{21}, are omitted from this plot for readability.

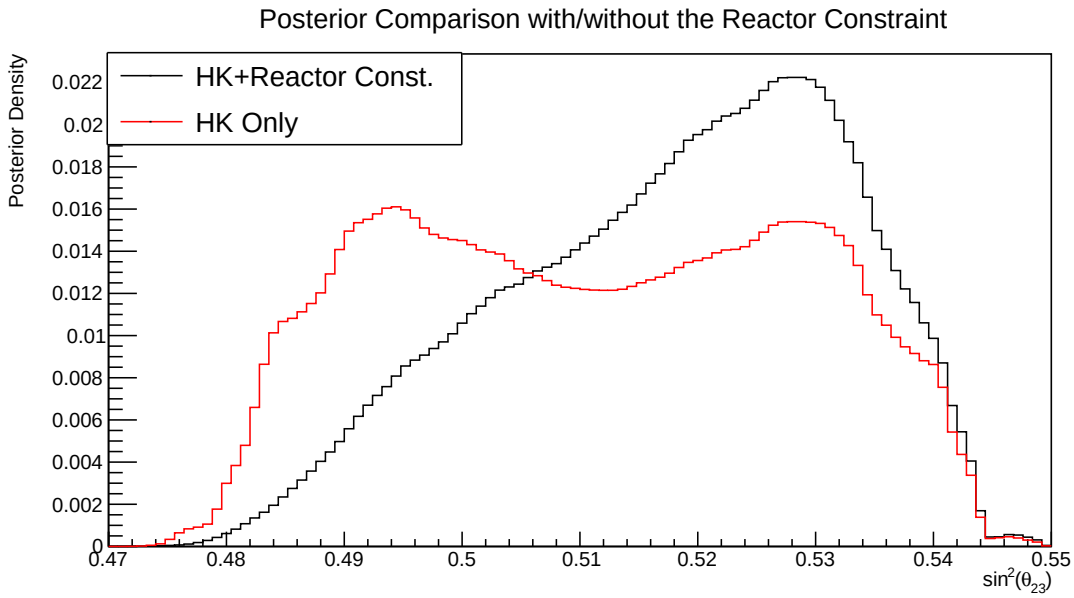

Figure 5.11: Comparison of $\sin^2(\theta_{23})$ between fits with and with the reactor constraint applied. The reactor constraint results in a shift in preference for the upper octant with the Bayes factor increasing from 2.10 to 6.15.

octant with the Bayes factor (UO/LO) increasing from 2.10 to 6.15 [Fig. 5.11]. On the Kaboth-Dunne scale [Tab. 3.1] this corresponds to less than 1σ but it is significant improvement with respect to the T2K result presented previously. In addition, this fit has been compared with PTheta and shows relatively good agreement [Fig. 5.12].

5.4 Future Analyses

Due to the time constraints from both the T2K analysis and timeline of this thesis and delays in the introduction of a refactored version of MaCh3 which is planned to be used across multiple experiments, this chapter is not a complete HK analysis. As the convergence bug was thought to be due to an issue within the 2020 MaCh3 package, all HK work was paused whilst waiting for the refactored version of MaCh3. This meant that many planned studies were put on hold including:

- further validations with the PTheta group;

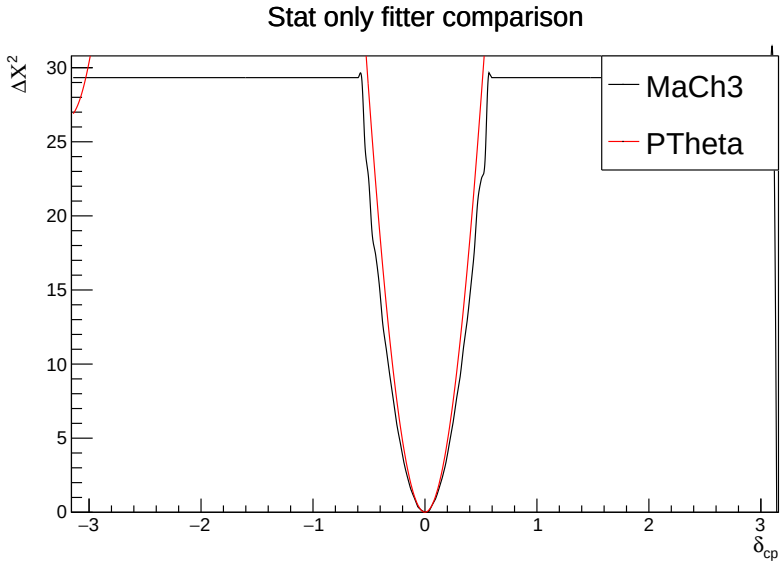

Figure 5.12: Comparison of PTheta and MaCh3 δ_{CP} distributions. Agreement is relatively good at low values of $\Delta\chi^2$ but becomes worse at larger values. This is simply a result of the lower number of steps in the MaCh3 chain at these values resulting in larger uncertainties

- HK fits run with increasing POT in order to see the impact of HK run lifetime on oscillation parameter sensitivities;
- proper introduction of the JAMS algorithm into the HK fitter to allow multi-modal adaption;
- analyses with a wider range of values of δ_{CP};
- studies into the impact of systematics on uncertainties.

Whilst many of these have been presented in an unfinished state, they are currently not complete enough to draw full conclusions from.

Chapter 6

MCMC Convergence

As detailed in section 3.6.2, Metropolis-Hastings requires step-size tuning in order to be utilised effectively. The adaptive mechanisms detailed in Alg. 2 and Alg. 3 can be highly effective but will often fail. This chapter will detail both the successes and potential pitfalls that can be encountered when using these methods.

6.1 Adaptive MCMC

6.1.1 Re-Introduction to Adaptive MCMC

MaCh3 currently performs MCMC throws using the prior covariance. For each parameter, i, the proposed new step y_i is given by

$$y_i \sim G\left(X_i^t,\ s_i \frac{2.64^2}{N} \sigma_{prior}\right) \tag{6.1}$$

where N is the number of parameters, and s is a vector of individual step sizes for each parameter that has been chosen manually. As was discussed in section 3.6.2, this is not necessarily the most efficient proposal function and it is often more efficient to use

$$y_i \sim G\left(X_i^t,\ s_i \frac{2.64^2}{N} \sigma_{\text{chain},\gamma}\right) \tag{6.2}$$

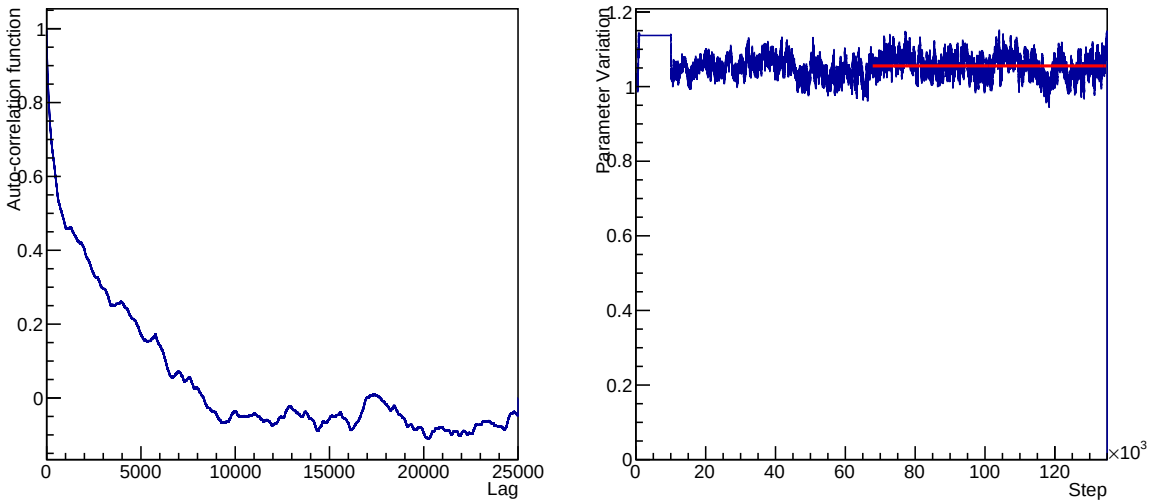

Figure 6.1: The trace (right) of a cross-section parameter when adaption was enabled for all parameters. Adaption is enabled from step 10,000 and the fit is totally un-tuned initially. Whilst the trace plot appears relatively well-behaved, showing a high degree of variance, this is slightly misleading. As can be seen from the autocorrelation (left) convergence is actually relatively slow with the autocorrelation only dropping to near 0 at relatively large values of lag. In addition, the expected range of this parameter is [0.8, 1.2] indicating that it has not converged on the true posterior.

where $\sigma_{\text{chain},\gamma}$ is the covariance of the chain so far and γ indicates the iteration number of the proposal function since it may only be updated every few steps.

6.2 Implementation in MaCh3

6.2.1 Implementation Procedure

The implementation procedure is relatively straight forward. Initial testing using adaption on the full parameter space [Fig. 6.1] indicated that substantial validation was required in order to ensure that adaptive fits converge to the correct posteriors. In addition, it was also found that a small subset of parameters are unsuitable for adaption. In order to test this, initial chains were run using a small subset of parameters and then gradually increasing the number of parameters used. The software implementation is relatively simple and roughly follows the steps in algorithm 3. The fit hyper-parameters

i.e. number of steps between adaptions, number of steps to adapt for, which parameters are adapting, etc. are configured via an external configuration file for ease of use. At low step counts the small numerical fluctuations, particularly from floating point calculations, may cause the matrix to become non-positive definite which is required for all 'true' covariance matrices. Two fixes can be used within MaCh3. The first simply adds a small correction to the diagonal repeatedly and checks if this has made the matrix positive definite. If this fails to work a more in-depth procedure is used. It can be shown that, for any non-positive definite matrix, C, the closest positive-definite symmetric matrix, C_P under the Frobenius norm ($\|C\|_F^2 = \sum_{(i,j)} C_{ij}^2$) can be computed through a relatively simple procedure, the full details of which can be found in [143].

6.2.2 Code Considerations

MaCh3 assigns each set of systematics in its own "covariance" object, all of which inherits some basic properties from a covariance base class. It was determined that the simplest method for implementing the adaptive scheme was to inject it into the covariance base class. This setup, controlled by a configuration file allows the user to specify the following

- Systematics that use adaption
- Whether an external matrix is to be used. This allows chains to use a "frozen" covariance matrix from a converged Markov Chain rather than needing to begin adaption from that start
- Number of steps before the adaptive covariance matrix is frozen
- Number of steps before throws are performed using the adaptive matrix
- Number of steps between proposal function updates
- "Burn-in" length. An initial set of steps where no steps are included in the covariance calculation. This allows for early numerical instability in the chain to

be ignored.

In addition, prior to the implementation of these methods the likelihood calculation for each systematic, whilst identical in all cases, was not included in the base class. By moving it to the base-class the adaptive step proposal could then be performed through:

1. Propose new step
2. Check parameter boundaries
3. Accept/reject new step
4. Update Covariance Matrix to include new information

The boundary check is a virtual method in the base-class and a concrete implementation in the derived systematics objects. This allowed for much easier code maintenance and hence made testing and bug-fixing substantially simpler.

6.3 Exploratory Fits

6.3.1 Analysis Strategy

In all tests listed below the following analysis procedure was applied:

1. Adaptive and non-adaptive fits were run using the same initial step sizes
2. The adaptive fit was run with slightly un-tuned step sizes
3. If this fit convergences correctly a "nightmare" fit was run with step sizes set to well beyond their manually tuned values

This allowed for the assessment of the convergence properties of the chain with decreasing amounts of manual input.

6.3.2 Fit with Limited Systematics

As a proof of concept, adaptive style fits were run with a reduced systematics set. The spectral function cross-section systematics [144] were selected as they have uncorrelated priors but are known to be correlated post-fit resulting in a potential improvement from the adaptive method. As can be seen in figure 6.5 adaption has the effect of tuning these parameters rapidly as the length of the chain increases. In practice, this means that the fit will converge far faster than a manually tuned fit. Further "nightmare" studies were performed on this parameter by allowing the initial pre-adaptive step sizes to be semi-randomized and thus highly inefficient. The effect of adaption on this space is very clear from the aforementioned trace plot which shows that the adaption results in a clear and immediate improvement in MCMC efficiency [Fig. 6.3]. In addition the acceptance rate rapidly trends towards the optimal 23.4% expected acceptance rate which is normally very tricky to obtain in MaCh3 using manual tuning [Fig. 6.2].

The addition of highly correlated systematics resulted in clear and immediate issues [Fig. 6.4]. The CC Coherent normalization parameters on carbon and oxygen have a highly correlated prior (0.9999).[1] It has been observed that, when step size tuning manually, highly correlated systematics require identical step sizes in order to converge effectively. Adaption often causes the step sizes to deviate for these parameters and hence results in poorly performing adaptive step size tuning. This effect can be minimised by simply increasing the number of parameters with uncorrelated priors. The exact reasons for this are not well understood but it seems that the additional contributions from non-correlated systematics have a dampening effect on these problem parameters.

The final test of this regime was the comparison of a limited systematics chain with a non-adaptive chain [Fig. 6.6].The result was a dramatic improvement in both trace and auto-correlation diagnostics after adaption was applied. In addition the effective sample size [Eqn. 3.24] (weighted by the number of steps in both fits) was compared

[1] The decision not to merge these parameters is primarily historic.

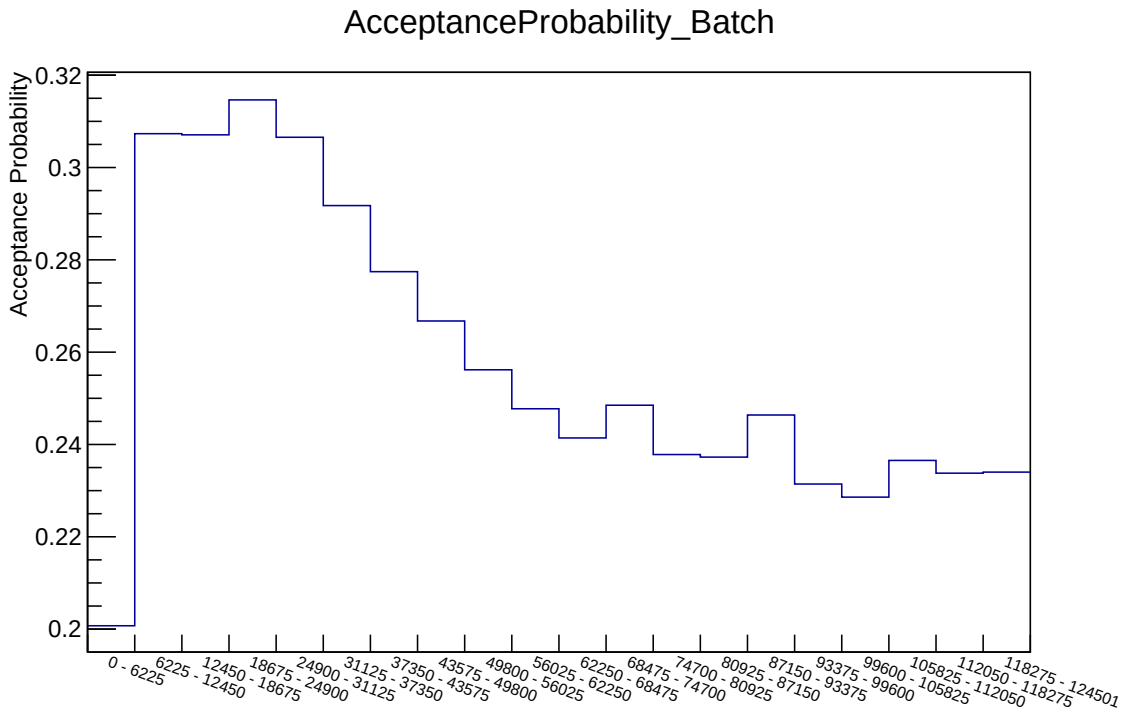

Figure 6.2: The acceptance probability for a Markov chain with an initial well tuned start. Adaption is enabled at step 10,000 which can be seen by a trend towards the optimal acceptance rate of 23.4%.

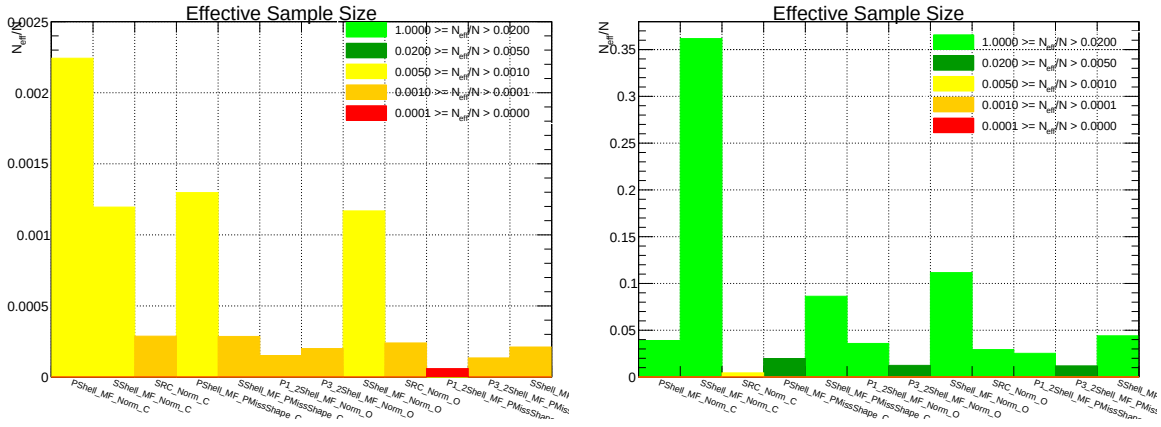

Figure 6.3: Effective sample size (ESS) for Markov chains without adaptive tuning (left) and with adaptive tuning (right) for a small selection of cross-section parameters. The adaptive tune results in an at least order of magnitude increase in effective sample size improvement in the adaptive chain.

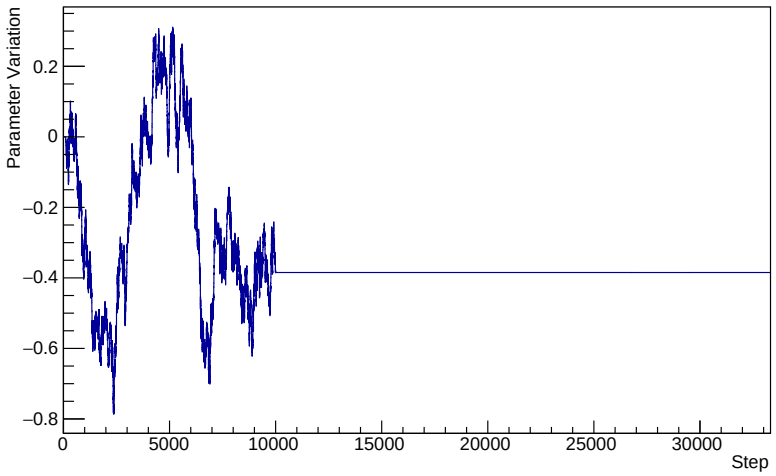

Figure 6.4: Trace from a cross-section parameter from a fit using the initial reduced parameter set + the correlation CC coherent normalisation systematics. Adaptive throws start from 10000 which is marked by a complete absence of accepted steps in the fit. This is thought to be due to large correlations between the CC Coherent normalisation systematics and has been observed with other parameters within the fit.

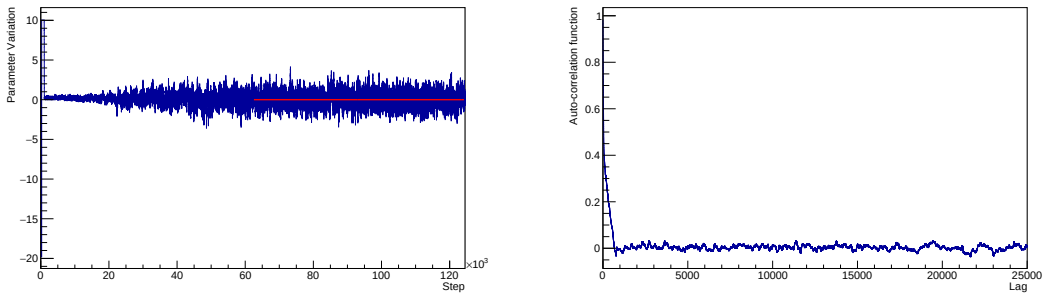

Figure 6.5: (Left) The trace of a parameter from a limited systematics fit with "nightmare" initial step sizes. The adaptive process can be seen from the increase in the range the chain explores. (Right) Autocorrelation for the same parameter. Notably, the autocorrelation approaches 0 very quickly indicating that the chain is highly Markovian.

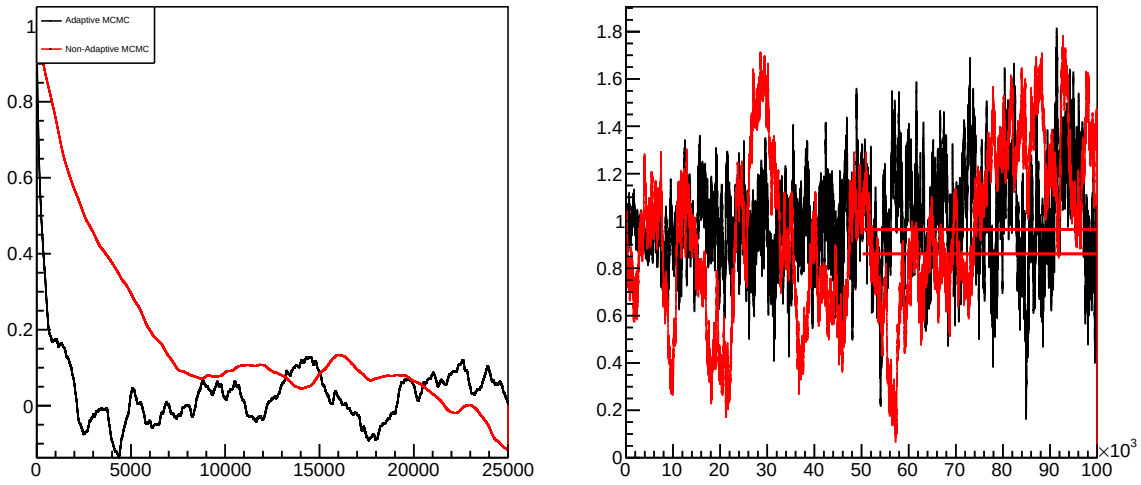

Figure 6.6: Comparison of the traces for a short adaptive chain (black) vs a non-adaptive chain (red) which has been manually step-size tuned. It is clear that the adaptive chain reaches an autocorrelation of 0 far faster than the non-adaptive chain and has much more rapid variance. The difference in parameter ranges for low step numbers is a result of the early exploration of the space required by adaptive MCMC before it convergences on the "true" covariance matrix. The horizontal red lines are an artifact of the plotting software showing best fit for each chain, unfortunately the chains required for these plots have long since vanished at the time of writing and so removal of these lines is no longer possible.

and, as can be seen in figure 6.7, the adaptive chain provides a larger effective sample size for most parameters. This is an overall encouraging result and shows that adaption is at least as effective as manual tuning for MaCh3 cross-section parameters even with these correlated parameters limiting its effectiveness.

6.3.3 PCA

It was thought that prior-based principal component analysis (PCA)[2] would be an effective method for minimising correlations and thus allow the adaptive algorithm to converge faster. However, the current implementation of PCA in MaCh3 does not normalise the input covariance matrix and is thus dominated by a small subset of parameters; this results in parameters that are thrown in PCA space being moved

[2]performing PCA on the input covariance matrix

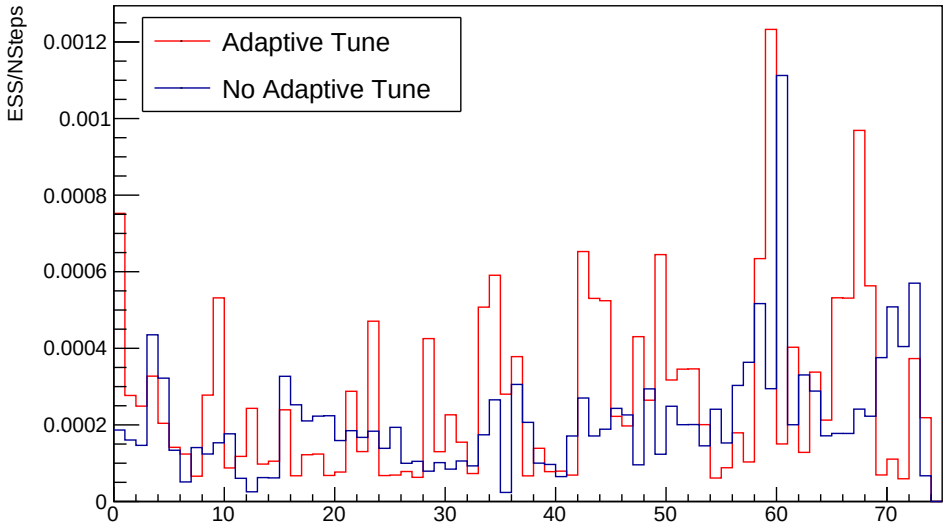

Figure 6.7: Effective sample size (ESS) for each parameter in the cross-section model. The x-axis refers to the internal index with MaCh3 assigned to each cross-section parameter. As can be seen in the figure there is a general improvement for most cross-section parameters after adaption.

towards their physical boundaries and hence results in a poor acceptance rate [Fig. 6.8] since steps with parameters outside of their allowed range are automatically rejected. Due to time constraints, a reimplementation of the PCA algorithm to normalise these parameters was not feasible.

6.3.4 Block Matrix Approach

As discussed in [117] adaption often takes longer to converge for larger matrices with multiple sets of relatively uncorrelated parameters. This can be seen clearly when fits are run using the full cross-section and flux model. As discussed in section 4.3, MaCh3 combines the cross-section and flux uncertainties into a single matrix. Adaptive fits run using correlations between both sets of parameters tend to result in low acceptance rates and can converge to local minima rather than the true posterior. The simple solution to this is to throw from a block matrix instead, where each block represents a particular parameter set, for example flux and cross-section.

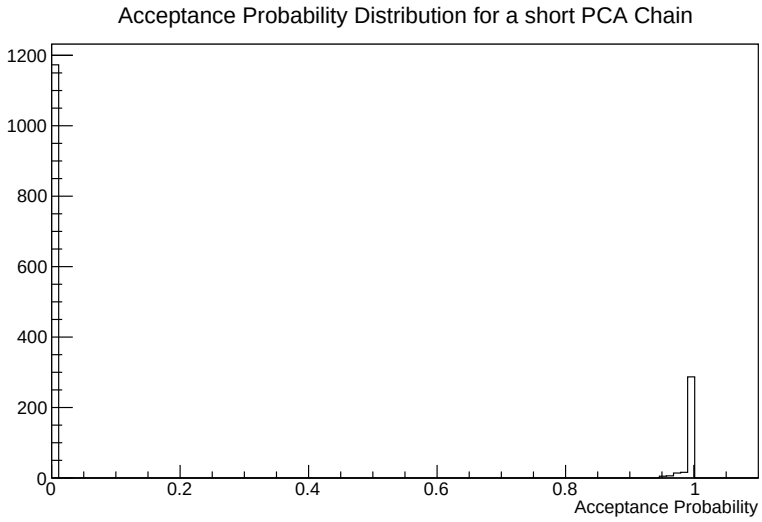

Figure 6.8: Distribution of acceptance probabilities for a short adaptive chain with PCA applied. The high proportion of steps with 0 acceptance probability indicates the chain is proposing a large number of steps to a highly improbable region of space.

6.3.5 ND Only fits

The natural next step when introducing adaption to MaCh3 was to see how effective it is when applied to an near detector (ND) only fit. Recall from section 4.1 that MaCh3 performs two kinds of T2K-only fit, ND and ND+far detector (FD). This first fit only requires the near detector and cross-section systematics models. Due to their relatively Gaussian nature, ND systematics parameters are a natural candidate for adaption. In order to produce these fits efficiently, ND data was taken from runs 2a and 9b. This choice was somewhat arbitrary and simply used the "reduced ND280 dataset" recommended in the MaCh3 configuration file. The cross-section/flux matrix used the block-matrix approach whilst the ND detector matrix was left as a single matrix. Initial tests were done using the block matrix method on the ND detector matrix with each block corresponding to a different oscillation channel, but, due to significant post-fit correlations between oscillation channels this was found to be ineffective.

Fits with ND280 parameters were run in several configurations:

- Non-adaptive;

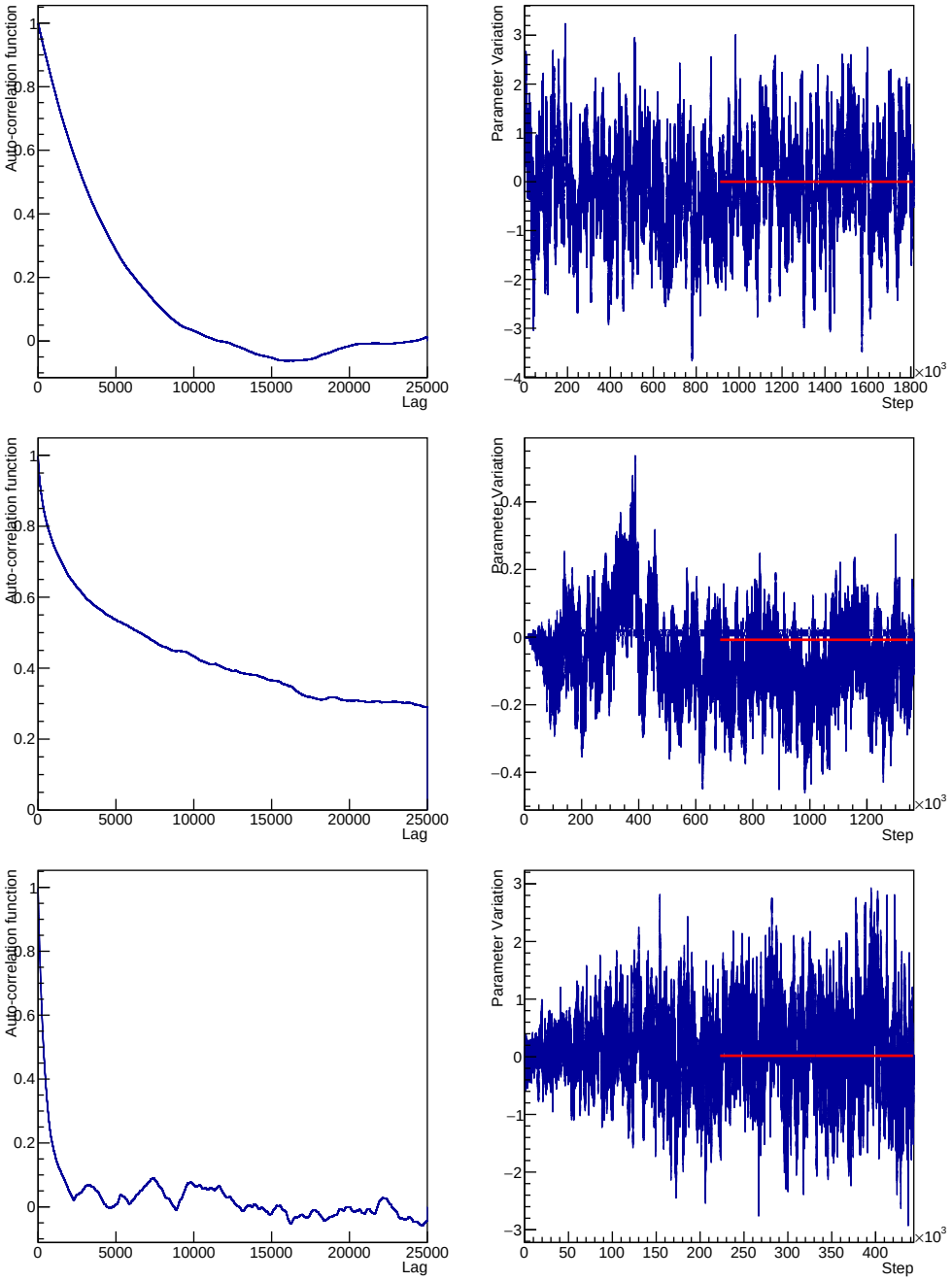

Figure 6.9: Diagnostics for a variety of adaptive MCMC chains from an ND only fit. From top to bottom these are non-adaptive, adaption from the star and, adaption starting at step 10000 with a well tuned start

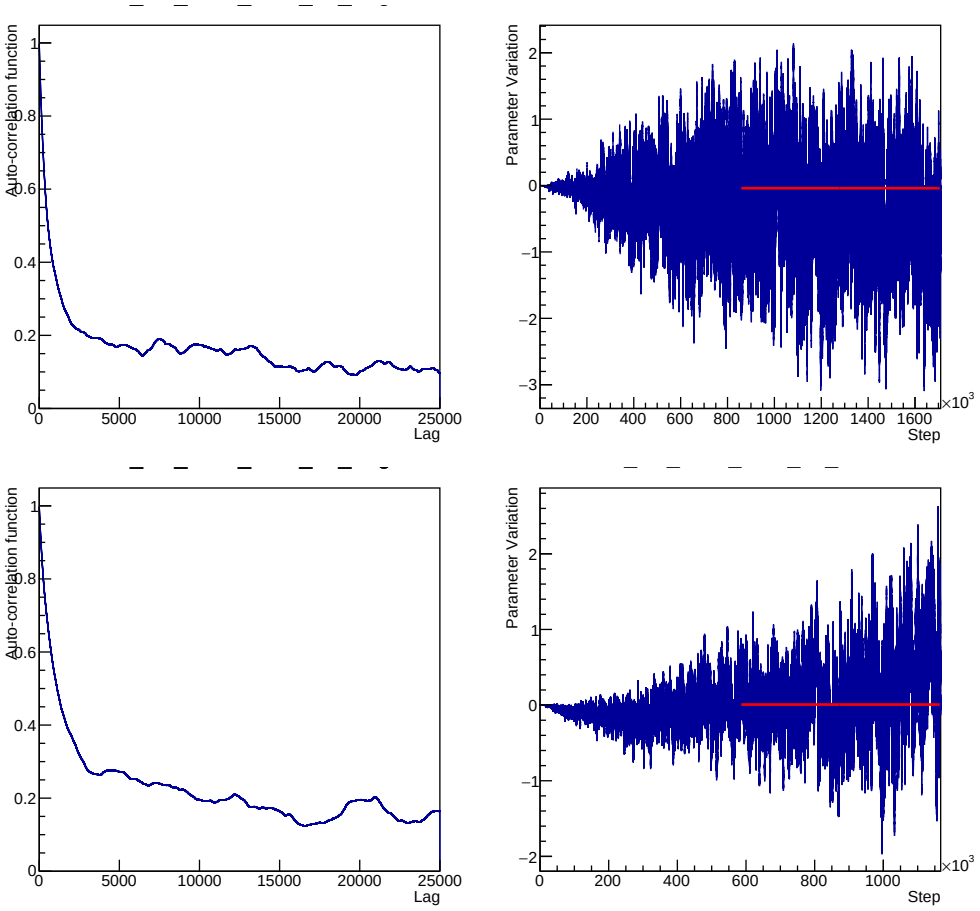

Figure 6.10: Diagnostics for a variety of adaptive MCMC chains from an ND only fit. Both chains have "nightmare" starts with adaption staring after step 1000 (top) and step 10000 (bottom).

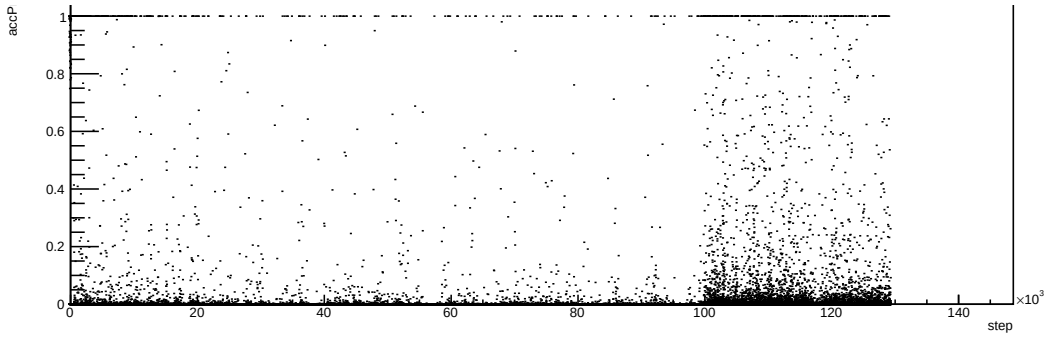

Figure 6.11: The acceptance probability (ratio of likelihood of proposed:current step) for each step in a Markov Chain with adaption enabled for all parameters is shown. Adaptive throws are enabled at step 1000 and is marked by a stark rise in states with acceptance probabilities outside of {≈0,1}.

- Adaptive with adaption starting immediately;
- Well tuned fit with adaptive starting from step 10000;
- Nightmare start Adaptive with adaption starting from step 1000;
- Nightmare start Adaptive with adaption starting from step 10000.

The MCMC diagnostics for the multi-pion axial Bodak-Yang correction parameter for each of these fits are shown in figures 6.9 and 6.10.[3] These fits are illustrative of some general trends seen when developing this technique. The first is that adaption from step 0 consistently seems to drive fits away from the "true" nominal. Since these fits are Asimov the expected value for each parameter is known, and in the case of this dial it should converge to a posterior centred at 0. The other adaptive chains seem to perform comparatively well, with the initially well tuned chain having a slightly better auto-correlation than the nightmare parameters. This effect may be amplified by the early low acceptance region in the nightmare fits combined with the lower number of steps in the non-nightmare chain and future studies should account for this when generating auto-correlations.

[3]A cross section parameter was chosen over an ND parameter since the ND fit has been really well tuned manually and thus the impact of adaption isn't as visible.

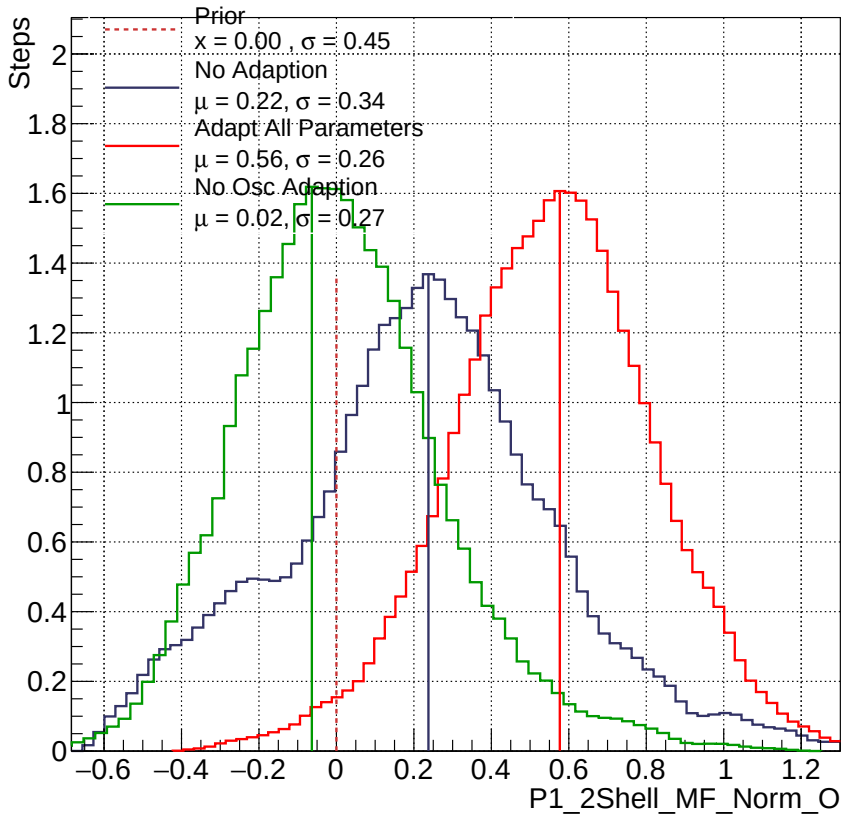

Figure 6.12: The expected posterior for an oscillation parameter in the case that all parameters including oscillation are allowed to adapt, when adaption is only enabled for systematics (cross-section, ND, FD) or when no adaption is enabled. The result is that the fits seem to converge to different distributions. Not only does remove the "nice" convergence seen in the ND-only fit, it shows that FD+Oscillation parameters may require different tuning when adaptive is applied.

6.3.6 Problematic Parameters

The final set of parameters that need to be added are those related to oscillation and far detector systematics. As can be seen in figure 6.12, running a full FD+ND fit with adaption switched on for all parameter sets including/excluding oscillation results in convergence to final posteriors that differ from the non-adaptive chain and from each other. This appears to be due to two features of the oscillation posterior:

- the multi-modality of Δm^2_{32};

- the cyclical boundaries for δ_{CP}.

The first is a known issue with MCMC convergence since the posterior at each mode of Δm^2_{32} will, in general, have a different geometry. As a result, both modes will have a different covariance matrix and hence the current method of finding an adaptive matrix is untenable since it simply obtains the average covariance across both modes (weighted by the likelihood of accepting steps between each mode). Several algorithms exist that claim to combat this, from the relatively mundane locally adaptive MCMC (simply have a covariance matrix for each mode), to the far more in-depth like JAMS (Jumping Adaptive Metropolis Sampler) which modifies Metropolis-Hastings to use multiple proposal functions. More positively, the effectiveness of this method opposed to poor tuning, even on this problematic space, can be seen in figure 6.11. The transition between adaption and non-tuned steps can be seen at step 100,000 with a stark rise in steps that aren't being immediately rejected.

6.4 Summary and Future Prospects

Adaptive MCMC is an extremely useful tool that can outperform manual step size tuning in many instances. As seen with the oscillation and FD parameters, it is not a 'silver bullet' that will fix all MCMC convergence issues, but, it can remove a lot of the manual effort previously required to tune the fitter. In addition the efficiency gains from these methods due to the relatively lower autocorrelations reduce the required number of steps to produce a fit. This is vital for future experiments (HK, DUNE) hoping to reach 5σ credible intervals since the required number of steps to do this is several orders of magnitude larger than the current number used at T2K.

Chapter 7

Summary

Markov Chain Monte Carlo provides a powerful and uniquely precise way to analyse complex statistical likelihoods. The analyses and techniques presented in this thesis show the current and future analyses that use this technique to provide new and exciting results in the field. The primary analysis represents the latest result from the T2K experiment. The key improvements to this analysis are the increase in ν-mode POT from 19.664×10^{20} to 21.428×10^{20} and new SK detector covariance matrix.

When fitting with the reactor constraint, CP-conserving values of δ_{CP} are excluded at the 90% credible interval. Mild preference for the upper octant and normal hierarchy is shown, although the Bayes factors for these preferences do not reach the level where this is a significant result (see Table 4.9). When fitting with a flat prior in δ_{CP}, we exclude a Jarlskog invariant of 0 (implying CP-conservation) at the 2σ credible interval, however, this statement does not hold when using a prior flat in $\sin\delta_{CP}$, so the statement on CP-conservation exclusion from the Jarlskog invariant is also at the 90% level.

The Hyper-Kamiokande sensitivity studies demonstrate the limits of MCMC techniques currently being used and provide information on the limitations of MCMC as a technique. The very limited studies performed show that high statistics fits are possible with heavy restrictions applied to the analysis. Significant work is required to run full sensitivity studies within MaCh3, but this early work has found many of the potential

problems.

Finally new MCMC techniques have been trialled which aim to fix many of the convergence issues seen in current and future oscillation analyses. These techniques are now useable in some limited analyses and potential issues with regards to implementation in MaCh3 are now beginning to be understood.

Bibliography

[1] P.A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[2] S. Rahaman et al. "Q values of the 76Ge and 100Mo double-beta decays". In: *Physics Letters B* 662.2 (2008), pp. 111–116. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2008.02.047. URL: https://www.sciencedirect.com/science/article/pii/S0370269308002554.

[3] W. Pauli. "Dear radioactive ladies and gentlemen". In: *Phys. Today* 31.9 (1978), p. 27.

[4] C. Jensen. *Controversy and consensus: nuclear beta decay 1911–1934*. Vol. 24. Springer, 2012.

[5] G. Rajasekaran. "Fermi and the theory of weak interactions". In: *Resonance* 19 (2014), pp. 18–44.

[6] C. Cowan Jr et al. "Detection of the free neutrino: a confirmation". In: *Science* 124.3212 (1956), pp. 103–104.

[7] E.C. Anderson. "The Reines-Cowan Experiments". In: ().

[8] F. Reines et al. "Detection of the Free Antineutrino". In: *Phys. Rev.* 117 (1 1960), pp. 159–173.

[9] G. Danby et al. "Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos". In: *Physical Review Letters* 9.1 (1962), p. 36.

[10] F Reines and C. Cowan. "Free Antineutrino Absorption Cross Section. I. Measurement of the Free Antineutrino Absorption Cross Section by Protons". In: *Phys. Rev.* 113 (1 1959), pp. 273–279.

[11] R. Davis. "A review of the Homestake solar neutrino experiment". In: *Progress in Particle and Nuclear Physics* 32 (1994), p. 13.

[12] S. Bilenky and B. Pontecorvo. "Lepton mixing and neutrino oscillations". In: *Physics Reports* 41.4 (1978), pp. 225–261.

[13] Z. Maki, M. Nakagawa, and S. Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Progress of Theoretical Physics* 28.5 (1962), pp. 870–880. ISSN: 0033-068X.

[14] M. Zralek. "From kaons to neutrinos: Quantum mechanics of particle oscillations". In: *arXiv preprint hep-ph/9810543* (1998).

[15] Martin L Perl. *The Discovery of the tau lepton*. Tech. rep. Stanford Linear Accelerator Center, 1992.

[16] K. Kodama et al. "Observation of tau neutrino interactions". In: *Physics Letters B* 504.3 (2001), pp. 218–224. ISSN: 0370-2693.

[17] S.M. Bilenky and S.T. Petcov. "Massive neutrinos and neutrino oscillations". In: *Reviews of Modern Physics* 59.3 (1987), p. 671.

[18] K. Zuber. *Neutrino physics*. CRC press, 2020.

[19] G.C. Branco and M.N. Rebelo. "Building the full Pontecorvo-Maki-Nakagawa-Sakata matrix from six independent Majorana-type phases". In: *Physical Review D* 79.1 (2009). ISSN: 1550-2368.

[20] J. Ellis. "Antimatter matters". In: *Nature* 424.6949 (2003), pp. 631–634.

[21] M. Maltoni and A.Y. Smirnov. "Solar neutrinos and neutrino physics". In: *The European Physical Journal A* 52.4 (2016), p. 87.

[22] K. Abe et al. "Solar neutrino measurements in Super-Kamiokande-IV". In: *Physical Review D* 94.5 (Sept. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.94.052010. URL: http://dx.doi.org/10.1103/PhysRevD.94.052010.

[23] G. Fogli and E. Lisi. "Evidence for the MSW effect". In: *New Journal of Physics* 6.1 (2004), p. 139.

[24] X. Qian and P. Vogel. "Neutrino mass hierarchy". In: *Progress in Particle and Nuclear Physics* 83 (2015), pp. 1–30. ISSN: 0146-6410.

[25] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 1998), pp. 1562–1567.

[26] Y. Suzuki. "The Super-Kamiokande Wxperiment". In: *The European Physical Journal C* 79 (2019), pp. 1–18.

[27] A. Bellerive et al. "The Sudbury Neutrino Observatory". In: *Nuclear Physics B* 908 (2016), pp. 30–51.

[28] E. C. G. Sudarshan and R. E. Marshak. "Chirality Invariance and the Universal Fermi Interaction". In: *Phys. Rev.* 109 (5 1958), pp. 1860–1862.

[29] M. Peskin. *An introduction to quantum field theory*. CRC press, 2018.

[30] L. Landau. "On the conservation laws for weak interactions". In: *Nuclear Physics* 3.1 (1957), pp. 127–131. ISSN: 0029-5582.

[31] P.F. Pérez and M.B. Wise. "On the origin of neutrino masses". In: *Phys. Rev. D* 80 (5 2009).

[32] P.W. Higgs. "Spontaneous symmetry breakdown without massless bosons". In: *Phys. Rev.* 145 (4 1966), pp. 1156–1163.

[33] A. Davidson and K.C. Wali. "Universal seesaw mechanism?" In: *Physical Review Letters* 59.4 (1987), p. 393.

[34] L.J. Wen, J. Cao, and Y.F. Wang. "Reactor neutrino experiments: present and future". In: *Annual Review of Nuclear and Particle Science* 67.1 (2017), pp. 183–211. ISSN: 1545-4134.

[35] F.P. An et al. "Observation of electron-antineutrino disappearance at Daya Bay". In: *Physical Review Letters* 108.17 (2012), p. 171803.

[36] A. Giaz. "Status and perspectives of the JUNO experiment". In: *arXiv preprint arXiv:1804.03575* (2018).

[37] K. Akita and M. Yamaguchi. "A precision calculation of relic neutrino decoupling". In: *Journal of Cosmology and Astroparticle Physics* 2020.08 (2020), p. 012.

[38] Y. Lyu. *Characterization of the PeV astrophysical neutrino energy spectrum with IceCube using down-going tracks*. 2021. eprint: 2107.14298.

[39] F Halzen and D. Hooper. "IceCube-Plus: an ultra-high-energy neutrino telescope". In: *Journal of Cosmology and Astroparticle Physics* 2004.01 (2004), p. 002.

[40] P. Antonioli et al. "SNEWS: The SuperNova early warning system". In: *New Journal of Physics* 6.1 (2004), p. 114.

[41] A. Burrows. "Neutrinos from supernova explosions". In: *Annual Review of Nuclear and Particle Science* 40.1 (1990), pp. 181–212.

[42] K. Nishikawa and T. Kajita. "Atmospheric Neutrinos". In: *Advances in High Energy Physics* 2012 (2012), p. 504715.

[43] K Abe et al. "The T2K experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 659.1 (2011), pp. 106–135.

[44] A.A. Aguilar-Arevalo et al. "The MiniBooNE detector". In: *Nuclear instruments and methods in physics research section a: accelerators, spectrometers, detectors and associated equipment* 599.1 (2009), pp. 28–46.

[45] C. Athanassopoulos et al. "Evidence for $\nu_\mu \to \nu_e$ oscillations from the LSND experiment at the Los Alamos Meson Physics Facility". In: *Physical Review Letters* 77.15 (1996), p. 3082.

[46] B. Abi et al. "Long-baseline neutrino oscillation physics potential of the DUNE experiment". In: *The European Physical Journal C* 80.10 (2020), pp. 1–34.

[47] K. Abe et al. "Hyper-Kamiokande design report". In: *arXiv preprint arXiv:1805.04163* (2018).

[48] JUNO collaboration et al. "JUNO physics and detector". In: *Progress in Particle and Nuclear Physics* 123 (2022), p. 103927.

[49] M. Aker et al. “The design, construction, and commissioning of the KATRIN experiment”. In: *Journal of Instrumentation* 16.08 (2021), T08015.

[50] N. Abgrall et al. “The large enriched germanium experiment for neutrinoless double beta decay (LEGEND)”. In: *AIP Conference Proceedings*. Vol. 1894. 1. AIP Publishing. 2017.

[51] A. Li, theKamLAND-Zen Collaboration, et al. “A Bayesian approach to neutrinoless double beta decay analysis in KamLAND-Zen”. In: *Journal of Physics: Conference Series*. Vol. 1468. 1. IOP Publishing. 2020, p. 012201.

[52] M.J. Dolinski, A.W.P. Poon, and W. Rodejohann. “Neutrinoless double-beta decay: status and prospects”. In: *Annual Review of Nuclear and Particle Science* 69 (2019), pp. 219–251.

[53] C.E. Patrick. “The MINERvA Experiment”. In: *Measurement of the Antineutrino Double-Differential Charged-Current Quasi-Elastic Scattering Cross Section at MINERvA* (2018), pp. 47–90.

[54] T. Odagawa et al. “Prospect and status of the physics run of the NINJA experiment”. In: *Proceedings of Science (NuFact2019)* 144 (2020).

[55] R.L. Workman et al. “Review of Particle Physics”. In: *PTEP* 2022 (2022), p. 083C01.

[56] S. Nagamiya. “Introduction to J-PARC”. In: *Progress of Theoretical and Experimental Physics* 2012.1 (2012), 02B001.

[57] T. Nakadaira et al. “T2K target”. In: *AIP Conf. Proc.* Vol. 981. 1. Citeseer. 2008, pp. 290–292.

[58] N. Abgrall et al. “Pion emission from the T2K replica target: method, results and application”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 701 (2013), pp. 99–114.

[59] T. Sekiguchi et al. “Development and operational experience of magnetic horn system for T2K experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 789 (2015), pp. 57–80.

[60] Y. Yamadaa. “The T2K program”. In: (2008).

[61] F. Retière, T2K collaboration, et al. “The neutrino near detector complex of the T2K experiment”. In: *Physics Procedia* 37 (2012), pp. 1231–1240.

[62] The T2K Collaboration. “Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations”. In: *Nature* 580.7803 (2020), pp. 339–344.

[63] S. Igarashi et al. “High-power beam operation at J-PARC”. In: *Proc. HB* 18 (2018), pp. 147–152.

[64] M. Ikegami. "Beam commissioning and operation of the J-PARC linac". In: *Progress of Theoretical and Experimental Physics* 2012.1 (2012), 02B002.

[65] K. Yamamoto, M. Kinsho, et al. "Design and actual performance of J-PARC 3 GeV rapid cycling synchrotron for high-intensity operation". In: *Journal of Nuclear Science and Technology* 59.9 (2022), pp. 1174–1205.

[66] T. Koseki et al. "Beam commissioning and operation of the J-PARC main ring synchrotron". In: *Progress of Theoretical and Experimental Physics* 2012.1 (2012), 02B004.

[67] A.K. Ichikawa. "Design concept of the magnetic horn system for the T2K neutrino beam". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 690 (2012), pp. 27–33.

[68] K. Abe et al. "T2K neutrino flux prediction". In: *Phys. Rev. D* 87 (1 2013), p. 012001. DOI: 10.1103/PhysRevD.87.012001. URL: https://link.aps.org/doi/10.1103/PhysRevD.87.012001.

[69] P. Podlaski. "NA61/SHINE overview". In: *arXiv preprint arXiv:2402.10973* (2024).

[70] L. Berns et al. "Flux Prediction and Uncertainty with NA61/SHINE 2010 Replica Target Measurements". In: *T2K Technical Report* 401 (2022).

[71] K.T. McDonald. *An Off-Axis Neutrino Beam*. 2001. arXiv: hep-ex/0111033 [hep-ex].

[72] Tomislav Vladisavljevic. *Predicting the T2K neutrino flux and measuring oscillation parameters*. Springer Nature, 2020.

[73] K. Matsuoka et al. "Design and performance of the muon monitor for the T2K neutrino oscillation experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 624.3 (Dec. 2010), pp. 591–600. ISSN: 0168-9002. DOI: 10.1016/j.nima.2010.09.074. URL: http://dx.doi.org/10.1016/j.nima.2010.09.074.

[74] T. Ovsiannikova et al. "The new experiment WAGASCI for water to hydrocarbon neutrino cross section measurement using the J-PARC beam". In: *Journal of Physics: Conference Series* 675.1 (2016), p. 012030. DOI: 10.1088/1742-6596/675/1/012030. URL: https://dx.doi.org/10.1088/1742-6596/675/1/012030.

[75] M. Antonova et al. *Baby MIND: A magnetised spectrometer for the WAGASCI experiment*. 2017. arXiv: 1704.08079 [physics.ins-det].

[76] K. Abe et al. "Measurement of the muon neutrino inclusive charged-current cross section in the energy range of 1–3 GeV with the T2K INGRID detector". In: *Phys. Rev. D* 93 (7 2016), p. 072002.

[77] M. Otani et al. "Design and construction of INGRID neutrino beam monitor for T2K neutrino experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 623.1 (2010). 1st International Conference on Technology and Instrumentation in Particle Physics, pp. 368–370. ISSN: 0168-9002.

[78] K. Abe et al. "Measurements of the T2K neutrino beam properties using the INGRID on-axis near detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 694 (2012), pp. 211–223. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2012.03.023.

[79] T. Kikawa, T2K collaboration, et al. "Development of the new T2K on-axis neutrino detector, INGRID proton module". In: *Nuclear Physics B-Proceedings Supplements* 229 (2012), p. 451.

[80] K. Abe et al. *T2K ND280 Upgrade – Technical Design Report*. 2020. arXiv: 1901.03750 [physics.ins-det].

[81] D. Allan et al. "The electromagnetic calorimeter for the T2K near detector ND280". In: *Journal of Instrumentation* 8.10 (2013), P10019.

[82] P-A. Amaudruz et al. "The T2K fine-grained detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 696 (2012), pp. 1–31.

[83] N. Abgrall et al. "Time projection chambers for the T2K near detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 637.1 (2011), pp. 25–46.

[84] Y. Giomataris. "Development and prospects of the new gaseous detector "Micromegas"". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 419.2-3 (1998), pp. 239–250.

[85] S. Aoki et al. "The T2K side muon range detector (SMRD)". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 698 (2013), pp. 135–146.

[86] S. Assylbekov et al. "The T2K ND280 off-axis pi–zero detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 686 (2012), pp. 48–63.

[87] A. Dergacheva et al. "3D SuperFGD detector for the T2K experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1041 (2022), p. 167219.

[88] D. Attié et al. "Analysis of test beam data taken with a prototype of TPC with resistive Micromegas for the T2K Near Detector upgrade". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1052 (2023), p. 168248.

[89] D. Attié et al. "Characterization of resistive Micromegas detectors for the upgrade of the T2K Near Detector Time Projection Chambers". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1025 (Feb. 2022), p. 166109. ISSN: 0168-9002. DOI: 10.1016/j.nima.2021.166109. URL: http://dx.doi.org/10.1016/j.nima.2021.166109.

[90] A. Korzenev et al. "A 4π time-of-flight detector for the ND280/T2K upgrade". In: *Journal of Instrumentation* 17.01 (2022), P01016.

[91] DuPont. *What is Tyvek?* URL: https://www.dupont.com/what-is-tyvek.html.

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

[93] M. Shiozawa. "Reconstruction algorithms in the Super-Kamiokande large water Cherenkov detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 433.1-2 (1999), pp. 240–246.

[94] B.M. Bolotovskii. "Vavilov–Cherenkov radiation: its discovery and application". In: *Physics-Uspekhi* 52.11 (2009), p. 1099.

[95] K. Abe et al. "First gadolinium loading to Super-Kamiokande". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1027 (2022), p. 166248. ISSN: 0168-9002.

[96] F. Di Lodovico, Hyper-Kamiokande Collaboration, et al. "The Hyper-Kamiokande experiment". In: *Journal of Physics: Conference Series*. Vol. 888. 1. IOP Publishing. 2017, p. 012020.

[97] T. Zhua. "Long-baseline neutrino oscillation sensitivities with Hyper-Kamiokande and impact of Intermediate Water Cherenkov Detector". In: (2022).

[98] Y. Kudenko. "Hyper-Kamiokande". In: *Journal of Instrumentation* 15.07 (2020), p. C07029.

[99] T. Seidenfeld. "R. A. Fisher's Fiducial Argument and Bayes' Theorem". In: *Statistical Science* 7.3 (1992), pp. 358–368. ISSN: 08834237.

[100] A.N. Kolmogorov and A.T. Bharucha-Reid. *Foundations of the theory of probability: Second English Edition*. Courier Dover Publications, 2018.

[101] J. Wakefield et al. *Bayesian and frequentist regression methods*. Vol. 23. Springer, 2013.

[102] J. Neyman. "Frequentist probability and frequentist statistics". In: *Synthese* 36.1 (1977), pp. 97–131.

[103] W.M. Bolstad and J.M. Curran. *Introduction to Bayesian statistics*. John Wiley & Sons, 2016.

[104] J. Berkson. "Bayes' Theorem". In: *The Annals of Mathematical Statistics* 1.1 (1930), pp. 42–56. ISSN: 00034851. URL: http://www.jstor.org/stable/2957673 (visited on 04/19/2024).

[105] U. Von Toussaint. "Bayesian inference in physics". In: *Reviews of Modern Physics* 83.3 (2011), p. 943.

[106] A. Gelman and C.R. Shalizi. "Philosophy and the practice of Bayesian statistics". In: *British Journal of Mathematical and Statistical Psychology* 66.1 (2013), pp. 8–38.

[107] J. Hammersley. *Monte carlo methods*. Springer Science & Business Media, 2013.

[108] M. Köppen. "The curse of dimensionality". In: *5th online world conference on soft computing in industrial applications (WSC5)*. Vol. 1. 2000, pp. 4–8.

[109] G. Huber. "Gamma Function Derivation of n-Sphere Volumes". In: *The American Mathematical Monthly* 89.5 (1982), pp. 301–302. ISSN: 00029890, 19300972.

[110] C.J. Geyer. "Practical Markov chain monte carlo". In: *Statistical science* (1992), pp. 473–483.

[111] J.R. Norris. *Markov chains*. Cambridge university press, 1998.

[112] J.S. Rosenthal. "Markov chain convergence: From finite to infinite". In: *Stochastic processes and their Applications* 62.1 (1996), pp. 55–72.

[113] G.L. Jones. "On the Markov chain central limit theorem". In: *Probability Surveys* 1.none (2004), pp. 299–320.

[114] S. Axler. *Measure, integration & real analysis*. Springer Nature, 2020.

[115] S. Chib and E. Greenberg. "Understanding the Metropolis-Hastings algorithm". In: *The A merican Statistician* 49.4 (1995), pp. 327–335.

[116] W.R. Gilks, S. Richardson, and D. Spiegelhalter. *Markov chain Monte Carlo in practice*. CRC press, 1995.

[117] G.O. Roberts and J.S. Rosenthal. "Examples of adaptive MCMC". In: *Journal of computational and graphical statistics* 18.2 (2009), pp. 349–367.

[118] H. Haario, E. Saksman, and J. Tamminen. "An adaptive Metropolis algorithm". In: *Bernoulli* (2001), pp. 223–242.

[119] C. Andrieu and J. Thoms. "A tutorial on adaptive MCMC". In: *Statistics and computing* 18 (2008), pp. 343–373.

[120] E. Pompe, C. Holmes, and K. Łatuszyński. "A framework for adaptive MCMC targeting multimodal distributions". In: (2020).

[121] J.S. Rosenthal et al. "Optimal proposal distributions and adaptive MCMC". In: *Handbook of Markov Chain Monte Carlo* 4.10.1201 (2011).

[122] M.K. Cowles and B.P. Carlin. "Markov chain Monte Carlo convergence diagnostics: a comparative review". In: *Journal of the American Statistical Association* 91.434 (1996), pp. 883–904.

[123] P. Dellaportas and G.O. Roberts. "An introduction to MCMC". In: *Spatial statistics and computational methods*. Springer, 2003, pp. 1–41.

[124] A.E. Brockwell and J.B. Kadane. "Identification of Regeneration Times in MCMC Simulation, with Application to Adaptive Schemes". In: *Journal of Computational and Graphical Statistics* 14.2 (2005), pp. 436–458. ISSN: 10618600.

[125] L. Hespanhol et al. "Understanding and interpreting confidence and credible intervals around effect estimates". In: *Brazilian journal of physical therapy* 23.4 (2019), pp. 290–301.

[126] J.M. Bernardo and R. Rueda. "Bayesian hypothesis testing: A reference approach". In: *International Statistical Review* 70.3 (2002), pp. 351–372.

[127] B. Efron et al. "Scales of evidence for model selection: Fisher versus Jeffreys". In: *Lecture Notes-Monograph Series* (2001), pp. 208–256.

[128] E. Atkin et al. *Measuring PMNS parameters in a joint T2K-NOvA analysis; Fitting workstream*. Tech. rep. TN-435. T2K-NOvA, 2023. eprint: https://t2k.org/docs/technotes/435.

[129] C.J. Geyer. "Introduction to Markov chain monte carlo". In: *Handbook of Markov chain monte carlo* 20116022 (2011), p. 45.

[130] R.Wendell T. Yoshida M. Wilking and K. Okumura. *A study of $\nu_\mu CC1\pi^+$ events at Super-Kamiokande*. Tech. rep. TN-388. T2K, 2020. eprint: https://t2k.org/docs/technotes/388.

[131] A.D. Missert, T2K Collaboration, et al. "Improving the T2K oscillation analysis with fiTQun: a new maximum-likelihood event reconstruction for Super-Kamiokande". In: *Journal of Physics: Conference Series*. Vol. 888. 1. IOP Publishing. 2017, p. 012066.

[132] R. Wendell et al. *Super-Kamiokande ν_μ multi-ring samples for the 2021 Oscillation Analysis*. Tech. rep. T2K, 2023.

[133] L. Haegel. *T2K near detector constraints for oscillation results*. 2017. arXiv: 1701.02559 [hep-ex].

[134] C. Riccio et al. *Constraining the flux and cross section models for the 2023 oscillation analysis using ND280 data*. Tech. rep. T2K, 2023.

[135] T. Vladisavljevic. "Constraining The T2K Neutrino Flux Prediction With 2009 Na61/Shine Replica-Target Data". In: (2018). DOI: 10.5281/ZENODO.1300546. URL: https://zenodo.org/record/1300546.

[136] R. Barlow and C. Beeston. "Fitting using finite Monte Carlo samples". In: *Computer Physics Communications* 77.2 (1993), pp. 219–228.

[137] K Abe et al. "Measurement of neutrino and antineutrino oscillations by the T2K experiment including a new additional sample of ν e interactions at the far detector". In: *Physical Review D* 96.9 (2017), p. 092006.

[138] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf. URL: https://doi.org/10.1093/ptep/ptaa104.

[139] T. Holvey et al. *MaCh3 2021 Run 1-10 Analysis and including Mutli-Ring Sample*. Tech. rep. T2K, 2024.

[140] P. Sedgwick. "Multiple significance tests: the Bonferroni correction". In: *Bmj* 344 (2012).

[141] F. Bench. *Study of ν and ν Oscillations in the Three-flavour PMNS Paradigm at the T2K Experiment: Determination of the CP-Violating Phase and the search for ν $\mu \rightarrow$ ν e Oscillations*. The University of Liverpool (United Kingdom), 2021.

[142] A.L. Moreno. *PyExotics*. 2023. URL: https://github.com/ALopezMoreno/pyExotics.

[143] N.J. Higham. "Computing a nearest symmetric positive semidefinite matrix". In: *Linear Algebra and its Applications* 103 (1988), pp. 103–118. ISSN: 0024-3795. DOI: https://doi.org/10.1016/0024-3795(88)90223-6. URL: https://www.sciencedirect.com/science/article/pii/0024379588902236.

[144] O. Benhar et al. "Spectral function of finite nuclei and scattering of GeV electrons". In: *Nuclear Physics A* 579.3 (1994), pp. 493–517. ISSN: 0375-9474.

Appendices

Appendix A

T2K Cross-Section Systematics Model

Overview

The T2K cross-section (neutrino interaction) model used in the oscillation analysis consists of 75 parameters which are binned in either outgoing lepton momentum (p) or p+outgoing lepton angle (θ). Systematics treatment is then further split in the following way:

- **Splines**: Systematic response is controlled by spline interpolation
- **Normalisation**: Same shift is applied regardless of sample bin
- **Function Shift**: Parameter is allowed to shift events between bins. Response for each bin is a function of individual event kinematics.

The treatment of each systematic is detailed in table A.1.

Systematic	Type	Generator Nominal	Prior	Range
CCQE				
M_A^{QE}	Spline	1.21	$\mathcal{N}(1.03, 0.06)$	$(0, \infty)$
Q2 supp. 0.25 – 0.50 GeV^2	Norm	1	$\mathcal{N}(1, 0.11)$	$(0, \infty)$
Q2 supp. 0.5 – 1.0 GeV^2	Norm	1	$\mathcal{N}(1, 0.18)$	$(0, \infty)$
Q2 supp. $>$ 1.0 GeV^2	Norm	1	$\mathcal{N}(1, 0.4)$	$(0, \infty)$
Spectral Function				
PShell MF Norm (C)	Spline	0	$\mathcal{N}(0, 0.2)$	$(-1, \infty)$
SShell MF Norm (C)	Spline	0	$\mathcal{N}(0, 0.45)$	$(-1, \infty)$
SRC Norm (C)	Spline	1	$\mathcal{N}(1, 2)$	$(0, \infty)$
PShell MF P–Miss Shape (C)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
SShell MF P–MissShape (C)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
P3 2Shell MF Norm (O)	Spline	0	$\mathcal{N}(0, 0.2)$	$(-1, \infty)$

Systematic	Type	Generator Nominal	Prior	Range
P1 2Shell MF Norm (O)	Spline	0	$\mathcal{N}(0, 0.45)$	$(-1, \infty)$
SShell MF Norm (O)	Spline	0	$\mathcal{N}(0, 0.75)$	$(-1, \infty)$
SRC Norm (O)	Spline	1	$\mathcal{N}(1, 2)$	$(0, \infty)$
P1 2Shell MF P–Miss Shape (O)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
P3 2Shell MF P–Miss Shape (O)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
SShell MF P–Miss Shape (O)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
Optical Potential (C)	Spline	0	Flat	$(0, 1)$
Optical Potential (O)	Spline	0	Flat	$(0, 1)$
Pauli Blocking				
Pauli Blocking (C, ν)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
Pauli Blocking (O, ν)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
Pauli Blocking (C, $\bar{\nu}$)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
Pauli Blocking (O, $\bar{\nu}$)	Spline	0	$\mathcal{N}(0, 1)$	$(-\infty, \infty)$
2p2h and MEC				
2p2h norm ν	Norm	1	Flat	$(0, \infty)$
2p2h norm $\bar{\nu}$	Norm	1	Flat	$(0, \infty)$
2p2h norm C→O	Norm	1	$\mathcal{N}(1, 0.2)$	$(0, \infty)$
2p2h Edep low E_ν	Spline	1	Flat	$(0, 1)$
2p2h Edep high E_ν	Spline	1	Flat	$(0, 1)$
2p2h Edep low $E_{\bar{\nu}}$	Spline	1	Flat	$(0, 1)$
2p2h Edep high $E_{\bar{\nu}}$	Spline	1	Flat	$(0, 1)$
PNNN Shape	Spline	0	$\mathcal{N}(0, 0.33)$	$(-1, 1)$
2p2h shape np (C)	Spline	0	$\mathcal{N}(0, 3)$	$(-1, 1)$
2p2h shape NN (C)	Spline	0	$\mathcal{N}(0, 3)$	$(-1, 1)$
2p2h shape np (O)	Spline	0	$\mathcal{N}(0, 3)$	$(-1, 1)$
2p2h shape NN (O)	Spline	0	$\mathcal{N}(0, 3)$	$(-1, 1)$
Single π Production				
C_A^5	Spline	1.01	$\mathcal{N}(1.06, 0.1)$	$(0, \infty)$
M_A^{Res}	Spline	0.95	$\mathcal{N}(1.3, 1.3)$	$(0, \infty)$
Non-resonant I-$\frac{1}{2}$ bkg., low P_π	Spline	1.3	$\mathcal{N}(1.3, 1.3)$	$(0, \infty)$
Non-resonant I-$\frac{1}{2}$ bkg.	Spline	1.3	$\mathcal{N}(1.21, 0.27)$	$(0, \infty)$
SPP π^0 Norm ν_μ	Norm	1	$\mathcal{N}(1, 0.3)$	$(0, \infty)$
SPP π^0 Norm $\bar{\nu}_\mu$	Norm	1	$\mathcal{N}(1, 0.3)$	$(0, \infty)$
Resonant π Production				
RES Eb (C, ν_μ)	Spline	0	$\mathcal{N}(25, 25)$	$(0, 50)$
RES Eb (O, ν_μ)	Spline	0	$\mathcal{N}(25, 25)$	$(0, 50)$
RES Eb (C, $\bar{\nu}_\mu$)	Spline	0	$\mathcal{N}(25, 25)$	$(0, 50)$
RES Eb (O, $\bar{\nu}_\mu$)	Spline	0	$\mathcal{N}(25, 25)$	$(0, 50)$
Final State Interaction				
π-FSI Quasi-Elastic	Spline	1.069	$\mathcal{N}(1.069, 0.313)$	$(0, \infty)$
π-FSI Quasi-Elastic, high	Spline	1.824	$\mathcal{N}(1.069, 0.313)$	$(0, \infty)$

Systematic	Type	Generator Nominal	Prior	Range
π-FSI Inelastic	Spline	1.002	$\mathcal{N}(1.002, 1.101)$	$(0, \infty)$
π-FSI absorption	Spline	1.404	$\mathcal{N}(1.404, 0.432)$	$(0, \infty)$
π-FSI charge exchange	Spline	0.697	$\mathcal{N}(0.697, 0.305)$	$(0, \infty)$
π-FSI charge exchange, high	Spline	1.8	$\mathcal{N}(1.8, 0)$	$(0, \infty)$
Nucleon FSI	Spline	0	$\mathcal{N}(0, 0.3)$	$(-1, 1)$
CC Coherent				
CC Coh. Normalisation (C)	Norm	1	$\mathcal{N}(1, 0.3)$	$(0, \infty)$
CC Coh. Normalisation (O)	Norm	1	$\mathcal{N}(1, 0.3)$	$(0, \infty)$
CC Multi-π				
Multi-π Multi TotXSec	Spline	0	$\mathcal{N}(0, 0.3)$	$(-\infty, \infty)$
Multi-π BY Vector	Spline	0	$\mathcal{N}(0, 0.3)$	$(-\infty, \infty)$
Multi-π BY Axial	Spline	0	$\mathcal{N}(0, 0.3)$	$(-\infty, \infty)$
Multi-π Multi Shape	Spline	0	$\mathcal{N}(0, 0.3)$	$(-\infty, \infty)$
CC BY DIS	Spline	0	$\mathcal{N}(0, 1.0)$	$(-\infty, \infty)$
CC DIS Mult-π Norm ν	Norm	1	$\mathcal{N}(1, 0.35)$	$(0, \infty)$
CC DIS Mult-π Norm $\bar{\nu}$	Norm	1	$\mathcal{N}(1, 0.65)$	$(0, \infty)$
CC Misc. Norms				
CC Misc	Norm	1	$\mathcal{N}(1, 1)$	$(0, \infty)$
NC Norms				
NC Coh.	Norm	1	$\mathcal{N}(1, 1)$	$(0, \infty)$
NC 1γ	Norm	1	$\mathcal{N}(1, 1)$	$(0, \infty)$
NC other near	Norm	1	$\mathcal{N}(1, 0.3)$	$(0, \infty)$
NC other far	Norm	1	$\mathcal{N}(1, 0.3)$	$(0, \infty)$
Coulomb Corrections				
CC norm ν	Norm	1	$\mathcal{N}(1, 0.02)$	$(0, \infty)$
CC norm $\bar{\nu}$	Norm	1	$\mathcal{N}(1, 0.01)$	$(0, \infty)$
Nue/Numu Uncertainty				
ν_e ν_μ	Norm	1	$\mathcal{N}(1, 0.0283)$	$(0, \infty)$
$\bar{\nu}_e$ $\bar{\nu}_\mu$	Norm	1	$\mathcal{N}(1, 0.0283)$	$(0, \infty)$
E_b Systematics				
EB dial (C, ν)	Shift	0	$\mathcal{N}(2, 6.0)$	$(-10, 15)$
EB dial (C, $\bar{\nu}$)	Shift	0	$\mathcal{N}(0, 6.0)$	$(-10, 15)$
EB dial (O, ν)	Shift	0	$\mathcal{N}(4, 6.0)$	$(-10, 15)$
EB dial (O, $\bar{\nu}$)	Shift	0	$\mathcal{N}(0, 6.0)$	$(-10, 15)$
$\alpha - q^3$	Shift	0	$\mathcal{N}(0, 1.0)$	$(0, 1)$

Table A.1: Summary table of cross-section systematic parameters grouped by interaction type. Priors are either flat (the identity everywhere) or a normal distribution with mean μ and error σ [$\mathcal{N}(\mu, \sigma)$]. And parameter responses are either functional (shift), spline-based or normalisation-based.

Appendix B

T2K-only OA2023 data fit results

Overview

This section contains the remaining T2K only plots (no reactor constraint) from the data fit detailed in chapter 4.

The following figures [Figs. B.2, B.1, B.3] detail the 1D posteriors from the T2K only data fit. The 2D appearance and disappearance contours are shown in figures B.4 and B.5 respectively.

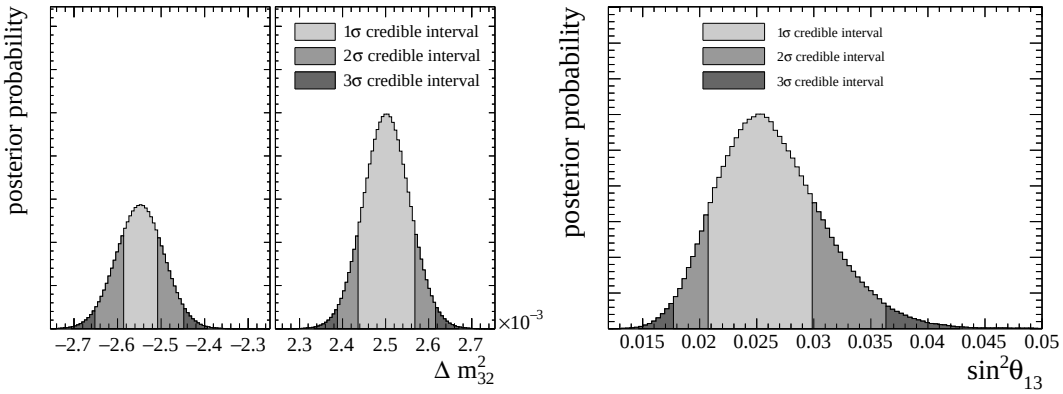

Figure B.1: T2K-only data fit posteriors for Δm^2_{32} (left) and $\sin^2(\theta_{13})$ (right).

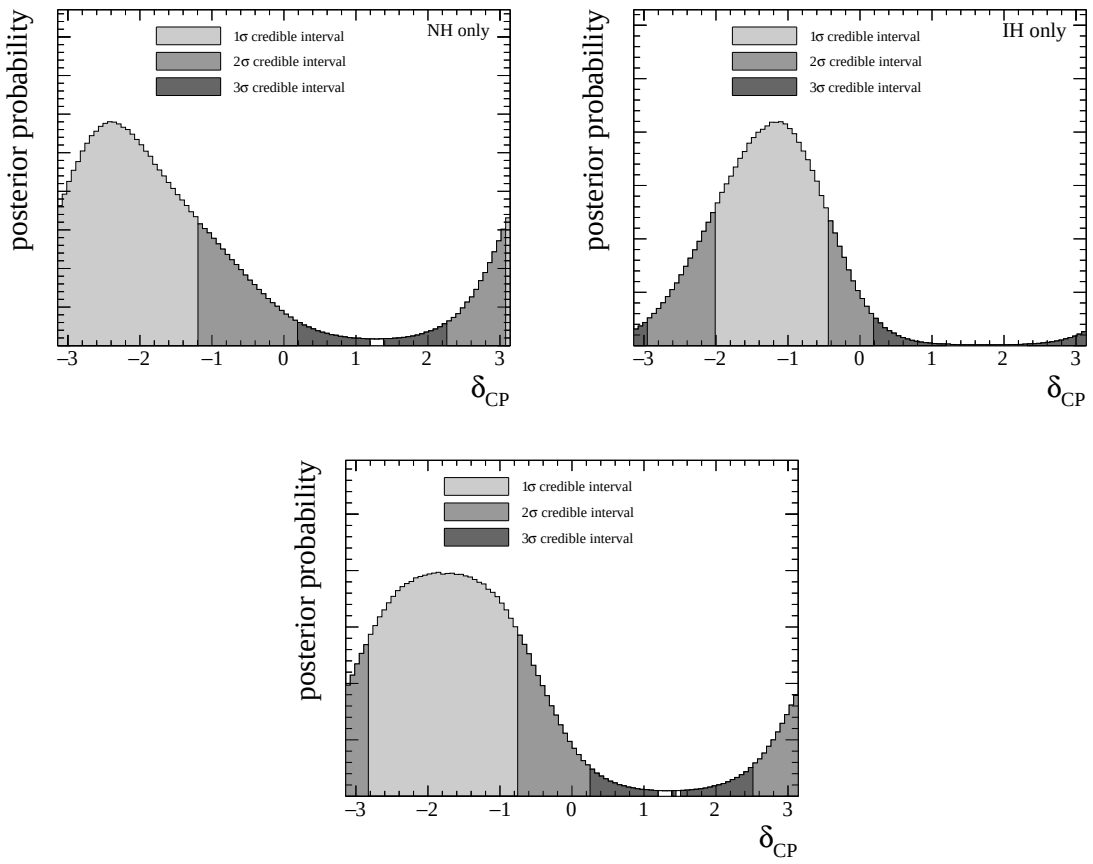

Figure B.2: T2K-only data fit showing δ_{CP} contour marginalised over normal ordering (top left), inverted ordering (top right) and including steps in both orderings (bottom) . In addition the PTheta MaCh3-Like contours are shown. These use the same sample binning scheme as MaCh3 and are expected to agree far better than the standard scheme.

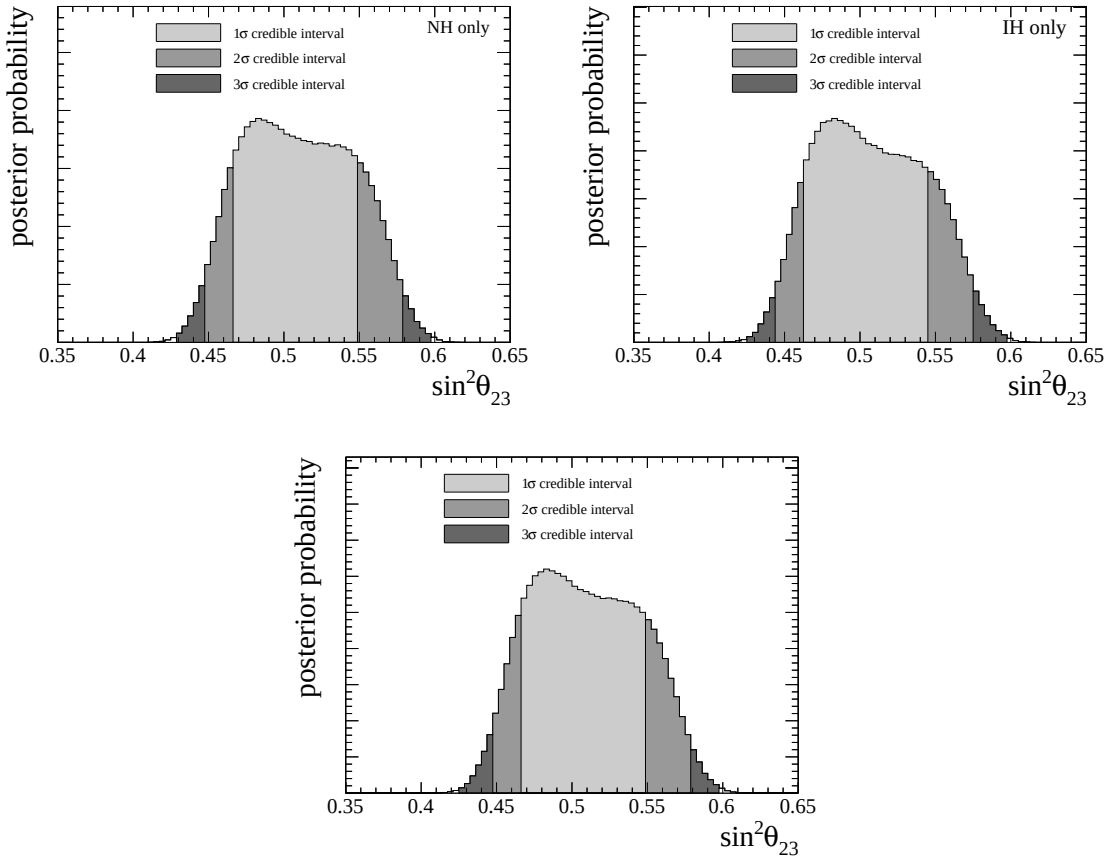

Figure B.3: T2K-only data fit showing $\sin^2(\theta_{23})$ contour marginalised over normal ordering (top left), inverted ordering (top right) and including steps in both orderings (bottom)

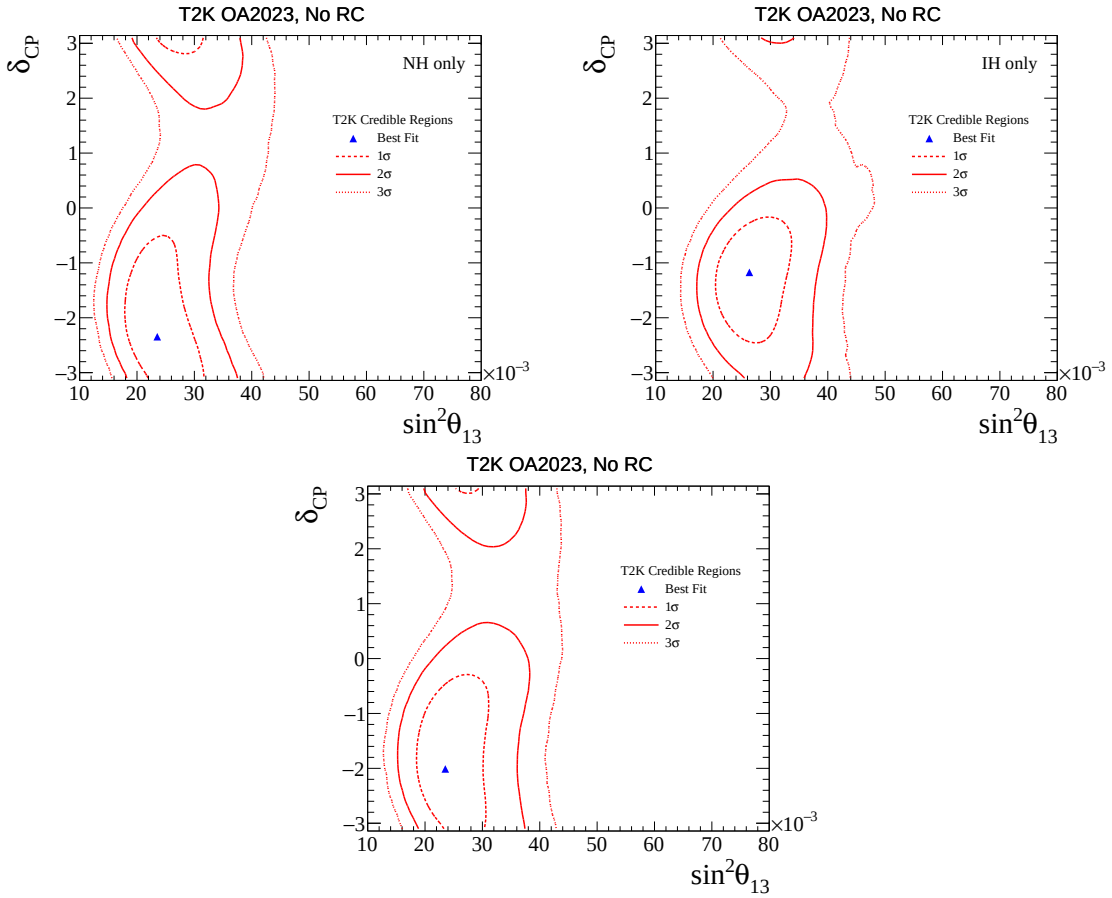

Figure B.4: Appearance contours from the T2K-only data fit marginalised over normal ordering (top left), inverted ordering (top right) and without ordering based marginalisation (bottom)

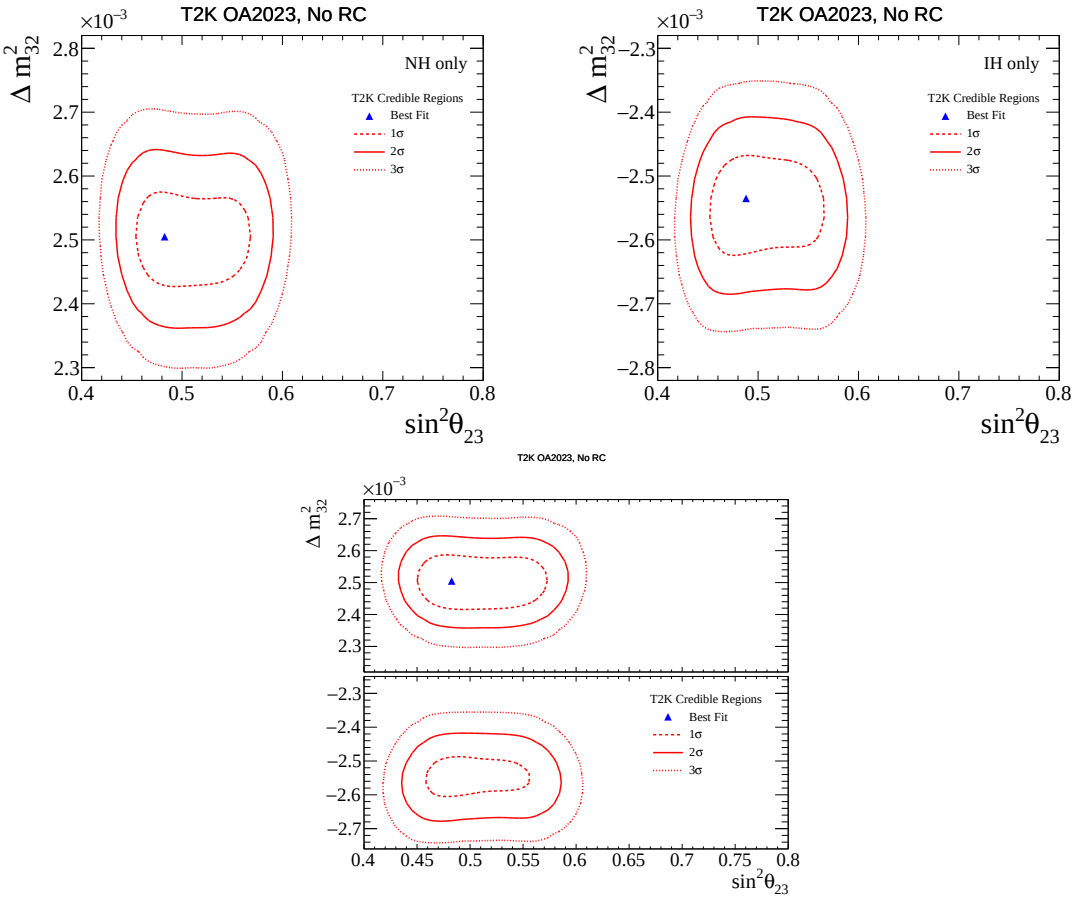

Figure B.5: Disappearance contours from the T2K-only data fit marginalised over normal ordering (top left), inverted ordering (top right) and without ordering based marginalisation (bottom)