Standard and Non-Standard Neutrino-Antineutrino Oscillation Analyses and Event Reconstruction Studies using Markov Chain Monte Carlo Methods at T2K

Artur Adam Sztuc
Imperial College London

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy in Physics

Abstract

T2K (from Tokai-to-Kamioka) is a long-baseline neutrino experiment using an off-axis neutrino beam to produce a narrow neutrino energy spectrum. T2K uses a near detector (ND280) to constrain the neutrino cross-section and beam flux systematics for the neutrino oscillation analyses, and provides some of the world's leading cross-section measurements. The neutrino beam is directed to, and detected at, Super-Kamiokande, a large 50 kT Water Cherenkov detector that is also being used to study atmospheric, cosmic, supernovae neutrinos and more.

We describe the oscillation analysis for the T2K experiment that uses both near and far detector data, using Markov Chain Monte Carlo techniques to construct the Bayesian posterior probability distributions for the oscillation parameters. The analysis of the full T2K run 1–9 data with the reactor constraint on $\sin^2\theta_{13}$ show that the CP conserving values of δ_{CP} are rejected with at least $2\,\sigma$ confidence and intervals between -2.95 and -0.50 with the best fit point -1.82 rad, whereas the $3\,\sigma$ intervals are $-\pi$ – 0.13 & 2.80 – π. The Normal Mass hierarchy has 88.9% of the total Markov Chain Monte Carlo steps, corresponding to the likelihood-ratio of the two hypotheses, Bayes Factor, of 8.0. This can be considered as "substantial" on the Jeffreys scale, which is an interpretation the strength of Bayes Factors. The best fit point of $\sin^2\theta_{23}$ is 0.537, preferring higher octant with 79.5% Markov Chain Monte Carlo steps. This corresponds to the Bayes Factor of 3.4, again "substantial" on the Jeffreys scale.

Neither run 1–9c nor full run 1–9d data are incompatible with the PMNS model. The $1\,\sigma$ intervals for $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ posterior probability distributions for both PMNS and non-PMNS analyses comfortably overlap. Far more data from the T2K's far detector, Super-K, is needed to see any potential deviations; we are currently statistically limited.

We also describe the development of a novel neutrino event reconstruction method for large Water Cherenkov detectors using similar Bayesian methods as in the oscillation analysis.

Declaration

The analyses described in this thesis are my own work for the T2K and Super-Kamiokande experiments, with the work and input of others properly cited and referenced where needed. The work in this thesis was not submitted for any other qualification and therefore was done for this thesis specifically.

Chapters 1 and 2 are technically a summary of the work of others; the through description of the neutrino experiments that this thesis relates to, and a brief theory and history behind the neutrino oscillation. Other people's work that contributed to the theory, or the commissioning and running of these experiments are appropriately referenced. Chapter 3 discusses the statistical methods used for the thesis' main topics, and describes my search for, and the implementation of novel methods in the T2K neutrino oscillation analysis.

Chapters 4 and 5 describe the Bayesian Oscillation Analyses done by me for the T2K experiment using, the data and other inputs generated by the T2K and Super-Kamiokande collaborations.

Chapter 6 describes my work on the tuning of the Bayesian Reverse-Jump Markov Chain Monte Carlo sampling method into neutrino event reconstruction framework in large Water-Cherenkov detectors. The method itself was not invented by me and the original implementation was originally started by another T2K collaborator.

Copyright Declaration

Artur Sztuc

Acknowledgements

This thesis and work done towards it would not be possible without the contribution of many people, whether from the HEP world or outside of it. First of all I would like to thank my supervisor, Professor Yoshi Uchida, who has been continuously providing the support, opportunities and faith in my abilities. This extends to the long baseline neutrino group at Imperial College London, past and present: Prof. Morgan Wascko, Dr. Per Jonsson, Dr. Phillip Litchfield, Dr. Patrick Dunne, Dr. Clarence Wret, Dr. Wing Ma, Dr. Toby Nonnenmacher, Edward Atkin, Wilf Shorrock, as well as Dr. Asher Kaboth and Dr. Will Parker from the Royal Holloway group. Your expertise, gradience and discussions have been invaluable. At the same time, I will never forget all the karaokes, coffee breaks, road trips, and the time spent like a family.

I am thankful to the MaCh3 group, the T2K and Super-Kamiokande experiments for providing me with the support and opportunities to do the work I love. All the analyses that I was involved in were a collaborative work of many people and many groups, and I really appreciate being a part of them. I would also like to thank the J-PARC group in Tokai, and the fantastic staff from the Kenkyuto office in Kamioka, for all the help with work, problems outside of Physics and administrative tasks during my LTA in Japan. Thank you Dr. Matthew Lawe and Dr. Trevor Stewart and Liam Ó Súilleabháin for all the onsen trips, beers, the time spent cycling between J-PARC and 7-Eleven.

My family and loved ones have been my backbone throughout the PhD. A special thank you to my parents, siblings and Cyreille for their patience, visits in Kamioka, long chats, and making sure that I do not go too crazy.

Finally, I would like to thank all the friends around Kamioka town for all the help, too many ochazukes, and making my time in the town one of the best experiences I have had. Thank you Akutsu-san, Sonoda-san, Takenaka-san, the 結城 family, Maekawa-san, Shimizu-san, Todani-san, Ishikura-san, Fujimura-san, Lauren Anthony, Jeong Yang, and many others that left a permanent mark on me. I will really miss going to local izakayas, exploring different onsens and getting lost on the mountain roads with you. Living in the mountains, the views, great local coffee and amazing people have really contributed to my time as a PhD student and to my day to day life.

Contents

List of Figures

Chapter 1

Neutrino Physics

1.1 Introduction

Neutrinos are neutral fermions predicted by Wolfgang Pauli in 1930 [1], and a part of the Standard Model (SM) that could help explain some of the biggest existential questions that physics poses. We know that as we create matter, we have to create equal amount of antimatter through pair-production. This, however, is not what we observe from the result of the Big Bang. We see an abundance of matter in the universe, and not much antimatter, meaning that the universe's current baryon number does not seem to be 0 as expected. Have the matter and antimatter been generated in equal proportions during the Big Bang, mass annihilation would have made the current abundance of variety in the universe impossible. The hypothetical process by which the universe started with an average baryon number of 0, and ended with mostly matter, is called Baryogenesis; still unresolved, and not understood.

The Standard Model cannot fully solve the problem of baryogenesis; there must be more physics beyond the Standard Model to explain this discrepancy. One of the possible processes allowing this, is violation of the Charge-Parity symmetry [2] – where the laws of physics would apply differently for particles and antiparticles. Neutrinos, in the standard model, are massless particles. There is more than enough of experimental evidence showing that this is not the case; neutrino flavour changes as it propagates

through space, a phenomena called "neutrino oscillation", which is a process that is only possible if neutrinos do indeed have a mass. Since neutrinos have mass that is linked with their flavour that oscillates, these oscillations could violate CP-symmetry, possibly contributing to the solution of the matter-antimatter asymmetry, and maybe even contribute - in the far future - to the development of a Grand Unified Theory (GUT), linking the electromagnetic, weak and strong force [3].

There are more still-unanswered questions in neutrino physics, some of which this thesis will contribute as well. Are the neutrino mass eigenstates arranged in a normal or inverted order? Answering this question would then help us with choosing appropriate Grand Unified Theory (GUT), and help with determining our current sensitivity to answering another question; are neutrinos their own antiparticles (Majorana-like), or are neutrinos and antineutrinos separate particles (Dirac-like)? If neutrinos are Majorana-like, it would be potentially far more difficult to resolve the Dirac vs Majorana neutrino. If neutrinos are, however, Majorana-like, the possibility of a fourth, heavy neutrino would be possible through the Seesaw mechanism (which tries to explain why neutrinos are so much lighter than the charged fermions), getting us closer to GUT [4].

Both the neutrino mass hierarchy and the CP-violating phase are some of the main goals the T2K experiment, and are being measured through the neutrino oscillations; the main topic of this thesis.

1.2 Discovery of neutrinos

Neutrinos were first proposed as the solution to anomalous results from β-decay experiments, in the early 20th century. James Chadwick found in 1914 that the β-decay of a nucleus does not result in fixed electron energy - as postulated given the then-agreed-upon proton-electron nuclear model [7]. This was later confirmed by C. Ellis and W. Wooster, who preformed more detailed experiments and confirmed that energy conservation seems to be broken [8].

In 1930, Wolfgang Pauli proposed the "neutron", an electrically neutral particle with spin 1/2, and with a mass of the same order of magnitude as an electron and no

greater than 0.01 of the proton's mass. That particle was meant to be emitted alongside the electron in β-decays, with the sum of energies being constant - potentially solving problems with the violation of energy conservation [1]. In 1932, J. Chadwick performed another β experiment, where he discovered a heavy neutral particle, which he called the "neutron".

This was not the light neutron Pauli had imagined, and it was not until Enrico Fermi developed the idea further. He added the newly discovered heavy neutron[1] to the nuclear model, and renamed Pauli's light neutron to "neutrino"[9]. He assumed the neutrino has no mass, and that an electron and antineutrino pair was produced during β-decay, through $n \to p + e^- + \bar{\nu}$.

It was a long wait for the first observation of a neutrino; in 1953 F. Reines and C.L. Cowan detected neutrinos for the first time at the Savannah River nuclear reactor[2] via inverse-β decay, $\bar{\nu} + p = e^+ + n$. The experiment consisted of a liquid scintillator loaded with $CdCl_2$ as the target material, with 110 photo-multiplier tubes (PMTs) to detect γs from positron-electron annihilation [11, 10]. The results from the Savannah River experiment eventually got F. Reines and C. Cowan a Noble price in 1995.

The existence of the muon flavored neutrino, ν_μ, was confirmed with the first ever accelerator beam neutrino experiment in 1962. L.M. Lederman, M. Schwartz, J. Steinberger et al. carried out an idea from B. Pontecorvo to use a beam of π^+, decaying via $\pi^+ \to \nu_\mu + \mu^+$ and $\pi^+ \to \nu_e + e^+$ [13]. The neutrino beam transversed heavy iron shielding to stop any charged particles, and finally through a 10 ton aluminium spark chamber where neutrino interactions produced charged particles - the designated neutrino detector. The pion decay involving electrons was expected to be heavily suppressed, and if ν_e and ν_μ were the same, we would expect to see equal numbers of μ^- and e^- in the detector. Instead, a large excess of μ^- and a background-consistent number of e^- was observed, confirming that ν_μ exists and is a different particle from ν_e. These results were later confirmed at CERN in 1964 with a greater beam intensity and lower ν_e contamination [14].

[1] The particle we still know today as the "neutron"

[2] Originally they wanted to use a nuclear bomb explosion as a source of neutrinos. It was only later that it occurred to them that the already-existing nuclear reactors make for a somewhat safer experiment

Years 1975-1977 brought another big discovery; an even heavier charged lepton, τ, was produced using an electron-positron collider at the Stanford Linear Accelerator Center (SLAC) laboratory [15]. Anomalous events, $e^+ + e^- \rightarrow e^\pm + \mu^\mp +$ *missing energy*, were found at the centre-of-mass energies of 4 GeV and above, where the spectra of the missing energy indicated two undetected particles. Of course, given that all the other charged leptons had the same flavoured neutrino companion, it was expected that the ν_τ must exist. This was proved to be true by the DONUT collaboration at Fermilab in 2000, where four ν_τ events were found by looking for τ at the interaction vertex in nuclear emulsion.

1.3 Neutrino Mixing

Neutrino Oscillation is a phenomenon where a neutrino with a certain flavour changes into a different flavour after propagating some distance. Neutrinos must have a non-zero mass to oscillate. This stands against the Standard Model, which assumes neutrinos to be Dirac-like, left-handed particles with zero mass.

The Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix links the three neutrino flavour eigenstates with three mass eigenstates, and it is the most commonly used model in the experiments that involve neutrino oscillations [6, 40]. The historical aspects of this matrix are discussed more in section 1.5. The formalisation starts by defining three mass eigenstates, ν_1, ν_2 and ν_3, where each mass eigenstate is a superposition of all the flavour states;

$$|\nu_i\rangle = \sum_\alpha \mathrm{U}_{\alpha i}|\nu_\alpha\rangle, \tag{1.1}$$

where $\mathbf{U}$ is the PMNS mixing matrix, linking the mass and flavour states together;

$$\mathbf{U} = \begin{pmatrix} \mathrm{U}_{e1} & \mathrm{U}_{e2} & \mathrm{U}_{e3} \\ \mathrm{U}_{\mu 1} & \mathrm{U}_{\mu 2} & \mathrm{U}_{\mu 3} \\ \mathrm{U}_{\tau 1} & \mathrm{U}_{\tau 2} & \mathrm{U}_{\tau 3} \end{pmatrix}. \tag{1.2}$$

This 3×3 unitary mixing matrix is often separated into three separate ones, in terms of the mixing angles, θ_{13}, θ_{23}, and δ_{CP}, the Charge-Parity violating phase. The

split matrix is:

$$\mathbf{U} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix}}_{\text{Atmospheric, Accelerator}} \underbrace{\begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -\sin\theta_{13}e^{i\delta_{\text{CP}}} & 0 & \cos\theta_{13} \end{pmatrix}}_{\text{Reactor, Accelerator}} \underbrace{\begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Solar, Reactor}}. \tag{1.3}$$

This split is done for convenience, as it shows which neutrino's origin constrain these parameters the best. The solar experiments, for example, measured the oscillations of ν_e to ν_μ, which is most sensitive to θ_{12}. Many experiments are trying to measure the PMNS parameters with neutrinos from multiple origins, for example, Super-K measures the solar and atmospheric neutrinos, with some potential for the reactor neutrinos [132].

A neutrino propagates through space as a superposition of the three mass eigenstates, so eq. 1.1 can be rewritten in terms of mass states by taking the Hermitian conjugate of the mixing matrix and propagating it with time in a form of a plane wave (in other words, a solution to the time-dependent Schrödinger equation):

$$|\nu_\alpha(t)\rangle = \sum_i \mathrm{U}^*_{\alpha i}|\nu_i\rangle e^{-i\phi_i}. \tag{1.4}$$

The neutrino now propagates through space with time, and thanks to the PMNS mixing matrix there is a probability that it will change its flavour as it travels. The amplitude of a neutrino with given flavour α being observed with a different flavour β after propagation is $\langle\nu_\beta|\nu_\alpha(t)\rangle$, and so the probability of a neutrino oscillating can be written;

$$P(\nu_\alpha \to \nu_\beta) = \sum_{i,j} \mathrm{U}^*_{\alpha i}\mathrm{U}_{\beta i}\mathrm{U}_{\alpha j}\mathrm{U}^*_{\beta j}e^{-i(\phi_j-\phi_i)}, \tag{1.5}$$

where the $\phi_j - \phi_i$ term is equivalent to $E_j t - E_i t - p_j x + p_i x$. We can assume that a neutrino is a relativistic particle, and we approximate that

$$p_i = \sqrt{E_i^2 - m_i^2} \approx E_i - \frac{m_i^2}{2E_i}. \tag{1.6}$$

We can make a further approximation, where the neutrino mass eigenstates are created with the same energy (momenta), and that $x = t = \mathrm{L}$ (where L is the distance travelled by the neutrino) to give

$$\phi_j - \phi_i = \frac{\Delta m^2_{ij}\mathrm{L}}{2E}. \tag{1.7}$$

Going back to eq. 1.5, we can now substitute the PMNS parameters and 1.7 and expand all the terms to get the probability of ν_α oscillating to ν_β,

$$P(\nu_\alpha \to \nu_\beta) = \langle \nu_\beta | \nu_\alpha \rangle - 4 \sum_{i>j} \mathbb{R}(\mathrm{U}^*_{\alpha i}\mathrm{U}_{\beta i}\mathrm{U}_{\alpha j}\mathrm{U}^*_{\beta j}) \sin^2\left(\frac{\Delta m^2_{ij}\mathrm{L}}{4E}\right) \tag{1.8}$$

$$+2 \sum_{i>j} \mathbb{I}(\mathrm{U}^*_{\alpha i}\mathrm{U}_{\beta i}\mathrm{U}_{\alpha j}\mathrm{U}^*_{\beta j}) \sin\left(\frac{\Delta m^2_{ij}\mathrm{L}}{4E}\right), \tag{1.9}$$

for the neutrino case, and the imaginary part is taken away in the antineutrino case. The $\langle \nu_\beta | \nu_\alpha \rangle$ term is the Kronecker delta $\delta_{\alpha\beta}$, equal to 1 if $\alpha = \beta$ and 0 otherwise.

It is easier to derive the neutrino oscillation probability using the the two-flavour neutrino approximation, where the unitary mixing matrix $\mathbf{U}$ becomes much simpler and reduces to

$$\mathbf{U} = \begin{pmatrix} \mathrm{U}_{\alpha 1} & \mathrm{U}_{\alpha 2} \\ \mathrm{U}_{\beta 1} & \mathrm{U}_{\beta 2} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{1.10}$$

whereas the oscillation probability reduces to

$$P(\nu_\alpha \to \nu_\beta) = 2\sin^2(2\theta)\sin^2\left(\frac{1.27\Delta m^2_{ij}L}{E}\right), \tag{1.11}$$

where L is given in km, Δm^2_{ij} in eV^2 and E in GeV. The number 1.27 is picked up when we take $\hbar$ and c into the account.

Most long-baseline accelerator neutrino experiments use the full three-flavour oscillation hypothesis and are concerned with two oscillation probabilities; electron neutrino (and electron antineutrino) appearance from a ν_μ beam, $P(\nu_\mu \to \nu_e)$, and the ν_μ survival probability (also called "ν_μ disappearance"),

$$\begin{aligned} P(\overset{(-)}{\nu}_\mu \to \overset{(-)}{\nu}_e) \simeq \; & \sin^2\theta_{23}\sin^2 2\theta_{13}\sin^2\frac{\Delta m^2_{32}L}{4E} + \sin^2 2\theta_{13}\cos^2\theta_{23}\sin\frac{\Delta m^2_{21}L}{4E} \\ & \overset{(+)}{-} \cos\theta_{13}\sin 2\theta_{12}\sin 2\theta_{23}\sin 2\theta_{13}\sin\frac{\Delta m^2_{32}L}{4E}\sin\frac{\Delta m^2_{21}L}{4E}\sin\frac{\Delta m^2_{32}L}{4E}\sin\delta_{\mathrm{CP}} \\ & + \cos\theta_{13}\sin 2\theta_{12}\sin 2\theta_{23}\sin 2\theta_{13}\sin\frac{\Delta m^2_{32}L}{4E}\sin\frac{\Delta m^2_{21}L}{4E}\cos\frac{\Delta m^2_{32}L}{4E}\cos\delta_{\mathrm{CP}} \\ & + (\text{solar, matter effect terms}), \end{aligned} \tag{1.12}$$

and

$$P(\overset{(-)}{\nu_\mu} \to \overset{(-)}{\nu_\mu}) \simeq \; 1 - 4\sin^2_{23}\cos^2_{13}\cos^2_{23}\sin^2\frac{\Delta m^2_{32}L}{4E} - 4\sin^2_{23}\cos^2_{13}\sin^2_{23}\sin^2_{13}\sin^2\frac{\Delta m^2_{32}L}{4E} + (\text{solar, matter effect terms}). \tag{1.13}$$

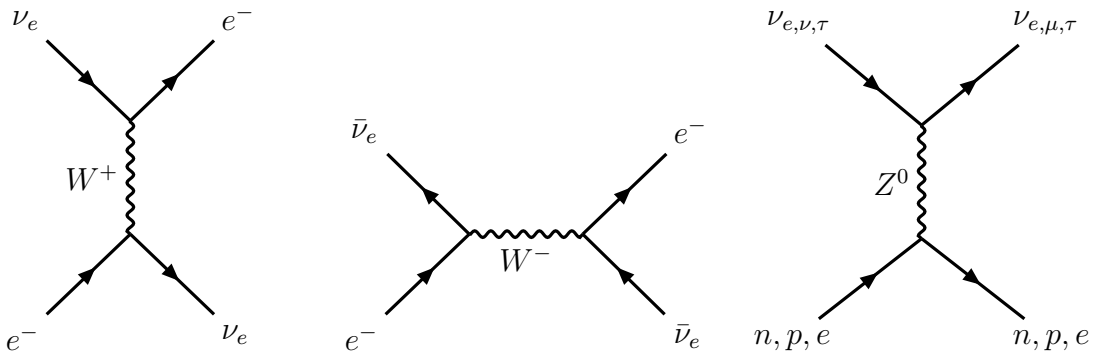
Most sensitivity to δ_{CP} comes from the appearance channel that has a dependence on $\sin\delta_{\text{CP}}$ and $\cos\delta_{\text{CP}}$. The appearance channel also gives a small sensitivity to the sign of the Δm^2_{32} and the octant of θ_{23}, whereas the disappearance channel gives a first-order sensitivity to the magnitude of Δm^2_{32} and θ_{23}.

1.4 Interactions in matter

The oscillation probability equations 1.12 and 1.13 contain the "matter effect" terms. The matter effect is very important as many neutrino experiments detect neutrinos that traverse through the Earth's soil, the Sun's core and outer layers etc. from the beam origin to the detector over a large distance. Figure 1.1 shows the two types of neutrino interaction with electrons in matter, one through the charged-current W boson exchange, and the neutral-current interaction with Z^0 boson exchange.

Each interaction modifies the potential that neutrinos propagate through – and this potential is different for electron neutrinos, since they can additionally interact via the charged-current interaction. This additional potential is felt through the modified Hamiltonian, $V_{CC} = \pm\sqrt{2}G_F N_e$, where G_F is the Fermi constant, and N_e is the electron density in matter. With the Earth being made of matter only, and not antimatter, an electron neutrino that propagates through matter gets a positive sign of V_{CC}, whereas anti-neutrino – negative. Without taking this into the account in the 3-flavour oscillation model, it would be easy to get a fake "CP-violation" effect.

This is called the Mikheyev–Smirnov–Wolfenstein (MSW) [128, 129] effect, and it appears in the oscillation probability as an extra dependence on the sign of Δm^2_{32} and the distance travelled L, separate from the L/E terms from the plane wave equations.



(a) ν_e on e^- charged-current interaction

(b) $\bar{\nu}_e$ on e^- charged-current interaction

(c) Neutral-current interaction for neutrino of any flavour

Figure 1.1: Feynman diagrams for charged-current (a, b) and neutral-current (c) neutrino interactions with matter

This means it is possible to gain extra sensitivity to the mass hierarchy by extending the distance travelled in matter by neutrinos whilst fixing the L/E ratio in, say, long-baseline accelerator experiments – even if δ_{CP} is 0. Chapter 5 will illustrate how the matter effect affects the T2K oscillation analysis, and compare the T2K matter effect against the NOvA experiment.

1.5 Evidence for neutrino oscillation

Ray Davis designed and started building an experiment in the depths of the South Dakota Homestake mine, almost 1.5km underground, around 1965. The experiment was designed to measure the solar neutrino flux, which was back then predicted by a well-established solar model. A 380 cubic litres tank was filled with a dry cleaning liquid rich with chlorine, and an electron neutrino would interact with that chlorine via

$$\nu_e + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar} + e^-,$$

to produce an unstable isotope of argon. R. Davis isolated the accumulated argon isotopes by bubbling the detector with helium gas and chemically purifying it. The isolated and purified argon sample would then be inserted into a small gas proportional counter to find the number of neutrino interactions. The experiment ran between 1970-1994 and

from the early start has seen around 1/3 of the expected electron neutrino flux. The first results, published in 1968 [20], were widely criticized. Both the calculated prediction of the solar flux and the experiment itself were blamed for the data-model expectation discrepancy. The calculations were checked and confirmed to be correct, whereas R. Davis insisted that there was nothing wrong with the experiment itself.

The Homestake experiment was the first to measure the discrepancy between the expected and measured solar neutrino flux, creating the famous "solar neutrino problem". Around the same time Bruno Pontecorvo predicted the oscillation of neutrinos [5, 6], which could solve the discrepancy. This required neutrino to have a mass, and the idea of a massless neutrino was still the dominant one in the physics community; therefore, not much attention was paid to Pontecorvo at that time.

A new 1 kton Water-Cherenkov experiment, Kamiokande, started taking data in the Kamioka mine in Japan around 1983, and published its own solar neutrino measurements around one year after the Homestake experiment, in 1989 [21]. The results echoed what the Homestake experiment has seen; there was an apparent deficit in the solar neutrino flux.

In 2002 the SNO collaboration has confirmed the deficit in a model-independent way. They measured the neutrino flux by using heavy water as a target material (Kamiokande used ultra-purified water), allowing for three interactions;

$$\nu_e + d \rightarrow e^- + p + p, \tag{1.14}$$

$$\nu_x + d \rightarrow \nu_x + p + n, \tag{1.15}$$

$$\nu_x + e^- \rightarrow \nu_x + e^-, \tag{1.16}$$

meaning that SNO could measure not only the electron neutrino flux, but measure the total neutrino flux as well. When measured the signal from interaction 1.14 alone, SNO observed only third of the expected neutrinos. The total neutrino flux, when measured over all the above interactions, was however consistent with the solar model expectation for electron neutrino flux. The ν_e flux was measured to be $1.76^{+0.14}_{-0.14} \times 10^6\,\mathrm{cm^2 s^{-1}}$, whereas the $\nu_{\mu,\tau}$ flux was $3.41^{+0.93}_{-0.90} \times 10^6\,\mathrm{cm^2 s^{-1}}$; the sum of which was consistent with the expected ν_e-only flux under no-oscillations hypothesis, $5.21^{+0.35}_{-0.35} \times 10^6\,\mathrm{cm^2 s^{-1}}$. The paper was titled "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions

in the Sudbury Neutrino Observatory"[22], with no further comment needed.

Thus, the era of large neutrino oscillation experiments has began. The experiments can be split into a few distinct groups, depending on the origin of the neutrinos that the experiments are measuring, and these are; atmospheric neutrinos, reactor neutrinos and accelerator neutrinos.

Atmospheric Neutrinos

Cosmic rays interacting with the Earth's atmosphere cause hadronic showers, resulting in charged pions. These pions, with energy widely peaked around 1 GeV, will decay further to muons and neutrinos via

$$\pi^{\pm} \rightarrow \mu^{\pm} + \overset{(-)}{\nu}_{\mu}. \tag{1.17}$$

Hints of ν_μ disappearance (ν_μ oscillating into different flavour) were first seen in IMB [23] and Kamiokande-II [24] experiments, but it was Super-Kamiokande collaboration, in 1998, that found a solid evidence of atmospheric neutrino oscillation [25]. Figure 1.2 shows the deficit of muon neutrinos at high L/E, as expected under $\nu_\mu \rightarrow \nu_\tau$ prediction shown with a dashed line. Neutrinos at low L/E had more probability to be generated in the atmosphere above Super-Kamiokande, having less time to oscillate, whereas the high L/E correspond to the neutrinos that travelled through the earth to reach the Super-Kamiokande tank.

Atmospheric neutrino experiments are the most sensitive to two PMNS-model oscillation parameters; Δm^2_{32} and $\sin^2\theta_{23}$. It is often common to call these "atmospheric parameters", even though nowadays these can also be successfully probed by long-baseline accelerator neutrino experiments. The comparison of the results between various atmospheric and accelerator neutrino experiments is shown in figure 1.3, where we see an excellent agreement between them.

Reactor neutrinos

Section 1.2 described the first neutrino discovery, which was done with a reactor experiment. Reactor neutrino experiments are usually short-baseline, on the order of $\sim$1 km,

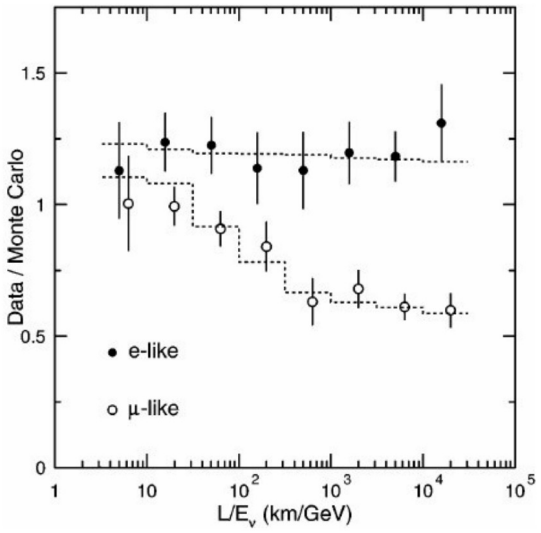


Figure 1.2: Results from the SK atmospheric analysis: ratio of the fully-contained events over the Monte-Carlo events for e-like and μ-like events for different reconstructed L/E. The dashed lines show the predicted ratios under the $\nu_\mu \to \nu_\tau$ hypothesis. Figure taken from [25].

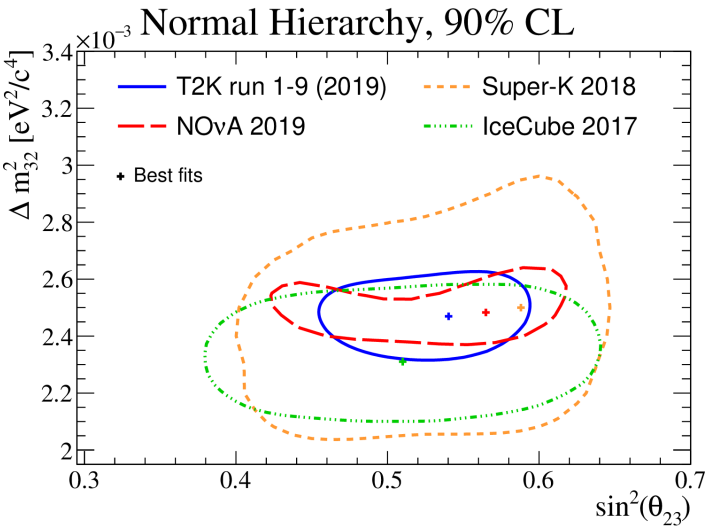


Figure 1.3: The latest comparison of the 90% confidence levels for "atmospheric parameters", Δm^2_{32} against $\sin^2\theta_{23}$ between various atmospheric and long-baseline accelerator neutrino oscillation experiments. This is assuming normal mass hierarchy. Figure taken from [48]

and measure θ_{13} through $\bar{\nu}_e$ disappearance. These measurements are important for this thesis; results from the reactor neutrino experiments are often used to enhance the sensitivity of long-baseline accelerator neutrino experiments by providing an external constraint on θ_{13}. The long baseline accelerator experiments currently do not have a high sensitivity to the values of θ_{13}, so having an external measurements is crucial — especially since θ_{13} has a large impact on the electron neutrino appearance in a muon neutrino beam.

Daya Bay was the first experiment to give an over $5\,\sigma$ evidence of non-zero θ_{13} in 2012 [29], after indications from Double Chooz [28], T2K [26] and MINOS [27]. This was confirmed later that year by the RENO collaboration with $4.9\,\sigma$ significance [30], and finally by T2K with $7.3\,\sigma$ in 2014 [31]. The current world-leading constriants on θ_{13} still come from the reactor neutrino experiments.

KamLAND is a Japanese experiment in the place of the old Kamiokande (which, in turn, was surpassed by Super-Kamiokande), and measures reactor antineutrinos on a long-baseline (among other sources of neutrinos). KamLAND is surrounded by ∼53 nuclear reactors with various baselines averaging to around 180 km. It consists of a balloon filled with 1 kton of liquid scintillator, surrounded by ∼1900 PMTs. The long baseline allowed KamLAND to measure Δm^2_{21} and $\tan^2\theta_{12}$ with the data collected between 2002 and 2005 [32, 33] and shown in figure 1.4.

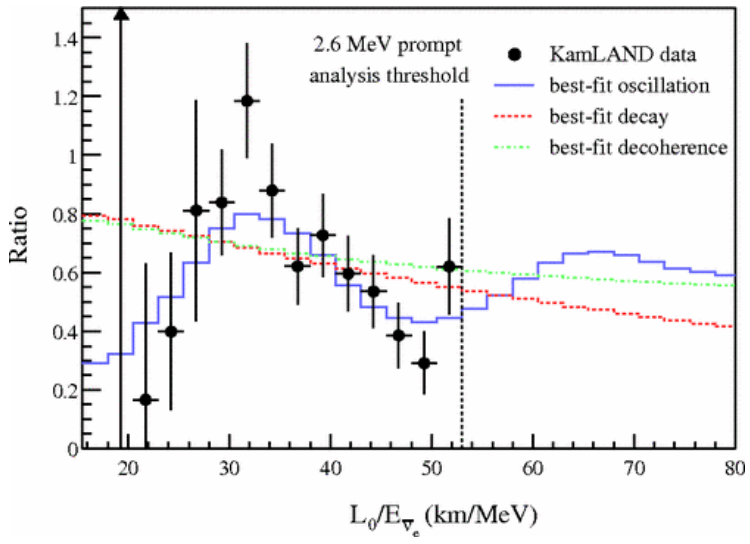


Figure 1.4: The reactor neutrino events seen for different L/E, with the best-fit oscillated spectrum. Figure taken from [33]

Accelerator neutrinos

The first long-baseline accelerator neutrino experiment was based in Japan, the KEK-to-Kamioka (K2K) experiment. The neutrino source was somewhat similar to the atmospheric neutrinos; a beam of 12 GeV protons from proton-synchrotron, incident on a beam-dump, would produce predominantly charged pions which would decay producing mostly muon neutrinos. Path of the positively and negatively charged pions could be diverted with magnetic horns. The magnetic horns generate a toroid-shaped magnetic field, which focus the right-sign and deflect the wrong-sign mesons – the focused $\pi^+(\pi^-)$ then decay into $\nu_\mu(\bar{\nu}_\mu)$ [130]. The neutrino beam would pass through a set of near detectors (including a 1 kton Water-Cherenkov tank) around 300 m away from the beam target, and finally reach the far detector, Super-Kamiokande, 250 km away from the KEK facility. A schematic diagram showing the experimental setup is shown in figure 1.5.

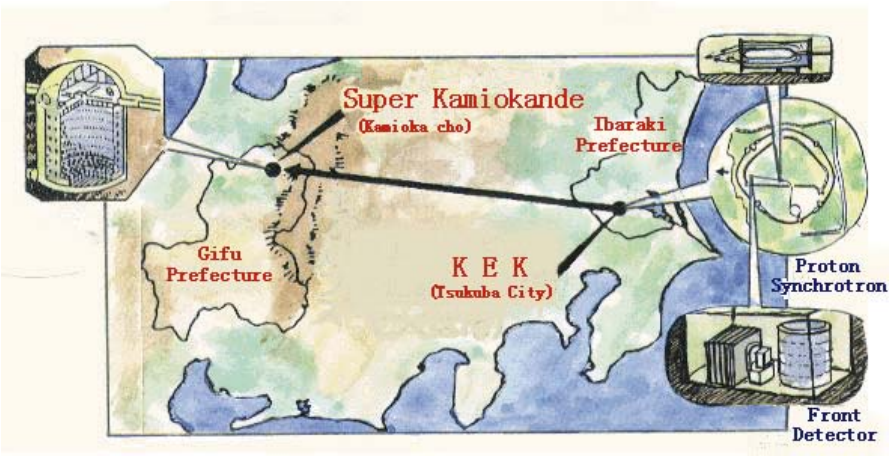


Figure 1.5: A bird's-eye view drawing showing K2K experimental setup. KEK facility was in Tsukuba city, Ibaraki prefecture, where the neutrinos were generated from pion decay with use of proton synchrotron. The neutrinos would travel ∼250 km through soil to finally reach Super-Kamiokande near Kamioka town in Gifu prefecture. Figure taken from [50].

The K2K experiment collected data from 1999 to 2004, detecting 112 accelerator neutrinos with $158^{+9.2}_{-8.6}$ expected under the no-oscillations hypothesis, rejecting it at $4.3\,\sigma$ and confirming the results from atmospheric neutrino experiments [34]. Figure 1.6 shows the data; the number of detected neutrinos per energy bin, with red and blue lines showing the best-fit spectrum to the data and the unoscillated expectation, respectively.

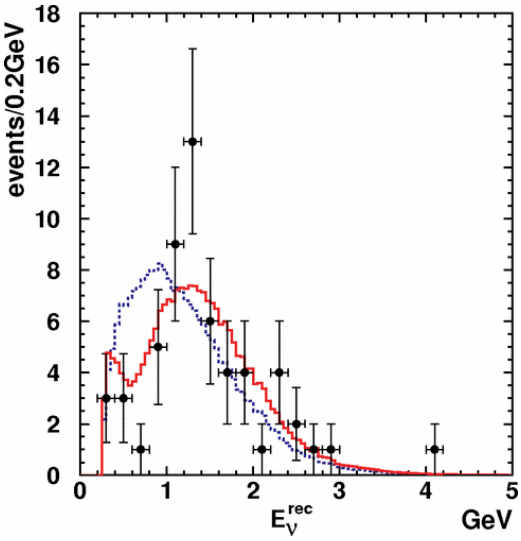


Figure 1.6: Results from the K2K experiment; best fit spectrum (red) on top of the data, together with the unoscillated expectation (blue). Figure taken from [34].

Many accelerator neutrino experiments followed K2K, and all of them had a similar setup; a proton beam incident on a target, generating mesons that would decay into mostly muon neutrinos. One example is MINOS; an experiment with a muon neutrino beam from NuMI beamline in Fermilab, 0.98 kton near detector placed ∼1 km away from the proton beam target, and a 5.4 kton far detector placed 735 km away in northern Minnesota. Both detectors were magnetized, segmented steel-scintillator calorimeters. The experiment, although with different baseline, neutrino energies and detector technology, has also confirmed that muon neutrinos oscillate [35].

There are a few accelerator neutrino experiments currently in operation; this thesis involves one of them, the T2K experiment, a direct successor of the K2K experiment (although with different beamline and near detectors). NOvA is a direct successor of MINOS, still using Fermilab's NuMI beam and with the far detector sitting 810 km from the proton beam target. These experiments, with higher statistics, improved detector technologies and knowledge from the previous experiments have broader goals; to measure δ_{CP}, $\sin^2\theta_{23}$, $|\Delta m^2_{32}|$, the mass hierarchy, and neutrino cross-sections on various targets, and more. They not only measure muon neutrino disappearance, but also electron neutrino appearance in a muon neutrino beam; thanks to the higher statistics with more intense beams and more far detector target material. These experiments will be then succeeded by Hyper-Kamiokande and DUNE, which are both entering the construction

stage.

Other neutrinos

Although not related much to this thesis, there are other sources of neutrinos, which are interest mostly in astrophysics; in these cases neutrinos are usually used as a tool, rather than a direct subject of research. A good example of this is the search for supernovae; star cores generate an enormous burst of neutrinos as they collapse. These neutrinos can reach the Earth before the visible light, and be detected with most large neutrino detectors in a form of an event pileup. This gives a possibly of giving directional information to the wider astronomical community before the event is detectable with light-based telescopes. An example of this is Supernova SN 1978A, detected by Kamiokande-II [36]. IceCube actively probes for very high-energy neutrino sources in a collaboration with cosmic-ray research [37] in a neutrino energy region of many PeV. Neutrinos are also being used for Dark-Matter detection, both as a direct and indirect signal [38]. There are possible applications in the tomography of the Earth with geoneutrinos [39], and even profiling of nuclear reactors [131].

1.6 The Current State of Neutrino Oscillations

Neutrino physics has come a long way since the first observation of neutrinos in 1953. We now know that neutrinos are not massless and they change their flavour as they propagate through space (a.k.a. "neutrino oscillation" described earlier). The evidence comes from a large variety of experiments; ones involving nuclear reactors, huge detectors buried under mountain peaks and Antarctic ice sheets, submerged in lakes and seas, built from gas, solid and liquid. Although the properties of neutrinos are still being measured, these particles are already being used as a tool to search, for example, for supernovae and dark matter.

Although this thesis is only concerned with the neutrino oscillation phenomena via long-baseline accelerator experiments, a large variety of neutrino experiments use neutrinos as a tool for discoveries, and also need to know the PMNS parameters for their

research. Below are the current states of the PMNS parameters, θ_{13}, θ_{23}, θ_{12}, Δm^2_{32}, Δm^2_{21} and δ_{CP}, at the time of writing this thesis;

$\boldsymbol{\theta_{13}}$: Reactor experiments with a km-baseline that measure antineutrino disappearance through $\mathrm{P}(\bar{\nu}_e \rightarrow \bar{\nu}_e)$ are the most sensitive to θ_{13}. It is also possible to measure θ_{13} with accelerator neutrino experiments through $P(\nu_\mu \rightarrow \nu_e)$.

The most recent update on θ_{13} comes from the Particle Data Group (PDG) 2019, which released the average best-fit and confidence levels from a combination of results from three different reactor neutrino experiments: Daya Bay, Reno and Double Chooz, most of which updated their results in 2018. The final θ_{13} constraint is $\sin^2\theta_{13} = 0.0218^{+0.0007}_{-0.0007}$ [18]. This value is still to be used in the coming T2K oscillation analysis, and in this thesis an older value from PDG 2018 is used: $\sin^2\theta_{13} = 0.0212^{+0.0008}_{-0.0008}$.

$\boldsymbol{\theta_{12}}$: Often called a "solar parameter" (together with Δm^2_{21}), can be also measured with reactor neutrinos with over 100 km baseline. Solar neutrino experiments have the best sensitivity to θ_{12} through $\mathrm{P}(\nu_e \rightarrow \nu_{\mu,\tau})$ over a long baseline.

The currently most up-to-date constraint on θ_{12} comes from a 2016 joint fit to all of the solar data from SNO and SK collaborations, together with the measurements from the KamLAND experiment. The value of θ_{13} in this joint fit is constrained with the results from reactor neutrino experiment, described above. The θ_{13}-constrained joint fit gives $\sin^2\theta_{12} = 0.307^{+0.013}_{-0.012}$ [17].

$\boldsymbol{\theta_{23}}$: This is one of the parameters that the long-baseline accelerator neutrino experiments have a large sensitivity via ν_μ disappearance, $P(\nu_\mu \rightarrow \nu_\mu)$, together with the atmospheric neutrino experiments via $\mathrm{P}(\nu_\mu \rightarrow \nu_\tau)$.

The best constraint is currently obtained from the T2K experiment, although the PDG does its own fit for different mass hierarchies and θ_{23} octants given by the T2K, NOvA, MINOS and IceCube DeepCore experiments. The most up-to-date value of θ_{23} from the T2K experiment (possibly not published yet by the time of submission of this thesis) from the Bayesian analysis is $\sin^2\theta_{23} = 0.537^{+0.027}_{-0.036}$ with the reactor experiment constraint on θ_{13}, and is further described in chapter 4.

$\boldsymbol{\Delta m^2_{21}}$: Similarly to θ_{12}, this parameter can be constrained via both solar and reactor neutrino experiments via $\mathrm{P}(\bar{\nu}_e \rightarrow \bar{\nu}_e)$.

The latest update of the PDG official limit on Δm^2_{21} was from the 2013 paper from KamLAND, which performed a fit with θ_{13} constraint from the accelerator (T2K + MINOS) and reactor (Daya Bay, RENO, Double Chooz) neutrino experiments. The data used included solar data and KamLAND's geoneutrino data, to extract $\Delta m^2_{21} = 7.53^{+0.18}_{-0.18} \times 10^{-5}\,\mathrm{eV}^2$ [19]. The more recent analysis from the Super-Kamiokande collaboration, that also included reactor, solar and KamLAND constraint, came out with similar limits of $\Delta m^2_{21} = 7.49^{+0.19}_{-0.18} \times 10^{-5}\,\mathrm{eV}^2$ [17].

$\boldsymbol{\Delta m^2_{32}}$: The Δm^2_{32} mass-splitting term, just like θ_{23}, is best measured with atmospheric and accelerator neutrino experiments. The atmospheric experiments have a wide range of L/E depending on the angle the neutrino came from, whereas the accelerator experiments have a fixed L/E.

The T2K experiment still provides some of the best constraints on Δm^2_{32}, which will continue to get better as we collect more data, especially in ν-beam mode. The most recent constraint from the T2K Bayesian analysis (possibly still not published by the time of submission of this thesis) is $\Delta m^2_{32} = 2.46^{+0.08}_{-0.09} \times 10^{-3}\,\mathrm{eV}^2$ when marginalized over both mass hierarchies, and will be explained more in chapter 4. The PDG 2019 has done its own joint fit by combining results from various accelerator, reactor and atmospheric neutrino experiments to extract values of $\Delta m^2_{32} = 2.444^{+0.034}_{-0.034} \times 10^{-3}\,\mathrm{eV}^2$ for normal mass hierarchy and $\Delta m^2_{32} = -2.55^{+0.04}_{-0.04} \times 10^{-3}\,\mathrm{eV}^2$ for the inverse.

$\boldsymbol{\delta_{\mathrm{CP}}}$: The CP-violating phase is negligible in the ν_e survival probability to the leading order [133], and is currently best probed with the ν_e appearance probability, $P(\nu_\mu \rightarrow \nu_e)$, using atmospheric and accelerator neutrino experiments.

δ_{CP} is still largely unknown; the CP-conserving values, between 0 and π, are only rejected with $\sim 2\,\sigma$ intervals. Constraining the value of δ_{CP} with the T2K experiment is the main topic of this thesis, and at the time of writing the document, the constraint is $\delta_{\mathrm{CP}} = -1.82^{+0.69}_{-0.57}\,\mathrm{rad}$. This is currently the world-best constraint on δ_{CP}, published in

the journal *Nature* [53]. The validations, analysis and results that led to these numbers are described in chapter 4.

Chapter 2

Tokai to Kamioka experiment

This chapter describes the Tokai to Kamioka experiment (T2K), a second-generation long baseline accelerator neutrino experiment based in Japan. It was commissioned in 2009, and started operations in 2010. T2K measures ν_μ and $\bar{\nu}_\mu$ disappearance together with ν_e and $\bar{\nu}_e$ appearance in ν_μ and $\bar{\nu}_\mu$ beams, through the $P(\nu_\mu \rightarrow \nu_\mu)$, $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_\mu)$, $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ oscillation probabilities. A proton synchrotron in the Japan Proton Accelerator Research Center (J-PARC) in Tokai (東海村) provides a proton beam at 30 GeV, which is directed onto a graphite target with five kicker magnets. The collisions on that target result in charged mesons (predominantly pions and kaons) which produce neutrinos as they decay whilst they propagate through a ∼96 m decay volume, which is very similar the K2K experiment described in section 1.5. As the charged mesons leave the graphite target, a set of magnetic horns bend them to either divert positively or negatively charged particles towards the tunnel; the so-called Forward-Horn-Current (FHC) results in a beam dominated by ν, and the Reverse-Horn-Current (RHC) results in a $\bar{\nu}$-enhanced beam. The decaying pion branching ratios strongly favour

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \text{ and } \pi^- \rightarrow \mu^- + \bar{\nu}_\mu,$$

although there is still ∼1.2% contamination of ν_e through

$$\pi^+ \rightarrow e^+ + \nu_e \text{ and } \pi^- \rightarrow e^- \bar{\nu}_e,$$

and much smaller contamination from the decaying muons. Downstream from the decay volume, the beam passes through a beam dump, muon monitors, soil, and a set of near

detectors. The two main near detectors are the INGRID (Interactive Neutrino GRID), which center is aligned with the beam center, and the ND280, which is positioned off the beam axis center, both positioned around 280 metres away from the target station. Section 2.1.2 will discuss the off-axis effect further. The neutrino beam then travels around 295 km (and down to 1.7 km underground) before it reaches the Super-Kamiokande (SK). SK is a 50 kton water Cherenkov detector located in the Kamioka mine with a 1 km of a mountain overburden, shielding it from cosmic rays. A schematic drawing showing the T2K setup is shown in figure 2.1.

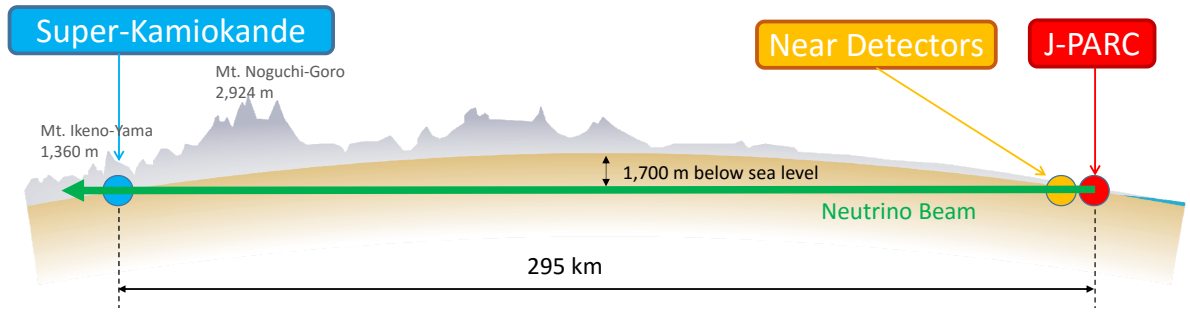


Figure 2.1: A schematic showing the T2K cross-section view of the neutrino beam travelling from the J-PARC laboratory through the near detectors to the Super-Kamiokande detector.

The main goals of the T2K experiment are:

1. Search for Charge-Parity violation in the neutrino sector (δ_{CP})
2. Measurement of Δm^2_{23} and $\sin^2\theta_{23}$ through ν_μ ($\bar{\nu}_\mu$) disappearance in a ν_μ ($\bar{\nu}_\mu$) beam
3. Measurement of $\sin^2\theta_{13}$ through the appearance of ν_e ($\bar{\nu}_e$) in a ν_μ ($\bar{\nu}_\mu$) beam
4. Measurement of the neutrino cross-sections with the near detectors

With the Hyper-Kamiokande proto-collaboration entering the construction stage, both T2K and SK are also being used as a testing ground for crucial Hyper-K R&D. A few inter-collaborations have emerged over the recent years; there is a continuous effort to produce joint analyses using the accelerator neutrinos from T2K and atmospheric neutrinos from SK, and another effort to produce joint results from the T2K and NOvA experiments. Although these joint efforts can increase the sensitivity to the PMNS parameters and be

a good validation check, they could also provide us with crucial pioneering techniques and solutions for the future experiments like Hyper-Kamiokande (which could potentially have atmospheric-accelerator neutrino analysis as its default), and any potential future Hyper-Kamiokande - DUNE joint efforts.

2.1 T2K Beamline

2.1.1 Beam

The beamline, built specifically for the T2K experiment, consists of three accelerators that together produce a high-intensity proton beam; a linear accelerator (LINAC), Rapid Cycling Synchrotron (RCS), and the Main Ring (MR). First, a H^- beam is accelerated in the LINAC to around 400 MeV. The hydrogen beam is stripped of electrons with carbon stripping foils before entering the RCS, which further accelerates the remaining protons to around 3 GeV with a 25 Hz cycle and two bunches in each cycle. The protons then enter the last acceleration stage, the MR, where they are accelerated to ∼30 GeV. A schematic showing the MR accelerator is shown in figure 2.2.

Protons from the MR are extracted into the neutrino beamline, shown in figure 2.3, with five kicker magnets. Each proton extraction is defined as one "spill", with eight proton bunches per spill separated by 500 ns. The spill information is sent via a GPS system to the far detector to trigger the T2K data-taking window. This is very important, helping not only with reducing, say, cosmic background in the T2K data, but also to reduce the accelerator-neutrino background in the Super-Kamiokande atmospheric data.

Each spill extracted from the MR into the neutrino beamline is directed onto a proton target, a graphite rod that is 91.4 cm long and 2.6 cm in diameter. The graphite rod is cooled down with gas helium to prevent overheating from the proton spills. The high energy proton-graphite interactions result in high quantities of charged mesons, which first pass through the three magnetic horns operating at ∼250 kA. The magnetic horns are made of inner and outer coaxial conductors that focus either positively or negatively charged mesons; Forward-Horn-Current (FHC, ν-mode) and Reverse-Horn-Current

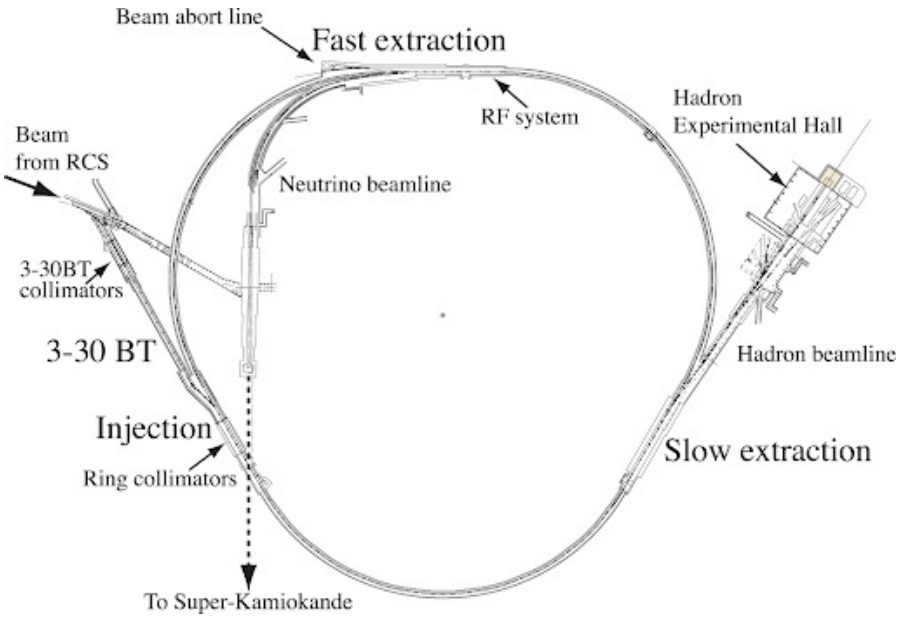
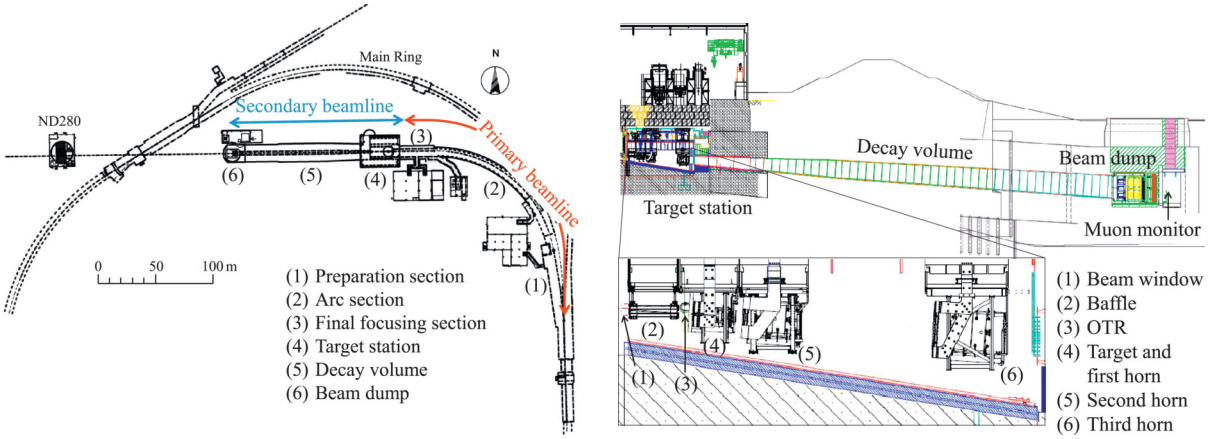


Figure 2.2: Birds-eye schematic view of the MR, with the proton beam shared between various experimental areas – including to the Neutrino Beamline, used by the T2K experiment. Figure taken from [134]



(a) Birds-eye view

(b) Side view

Figure 2.3: Schematic drawings showing the neutrino beamline, with the target station, magnetic horns, decay volume and the beam dump. Figure taken from [59].

(RHC, $\bar{\nu}$-mode), respectively. The focused mesons then travel through the 96 m decay volume, where they will decay into either neutrinos (with FHC mode) or anti-neutrinos (when in RHC mode), via;

$$\begin{array}{lcl} \pi^+ & \rightarrow & \mu^+ + \nu_\mu, \\ K^+ & \rightarrow & \mu^+ + \nu_\mu, \end{array} \quad \text{(FHC)} \qquad \begin{array}{lcl} \pi^- & \rightarrow & \mu^- + \bar{\nu}_\mu, \\ K^- & \rightarrow & \mu^- + \bar{\nu}_\mu. \end{array} \quad \text{(RHC)}$$

There is a non-negligible electron-neutrino contamination in the T2K muon-neutrino beam. The contamination at the peak energy of 0.6 GeV comes mostly from muons that decay into electrons and electron-neutrinos (before the muons reach the beam dump), with a small contribution form mesons decaying into electron-neutrino flavoured leptons. At energies above 1.5 GeV, most of the contamination comes from charged and neutral kaons decaying into pions and electron-flavoured leptons. These relevant decay modes are;

$$\begin{aligned} K^+ &\rightarrow \pi^0 + e^+ + \nu_e, \\ K^0_L &\rightarrow \pi^- + e^+ + \nu_e, \end{aligned}$$

and

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu.$$

The wrong-sign contamination is especially prominent in the RHC mode. Since the neutrinos are produced from a proton (rather than anti-proton) beam, we expect more positively charged mesons, resulting in more neutrinos than anti-neutrinos being created. It therefore makes sense that there is a higher neutrino background in anti-neutrino beam than the vice-versa. This is shown in figure 2.4, where we see the unoscillated beam composition for FHC and RHC modes as seen at ND280.

The beam dump sits at the end of the decay volume, stopping any hadrons that did not decay, and muons below ∼5 GeV. Muons above this threshold will mostly pass through the beam dump, and propagate through the muon monitors that were installed just behind it, around 120 m away from the carbon target. The muon monitors (MUMON) are made of an array of ionization chambers, and arrays of silicon PIN photodiodes, as shown in figure 2.5. Since the muons are produced alongside neutrinos from the π^+ two-body decay, the muon beam center aligns together with the neutrino beam center. Monitoring the distribution profile of the muon beam allows a better positioning and alignment of the neutrino beam direction and intensity. The resolution of the detector is

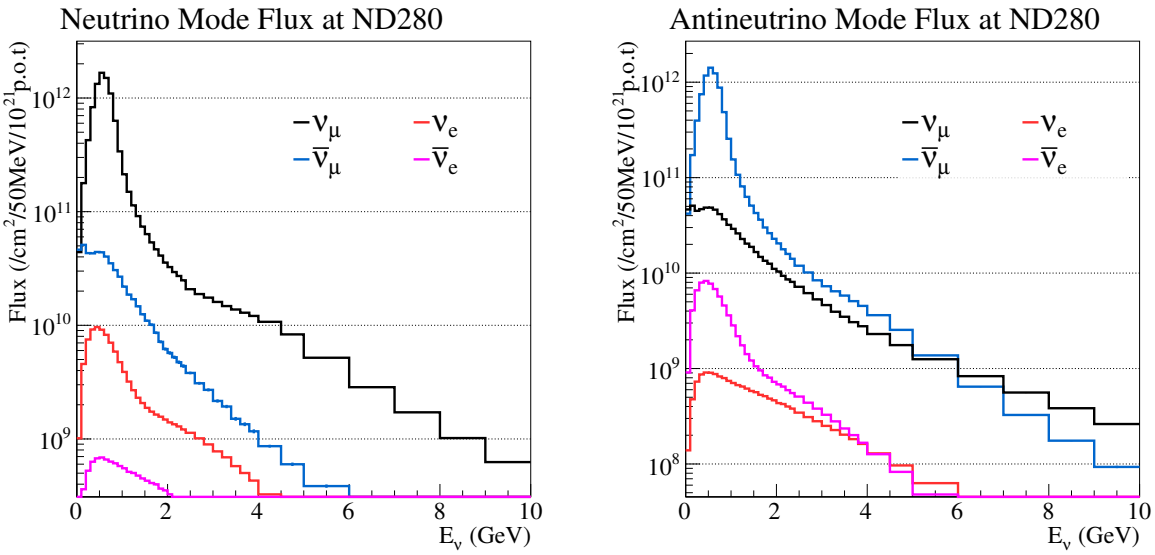


Figure 2.4: Neutrino beam composition before oscillations, as seen at the off-axis ND280 detector for FHC (left) and RHC (right) beam modes. The wrong-sign background is especially prominent in the RHC mode at higher neutrino energies. Figure taken from [49].

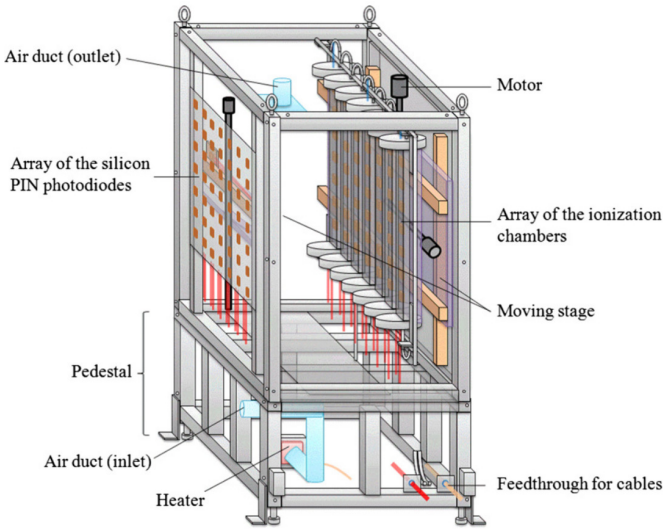


Figure 2.5: Muon monitors (MUMON) measures the neutrino beam profile and direction by looking at muons with energy high enough to pass the beam dump (higher than ∼5 GeV). Figure taken from [55].

around 3 cm on the muon beam profile, giving a precision on the neutrino beam direction of less than 0.25 mrad.

2.1.2 Beam off-axis angle

T2K is the first accelerator-neutrino experiment to target the neutrino beam off-axis from the far detector. The use of the off-axis offset for the neutrino beam to the far detector was first proposed at the AGS in Brookhaven [56] and then for the T2K experiment specifically [57]. The T2K far detector – beam center misalignment narrows down the neutrino energy spectrum band around the oscillation maxima (maximizing $P(\nu_\mu \rightarrow \nu_e)$ and minimizing $P(\nu_\mu \rightarrow \nu_\mu)$), and reduces backgrounds from the π^0 interactions. The off-axis angle used in the T2K beamline is 2.5° set in 2007 using results from other experiments.

If we consider the two-body pion decay $\pi^+(\pi^-) \rightarrow \mu^+(\mu^-) + \nu_\mu(\bar{\nu}_\mu)$, the outgoing neutrino energy, E_ν, can be approximated in terms of the initial charged pion energy, E_π, pion and muon masses, m_μ and m_π, and the initial pion – outgoing neutrino angle θ;

$$E_\nu = \frac{m_\pi^2 - m_\mu^2}{2(E_\pi - p_\mu \cos\theta)}. \tag{2.1}$$

If we select a specific angle θ, we will find that the neutrino energy is maximized when the pion energy is $E_\pi = p_\pi / \cos\theta$, giving the maximum neutrino energy per angle;

$$E_\nu^{\max} = \frac{m_\pi^2 - m_\mu^2}{2(E_\pi \sin^2\theta)}, \tag{2.2}$$

where the only real dependency for the neutrino energy peak is the angle between the incoming pions and outgoing neutrino. As the angle deviates from the collinearity between the pion and lepton, the range of neutrino energies becomes smaller and centered around $E_\nu^{\max}$. With the off-axis angle of 2.5°, the T2K neutrino energy is sharply peaked around 0.6 GeV. Given that the T2K baseline is 295 km, the 0.6 GeV energy peak minimizes the muon neutrino survival, $P(\nu_\mu \rightarrow \nu_\mu)$, whilst maximizing the electron-neutrino appearance, $P(\nu_\mu \rightarrow \nu_e)$. This is also shown in figure 2.6, where the predicted neutrino energy distribution at SK is aligned with $P(\nu_\mu \rightarrow \nu_\mu)$ and $P(\nu_\mu \rightarrow \nu_e)$, calculated over a range of neutrino energies.

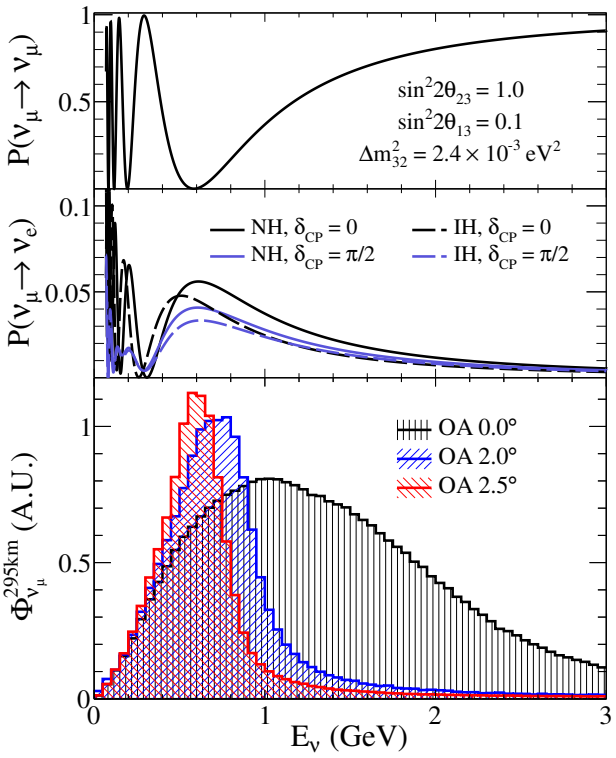


Figure 2.6: The relative neutrino flux energy bands for an on-axis (black) and two off-axis (blue and red) beams as seen at the SK detector, with the offset angle of 2.5° being currently used. Figure taken from [58].

2.2 Interactive Neutrino GRID (INGRID)

The on-axis Interactive Neutrino GRID detector is designed to measure the neutrino beam direction with a precision better than 1 mrad, by measuring the ν_μ beam profile. This is crucial for an off-axis experiment, where a small uncertainty on the neutrino beam angle can result in a large systematic uncertainty. The spacial width of the neutrino beam at 280 m from the proton beam target is 5 m, with the INGRID x–y dimensions being 10 m × 10 m [42]. The actual measured beam direction stability is well within the 1 mrad design target; ∼0.2 mrad [41]. The detector is made of of 16 separate 1.2 m × 1.2 m × 1.3 m cubic modules. Fourteen of these modules are positioned in a cross pattern around the beam centre, with two extra diagonal modules that are currently switched off. A schematic model of INGRID detector is shown in figure 2.7.

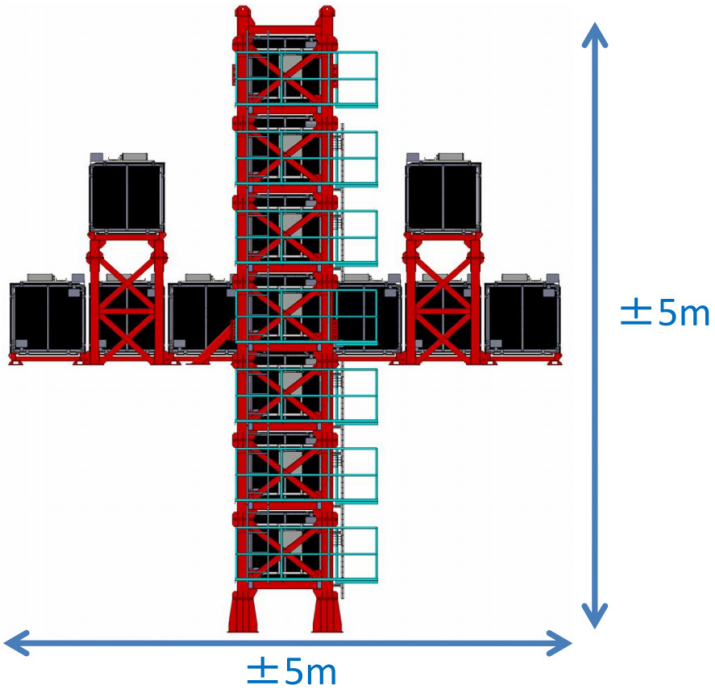


Figure 2.7: The on-axis INGRID near detector, 280 m from the carbon beam target. It measures the beam direction and profile, and contributes to the cross-section measurements. It is made of 16 cubic modules in a cross formation. Figure taken from [42].

Each module is made of 11 scintillator planes interleaved with 9 iron plates with a target mass of 7.1 tonnes, enough to monitor the beam stability on day to day basis. This is surrounded by scintillator veto plates, that help with rejection of charged particles

originating from the outside of the modules. The scintillator planes are made of 24 vertical scintillating bars, attached to 24 horizontal scintillating bars. Figure 2.8 shows the exploded view on INGRID module.

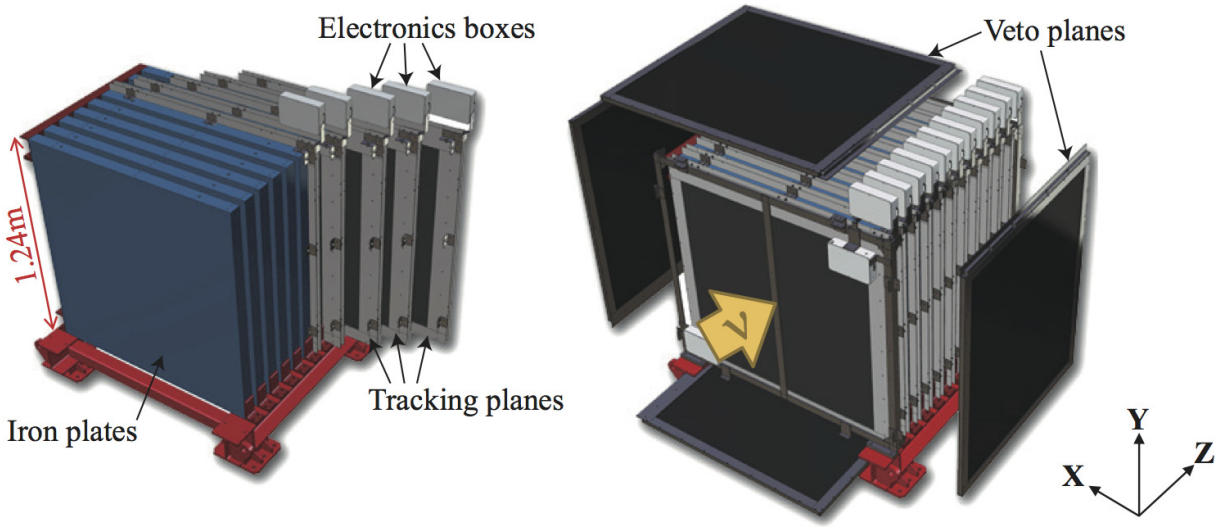


Figure 2.8: One INGRID module with its structure shown. It is composed of layers of iron plates (for target material) and scintillators for tracking. 16 of these modules form the INGRID near detector. Figure taken from [42].

2.3 ND280 detector

The primary T2K near detector, ND280, is positioned 2.5° off-axis, around 280 m away from the carbon target. Data from ND280 is used in the T2K Oscillation Analysis directly to constrain the neutrino flux and cross-section uncertainties.

The neutrino flux uncertainties are constrained by measuring the ν_μ spectrum before the neutrino oscillations, with the higher event rates thanks to the proximity to the carbon target. The ν_e spectrum is also measured, and included in the analysis as the irreducible background contamination in the beam coming from the kaon decay (this is especially important in measuring the ν_e and $\bar{\nu}_e$ appearance in the far detector). The rates of various neutrino interaction types are used to measure the neutrino cross-sections, for example, the Neutral Current interactions with outgoing π^0 are particularly important ν_e background at Super-Kamiokande. Although neutrino cross-sections on water are

most important for constraining interactions at Super-Kamiokande, ND280 also takes measurements on carbon and iron.

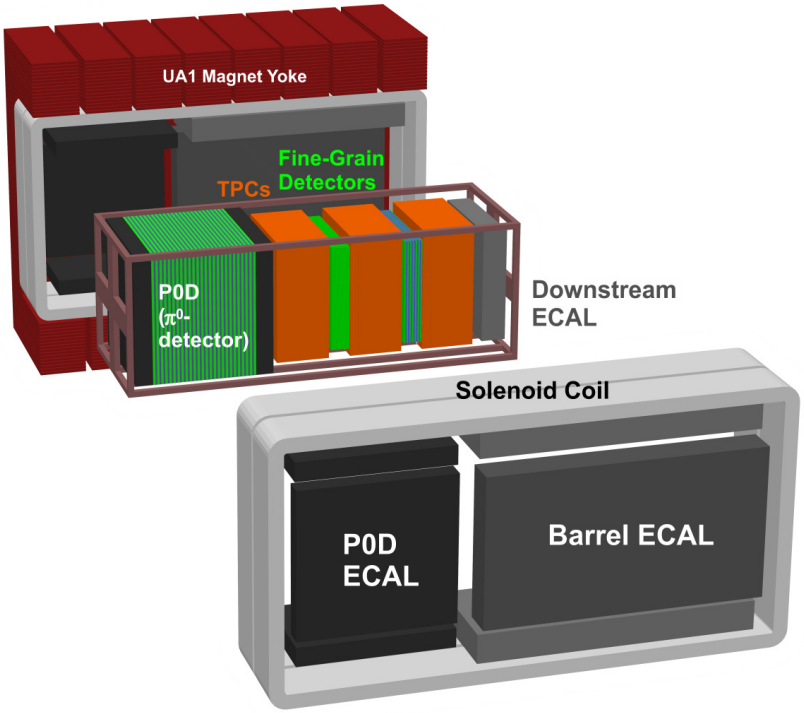


Figure 2.9: Exploded view of the off-axis ND280 detector with all the sub-detectors labelled. ND280 is crucial in the T2K oscillation analysis, drastically reducing the systematic uncertainties at Super-Kamiokande. Figure taken from [59].

ND280 is made of multiple sub-detectors to archive the goals above. A dedicated π^0 detector, PØD, sits on the upstream side of the detector, followed by the tracker. The tracker is made of three Time Projection Chambers and two Fine Grained Detectors. The tracker is surrounded by a set of electromagnetic calorimeters (ECal), and finally, by a magnet yoke, instrumented with Side Muon Range Detector (SMRD). The detector schematic diagram is shown in figure 2.9.

2.3.1 Magnet

All the sub-detectors are surrounded by a magnet repurposed from the UA1/NOMAD experiment at CERN, to measure the momenta of charged particles generated by the neutrino interactions happening inside of the ND280 [44]. The magnet's operating current

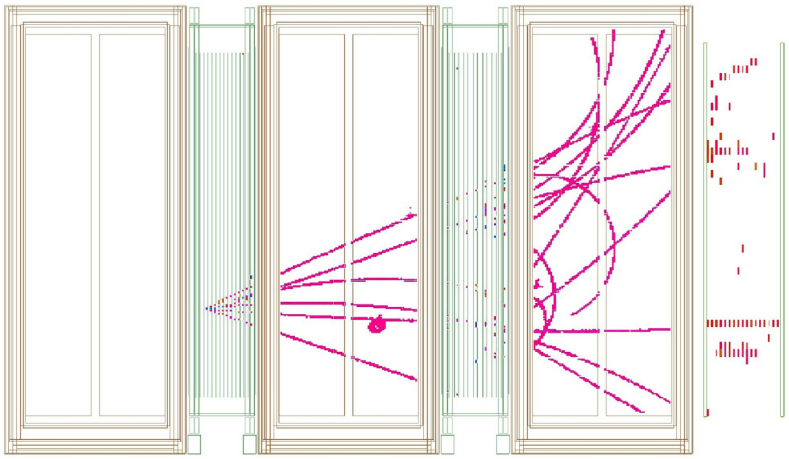


Figure 2.10: Deep inelastic scattering event originating from FGD1. The event transverses through and triggers most of the tracker, including both FGDs, two TPCs and ECal. Figure taken from [45].

is 2900 mA, generating a uniform horizontal magnetic field of 0.2 T.

The Magnetic Yoke is instrumented with layers of scintillators interleaved in the magnet air gaps, to form the SMRD. The purpose of SMRD is to measure the ranges of muons exiting the detector, and to veto any muonic interactions originating from within the walls of ND280. The whole magnet is made of two separate C-shaped parts, allowing it to be opened for the inner-detector access.

2.3.2 PØD

The π^0 detector (PØD), with its surrounding ECal, is the first ND280 sub-detector downstream from the neutrino beam. It measures the cross-sections of neutrino interactions producing neutral pions, NC1π^0, which is one of the largest background in the $P(\nu_\mu \rightarrow \nu_e$) oscillation analysis at Super-Kamiokande.

PØD is a 2.5 m cube that consists of panels of removable water target, needed to constraint the interactions at Super-Kamiokande. These panels are interleaved with brass and scintillator panes. The surrounding ECal is made of triangular scintillator plates and lead sheets. The total PØD target mass is 15.8 tonnes, including all the water target.

2.3.3 Fine Grained Detectors (FGDs)

The two Fine Grained Detectors, FGD1 and FGD2, are part of the tracker, which is positioned downstream after PØD and also consists of three Time Projection Chambers [45]. FGDs provide the ND280 with a target mass for neutrino interactions, and track charged particles that come from the interaction vertex. FGD1 is made of 186.4 cm long scintillator bars, orientated alternatingly in x and y directions perpendicular to the neutrino beam with 192 bars in each direction making one module. FGD1 has 15 of these modules, whereas FGD2 contains 7 modules interleaved with 6 water layers, 2.5 cm thick (the first water layer is drained due to an accident with the water-cooling system). The FGDs weigh 1.1 tonnes each, allowing sufficiently high neutrino interaction event rate. The cross-sections are measured from interactions on both water (oxygen) and plastic (carbon).

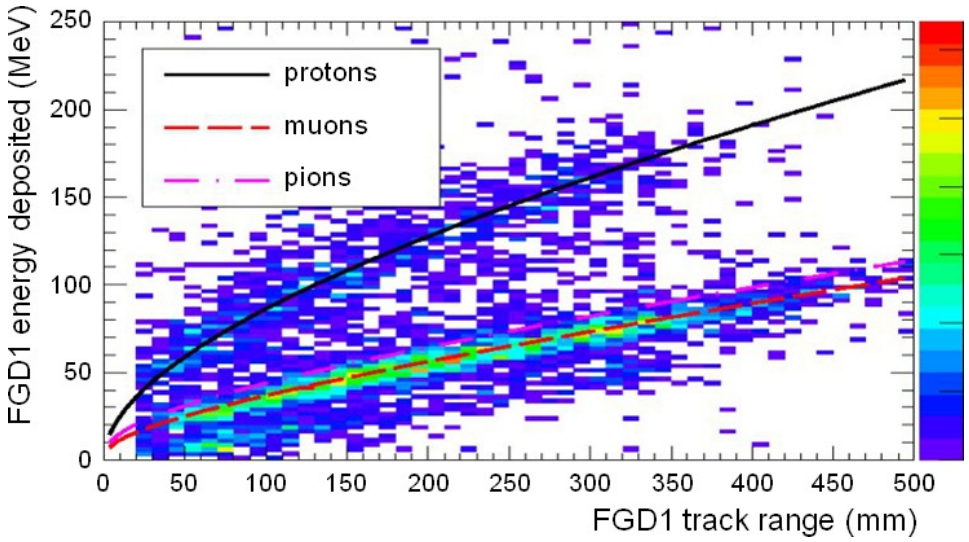


Figure 2.11: Distribution of the deposited energy along FGD1 for stopping particles, with the MC expectation for protons, muons and pions drawn. FGDs can easily distinguish protons from muons by looking at their energy loss distributions. Figure taken from [45].

The FGDs alone are capable of distinguishing muons from protons by looking at the reconstructed energy from the scintillating bars along the sub-detector, and comparing to the Monte Carlo expectation, as shown in figure 2.11. This is partially due to the fact that the recoil protons are short-lived, and usually fully contained within the FGD. The fine grained nature of the FGDs make it is easier to fully reconstruct such low-momentum

fully-contained protons.

2.3.4 Time Projection Chambers (TPC)

The are three identical TPCs within the tracker, with 3D tracking and energy reconstruction capabilities. Since the sub-detectors are operating in a magnetic field, good particle identification and momentum reconstruction are possible.

The TPCs are made of an inner and outer boxes, as shown in figure 2.12, with the inner box forming a field cage containing argon gas. The argon gas ionizes as the charged particles from neutrino interactions transverse through. The ionized electrons drift towards the bulk micromegas detectors, which both amplify and measure the accumulated charge. There are 24 micromegas detectors per TPC, aligned in a way to minimize the dead-zones.

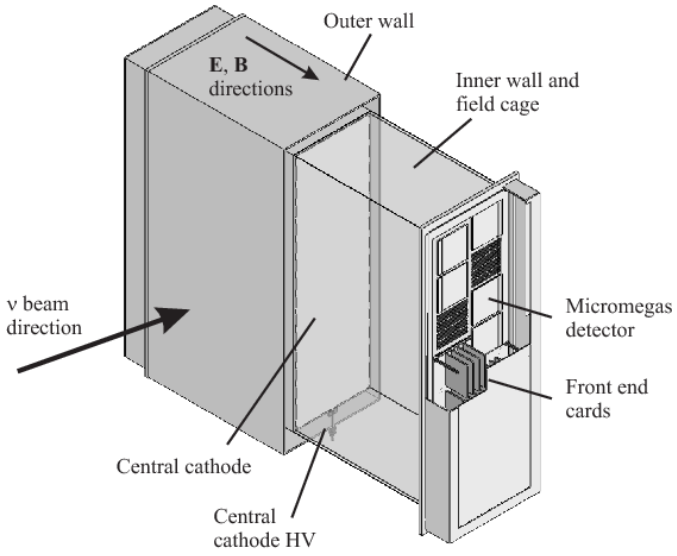


Figure 2.12: A cut schematic showing a time projection chamber for T2K. All three time projection chambers are identical. Figure taken from [45].

TPCs, just like the FGDs, have an excellent particle identity capabilities. This is done using the particle deposited energy per unit distance (energy loss) as a function of the reconstructed particle momentum. This is shown on figure 2.13, where the distribution of the energy loss is shown as a function of momenta, with MC prediction shown on top.

The archived resolution on the energy loss is 7.8±0.2%, allowing the distinction between electrons and muons.

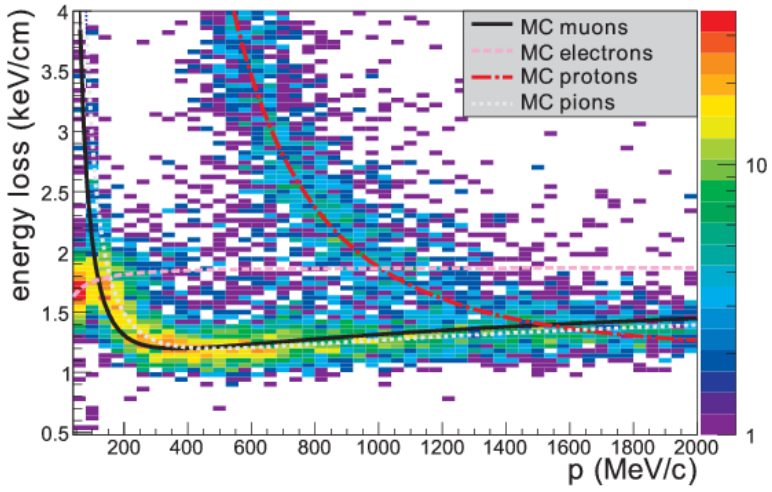


Figure 2.13: Distribution of the deposited energy against the particle momenta, for one of the TPCs, for positively charged particles from neutrino interactions. The MC expectation is drawn on top of the data. TPCs are capable of differentiating muons from electrons. Figure taken from [45].

2.3.5 Electronic Calorimeters (ECal)

Both PØD and the tracker are surrounded by the Electronic Calorimeters, which was the main contribution from the UK for the ND280 commissioning [46]. Each ECals module is made of scintillator bars sandwiched between lead sheets.

ECals can be split into three groups; PØD ECal, barrel ECal (surrounding the tracker) and the downstream ECal (positioned downstream from the tracker). ECals complement the event reconstruction by detecting photon showers with their energies and charged particles for their identification. One of main functions is to reconstruct the neutral pions from neutrino interactions, with 50% of photon shower from π^0 being captured.

2.4 Far Detector: Super-Kamiokande

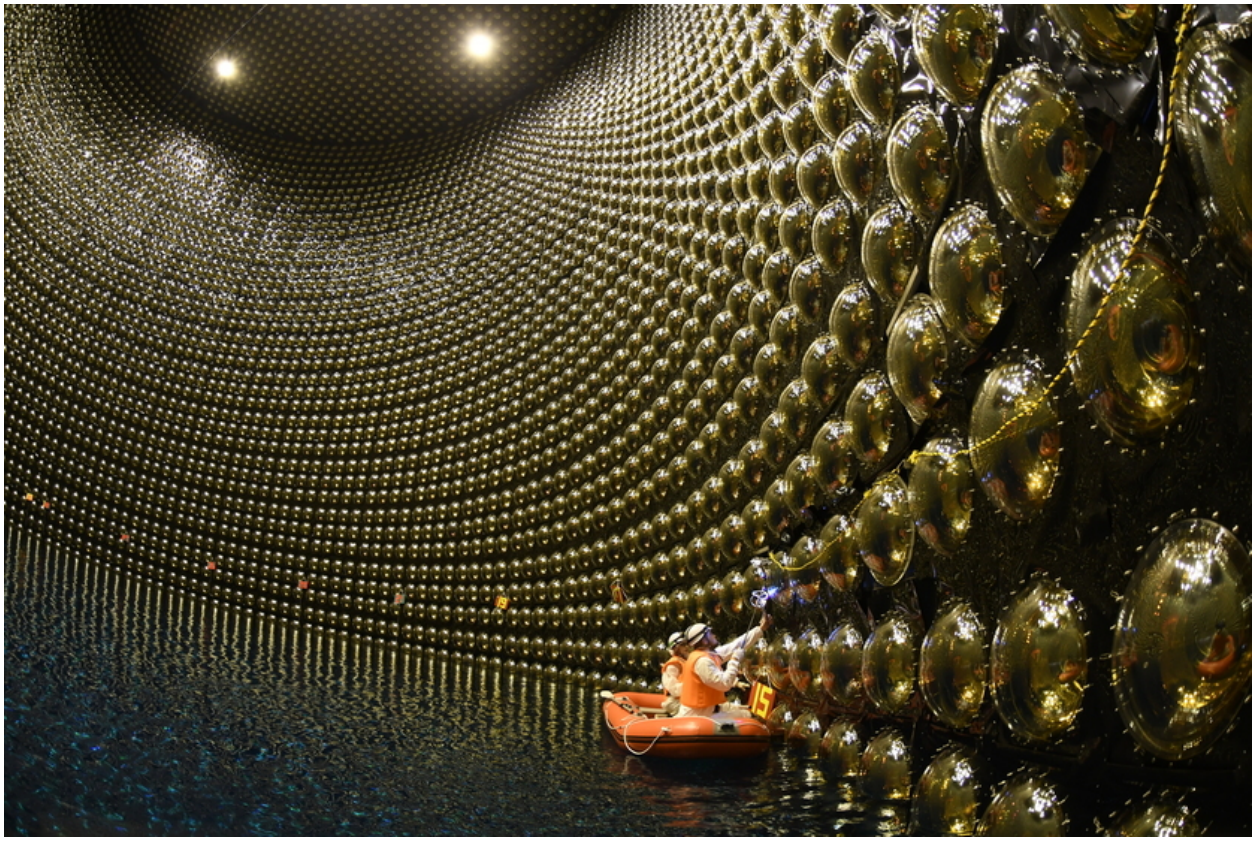


Figure 2.14: Inside of the Super-Kamiokande tank during the 2018 refurbishment work, with myself and Lauren Anthony taking PMT dynode direction measurements. Around three-quarters of water was drained by the time of taking that picture. Photograph taken by M. Nakahata.

Super-Kamiokande (SK) [47] is a successor of Kamiokande, previously described in 1.5. SK is a large water-cherenkov detector, and has been an experiment in its own right since 1996. It is being used as the far-detector for the T2K experiment since 2010, with 50 kton of ultra-pure water (27.2 kton of fiducial volume) as the target material. It is placed in the Kamioka mine in the Do (土) village, in the far north of the Gifu prefecture. Being inside the base of the Ikenoyama mountain, SK has roughly about 1 km of soil overburden, shielding it from cosmic muons.

The detector is split into two optically-separated sections, shown in figure 2.15. The outer-detector (OD) contains ∼1,900 outward-looking PMTs to veto the interactions originating outside of the detector. The inner-detector (ID) contains ∼11,000 inward-

looking PMTs to detect the cherenkov-light from the charged particles transferring the inner volume. The tank is 41 m tall with 39 m diameter, and the inner detector is 36 m tall with the diameter of 34 m. Originally built for proton-decay search, it is now used for a variety of studies, including neutrinos originating from solar, atmospheric, geological, reactor, dark-matter and accelerator sources.

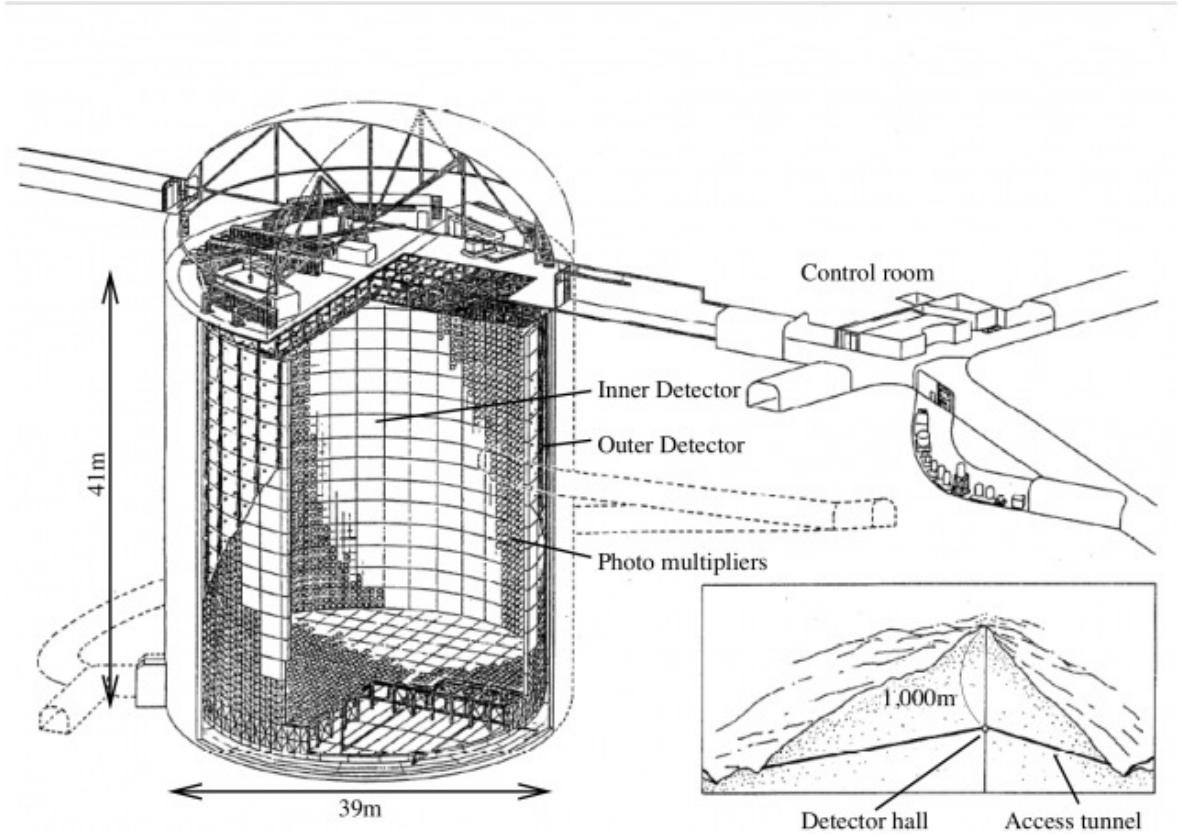


Figure 2.15: Schematic diagram showing the Super-Kamiokande tank with the inner and outer detectors. The inner and outer detectors are optically separated with black sheet (inside the inner) and white (inside the outer) Tyvek material. Figure taken from [60].

Super-Kamiokande underwent a large refurbishment in the summer 2018, with the prime goal of fixing known leaks in the tank. There is a plan for doping the SK water with Gadolinium to tag neutrons from neutrino interactions. Given the not-fully understood effect of Gadolinium on human body, and a large historical catastrophe related with leaks in mines in the area, it was crucial to fully seal the leaks in SK before the doping can start. The first tank opening since 2013 also gave an opportunity to do some general maintenance and install new calibrations systems.

2.4.1 Cherenkov Radiation

Super-Kamiokande detects the cherenkov light emitted by relativistic charged particles produced by neutrino interacting with the water target. As a charged particle is traversing with relativistic speed through a medium with refractive index n, it will emit electromagnetic shock wave if the particle speed speed is

$$v > \frac{c}{n}, \tag{2.3}$$

where c is the speed of light in vacuum. The light will be emitted along the particle's trajectory with the opening angle;

$$\cos\theta_C = \frac{1}{\beta n}, \tag{2.4}$$

also referred as "cherenkov angle", where $\beta = v/c$. Since refractive index of water n is 1.34 and β is effectively 1 at relativistic speed, the cherenkov angle in water is ~42° (schematic drawing of cherenkov radiation shown in figure 2.16). The required momentum to reach the cherekov threshold in water medium for an electron is 0.57 MeV/c, and 118 MeV/c for a muon.

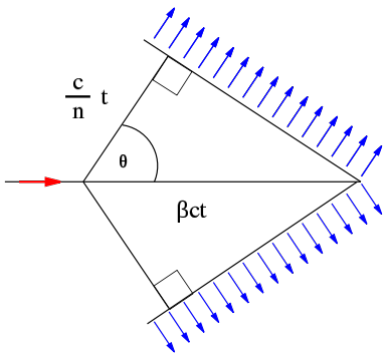
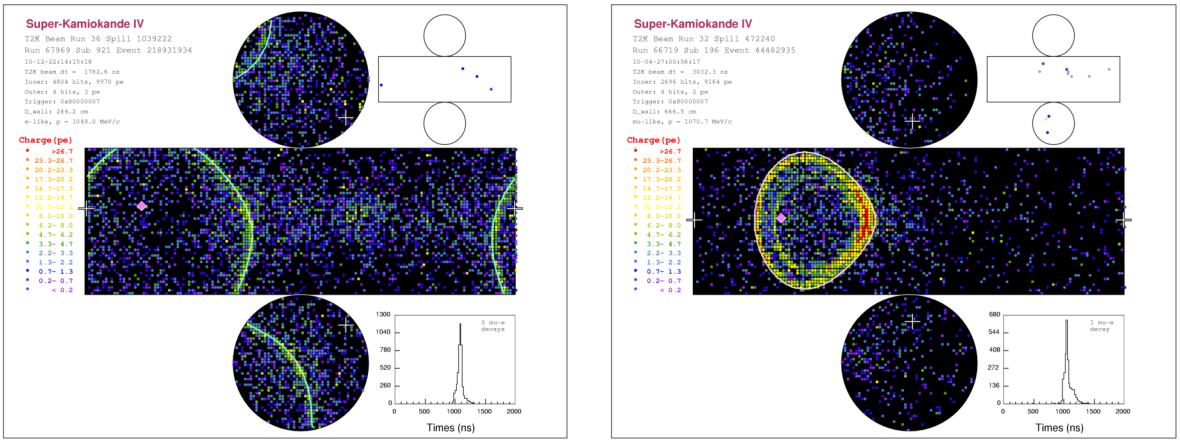


Figure 2.16: Diagram showing cherenkov radiation. Particle with relativistic speed βv propagates through with a refractive index n, emitting cherenkov light at angle θ. Figure created by A. Horvath

Electrons, being very light, scatter as they transverse through water, and often cause electromagnetic showers. The result of this are blurred rings, with no well-defined contours. Muons, being much heavier, do not scatter much. They still loose energy and therefore progressively emit lest photons as they transverse, resulting in a rather unique

signal; muon cherenkov rings have a sharp outer edge, and blurred inside edge (due to the energy loss). Both electron and muon signals are shown in figure 2.17.



(a) Electron-like ring (b) Muon-like ring

Figure 2.17: Super-Kamiokande unrolled event display with examples of muon and electron-like events. The electron-like events have a more diffused signal, due to electron interacting more as it travels. Muon-like events have sharp cherenkov-ring outer boundary, with diffused inner part due to energy loss as the particle transverses through the detector.

Chapter 3

Bayesian Statistics and Markov Chain Monte Carlo methods

This thesis describes different analyses of the ND280 and SK data using Bayesian Markov Chain Monte Carlo (MCMC) methods. The output of these analyses are the Bayesian posterior density distributions, which define the probability distributions across the parameter space – effectively tell us about the probability of the sampled parameters having some specific values across the parameter's range. The terms "posterior probability" and "posterior density" will be used interchangeably. To get an estimate of the parameters of interest, we need to integrate the probability distribution over all the nuisance parameters (i.e. the systematic parameters). This marginalization process is very commonly used in various forms. It is common to 'scan over' the likelihood of some parameter of interest, where for every 'scan point' we would generate many random throws of the nuisance parameters, and average out the likelihoods for these throws. By providing the full posterior density over all the parameters, MCMC methods essentially marginalize over all the parameters with respect to each other. A pedagogical description which is aligned to the treatment used here is available in [64]. Although this is a computationally expensive process, it allows for many additional studies, with some T2K-specific ones described in chapters 4 and 5.

This chapter will briefly describe the Bayesian formalism in the context of the T2K experiment, and follow up with more detailed descriptions of various MCMC algorithms

used, together with their efficiency-tuning methods, and some possible improvements.

3.1 Bayes Theorem for T2K

In the Bayesian formalism, the probability is interpreted as level of our belief of a hypothesis being true. This level of belief is proportional not only on the probability of observing the data, but on our prior knowledge of the model as well (the words 'model' and 'hypothesis' will be used interchangeably). In the full Bayes theorem, the posterior probability is written as

$$P(H|\{D\},I) = \frac{P(\{D\}|H,I) \times P(H|I)}{P(\{D\}|I)}, \tag{3.1}$$

where $P(H|\{D\},I)$ is the posterior probability distribution; our belief that the hypothesis H, given a dataset $\{D\}$, and any additional information I, is true. $P(\{D\}|H,I)$ is the probability of observing the data given our hypothesis, and is normally referred to as the "likelihood". $P(H|I)$ is our prior information about the hypothesis H given any extra information I, whereas $P(\{D\}|I)$ is the prior probability of the dataset. The latter is usually interpreted as the normalization term, that makes the posterior probability distribution sum up to 1, and is often re-written as the integral of the numerator with respect to the hypothesis. This is usually difficult to compute, and very often not useful; from the computational point of view, the joint posterior distribution can be normalized after the "fit", and from the experimental point of view, the data was already observed; it is constant, and therefore the term is 1. An additional motivation for this will be given in section 3.2. Bayes' theorem can therefore be rewritten simply as

$$P(H|\{D\},I) \propto P(\{D\}|H,I) \times P(H|I). \tag{3.2}$$

In T2K, the hypothesis H is the predicted number of neutrino events observed at the near and far detectors. This is constructed from various parameters defining; the near/far detector efficiencies, neutrino flux uncertainties, the probability of a neutrino interacting (cross-sections) and finally, the oscillation probabilities given the PMNS model.

The likelihood is often simply written as $\exp(-\chi^2/2)$, where the T2K negative log-

likelihood, with the ND280 and SK data split into energy bins, can be written;

$$\begin{aligned} -\ln(\mathcal{L}) = & \sum_{i}^{ND280bins} N_i^{\mathrm{ND,p}} - N_i^{\mathrm{ND,d}} + N_i^{\mathrm{ND,d}} ln[N_i^{\mathrm{ND,d}}/N_i^{\mathrm{ND,p}}] \\ & + \sum_{i}^{SKbins} N_i^{\mathrm{SK,p}} - N_i^{\mathrm{SK,d}} + N_i^{\mathrm{SK,d}} ln[N_i^{\mathrm{SK,d}}/N_i^{\mathrm{SK,p}}], \end{aligned} \tag{3.3}$$

where $N_i^{\mathrm{ND,p}}$ and $N_i^{\mathrm{ND,d}}$ are the ND280 predicted number of events and the data for the i^{th} energy bin, respectively, with $N_i^{\mathrm{SK,p}}$ and $N_i^{\mathrm{SK,d}}$ being the SK counterparts [61, 62]. The ND280 predicted number of events per energy bin, $N_i^{\mathrm{ND,p}}$, is dependent on the neutrino flux uncertainties, the ND280 detector uncertainties and the neutrino cross-sections. The SK prediction counterpart, $N_i^{\mathrm{SK,p}}$, is additionally dependent on the neutrino oscillation probabilities and the SK detector uncertainties[1]. The correlations between ND280 and SK data are taken into account by the fact that the predicted number of events for both detectors are generated using shared flux and cross-section systematic parameter values.

In the T2K experiment, there are few "groups" of parameters forming the "hypothesis" model. These parameters correspond to the usual formula for the observed accelerator neutrino event rate,

$$\mathrm{N} = \Phi \times \sigma \times P(\nu_\alpha \rightarrow \nu_\beta) \times \epsilon \tag{3.4}$$

where Φ is the neutrino beam flux, σ are the neutrino cross-sections, and ϵ are the detector efficiencies. In T2K, all these sets of parameters have their own working groups that are are responsible for developing the model, and progressing it onto the analysis - partially in the form of covariance matrices, with the correlated uncertainties for each parameter. These correlation matrices form Gaussian priors in equation 3.2;

- Flux: 100 parameters
- ND280 detector: 500 parameters (250 per FGD)
- Cross-section: ∼32 parameters
- SK detector: 44 parameters

[1] Of course, the SK prediction is not dependent on the ND280 detector uncertainties directly. These efficiencies will, however, have an effect on the cross-section and flux uncertainties, indirectly influencing the prediction at SK.

Of course, the six oscillation parameters have their priors too. Although these are provided in the analysis in the form of a covariance matrix too, and the overall treatment is the same as the nuisance parameters, δ_{CP}, $\sin^2\theta_{13}$, $\sin^2\theta_{23}$ and Δm^2_{32} have flat priors (i.e. always equal to 1). With these covariances, the negative log-prior probability is given by

$$-\ln(P(H|I)) = \sum_{p}^{covariance} \frac{1}{2} \sum_{i}^{p} \sum_{j}^{p} \Delta p_i (V_p^{-1})_{i,j} \Delta p_j. \tag{3.5}$$

The posterior probability in T2K is therefore given by the sum of 3.3 and 3.5, and, confusingly, is always referred to as "the likelihood". We will assume this terminology for the rest of this thesis.

3.2 Markov Chain Monte Carlo

There are various ways of marginalization; integrating over the systematic parameters to obtain the parameters of interst. One of these is through Markov Chain Monte Carlo (MCMC), which generates a distribution of "samples" across the n-dimensional parameter space, where the density of these samples corresponds to the posterior probability distribution.

Sampling from the hypothesis takes the form of a guided random walk; the MCMC algorithm "steps" around the parameter-space, with the probability of stepping in a certain direction being dependent on the likelihoods at different points in the parameter-space. Each step is saved in a sequence, or *chain*, generating the distribution of samples. There are many different types of MCMC, but they all have shared properties once enough steps have been made such that the likelihood is well-represented by the distribution;

1. **Stationarity**: the samples' density distribution must be stationary. This means, after some number of MCMC steps, the produced posterior density does not change with subsequent MCMC steps. We say that the outcome posterior distribution is stationary (or in equilibrium). In practice, it means a MCMC chain has reached the favourable part of the phase-space, and generated enough MCMC samples across it.

2. **Reversibility**: The MCMC chain is reversible in time, meaning the probability of obtaining certain density distribution at time t is the same as at time t'. This is only possible if the stationary state has been reached.

The above conditions imply that once the MCMC chain has reached the proximity of the underlying stationary distribution (see "Burn-in", section 3.2.2), all the subsequent steps sample from the same posterior distribution. Great care must be taken to satisfy this "detailed balance" when implementing a MCMC algorithm.

3.2.1 Metropolis-Hastings MCMC

The T2K Bayesian analysis framework uses the Metropolis-Hastings MCMC [65] algorithm, which is initialized by arbitrarily choosing an initial point on the parameter-space as the first step in the Markov Chain. The algorithm then loops over a few operations, with the number of iterations corresponding to the desired length of the Markov Chain. The following tasks are performed at each iteration (MCMC step):

1. **Propose a new step** on the parameter space. If this is the first step, parameters can be randomized or thrown from their priors. Otherwise, each parameter p_i is deviated from its current value;

$$p_i(t+1) = p_i(t) + f(p_i), \tag{3.6}$$

where is $f(p_i)$ the step-size, usually a function of the parameter itself, for example, a random Gaussian number taken from the parameter's prior uncertainty, multiplied by user-determined scaling constant. This is done sequentially for all the parameters.

2. **Evaluate the likelihoods** for both the current, $\vec{p}(t)$, and the proposed, $\vec{p}(t+1)$, parameter vectors. We will call them $\ln(\mathcal{L}_{\text{old}})$ and $\ln(\mathcal{L}_{\text{new}})$.

3. **Compare the likelihoods** against each other. This is done differently in various MCMC algorithms. In Metropolis-Hastings it is

$$A = \min[1, \exp\{\ln(\mathcal{L}_{\text{old}}) - \ln(\mathcal{L}_{\text{new}})\}] \tag{3.7}$$

where A is the probability of accepting the step, which is always between 0 and 1.

4. **Accept or reject the step**, based on the probability A. This is done by throwing a random number, R, to see what the next MCMC step will be.

   - If $A \geq R$, the proposed step, $p_i(t+1)$, is appended to the end of the chain.
   - If $A < R$, the current step, $p_i(t)$, is appended to the end of the chain.

   This means that the chain, whether the new step is accepted or not, is always getting longer; when the proposed step is rejected, the old parameter values are repeated and appended to the end of the chain.

This has a few interesting implications on how the MCMC works. Since we calculate the probability of accepting the new parameter values based on the $\ln(\mathcal{L})$ difference ($\mathcal{L}$ ratio), it means there is a chance of accepting a step that is in a less-favourable area of parameter-space. This is very important; MCMC is not a minimizer, it is a sampler that steps around the parameter-space, reconstructing the underlying posterior density; whether likely or unlikely.

Furthermore, when MCMC steps are rejected, the parameter values are repeated in the chain; so in a way, the parameter-space is not being-explored at that instance. On the other end of the extreme, if we accept all the steps, the MCMC sampler performs a true random-walk, which is highly inefficient. There is a certain balance needed for the optimal phase-space exploration. The larger the step-size, the larger the $\ln(\mathcal{L})$ difference in equation 3.15, rejecting more steps. The smaller the step-size, the smaller the $\ln(\mathcal{L})$ difference, accepting more steps. There are a lot of debates in the literature about the perfect acceptance ratio for the most efficient phase-space exploration, with the most widely accepted described in [63].

As the number of dimensions (parameters free in the sampler/fit) is approaching infinity, the perfect acceptance ratio is found to be 0.234. Although this number is the asymptotic limit, it is close even down to 5 dimensions (for a hypothesis with 1 dimension, the perfect acceptance ratio is 0.44) [64].

An example from a toy study is shown on the next page. Chains that are well mixing, with step-sizes that are too small and too large are shown in figures 3.1, 3.2 and

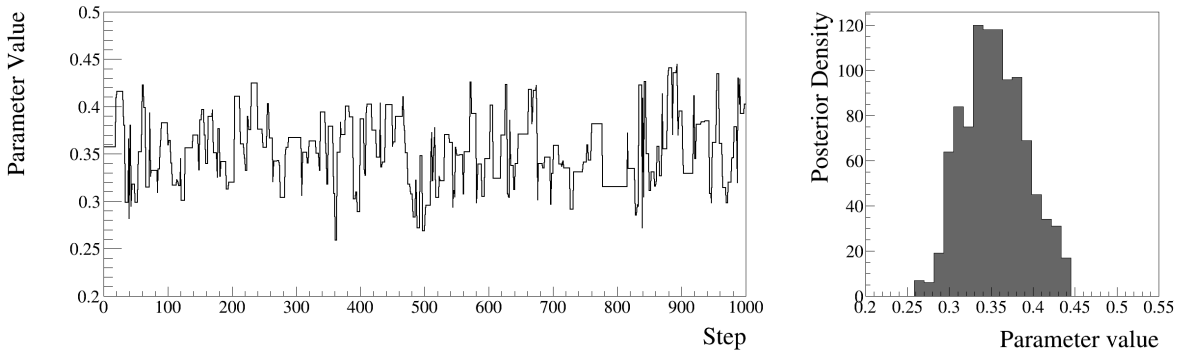


Figure 3.1: Trace (left) and the binned posterior density (right) from a well-mixing MCMC chain, with acceptance ratio ∼0.23.

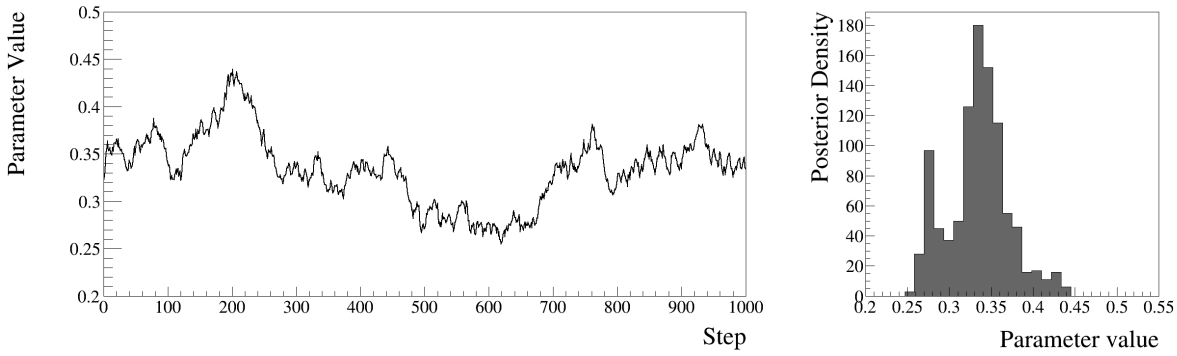


Figure 3.2: Trace (left) and the binned posterior density (right) from a badly-mixing MCMC chain, with too small step-sizes and a high acceptance ratio.

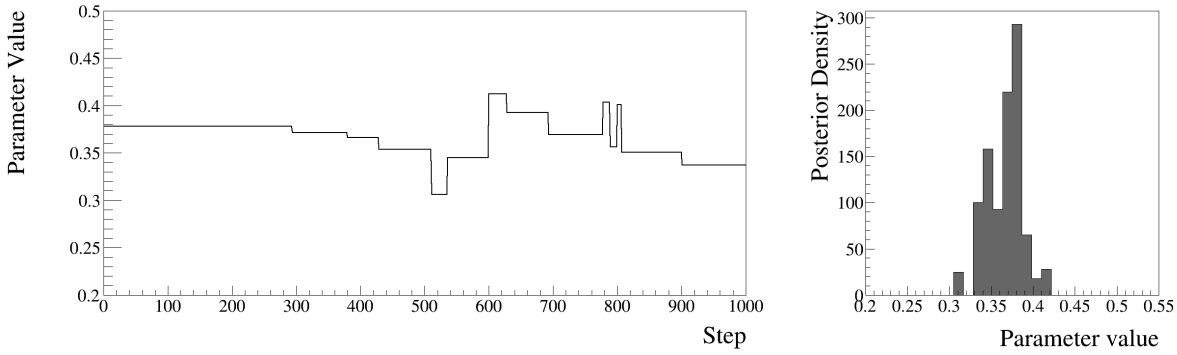


Figure 3.3: Trace (left) and the binned posterior density (right) from a badly-mixing MCMC chain, with too high step-sizes and a low acceptance ratio.

3.3, respectively. The parameter values from each point on these chains are binned to get the posterior densities (left hand side on each plot). It is easy to see that the chain with an acceptance ratio of ∼0.23 produces the best results. Crucially, this only affects the efficiency of the sampler; if we were to run each one of them sufficiently long time,

the posteriors from all three chains would converge to a satisfactory, stationary posterior distributions.

We can now also see why the denominator in the equation 3.1 is not important; first, we can easily normalize the posterior after the sampling (simply divide each bin by the total area of the posterior), and, since our prior knowledge on the dataset does not change step-to-step (i.e. the data is fixed), this term would cancel-out in equation 3.15 anyway.

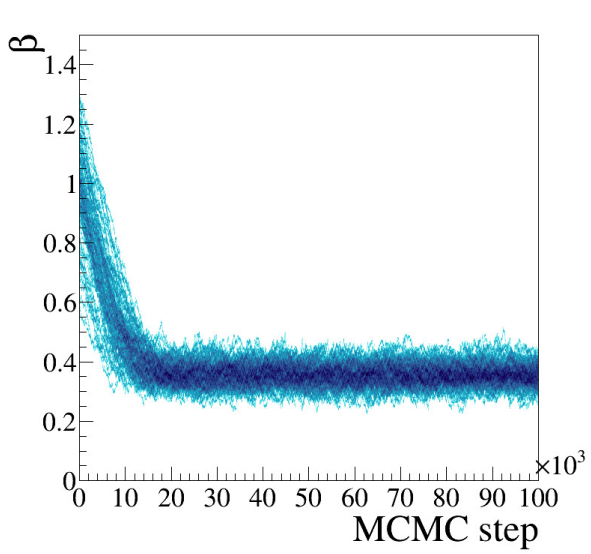
3.2.2 Burn-in

When the MCMC sampling begins, the initial parameter values are usually far from optimal, and often far from the true stationary posterior distribution. As the MCMC sampling progresses, the produced steps are first slowly moving towards the high-density area, before the stationary distribution can be sampled effectively. This stage of chain-converging is called burn-in, where the initial MCMC steps are often discarded; they do not become part of the posterior distribution.

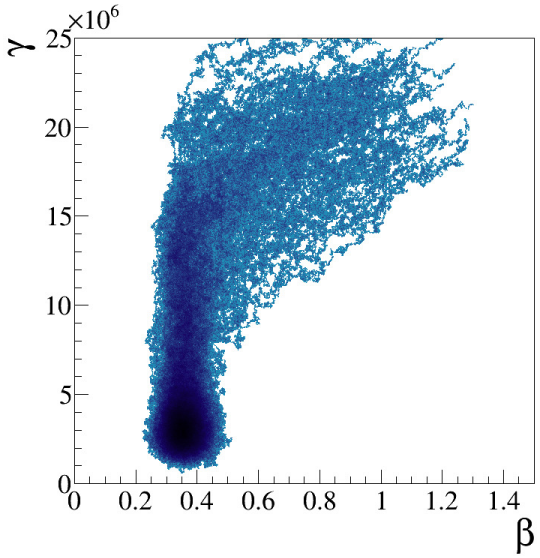
Figure 3.4 shows parameter variation per step (trace) in 3.4a, where it can be clearly seen that the burn-in lasts around ∼30,000 MCMC steps. The sub-figure 3.4b shows two parameters projected onto each other, where the burn-in seems to be different for each parameter. This poses an interesting challenge; since the selected burn-in means removing the initial steps in all dimensions, a burn-in value that works for all the parameters needs to be chosen. In practice a few methods are commonly used; manually scan through all the parameters and select the worst-case burn-in, use log-likelihood traces for selecting the burn-in, automatically discarding, say, 50% of the chain regardless of its length, or manual check if the resultant posterior distribution is stable when adding more steps.

3.2.3 Autocorrelations

The derivation of the optimal acceptance ratio in [63] has two important assumptions; first, the stationary distribution is a multivariate Gaussian. This is not always true;



(a) MCMC traces for some parameter, from 100 chains with a burn-in included. We see that the chain is converging in the stationary-distribution area after ∼30,000 MCMC steps.



(b) MCMC traces for two parameters projected onto each other, with the burn-in included. It is interesting to note that both parameters do not have the same burn-ins

Figure 3.4: Parameter traces produced from 100 MCMC chains in a toy framework, with the burn-in included. A special care has to be taken when when selecting the burn-in to make sure no steps from outside of the stationary distribution are included; usually it is better to over-estimate the burn-in.

some of the T2K cross-section parameters are very non-Gaussian, are not symmetric and have local maxima in the posterior. The second assumption is that the step-sizes for all the parameters are proportional to the variance of the full stationary distribution. The second assumption has large implications for complex models with many parameters, like in the T2K likelihood.

It is rare to have the full knowledge of the stationary distribution before running the MCMC analysis (if that was the case, we would not need to analyse anything to start with). Optimizing step-sizes for models with many correlated parameters, with respect to each other, is usually very difficult, and often based on trial-and-error. One important

metric to make this job easier is called "autocorrelation":

$$A_k = \frac{\sum_{i=1}^{N-k}(X_i - \bar{X})(X_{i+k} - \bar{X})}{\sum_{i=1}^{N}(X_i - \bar{X})^2}, \tag{3.8}$$

which measures how correlated parameter X is with the delayed copy of itself, X_k. This, calculated as a function of delay itself, tells us of how correlated the consecutive MCMC steps are with each other, and effectively how many MCMC steps we need to take to draw a random sample from the posterior. Even with the perfect acceptance ratio of 0.234, if the autocorrelations for various model parameters are very different from each other, then they will not explore the phase-space with the same efficiency.

Figure 3.5 shows the autocorrelation dropping to 0 after a lag of 500. That means we can obtain a fully independent random value for that parameter, drawn from the stationary posterior distribution, every 500 MCMC steps. This can also help us to determine the number of MCMC steps needed for a satisfactory stationary posterior distribution. Let us say we want at least 10,000 fully independent samples distributed across the parameter-space, with the density of samples representing the posterior distribution. With a lag of 500 from figure 3.5, we know that we need the MCMC chain to be at least five million steps long.

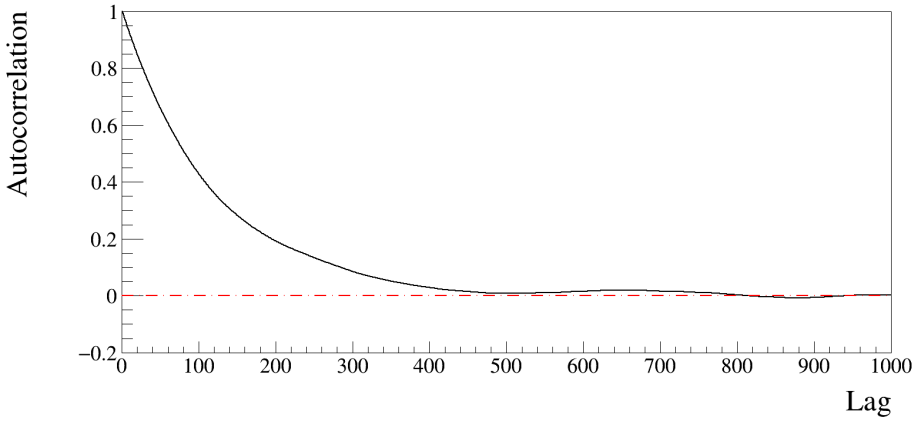


Figure 3.5: Example of an autocorrelation plot. It shows how many MCMC steps we need to effectively draw a random sample from the stationary posterior distribution; value called "lag". In this example, autocorrelation is 0 at lag $\sim$ 500, so we draw a completely random value from the posterior distribution every 500 MCMC steps.

When tuning the step-sizes for each parameter, the most efficient method is often

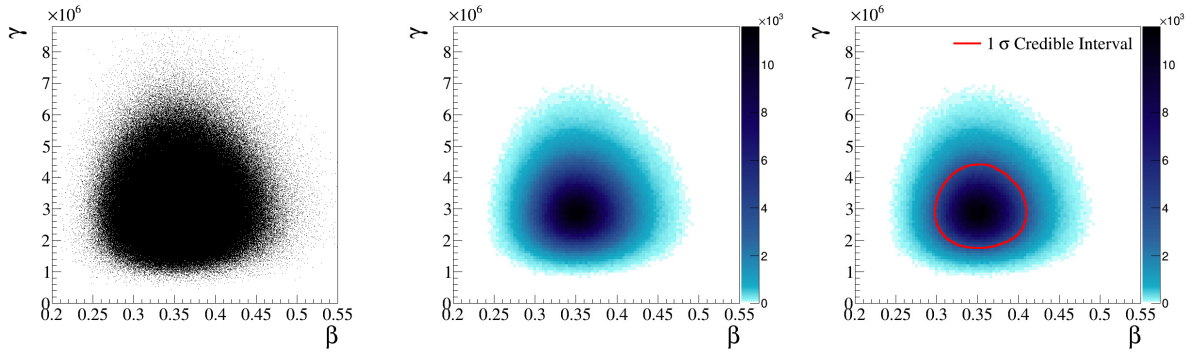
to first ensure that the autocorrelations are the same (or, more realistically, as similar to each other as possible), before moving onto trying to get the perfect acceptance ratio by tweaking all step-sizes equally with one multiplier.

3.2.4 Credible Intervals

Credible intervals are bounds on parameters which reflect our level of belief in their values. For example, $1\,\sigma$ credible interval interval on some parameter means that we have 68.23% belief that the parameter is within those bounds. It is rather easy to place such bounds given the binned posterior distribution; we sequentially add the contents of each posterior bin from the most to the least populous, until the integral of the selected bins, divided by the total integral, reaches the $1\,\sigma$ level. The principle is exactly the same with the 2D intervals, with the progression of creating the credible intervals shown in figure 3.6.

The Frequentist confidence intervals have a different construction; a χ^2 distribution for a parameter of interest is obtained through marginalization, and the intervals are selected by drawing a line across, for example, $\chi^2 = 1$ for $1\,\sigma$ given one degree of freedom. The confidence intervals are then the two points where that line crosses the χ^2 distribution. The exact χ^2 levels for different degrees of freedom per confidence interval are provided by the PDG in Table 39.2 of the Statistics section [18].

The interpretation of the confidence intervals is very different from the Bayesian credible intervals; it is, however, possible to convert the Bayesian intervals into Frequentist, if extra validations and checks are needed.



(a) Example of MCMC steps for two parameters projected onto each other. The density of the samples corresponds to the posterior density.

(b) The projected steps are binned to produce the binned posterior density - which is far easier to interpret than 3.6a.

(c) Credible intervals are made when the integral of progressively selected bins, divided by the total integral, reaches the desired interval.

Figure 3.6: The progress of making 2D credible intervals from a MCMC chain. Steps from two parameters are projected onto each other first, as in 3.6a, and then binned, showed in 3.6b, to produce the binned posterior density. The 68% Credible Intervals, shown in 3.6c, are produced by iterating from the highest bin to the lowest, halting when the integral between the selected bins and full posterior reaches the required interval (for example, $1\,\sigma$).

3.3 Markov Chains for 3 flavour analysis (MaCh3)

T2K has three far-detector neutrino oscillation analysis groups; two of them, VALOR (named after the neutrino groups that created it, Valencia-Lancaster-Oxford-Rutheford) and P-Theta (named after the kinematic parameter the fit is done in), use a combination of Frequentist and Bayesian approaches, whereas the third group, MaCh3, is fully Bayesian and uses MCMC to sample both ND280 and the SK data simultaneously. There is a fourth group, BANFF (Beam and ND280 flux extrapolation task force), which performs a Minuit fit to the ND280 data to constrain the neutrino cross-sections and flux. Results from BANFF are then propagated to VALOR and P-Theta in a form of covariance matrix, to be used in their fits for marginalization over the systematic parameters.

Work for this thesis was done using the Bayesian package, MaCh3. The principles of how it works are described in 3.2, with a few nuances specific to the T2K analysis described in this section, in 4.4.3 and 4.4.6.

3.3.1 Step-sizes

MaCh3 deals with both ND280 and SK data; meaning that both detectors' systematics have to be taken into the account, resulting in ∼750 parameters in the model predicting the neutrino event rates at SK. Some of these parameters are highly correlated (for example, the normalization flux parameters, anti-correlations between flux and cross-section parameters), and some of these parameters are not very well-predicted prior to the fit (i.e. the data has a big impact on these, say, the cross-section parameters).

Although this usually poses a challenge when it comes to tuning the MCMC chain, the flux and SK detector systematics are well modelled prior to the fit (for example, through calibration data), and we therefore have strong priors - in the form of covariance matrices per detector. These matrices are - very successfully - used for making sure all the parameters have good step-sizes with respect to each other. This cannot be said about the cross-section parameters; although they have physics-motivated priors and step-sizes, the data pulls them far away from their nominals and constrain them differently from each other. This makes the tuning difficult, and the "covariance" trick is not sufficient.

All the cross-section parameters therefore have their own own step-size dials, that are tuned with preliminary MCMC runs and autocorrelation checks.

It is quite similar for the oscillation parameters. Although some (for example, Δm^2_{21}, $\sin^2\theta_{12}$) have strong priors that the data does not alter, others often need tuning on a parameter-by-parameter base.

Since many parameters are well-described by their priors in the form of a correlation matrix, it is possible to further optimize the step-sizes with Cholesky decomposition. Decomposing the matrix allows more efficient exploration exploration of the phase-space, by transforming in into an uncorrelated space. This is done by splitting the matrix

$$M = LL^T, \tag{3.9}$$

where the lower triangular matrix, L, multiplied by its transpose, returns the original matrix M.

3.3.2 Model Selection, Δm^2_{32} and $\sin^2\theta_{23}$

The T2K experiment has some sensitivity to the mass hierarchy (the sign of Δm^2_{32}) and the octant of $\sin^2\theta_{23}$ (below or above 0.5) through the $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ detection channels. It is therefore interesting to look at methods of Model Selection; to see which hypothesis is more probable. The Bayesian formalism and the nature of how MCMC works make it rather easy.

Bayesian formalism defines the Bayes Factors, which tell us how many times one model is more probable than the other. Using equaton 3.2, we can write

$$\frac{P(H_1|\{D\},I)}{P(H_2|\{D\},I)} = \frac{P(\{D\}|H_1,I)}{P(\{D\}|H_2,I)} \times \frac{P(H_1|I)}{P(H_2|I)}, \tag{3.10}$$

where the middle term defines the Bayes Factor (with the other ones usually called "posterior odds" and "prior odds"). The Bayes Factor, is therefore the ratio of the likelihood densities for both hypotheses, if the full posterior is marginalized over all these hypotheses.

In MaCh3, the marginalization over both mass hierarchies (positive and negative values of Δm^2_{32}) is done by placing a 50% chance for proposing a step with the opposite

sign on Δm^2_{32}. The value of 50%, is not a prior; it just defines the probability of throwing the "mass hierarchy" step, and is chosen arbitrarily. If the probabilities of proposing negative from positive, and from positive to negative Δm^2_{32} are equal, whether 20, 50 or 80%, then the prior is non-informative.

Given the proper marginalization over all the values of Δm^2_{32}, calculating the Bayes Factors is a post-sampling task. We can simply count the number of MCMC steps in the normal hierarchy, and divide by the number of steps in the inverted hierarchy to get the Bayes Factor for NH/IH. Likewise for the octant of $\sin^2\theta_{23}$; count the number of steps where $\sin^2_{23} > 0.5$ and divide by the number of steps in $\sin^2_{23} < 0.5$ to get a Bayes Factor for Upper/Lower octant.

3.4 Alternative MCMC algorithms for MaCh3

The Metropolis-Hastings algorithm works in T2K very well indeed; the analyses described in chapters 4 and 5 are both completed using the MaCh3 framework. The analysis described in chapter 4 required MCMC chains long enough to achieve contour stability at the $3\,\sigma$ level; resulting in an MCMC chain with 100 million steps. Each step requires a new $\ln(\mathcal{L})$ evaluation, which involves reweighing the ND280 and SK Monte Carlo prediction to the new parametrization, taking anywhere between 0.6–2 seconds per reweight. The analysis was running on roughly 50 NVIDIA graphic cards, 600 CPU cores and 800 Gb RAM for around one month.

As T2K adds more parameters to the hypothesis, and collects more data (and data samples) to analyse, the computing burden will only become larger. There are a few ways of speeding things up from the MaCh3 side, including;

1. Optimize the step-sizes better; reduce the autocorrelations and therefore the number of MCMC steps required.

2. Code optimization; speeding up the MC reweight via more optimized code.

3. Algorithmically; change, or update the MCMC sampler algorithm

This section will briefly describe the search for a better MCMC algorithm in the context of the T2K experiment, and the results of that search.

A small framework was written to test various MCMC algorithms for a T2K-like problem. The framework was capable of generating a large covariance matrix with any user-provided number of parameters, where one parameter was bi-modal (Δm^2_{32} central values and their uncertainties for both NH and IH were taken from the 2016 T2K analysis described in [41]). The two main goals were to find an algorithm that would result in a stationary distribution at lesser computational expense, and to remove the 50% chance on switching the sign of Δm^2_{32} that is hard-coded in the current MaCh3 implementation - it would be more robust if the algorithm itself were capable of finding the other mass hierarchy. A few algorithms that were initially showing the most interesting results are described below.

3.4.1 Parallel Tempered MCMC

Parallel Tempered MCMC (PTMCMC) [66, 67] is an interesting algorithm that in principle results in a fast-mixing chain that is capable of marginalizing over separated posterior modes. This is done by running multiple chains in parallel, and proposing position swaps between them, where only one chain is actually used to build the posterior. Furthermore, some of these chains are tempered (the terms "tempering" and "annealing" will be used interchangeably from now); the likelihood function has an extra term that effectively spreads the sampled posterior density. In general, the likelihood $\exp(-\chi^2/2)$ can be re-written as

$$e^{-\frac{\chi^2}{2kT}}, \tag{3.11}$$

where k is an arbitrarily chosen constant (say, 1), and T is the "temperature" of the chain, responsible for spreading the underlying posterior density. As we increase the temperature, the MCMC chain explores a wider area of phase-space, with the modified acceptance probability;

$$A = \min[1, \exp\{\beta \times (\ln \mathcal{L}_{\text{old}} - \ln \mathcal{L}_{\text{new}})\}], \tag{3.12}$$

where $\beta = 1/kT$. With $\beta < 1$ for the annealed chains, the log-likelihood difference seen by the MCMC chain per step should be smaller, and the probability of accepting the

step higher. With this modified phase-space exploration, it may be possible to probe separated posterior modes (if β is small enough).

Of course, we still want to recreate the unmodified posterior density. Parallel Tempering runs multiple MCMC chains in parallel, with progressively higher temperature, starting from the unmodified MCMC chain. These chains are allowed to swap positions with each other, with the swap acceptance probability being;

$$A = \min[1, \exp\{(\beta_1 - \beta_2) \times (\ln \mathcal{L}_2 - \ln \mathcal{L}_1)\}], \tag{3.13}$$

where the 1 and 2 subscripts represent two random chains (with different temperatures) selected during the MCMC step. This probability ensures that as chains with various underlying likelihood distributions are swapping, none of them will end up in an area of phase-space that is not likely enough given their temperatures.

With PTMCMC, the chain that is not annealed mixes with the annealed chains; making its parameter-space exploration faster, and potentially discovering separated posterior density modes via the annealed chains. Unfortunately, as tests in a testing framework progressed, it became clear very quickly that this method is only efficient if the hypothesis has a low number of dimensions. After $\sim$20 parameters, hundreds of parallel chains are needed for convergence, making this method computationally worse for T2K.

3.4.2 Ensemble sampling with affine invariance

J. Goodman and J. Weare proposed a new type of ensemble samplers, with their efficiencies that are unaffected by the correlations between the sampled parameters (i.e. affine invariant) [68, 69]. Instead of building a single chain, the ensemble samplers are initialized with already pre-existing distribution of samples (in this case, often called "walkers"). These samples are usually initially distributed uniformly across the parameters' priors, and each MCMC step involves a linear transformation between the samples, moving the ensemble closer to the stationary distribution.

A few separate methods are available here; we can either start with a small amount of "walkers", each one of them developing a MCMC chain via linear transformations, or start with a large ensemble, where the last MCMC step from each walker is used to build

the final posterior density. The latter is, in many ways, similar to the Genetic algorithms, Differential Evolution, Nested Sampling and so on (differential evolution was tested for T2K too, with similar results to section 3.4.1).

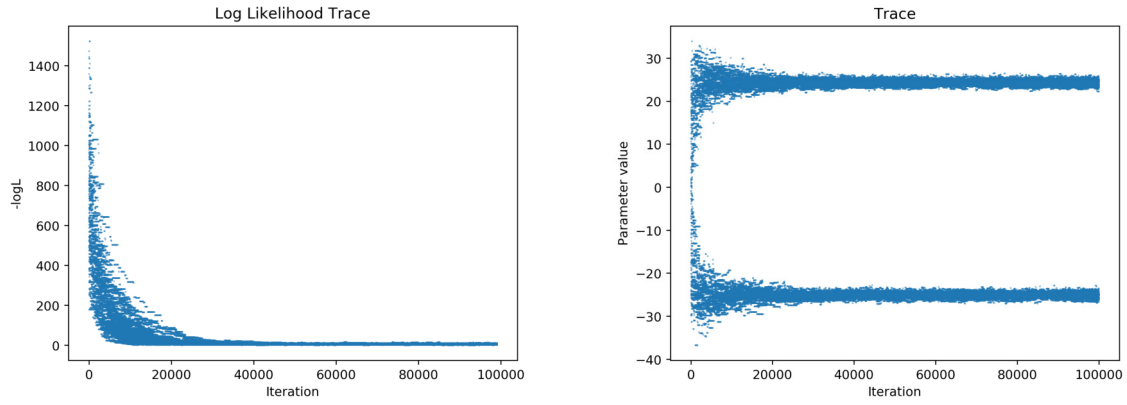
For every MCMC step, each walker is updated via a linear transformation called "stretch", given a secondary randomly-chosen walker that we can call "donor";

$$p_i(t+1) = p_i(t) + Z[p_i(t) - p_j(t)], \tag{3.14}$$

where the subscript i, j represent i$^{\text{th}}$ and j$^{\text{th}}$ walker, $p_j(t)$ is a randomly chosen donor for each step, and Z is a random variable; either fixed step-size, or randomly drawn from some symmetric distribution (for example, Gaussian). This is repeated for all the walkers, and each stretch is accepted with the probability;

$$A = \min[1, Z^{N-1} \exp\{\ln(\mathcal{L}_{\text{i, old}}) - \ln(\mathcal{L}_{\text{i, new}})\}], \tag{3.15}$$

where N is the number of parameters (dimensions).



(a) Log likelihood trace. There are 10,000 walkers, with 1,000 evolutions each.

(b) Trace for a parameter imitating Δm^2_{32}. As the ensemble is converging after 1,000 evolutions, it recreates the expected posterior density.

Figure 3.7: Traces from the ensemble MCMC sampler, with 10,000 walkers and 1,000 evolutions. At the beginning the ensemble is spread uniformly across the parameter space, slowly converging to the stationary distribution after each evolution.

At the start of testing with a lower number of parameters, the performance of the ensemble sampler was rather impressive. It managed to generate many samples across

the parameter-space, more than enough for high-statistics analysis. Furthermore, the autocorrelations were always negligible, with very minimal tuning.

One of the main issues with the algorithm was that although it could find both modes of Δm^2_{32}-like parameter given the sufficient distribution of initial samples, it did not marginalize properly over the mass hierarchy. This means the 50% swap proposal would still need to be in place. The fact that this algorithm did not perform well when it was sampling from more than 50 parameters was the final nail in the coffin. This, unfortunately, held true for the Differential Evolution algorithm too. Ensemble samplers, although very-well performing in highly non-linear problems, do not seem to be efficient in high dimensionality.

3.4.3 Hamiltonian MCMC

The Hamiltonian MCMC (HMCMC) algorithm has a potential to solve the dimensionality and autocorrelation problems by generating posterior samples that are fully uncorrelated, whilst having a high optimal acceptance ratio. This means it can draw fully random samples from the stationary posterior dirstribution, and it does so by utilizing the Hamiltonian dynamics to both propose a new step, and to calculate the acceptance probability [72].

HMCMC proposes new steps not by randomly choosing a direction with fixed step-sizes, but by giving the current sample some 'kinetic energy' in a random direction, and letting it evolve using Hamiltonian dynamics. The evolution is stopped at some time $(t + \epsilon)$, giving the evolved position (parameter values) to the newly generated step.

In Hamiltonian MCMC the total energy of the system is conserved by trying making sure that the Hamiltonian stays constant throughout the evolution;

$$H = U(\boldsymbol{x}) + K(\boldsymbol{p}), \tag{3.16}$$

where $U(\boldsymbol{x})$ is the potential energy — the $\log(\mathcal{L})$ — given some some parameter values $\boldsymbol{x}$, and $K(\boldsymbol{p})$ is the kinetic energy — $|\boldsymbol{p}|^2/(2m)$ — with a new set of momentum parameters that are being evolved along the positions (parameter values). From that, we can write

simple equations of motion in the form of partial derivatives, to see how $\boldsymbol{x}$ and $\boldsymbol{p}$ behave over time;

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \tag{3.17}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i}. \tag{3.18}$$

The symbol i represents the dimension (i.e. the selected parameter), and H is the standard Hamiltonian from equation 3.16. With the previously defined kinetic and potential energies, we can rewrite equations 3.17 and 3.19 as

$$\frac{dx_i}{dt} = \frac{p_i}{m_i}, \tag{3.19}$$

$$\frac{dp_i}{dt} = -\frac{\partial U}{\partial x_i}, \tag{3.20}$$

where the potential $U(\boldsymbol{x})$ is the negative log likelihood at $\boldsymbol{x}$ and m_i is the "mass term". Partial derivatives must be computed for all the sampled parameters, which is a computationally expensive process. The mass term can be chosen arbitrarily to be, say, 1, or based on some prior to speed up the mixing, by taking the parameters' correlations into account. These equations of motion define the mapping from a state $(\boldsymbol{x}, \boldsymbol{p})_t$ at time t to a new state at time $t + \epsilon$, $(\boldsymbol{x}, \boldsymbol{p})_{t+\epsilon}$; a new MCMC step.

In practice we have to discretize the Hamiltonian equations of motion by discretizing the time, using a number L of small steps of size ε. We then iterate through $\varepsilon_1, \varepsilon_2, ..., \varepsilon_L$, updating $\boldsymbol{x}$ and $\boldsymbol{p}$ at each step to evolve the Hamiltonian trajectory. There are many methods of discretization, with the most widely used being the "leapfrog" method shown below;

$$p_i(t + \varepsilon/2) = p_i(t) - (\varepsilon/2)\frac{\partial U}{\partial x_i}(x(t)), \tag{3.21}$$

$$x_i(t + \varepsilon) = x_i(t) + \varepsilon\frac{p_i(t + \varepsilon/2)}{m_i}, \tag{3.22}$$

$$p_i(t + \varepsilon) = p_i(t + \varepsilon/2) - (\varepsilon/2)\frac{\partial U}{\partial x_i}(x(t + \varepsilon)), \tag{3.23}$$

where the Hamiltonian equations of motion are of the same form as in equations 3.17 and 3.19. We can therefore get from $t(0)$ to $t(\epsilon)$ by iterating through a number of $t(\epsilon/L)$ steps. Since the trajectory is based on small discrete steps, the Hamiltonian value in

practice does not stay perfectly constant; it fluctuates, and the discrerization step-size, ϵ, and the number of discretizations need to be tweaked to ensure similar Hamiltonian value at each end of the trajectory. Since the discretization cannot be perfect, a Metropolis-like probability is used to determine whether to accept or reject the HMCMC step, by computing the difference in the Hamiltonians.

It it quite easy to visualize how this works in practice. Figure 3.8 shows the progression of both parameter values (left) and momenta (middle) for one Hamiltonian trajectory (one HMCMC step), whilst the total energy of the system remains relatively stable (right), for two parameters against each other. Only the beginning and the end of the trajectory are valid MCMC steps (current and proposed steps, respectively). The actual values of the parameters' momenta are not interpretable, i.e. we cannot use them to extract more information from the fit.

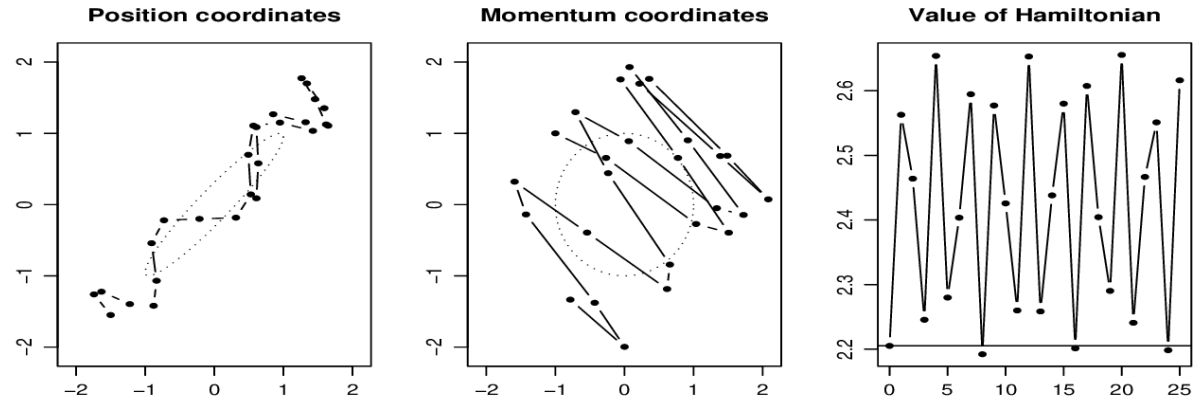
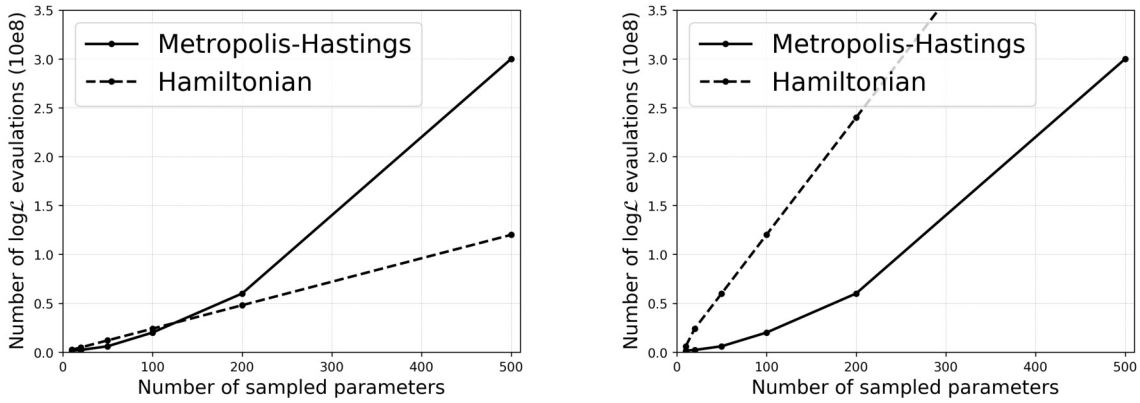


Figure 3.8: Visualization of how HMCMC works, with one trajectory on 2D parameter-space shown on the left, their corresponding momentum trajectory in the middle, and the evolution of the total Hamiltonian value on the right. We can see that within one trajectory, the sample's position transverses across the parameters' entire error band, generating a fully independent point at the end of it. In other words, the information about the original position is lost, meaning the autocorrelation at lag 1 is zero. Figure taken from [70].

The number of discriminations, L, and the discretization size, ε, are the two main parameters to be tuned. Luckily this can be done somewhat separately; ϵ should be small enough not to vary the Hamiltonian value too much, and L should be large enough for one trajectory to cover entire phase-space, producing fully uncorrelated MCMC samples.

The mass term can be provided in a form of a correlated matrix, allowing HMCMC evolution in a fully uncorrelated phase-space for extra efficiency.

At the beginning of testing the HMCMC algorithm in the toy framework imitating a T2K-like model with variable number of model parameters, it showed far better efficiency scalability in comparison to the PTMCMC and ensemble MCMC samplers described in sections 3.4.1 and 3.4.2. As the number of sampled parameters was raised, the HMCMC efficiency was decreasing linearly; this was a huge improvement over the Metropolis-Hastings algorithm, which scales exponentially — making efficiency difference larger as the number of parameters is increasing.



(a) Metropolis-Hastings against HMCMC comparison of the number of $\log(\mathcal{L})$ evaluations needed for stationary posterior density, with 10 leapfrogs per Hamiltonian evolution.

(b) Metropolis-Hastings against HMCMC comparison of the number of $\log(\mathcal{L})$ evaluations needed for stationary posterior density, with 5 leapfrogs per Hamiltonian evolution.

Figure 3.9: Metropolis-Hastings against HMCMC comparison of the number of $\log(\mathcal{L})$ evaluations needed for stationary posterior density. Although larger number of leapfrog steps per trajectory are more computationally expensive, they produce a far less auto-correlated chain, reducing its needed total length.

This is shown in figure 3.9, where the number of needed $\log(\mathcal{L})$ evaluations for a stationary distribution (10,000 fully uncorrelated MCMC steps) was compared for Metropolis-Hastings against HMCMC, for an increasing number of parameters. In 3.9a we can see that HMCMC, although less efficient in lower numbers of dimensions, scales linearly when it comes to the required computational burden, causing it to perform better in a higher number of dimensions. This is only with the sufficient number of leapfrog

steps, which raises an interesting conclusion.

HMCMC is only more efficient if the number of leapfrog discretizations is high enough to generate fully uncorrelated steps (i.e. autocorrelation is $\sim$0 at lag $\sim$1). This is somewhat counter-intuitive; when we decrease number of leapfrog discretizations, the computing time per HMCMC step (trajectory) decreases significantly, but the resulting autocorrelation makes it far less efficient. When the Hamiltonian tracks are longer, however, the autocorrelations become negligible and this outweighs the extra computational time per-track because far less MCMC steps are needed to sufficiently sample the parameter-space.

The toy framework, in which these MCMC studies were performed, made it easy to optimize the discretization length ε and the number of discretizations L; the covariance matrix, from which all the algorithms were sampling, was also used for optimizing the step-sizes. This means the comparison is for well-optimized algorithms, which is usually difficult in practice due to the not-perfectly-known effect of the data pulling the model parameters.

3.4.4 HMCMC implementation in MaCh3

The HMCMC algorithm was implemented in the MaCh3 framework, with full tuning for systematic parameters left for future analysers. The implementation of the gradient calculator was a rather simple one; the parameter values are slightly deviated, and the differences between their likelihoods and their parameter values were used to obtain gradients for each parameter. An example of a gradient and $\log(\mathcal{L})$ scan over one of the T2K flux parameters is shown in figure 3.10.

The HMCMC algorithm was tested on the oscillation parameters first, with all the nuisance parameters fixed to their pre-data-fit nominal predictions. A fake dataset called "Asimov A", generated with the nominal T2K MC tuned to oscillation parameterd from older analyses, was used for the first HMCMC runs. Figure 3.11a shows five HMCMC trajectories, each with a different colour, together with the trajectories' leapfrog points, on top of the posterior probability obtained from sampling with the standard Metropolis-

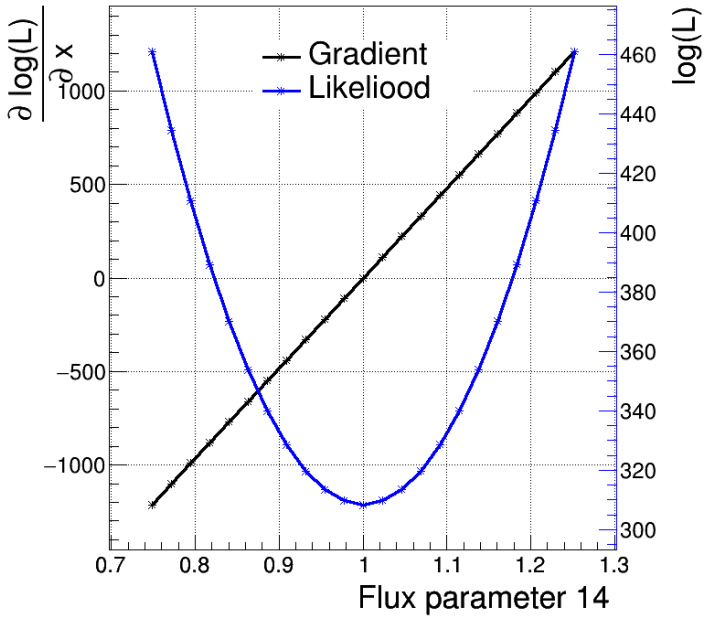
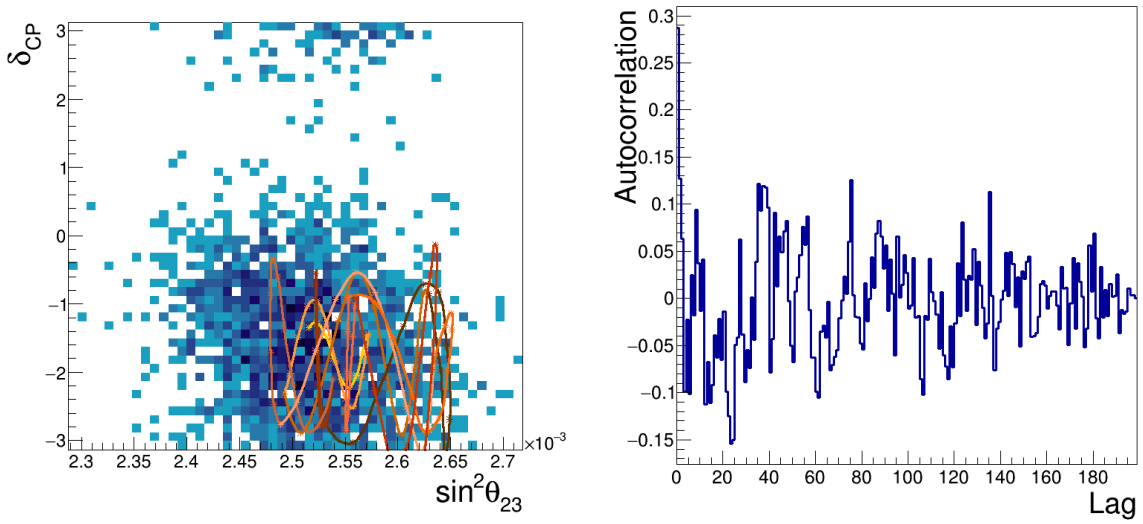


Figure 3.10: Gradient scan on top of a $\log(\mathcal{L})$ scan for one of the T2K flux parameters, with the "Asimov A" fake dataset described in chapter 4. The gradients match the expectation from the $\log(\mathcal{L})$ scan. These gradient scans were completed for all the 750 T2K parameters.

Hastings MCMC. Figure 3.11b shows the autocorrelation from 1000 HMCMC steps, where the autocorrelation seems to drop to $\sim$0 when the lag is $\sim$1.

Although HMCMC seems to be working well for the oscillation parameters, more work needs to be done for the nuisance parameters. The parametrization changes between the analyses as more data is added, and HMCMC is particularly sensitive to the step size (ε, L, mass matrix) tuning. One way to make the optimization easier is through some form of adaptive step-size tuning. For example, a short Metropolis-Hastings chain could be run to find the preliminary $1\,\sigma$ contours for each parameter. The resultant preliminary posterior distribution could be used as a mass-matrix in HMCMC. If the parameters follow perfect Gaussian distributions, such tuning would be maximally optimal. In reality, even a rough estimation of the matrix for non-Gaussian parameters can help significantly with the step-size tuning and the sampler's efficiency.

Furthermore, more issues might arise if some parameters have non-linear likelihood surfaces; the discretization length ε will have to be small enough so that the trajectory



(a) (b)

Figure 3.11: HMCMC evolutions on δ_{CP} vs $\sin^2\theta_{23}$ space, with all the oscillation parameters free in the model, and nuisance parameters fixed to their pre-data-fit nominal expectations. The algorithm does produce HMCMC tracks, with five tracks shown in (a) with the underlying binned posterior produced using standard Metropolis-Hastings. The autocorrelation values drops to ∼0 after only one HMCMC step, shown for δ_{CP} in (b), producing fully uncorrelated samples with each HMCMC step.

can reconstruct such irregularities well. Consequently, the number of leapfrog-steps L would have to be raised to ensure un-autocorrelated chain progression. Finally, the mass hierarchy will have to be dealt with in exactly the same way as it is currently; with the proposed step of swapping the sign of Δm^2_{32}, with a standard Metropolis-Hastings acceptance update.

The main outcome of this analysis is in figure 3.11. HMCMC is a promising algorithm in a high-dimensional space, that could reduce the computational burden. This could be especially important for producing highly significant results (for example, with $3\,\sigma$ intervals), or if the number of systematic parameters is raised.

Chapter 4

T2K Bayesian Oscillation Analysis

The T2K neutrino oscillation analysis involves inputs from various internal working groups, whether it is cross-section or detector systematics modelling, often using external data. These inputs go into four main analysis teams; BANFF, VALOR, P-Theta and MaCh3 (all first mentioned in section 3.3). Although this thesis describes the results using MaCh3, the frameworks were validated between each other before the main data fits and the final results compared against each other, with some of these validations and comparisons shown in this thesis.

At the beginning of the analysis, both BANFF and MaCh3 perform a fit to the ND280 data, given inputs from the flux, cross-section and near-detector groups. The results from BANFF then become inputs for P-Theta and VALOR, whereas the MaCh3 ND results are used for validations and additional studies. These inputs help constrain the systematics at SK for the oscillation analysis, since the ND280 sits in the same neutrino beam, collecting higher-statistics data. The effect of the ND280 data constraining the systematic errors on the SK data samples is shown in table 4.1, taken from [73].

The neutrino flux and cross section systematic uncertainties are described in 4.1.1 and 4.1.2 respectively. The ND280 data samples are described in section 4.2.1, and the SK data samples mentioned in table 4.1 are described in section 4.2.2, both with their detector uncertainties.

	FHC 1Rμ	RHC 1Rμ	FHC 1Re	RHC 1Re	FHC CC1π
No ND280	15.0%	13.0%	15.0%	13.7%	20.1%
With ND280	5.5%	4.4%	8.8%	7.3%	17.8%

Table 4.1: Effect of the ND280 data on the event rate systematic uncertainty constraints of the three Forward-Horn-Current (FHC) and two Reverse-Horn-Current (RHC) data samples. Numbers in the table taken from [73].

4.1 Systematic uncertainties

The systematics for the T2K oscillation analysis are split into the SK and ND280 detector efficiencies, neutrino flux uncertainties and the neutrino-matter cross-section uncertainties. Additionally, there are uncertainties on oscillation parameters that the T2K experiment is not sensitive to, namely, $\sin^2\theta_{12}$ and Δm^2_{12}. Each parameter (systematic or not) needs to have a prior distribution in Bayesian analyses, as described in chapter 3. Most of these priors are correlated Gaussian that come in the form of covariance matrices, separate for each group of systematics. This means there are no prior correlations between separate systematic groups, for example, between cross-section and flux uncertainties. Most covariance-based priors give a penalty to the T2K likelihood function, varying the penalty as the parameter is deviated. The parameters with non-informative, flat priors do not change the likelihood-penalty depending in its value.

4.1.1 Neutrino beam flux

The neutrino flux at T2K is simulated by modelling the proton beam interacting with the graphite target. These simulations are tuned using external data from the hadron production experiment, NA61/SHINE (SPS Heavy Ion Neutrino Physics Experiment) at CERN [74, 75], which uses a 30 GeV proton beam incident on a thin[1] T2K target replica.

The FLUKA [77] software is used to simulate the hadronic interactions on the carbon target (tuned with the above NA61/SHINE data) producing the charged mesons. The surrounding geometry is simulated with GEANT3 [78], with the secondary interactions

[1] A full T2K target replica, described in [76], will be used for the next analyses

outside of the target simulated in GCALOR [79], and tuned to the experimental data. The secondary particles are tracked in the simulation through the decay tunnel as they decay into neutrinos and antineutrinos.

The flux systematic errors are a result of uncertainties in the hadron production model, the secondary interactions outside of the carbon target, beam profile and alignment, the magnetic horn current and field, uncertainties in the geometrical and material modelling, and the amount of data collected, in units of protons-on-target (POT). The flux uncertainties are around 10% at the flux peak, with the main contribution coming from the hadron production model, and at low energies the uncertainties coming from the modelling of both the pion production and secondary interactions outside of the target.

The neutrino flux uncertainties consist of one hundred parameters, split equally between flux uncertainties at ND280 and SK. For each detector the flux uncertainties are further split into FHC and RHC modes, and each mode for each detector contains uncertainties for all four neutrino species, ν_μ, $\bar{\nu}_\mu$, ν_e and $\bar{\nu}_e$. Every detector-beam mode-species uncertainty is binned in the true neutrino energy. The number of energy bins is chosen to be broadly proportional to the magnitude of the flux per group itself, meaning that, say, FHC-ν_μ will have the same number of energy bins as RHC-$\bar{\nu}_\mu$. Both are defined as *right-sign*, since the neutrino flavour corresponds to its neutrino beam mode. The *wrong-sign* would be defined as, for example, $\bar{\nu}_\mu$ in FHC mode, and the wrong-sign background having less bins. The energy bin width is smaller at the neutrino oscillation maxima, around 0.6 GeV, and courser away from this maxima. With this in mind, table 4.2 shows the neutrino flux energy binning for right-sign and wrong-sign flavours.

Neutrino flavour	Sign	Neutrino energy binning boundaries [GeV]	No° of bins
μ	right	0, 0.4, 0.5, 0.6, 0.7, 1, 1.5, 2.5, 3.5, 5, 7, 30	11
μ	wrong	0, 0.7, 1, 1.5, 2.5, 30	5
e	right	0, 0.5, 0.7, 0.8, 1.5, 2.5, 4, 30	7
e	wrong	0, 2.5, 30	2

Table 4.2: Neutrino flux energy binning (in GeV) for right-sign and wrong-sign neutrino flavours. All the bin numbers and widths are the same for FHC and RHC modes, and for ND280 and SK detectors.

All the 100 flux parameters in MaCh3 have Gaussian priors, with mean of 1 and widths corresponding the parameter's deviation from its covariance. These uncertainties are applied in a form of normalization, which scales the number of MC events. The flux parameters are all highly correlated, with the fractional covariance shown in figure 4.1, produced by the T2K beam group.

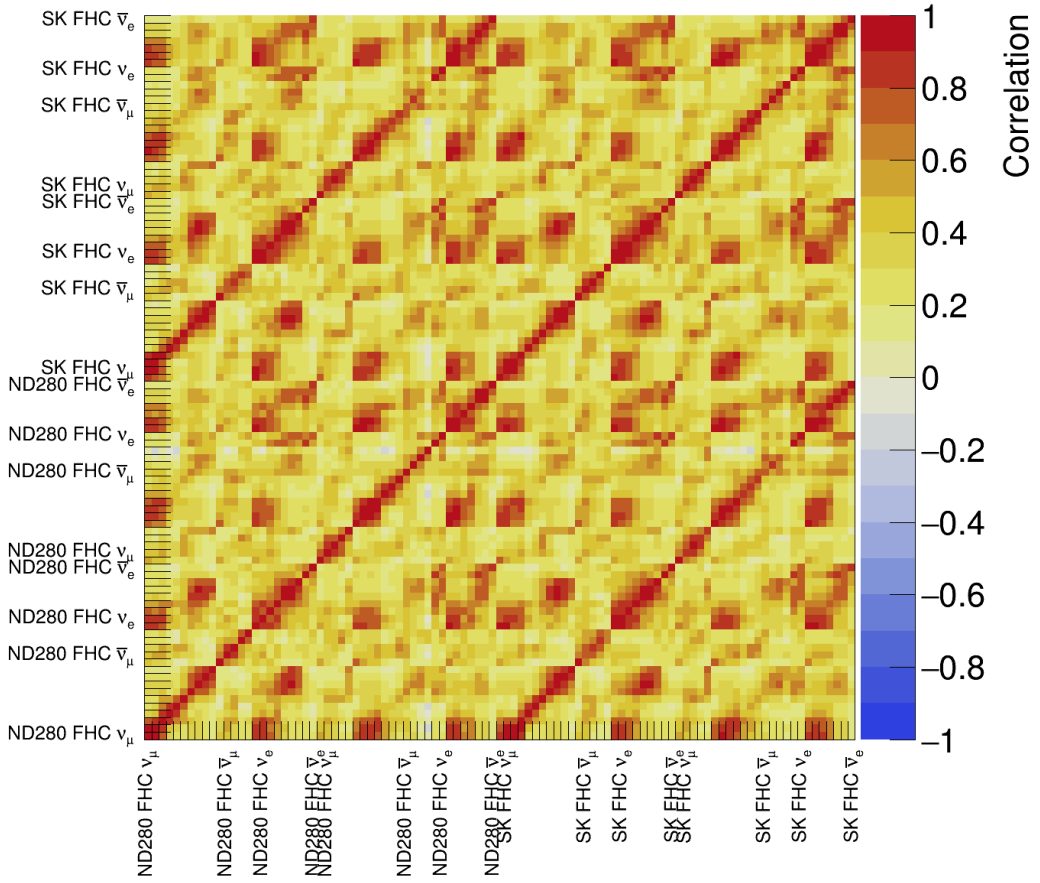


Figure 4.1: The beam flux fractional covariance matrix, binned in the neutrino energy for all the both ND280 and SK data samples (and detailed in table 4.2). The matrix is used as Gaussian priors in the joint MaCh3 fit.

4.1.2 Neutrino cross-sections

The neutrino interaction cross-sections on water are the dominant systematic uncertainty in most long-baseline accelerator neutrino experiments, and T2K is not an exception.

The neutrino-matter parametrization in T2K is updated regularly, and changes for the analysis in this chapter too. The NEUT neutrino event generator is used to simulate the neutrino interactions in the T2K detectors [81, 80].

Since in T2K the neutrino interactions take place in complex nuclei, interactions with multiple nucleons must be taken into account. The nucleus is modelled with the Relativistic Fermi Gas approximation (RFG), modified with the Random Phase Approximation (RPA) to take long-distance correlations into the consideration [83]. Neutrino interactions with multiple nucleons is included in the two-particle two-hole (2p2h) model, also called Meson Exchange Currents (MEC) [84, 82]. The 2p2h interactions can be easily misidentified as CCQE biasing the reconstructed neutrino energy [137], it is therefore important to include them in the T2K cross-section model. An example of a 2p2h interaction is shown in figure 4.2. The default T2K MC generated with NEUT is modelled with the Spectral Function (SF) model, meaning it needs to be reweighted to RFG and then again the relativistic RPA (MEC is already included in the pre-fit MC). External data from the MiniBoonNE [85, 86] and MINERvA [87] experiments was used by the T2K Neutrino Interaction Working Group (NIWG) [88] to tune the CCQE nuclear model parameters. The resonant pion interaction (RES) parameters were tuned using data from the ANL and BNLexperiments [89].

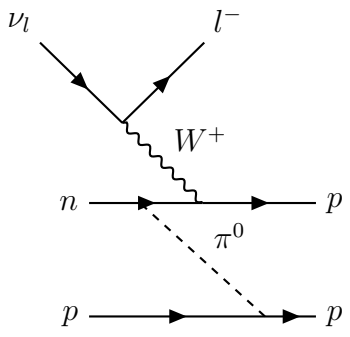
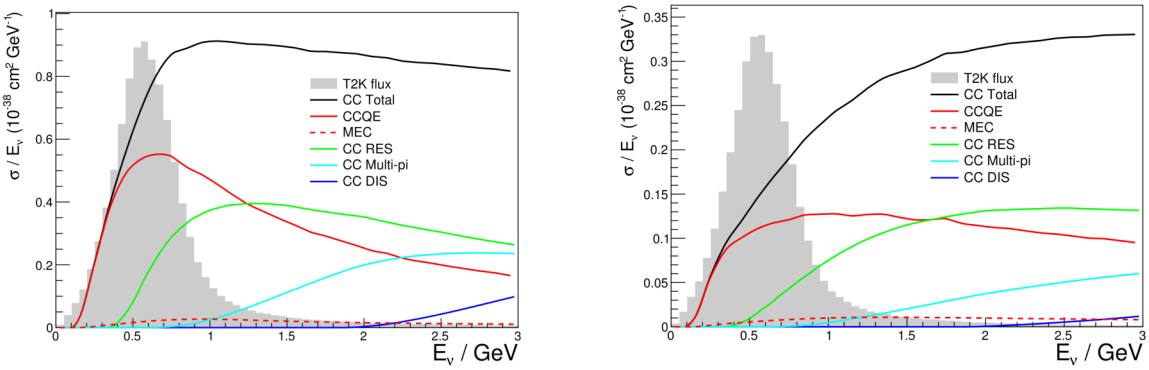


Figure 4.2: Example of a Feynman diagram for l-flavoured neutrino interacting with a pair of nucleons via the 2p2h process.

The binding energy of a nucleon that exits the nucleus, E_b,must be subtracted in the simulation. It is tuned using external electron scattering data [90]. After the first neutrino interaction with the nucleus, the secondary hadrons propagate through the rest of the nucleus, interacting on their way out. This effect is called the Final State Interactions

(FSI), and it is simulated through a cascade model where the hadrons are interacting with the nuclear medium classically, altering its final state momentum and energy. A free-particle path that is dependent on the nuclear density is taken into account, and constrained using particle-nucleon scattering data.

The interactions in T2K around the 0.6 GeV neutrino energy are dominated by charged-current quasielastic (CCQE) interactions, which are the most-well predicted by the current T2K cross-section model. NEUT also includes the 2p2h interactions, CC resonant pion production and deep inelastic scattering. The cross sections for each interaction type are shown in figure 4.3, for ν_μ and $\bar{\nu}_\mu$ separately.



(a) Muon neutrino interaction mode cross-sections across the T2K neutrino energy range.

(b) Muon antineutrino interaction mode cross-sections across T2K neutrino energy range.

Figure 4.3: Muon neutrino (left) and antineutrino (right) interaction mode cross-sections across neutrino energy range, overlaid on the T2K beam flux energy distribution.

The parameters describing these interactions are defined as either "normalization" or "shape" parameters. The normalization parameters are simple multipliers that weight all the events, regardless of the neutrino energy involved. The shape parameters are multipliers that are dependent on on the neutrino energies; either only affecting certain energies, or affecting different energies with different weights, modifying the resultant neutrino energy spectra *shape*. The shape parameters are implemented in a form of response functions implemented as splines, where a spline corresponding to a given energy will return a weight corresponding to the parameter value.

A total of 32 shape and normalization parameters were used in this analysis, tuning

the cross-sections for the above cross-section systematics. Most of the parameters have correlated Gaussian priors, derived either from theory [91], or from fits to external data, as described before. All the cross-section systematics are listed in table 4.3 with their pre-fit central values, prior errors, prior types and the parameter type (shape or normalization). Table 4.4 shows the cross-section pre-fit correlation matrix used as a multivariate prior.

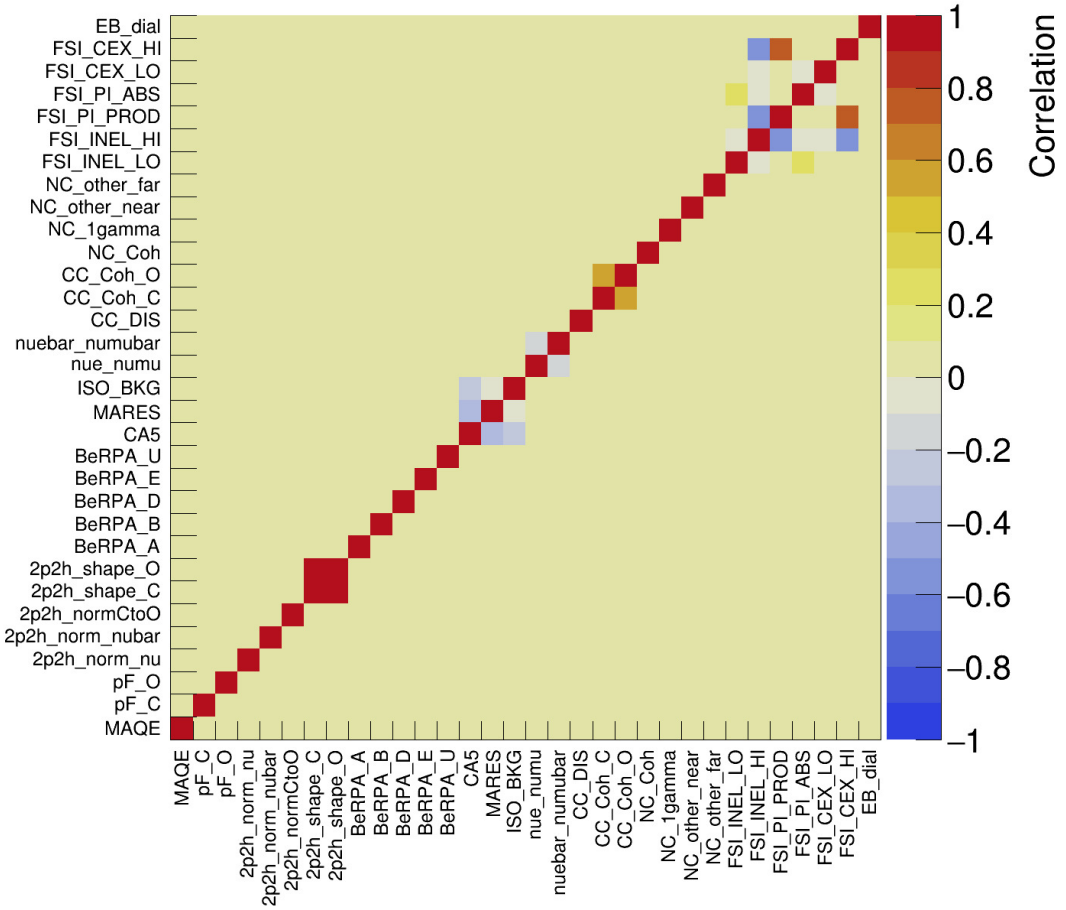


Figure 4.4: The cross-section pre-fit correlation matrix showing all the cross-section systematic parameters. Correlations between the parameters are mostly taken from theory and fits to external data, with most correlations unknown prior to the fit.

4.1.3 The ND280 Detector

The ND280 detector has 556 systematic normalization parameters binned in the outgoing muon momentum and angle, p_μ and $\cos(\theta_\mu)$. They are all are included in the joint-detector

Parameter name	Nominal value	Prior Type	Prior error	Parameter type
M_A^{QE}	1.2	Flat prior	—	Shape
$p_F{}^{12}C$	217	Flat prior	—	Shape
$p_F{}^{16}O$	225	Flat prior	—	Shape
2p2h ν norm	1	Flat prior	—	Normalization
2p2h $\bar{\nu}$ norm	1	Flat prior	—	Normalization
2p2h $^{12}Cto^{16}O$ norm	1	Gaussian	0.2	Normalization
2p2h ^{12}C norm	1	Gaussian	3.0	Shape
2p2h ^{16}O norm	1	Gaussian	3.0	Shape
BeRPA A	0.59	Gaussian	0.118	Shape
BeRPA B	1.05	Gaussian	0.210	Shape
BeRPA D	1.13	Gaussian	0.170	Shape
BeRPA E	0.88	Gaussian	0.352	Shape
BeRPA U	1.2	—	—	Fixed
C_A^5	1.01	Gaussian	0.15	Shape
M_A^{RES}	0.95	Gaussian	0.15	Shape
Isospin$\frac{1}{2}$ Back.	1.3	Gaussian	0.40	Shape
CC ν_e/ν_μ	1	Gaussian	0.03	Normalization
CC $\bar{\nu}_e/\bar{\nu}_\mu$	1	Gaussian	0.03	Normalization
CC DIS	1	Gaussian	0.4	Shape
CC Coherent ^{12}C	1	Gaussian	0.3	Normalization
CC Coherent ^{16}O	1	Gaussian	0.3	Normalization
NC Coherent	1	Gaussian	0.3	Normalization
NC 1 γ	1	Gaussian	1.0	Normalization
NC Other Near	1	Gaussian	0.3	Normalization
NC Other Far	1	Gaussian	0.3	Normalization
FSI Inelastic Low E.	0	Gaussian	0.41	Shape
FSI Inelastic High E.	0	Gaussian	0.34	Shape
FSI Pion Production	0	Gaussian	0.50	Shape
FSI Pion Absorption	0	Gaussian	0.41	Shape
FSI Ch. Exchange Low E.	0	Gaussian	0.57	Shape
FSI Ch. Exchange High E.	0	Gaussian	0.28	Shape
E_b dial	0	Gaussian	2.0	Shape

Table 4.3: Cross-section systematic parameters used for all the oscillation analyses presented in this thesis, showing the nominal value, the type of prior, the prior error, and type of the parameter.

(ND280 + SK) data analysis, similarly to the flux parameters, after optimization, cross-validations and fits to ND280-only data together with the BANFF group (first described in section 3.3).

The true systematic parameters for each sample, for example, TPC Charge ID Efficiency, are varied systematically and statistically to produce 2000 random variations. These randomly fluctuated samples are then selected according to their priors. The predicted number of events rates from each fluctuated sample are binned in the outgoing muon momentum and angle, for each FGD separately. The binning of these p_μ–$\cos(\theta_\mu)$ parameters is shown in table 4.4.

Binning in these variables allows high correlations between FGD1 and FGD2, and between the adjacent bins through a covariance matrix. The MC uncertainty is accounted for by adding statistical MC and MC-shifting covariance matrices. The statistical MC matrix is generated including Poisson fluctuations in the MC event weights, and adding these to the detector covariance martix. The MC-shifting covariance matrix tries to solve the Martini and Nieves 1p-1h interaction model difference, which results in a shift in the reconstructed lepton momentum in CCQE interactions [138]. The two near detector fitters mentioned in section 3.3, BANFF and MaCh3, are used for independent validations [94]. The ND280 covariance matrix with all the ND280 systematic parameters, produced from the 2000 fluctuated samples and the two above matrices, and used as prior in the T2K oscillation analysis, is shown in figure 4.6.

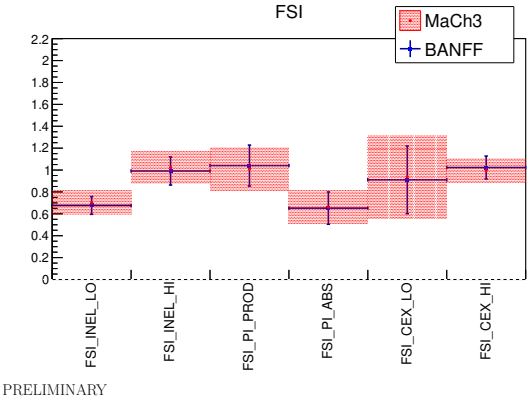
As a part of further ND280 analysis, BANFF and MaCh3 both output the covariance matrix from the full ND280 data analysis together with the constrained neutrino flux and cross-section uncertainties, which are being used for the two hybrid-Frequentist far-detector neutrino oscillation analysis groups. The Mach3 group does not use these matrices; the data analysis is re-done with the use of matrix in 4.6 as a prior, and the addition of the SK data, SK systematic uncertainties and extra SK-only cross-section parameters in a joint ND280-SK T2K data analysis.

Figure 4.5 shows example comparisons of the cross-section parameters between MaCh3 and BANFF fits to the ND280 data. There is an overall good, but not perfect agreement between the two fitters. The visible differences in, for example, the 2p2h parameters, were

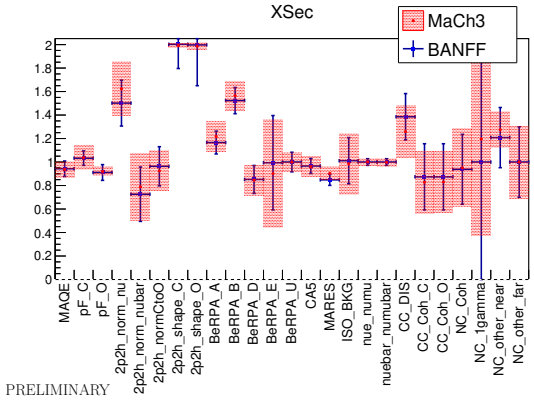
Systematic sample	p_μ bins (MeV)	$\cos(\theta_\mu)$ bins
FHC ν_μ CC 0	0, 1000, 1250, 2000, 3000, 5000, 30000	-1, 0.6, 0.7, 0.8, 0.94, 0.96, 1
FHC ν_μ CC 1 π	0, 300, 1250, 1500, 5000, 30000	-1, 0.7, 0.85, 0.9, 0.92, 0.96, 0.98, 0.99, 1
FHC ν_μ CC N π	0, 1500, 2000, 3000, 5000, 30000	-1, 0.8, 0.85, 0.9, 0.92, 0.96, 0.98, 0.99, 1
RHC $\bar{\nu}_\mu$ CC 1 track	0, 400, 900, 1100, 2000, 10000	-1, 0.6, 0.7, 0.88, 0.95, 0.97, 0.98, 0.99, 1
RHC $\bar{\nu}_\mu$ CC N track	0, 700, 1200, 1500, 2000, 3000, 10000	-1, 0.85, 0.88, 0.93, 0.98, 0.99, 1
RHC ν_μ CC 1 track	0, 400, 800, 1100, 2000, 10000	-1, 0.7, 0.85, 0.90, 0.93, 0.96, 0.98, 0.99, 1
RHC ν_μ CC N track	0, 1000, 1500, 2000, 3000, 10000	-1, 0.80, 0.90, 0.93, 0.95, 0.96, 0.97, 0.99, 1

Table 4.4: ND280 systematic uncertainties split per sample, p_μ and $\cos(\theta_\mu)$, making all the 556 uncertainties. These systematic uncertainties are duplicated, corresponding to FGD1 and FGD2.

attributed to marginalization effects near the parameters' physical constraints. Various validations (for example, studying systematic pulls and log-likelihood scans) have found that both fitters are behaving as expected [93].



(a) Comparison between BANFF (blue error bars) and MaCh3 (red blocks) for the FSI parameters.



(b) Comparison between BANFF (blue error bars) and MaCh3 (red blocks) for all the cross-section parameters except Final State Interactions.

Figure 4.5: Comparison between BANFF (blue error bars) and MaCh3 (red blocks) for the FSI parameters (left) and the rest of the cross-section parameters (right). There is a good agreement between the fitters, with the differences (for example, the 2p2h parameters) attributed to the marginalization effects near the parameters' physical constraints. Figures taken from [93].

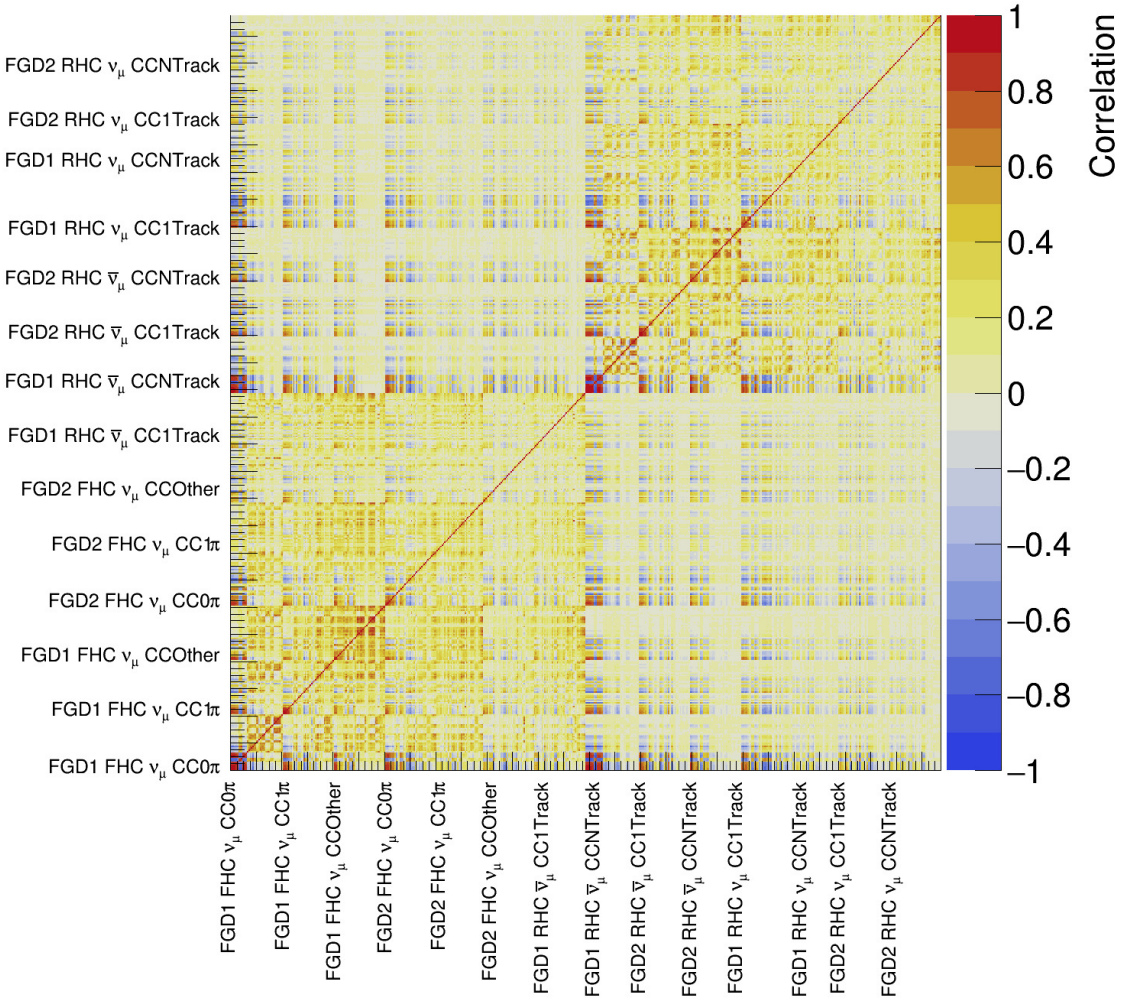


Figure 4.6: ND280 pre-fit correlation matrix showing all the ND280 detector systematic parameters binned in p_μ–$\cos(\theta_\mu)$. The matrix was produced by fluctuating the ND280 parameters, and recording the resultant event numbers in each bin from table 4.4

4.1.4 The Super-Kamiokande Detector

The SK detector systematic parameters are parametrized in a very similar way to the Flux and ND280 detector systematics; they are normalization parameters correlated between the various Super-K data samples. There are total of 45 parameters, with 44 of them split into ν and $\bar{\nu}$ for 1 ring e (1Re) and 1 ring μ (1Rμ), and ν 1Re-CC1π^+ samples. The last parameter, the SK Energy Scale, is applied to the MC reconstructed energy before the sample binning binning, taking the differences between the reconstructed energy between MC and Data into account. The SK Energy Scale is computed using SK control samples, covering a range between 30 MeV and 10 GeV [99].

The binning is listed in table 4.5, and the full covariance matrix is shown in figure 4.7. The final matrix contains extra uncertainties due to the FSI interactions the secondary interactions (SI) in the detector, as well as the photon-nucleon (PN) interactions. The errors from the detector-only and FSI+SI+PN matrices are summed in quadrature to give the final matrix.

The detector systematics are estimated by the SK group by performing a fit to the atmospheric[2], cosmic, and hybrid π^0 data samples [96, 98, 97], and the FSI+SI+PN matrix is generated by looking at the MC predictions with different deviations of FSI, SI and PN.

[2]The first investigation of a joint atmospheric and beam data oscillation analysis is described in a thesis by Xiaoyue Li, in [95].

Systematic sample	Mode	Bin edges (GeV)
1Re	Osc. ν_e CCQE	0, 0.35, 0.8, 1.25
	$\nu_\mu/\bar{\nu}_\mu$ CCQE	
	Intrinsic ν_e CCQE	
	All NC	
1Rμ	$\nu_\mu/\bar{\nu}_\mu$ CCQE	0, 0.4, 1.1, 30
	$\nu_\mu/\bar{\nu}_\mu$ CCnQE	All
	Intrinsic and Osc. ν_e CCQE	All
	All NC	All
1Re CC1π^+	Osc. ν_e CCQE	0.3, 0.8, 1.25
	$\nu_\mu/\bar{\nu}_\mu$ CCQE	
	Intrinsic ν_e CCQE	
	All NC	

Table 4.5: Super-K systematic uncertainties split per sample, mode and energy bins, making 44 out of the 45 systematic parameters. The last parameter, the SK energy scale, is derived from SK control samples. The 1Re and 1Rμ systematic uncertainties are duplicated, corresponding to FHC and RHC, with 1Re CC1π^+ sample only being present in FHC beam mode.

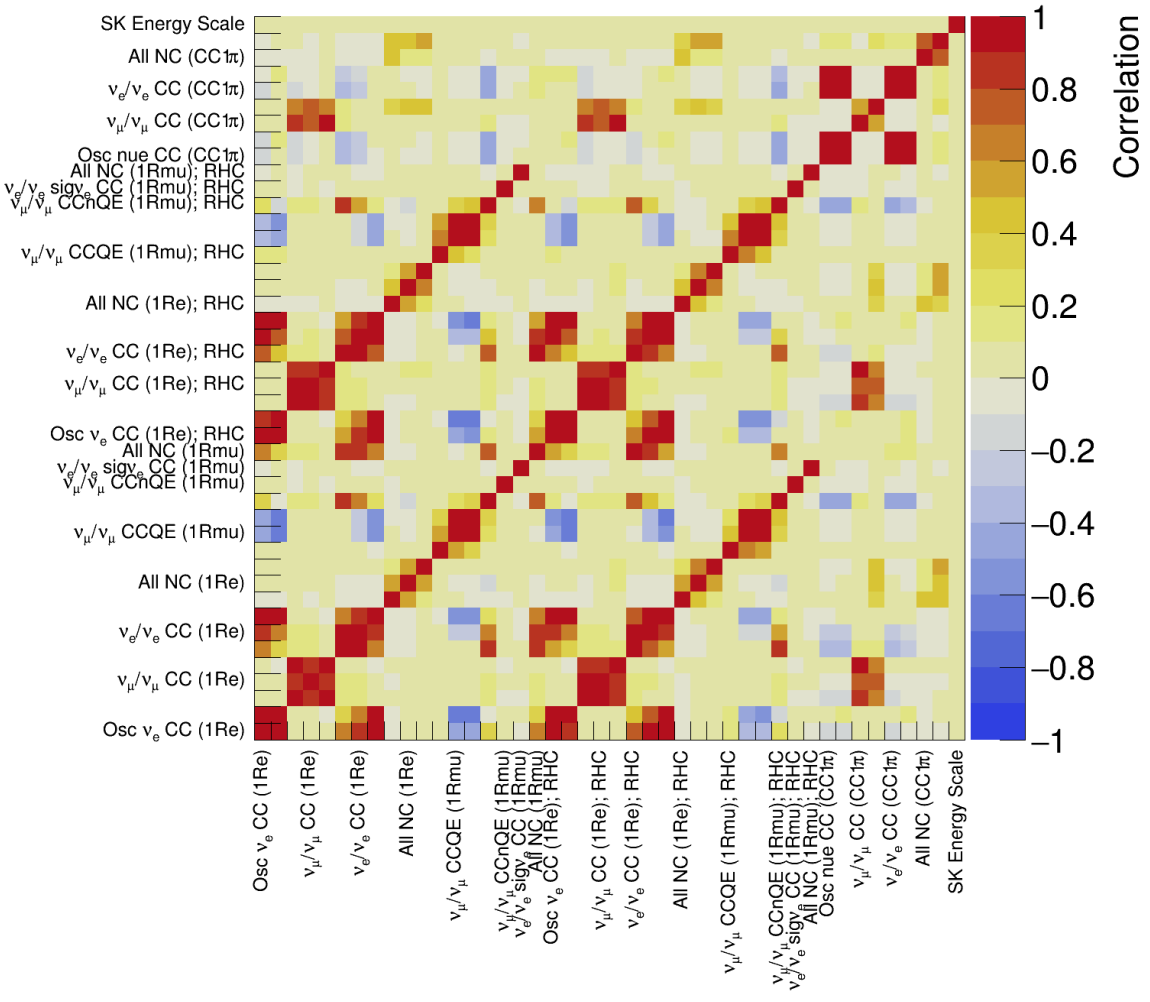


Figure 4.7: SK correlation matrix showing all the SK detector systematic parameters binned in the reconstructed ν energy.

4.2 Data

The data accumulated, measured in proton-on-target (POT), collected over the T2K running period up to this chapter's analysis is shown in figure 4.8, with 1.51×10^{21} POT in the FHC mode and 1.65×10^{21} POT in the RHC mode. This corresponds to the POT as seen by SK, which has almost 100% uptime, but with ND280 being exposed to less POT due to various detector troubles resulting in around 30% downtime. The neutrino beam is on even when ND280 experiences troubles during run-time, allowing SK to collect more data. The impact of this ND280 downtime on the analysis is small, but non-negligible nevertheless. T2K is still a statistically limited experiment, so collecting the neutrino data at SK is most important for the oscillation analysis. At the same time, extra data at ND280 would further constrain the systematic uncertainties, and help reducing potential biases from the neutrino cross-sections. This will become even more important as T2K gets closer to the precision-measurements era.

The ND280 data used for this analysis only include the beam runs 2–6, corresponding to 5.82×10^{20} POT in FHC mode, and 2.84×10^{20} in RHC mode, , all collected between 2010 and 2015. The SK data contains all the data spanning runs 1–9, with 14.94×10^{20} POT in the FHC mode and 16.35×10^{20} POT in the RHC mode. The differences in the SK POT and the total POT in figure 4.8 are due to the SK downtime.

4.2.1 ND280 event selections

The analysis described in this chapter involves fourteen ND280 data samples, seven per FGD, since each FGD has different target material — FGD1 mostly carbon, and FGD2 carbon and water — and because FGD1 has more recorded muon track information available for reconstruction, since it is more massive than FGD2. Each FGD has three samples in FHC and four samples in the RHC modes. The binning in the outgoing muon's p_μ–$\cos(\theta_\mu)$ for these samples is the same as for the ND systematics binning, described in 4.1.3. The correlated FGD1 and FGD2 samples are being fit simultaneously in the near detector fit, allowing for some sensitivity to both oxygen and carbon interactions individually, constraining oxygen interactions at SK. Each FGD has its own separate set

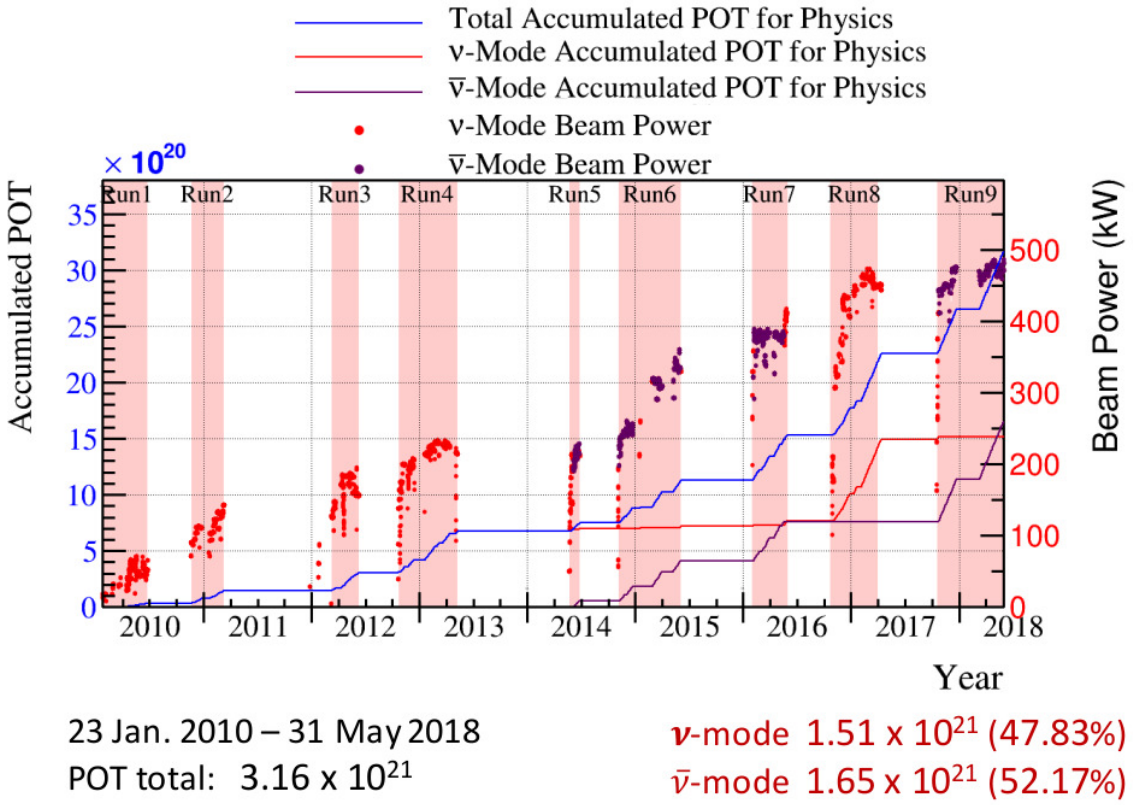


Figure 4.8: The accumulated total (blue line), FHC (red line) and RHC (violet line) POT, together with the beam power, on a timeline over the beam runs 1–9, corresponding to the years 2010–2018.

of selections due to the different target, and systematics and reconstruction [100, 101, 102].

FHC CC samples

There are three CC samples per FGD, named after the number of reconstructed pions: CC0π, CC1π and CCOther. The events reconstructed as CC0π have no secondary tracks recognized as charged pions, protons or positrons in a TPC, and no Michel electrons in the FGD. CC1π sample contains events that have either one charged pion in a TPC and one Michel electron, or one charged pion in a TPC and in an FGD. The events that do not satisfy the other two samples' criteria are tagged as CCOther. They all have to first undergo the same set of cuts to select CC-inclusive samples, before splitting further.

These cuts select a negatively-charged muon track that originated in one of the FGDs, and passed through a neighbouring TPC. It is only after these selections that additional cuts are being applied to separate the CC-inclusive samples into the three sub-samples [100]. The cuts are;

1. **Data Quality:** The event must originate within the correct time-window, inside the beam spill timing, and the whole spill must be within good ND280 data quality flag. Each bunch within a spill is treated independently to avoid pile-up.

2. **Total Multiplicity:** The reconstructed track must pass through at least one TPC.

3. **Quality and Fiducial Volume:** There must be at least one reconstructed track inside the FGD1 or FGD2 fiducial volumes. The track must have originated inside the FGD and passed through a TPC. The FGDs' fiducial volume cuts are listed in [100]. The TPCs is less reliable, and the particles needs to pass at least 18 layers of a TPC to reject short tracks.

4. **Upstream Background Veto:** Events are vetoed where the second-highest momentum track starts less than 150 mm upstream of the reconstructed muon, and if the secondary track originates in FGD2 when the primary track starts in FGD1. This is due to reconstruction failures where an event where the real muon originates far upstream the detector has its reconstructed vertex in one of the FGDs.

5. **Broken Track:** The reconstructed muon vertex must be less than 425 mm from the FGD's upstream edge, if there is a FGD-only track present. This is to reject candidates where the mis-reconstructed track is broken into two, where there should be only one.

6. **Muon PID:** The events that pass the above selections are considered as muon candidates. To be considered as muon, an additional PID selection is applied, based on the measured dE/dx in a TPC, where the deposited energy is compared against the expectation to distinguish between muons, electrons and protons.

When all the CC-inclusive cuts are satisfied, further selections can be applied to split the FHC data into CC0π, CC1π and CCOther samples. As the samples' names suggest,

these selections are based on the number of the reconstructed pions in the detector.The following cuts are applied to identify secondary tracks;

- **PID:** The reconstructed PID must be different from muon from the CC-inclusive cuts.
- **Time Bunching:** The secondary tracks must be in the same spill and beam bunch as the reconstructed muon.
- **Quality and Fiducial Volume:** The secondary track must originate inside the same FGD's fiducial volume as the reconstructed muon, and enter the downstream TPC. As with muons, at least 18 layers of the TPC must be triggered.

After passing the above criteria, the secondary track is identified as either a proton for the positive tracks, and pion or an electron for the negative tracks, using the energy deposited in the TPC. These events can be then split into the three samples, CC0π, CC1π and CCOther.

The wrong-sign background data is not included in the FHC samples. As shown in figure 2.4, the wrong-sign background in FHC mode is much smaller than in RHC mode, with a negligible contribution to the analysis.

RHC CC samples

The RHC mode has two signal samples, $\bar{\nu}_\mu$ CC1-Track and $\bar{\nu}_\mu$ CCN-Track, and two background samples, ν_μ CC1Track and ν_μ CCNTrack, per FGD. The RHC $\bar{\nu}_\mu$ samples have different CC-inclusive selections than the FHC samples because of the higher wrong-sign background, requiring different cuts to separate between $\bar{\nu}_\mu$ and ν_μ interactions [101, 102]. The "Data Quality" and the "Quality and Fiducial Volume" cuts are the same, whereas the "Total Multiplicity" cut requires that at least one track has to pass through TPC2. Additionally, the following cuts are included in the RHC $\bar{\nu}_\mu$ CC-inclusive samples:

- **Positive Multiplicity:** The track with the highest momentum must be positive, to reduce mis-identifying the wrong-sign background interactions as $\bar{\nu}_\mu$.

- **TPC1 Veto:** Reject events entering FGD1 coming from the upstream direction - PØD or the magnet.

- **TPC PID:** The energy deposited as the particle travels in the TPCs, dE/dx, is used to select a μ^+ track.

With the CC-inclusive selection passed, the $\bar{\nu}_\mu$ candidates are split into CC1Track and CCNTrack. The $\bar{\nu}_\mu$ CC1Track selections requires one positive muon final state, and no reconstructed pions. The $\bar{\nu}_\mu$ CCNTrack sample requires the event to have the reconstructed muon to have any number tracks, with at least one reconstructed pion.

The same initial CC-inclusive selection criteria are used for the wrong-sign background, RHC ν_μ CC interactions, with the negative multiplicity (meaning the highest momentum track must have negative charge) and μ^- TPC PID required. The splitting into CC1Track and CCNTrack is then the same as for the RHC $\bar{\nu}_\mu$ samples. The wrong-sign background selections are important because the SK detector does not have a charge-discriminating magnet, so it will not distinguish between positively and negatively charged leptons.

4.2.2 Super-K event selections

The analyses in this thesis are done using the SK data corresponding to runs 1–9. Comparisons against older analyses with less data are also shown, but they all have the same SK event selections applied. All the SK MC and data preparation, including making the data sample selections, is done by the T2K-SK group on T2K. This group consists of people that are mostly members of both the T2K and SK collaborations.

There are three distinct SK data samples; one-ring electron-like (1Re), one-ring muon-like (1Rμ) and one-ring electron CC1π^+-like (1ReCC1π^+). The examples of 1Re and 1Rμ events are shown in figure 2.17, with the 1ReCC1π^+ events often looking like 1Re to human eye, and are reconstructed by searching for PMT hits clustered in time (more on the reconstruction in chapter 6). The FHC mode contains all the three samples, and the RHC mode only contains the 1Re and 1Rμ samples, giving a total of five effective

SK data samples. Before the data is split into samples, it must pass through a series of general cuts that are based on the selected good beam spills, beam spill timing window, SK data quality and so on [104, 99].

All the events must be fully contained (FC) within the fiducial volume (FV). To be classified as FC, the event must have less than 16 hits in the outer detector. Furthermore, there are cuts on the minimum distance between the reconstructed vertex and the inner detector. wall (*dwall*) and the minimum distance between the vertex and the wall in the momentum direction (*towall*). These distances are different for the electron-like and muon-like rings. The event reconstruction at SK is done with the fiTQun software, and is described further in chapter 6.

1R*e* selection

To be classified as 1Re, the event must pass the following selection criteria:

1. **Fiducial Volume:** The event must satisfy the FC cut, have *dwall* above 60 cm, and *towall* above 170 cm.

2. **Ring Count:** The most likely number of rings reconstructed by the fiTQun multi ring fitter must be one.

3. **Ring ID:** The most likely PID of the ring reconstructed by the fiTQun single-ring fitter must be e.

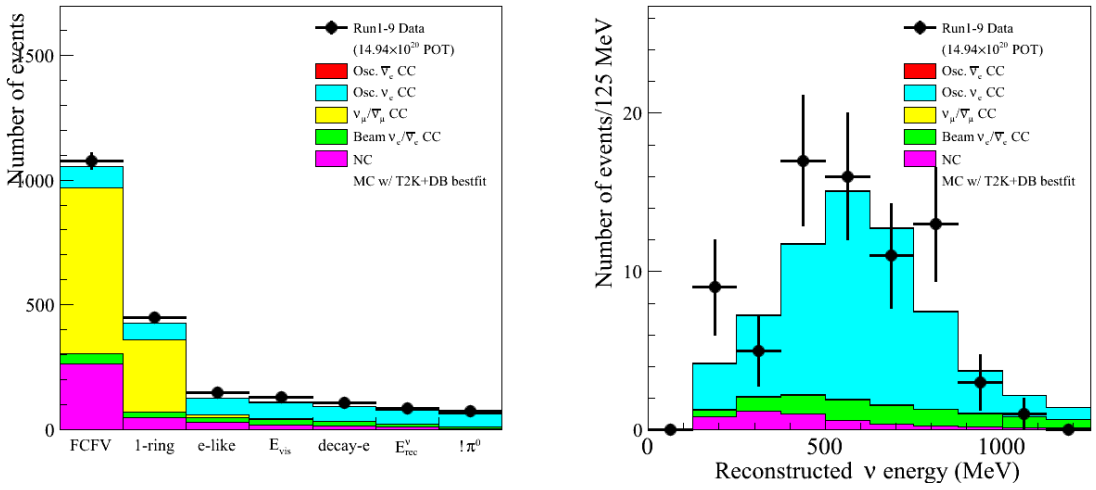
4. $\mathbf{E_{vis}}$**:** The visible energy must be above 100 MeV (in practice, one-ring electron momentum is used as the visible energy).

5. **Michel** $\boldsymbol{e}$**:** No detected Michel electrons, by looking at separate PMT hit clusters spaced in time.

6. $\mathbf{E_{\nu}^{rec}}$**:** The reconstructed neutrino energy must be below 1250 MeV

7. π^0 **Rejection:** The log likelihood ratio between π^0 and e hypotheses, $\ln(\mathcal{L}_{\pi^0}/\mathcal{L}_e)$, must be below $175 - 0.875 \times m_{\gamma\gamma}$.

The *dwall*, *towall* and the π^0 rejection cuts were empirically derived from the fits to the SK atmospheric data and the T2K MC separately, and optimized for the δ_{CP} measurements [98, 105]. The π^0 events can be easily mis-reconstructed as 1Re, especially at lower momenta, making the π^0 cut necessary. The 1Re selection sequence with the E_ν^{rec} distribution post-selections are shown in figure 4.9 for the FHC events.



(a) 1Re sample selections sequence with data overlaid against the MC.

(b) 1Re reconstructed energy for the events that passed all the selection criteria, with the data overlaid against the MC.

Figure 4.9: The 1Re selection sequence one the left, with the reconstructed ν energy spectrum for the events that passed all the selections on the right. The MC and data are from the FHC mode.

1Rμ selection

The 1Rμ sample selection criteria are quite similar to the 1Re sample, but without the upper limit on the muon reconstructed energy, and cuts on the momentum rather than the visible energy (which in practice is the same, since electron momentum is used as the visible energy in 1Re hypothesis):

1. **Fiducial Volume:** The event must satisfy the FC cut, have *dwall* above 50 cm, and *towall* above 250 cm.

2. **Ring Count:** The most likely number of rings reconstructed by the fiTQun multi ring fitter must be one.

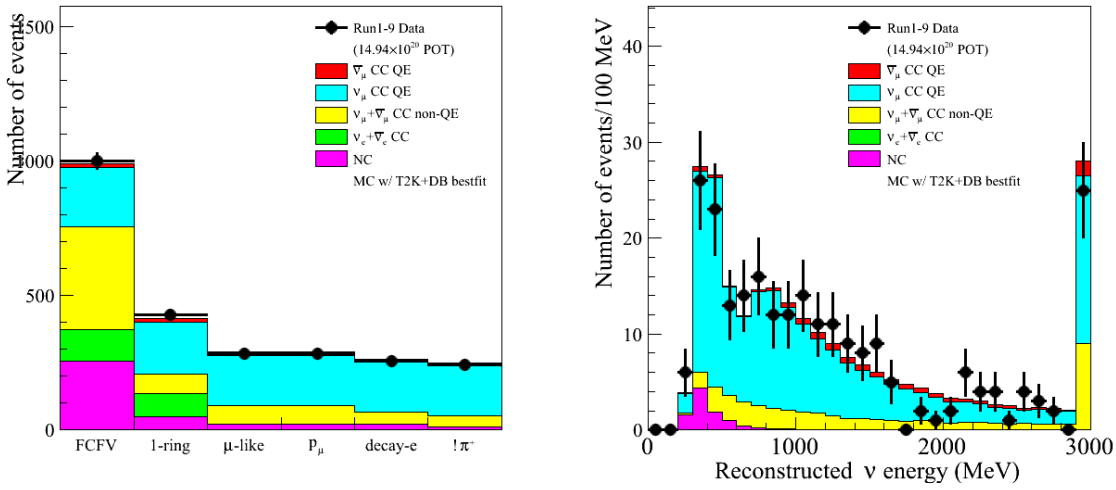
3. **Ring ID:** The most likely PID of the ring reconstructed by the fiTQun single-ring fitter must be μ.

4. **Momentum Cut:** The reconstructed μ momentum must be larger than 200 MeV/c.

5. **Michel *e*:** Less than two Michel electrons.

6. π^+ **Rejection:** The event must pass a linear 2D cut in $\ln(\mathcal{L}_{\pi^+}/\mathcal{L}_{\mu})$ (based on MC-derived cuts on the log-likeligood difference between the two hypotheses, which will be described further in section 6.2.2) and the reconstructed muon momentum.

Figure 4.10 shows the 1Rμ selection sequence and the reconstructed ν energy after the selections, for the FHC mode only.



(a) 1Rμ sample selections sequence with data overlaid against the MC.

(b) 1Rμ reconstructed energy for the events that passed all the selection criteria, with the data overlaid against the MC.

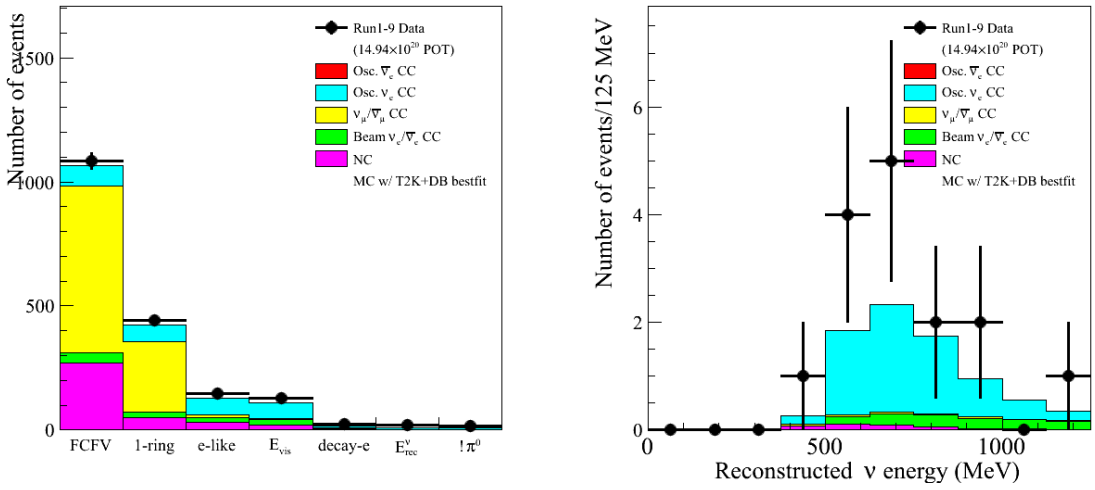
Figure 4.10: The 1Rμ selection sequence one the left, with the reconstructed ν energy spectrum for the events that passed all the selections on the right. The MC and data are from the FHC mode.

1Re CC1π^+

The 1Re CC1π^+ sample has very similar selection criteria as the 1Re sample:

1. **Fiducial Volume:** The event must satisfy the FC cut, have *dwall* above 50 cm, and *towall* above 270 cm.
2. **Ring Count:** The most likely number of rings reconstructed by the fiTQun multi ring fitter must be one.
3. **Ring ID:** The most likely PID of the ring reconstructed by the fiTQun single-ring fitter must be e.
4. $\mathbf{E_{vis}}$**:** The visible energy must be above 100 MeV.
5. **Michel** $\boldsymbol{e}$**:** The event must have one decay electron.
6. $\mathbf{E_\nu^{rec}}$**:** The reconstructed neutrino energy must be below 1250 MeV
7. π^0 **Rejection:** The log likelihood ratio between π^0 and e hypotheses, $\ln(\mathcal{L}_{\pi^0}/\mathcal{L}_e)$, must be below $175 - 0.875 \times m_{\gamma\gamma}$.

After the event is found to have a Michel electron in cut 5, the $\mathrm{E}_\nu^{\mathrm{rec}}$ reconstruction algorithm applied in fiTQun is different from the one applied in 1Re sample. It takes into account that CC1π^+ events have an outgoing Δ^{++} instead of a proton [103]. Figure 4.11 shows the 1ReCC1π^+ cut sequence, and the reconstructed neutrino energy spectrum after the selections for FHC mode.



(a) (b)

Figure 4.11: The 1ReCC1π^+ selection sequence one the left, with the reconstructed ν energy spectrum for the events that passed all the selections on the right, both with the data overlaid against the MC. The MC and data are for the FHC beam mode.

4.3 Validations and sensitivities

After the validations and ND280 data fits are done by MaCh3 and BANFF, and the SK data samples are prepared by the T2K-SK group, validations and sensitivities with the SK frameworks can be done. In MaCh3 the ND280 code and samples are progressed to the joint-fit executables that perform the validations, checks and fits using both the ND280 and SK data samples simultaneously. Separate pre-fit flux, cross-section matrices are used as priors in the fits. The post-fit BANFF matrix is usually only used in MaCh3 to produce tables with the predicted event rates to be compared against the other groups groups (and vice-versa, the other groups can use the pre-fit matrices, but only for some validations against MaCh3). This is in contrast to VALOR and P-Theta, which use the BANFF post-fit matrix that constrains the flux and neutrino cross-sections in all of their fits.

4.3.1 Predicted event rates

The first validation involves producing the predicted event rates for every SK sample using different oscillation parameters and the ND280-pre-fit – ND280-post-fit covariance matrix combinations. This is not just a validation tool between the three fitters; it also shows the effect of the ND280 fit on the SK samples, and shows what we should expect to change with respect to the previous analyses per sample, before doing the fit.

Table 4.6 shows the predicted event rates for each sample, where the SK MC was tuned to nominal parameters both before and after the ND280 fit for the oscillated hypotheses, and to the post-ND280 fit for the unoscillated hypothesis. The selected oscillation parameters, called Asimov A, are set to the results from the previous oscillation analyses [106] and to the value of θ_{13} taken from the PDG 2018 fit to the reactor neutrino oscillation experiments [18]. The Asimov A values are:

- $\sin^2\theta_{12} = 0.304$
- $\sin^2\theta_{23} = 0.528$
- $\sin^2\theta_{13} = 0.0212$

- $\Delta m^2_{21} = 7.53\times10^{-5}$ eV^2
- $\Delta m^2_{32} = 2.509\times10^{-3}$ eV^2
- $\delta_{CP} = -1.601$

It is clear from the table that the ND280 has a huge effect on the predicted oscillated event rates through the tuning of the systematic parameters. This is true for all the different oscillation parameter sets (Asimov A, Asimov B, unoscillated), with the full tables shown in appendix A. The predicted number of events usually match very well between the three fitters – especially the unoscillated predictions, since the three fitters use different oscillation probability calculators - with MaCh3 using the Prob3++ package [108] modified to be parallelised and to do the calculations on a GPU [107].

There are further validations between the three groups, done without the ND280 constraint, where each group sequentially deviates every systematic parameter separately by $\pm1\sigma$ and $\pm3\sigma$, and compares the resultant difference in the neutrino energy spectrum for each sample separately against the nominal prediction and between the groups.

Table 4.6: The predicted number of events for all five SK samples. The table contains the event rates for the pre-ND280 and post-ND280 fits with the oscillated Asimov A input parameters, and the post-ND280 fit unoscillated predictions. The neutrino event rates per sample are split into all neutrino flavours together with the right and wrong-sign events to show all the backgrounds. For example, the FHC 1Rμ sample's signal are the ν_μ events, with the largest background coming from the $\bar{\nu}_\mu$ events, and very little from the miss-reconstructed ν_e events.

Sample	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
Pre-ND280 fit, Asimov A							
FHC 1Rμ	234.74727	0.23317	15.38855	0.02300	0.05847	0.00028	250.45072
FHC 1Re	4.71632	7.78573	0.22836	0.35524	50.73694	0.34333	64.16593
FHC 1ReCC1π^+	0.51896	1.06192	0.02443	0.00970	6.17675	0.00685	7.79861
RHC 1Rμ	49.90198	0.08557	80.47949	0.05620	0.00325	0.00441	130.53090
RHC 1Re	0.92420	1.43065	1.56390	2.27242	2.57289	6.97070	15.73475
Post-ND280 fit, Asimov A							
FHC 1Rμ	255.98572	0.20220	16.08909	0.02094	0.07116	0.00027	272.36939
FHC 1Re	3.98182	9.19099	0.20280	0.38955	58.66736	0.36161	72.79413
FHC 1ReCC1π^+	0.48109	0.96424	0.02416	0.01074	5.38527	0.00739	6.87289
RHC 1Rμ	55.17938	0.07670	84.10914	0.04984	0.00379	0.00471	139.42356
RHC 1Re	0.81559	1.71799	1.37963	2.45093	2.97063	7.43509	16.76987
Post-BANFF fit, Unoscillated							
FHC 1Rμ	1193.75076	0.20262	31.18278	0.02095	0.00000	0.00000	1225.15712
FHC 1Re	5.08870	9.85319	0.22136	0.41113	0.00000	0.00000	15.57439
FHC 1ReCC1π^+	1.49812	1.03277	0.03689	0.01120	0.00000	0.00000	2.57898
RHC 1Rμ	114.79849	0.07686	343.46537	0.04998	0.00000	0.00000	458.39070
RHC 1Re	0.90775	1.83306	1.66775	2.59852	0.00000	0.00000	2.57898

4.3.2 Sensitivity fits

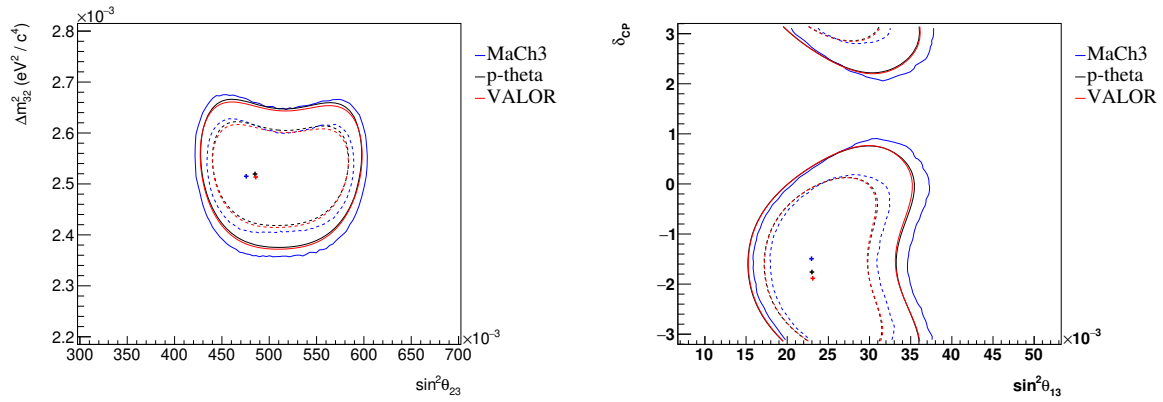
The sensitivity fits (called Asimov fits or Asimov sensitivity from here) were done to the run 1–9 SK MC to see what differences we should expect with respect to the previous analysis, and to further validate the three fitters before progressing to the data fits. The SK MC is reweighted to the correct POT, the nominal systematic parameter values, and a set of oscillation parameters – in this case only shown for Asimov A. This reweighted MC is then used as the data – meaning that we effectively fit the MC to a fixed representation of itself [109].

The Asimov fits were done both with and without the external constraint on θ_{13} from the short-baseline reactor neurino experiments (see section 1.5), using the Metropolis-Hastings MCMC algorithm described in chapter 3. Although the data fits are expected to look similar to the Asimov fits, the last update of the Asimov values was in 2016 for the θ_{13} and 2015 for the rest of the oscillation parameters – the T2K data without the reactor constraint also produces different result than the reactor experiments, furthering the difference between the Asimov and data fits.

T2K data only

Figure 4.12 shows the confidence interval contour comparisons between the three fitters for the paired atmospheric and appearance oscillations parameters, using the T2K Asimov data only – without the reactor constraint on θ_{13}. The agreement between the three analyses is relatively good, with the largest differences coming from MaCh3. It is not easy to pinpoint the source of these deviations, due to the number of differences between MaCh3 and the other two fitters. MaCh3 performs a fit simultaneously to all the oscillation and systematic parameters, including all of ND280. This allows for a non-Gaussian treatment of the systematics, which can then translate into interesting marginalization effects when only plotting, say, two parameters against each other. Furthermore, there is a large difference in the interpretation between the Frequentist best-fit-point and the Bayesian highest-posterior. The best-fit points defines the point of the highest likelihood, whereas the "highest-posterior" used in MCMC is based on the density of the MCMC steps on the parameter space, not the actual posterior probability for each MCMC step. This means

that the "highest-posterior" might change for a parameter when ploted against another, and is a subject to marginalization effects.



(a) Asimov A 68% and 95% confidence-level comparisons without the reactor constraint on θ_{13}: NH Δm^2_{32}–$\sin^2\theta_{23}$

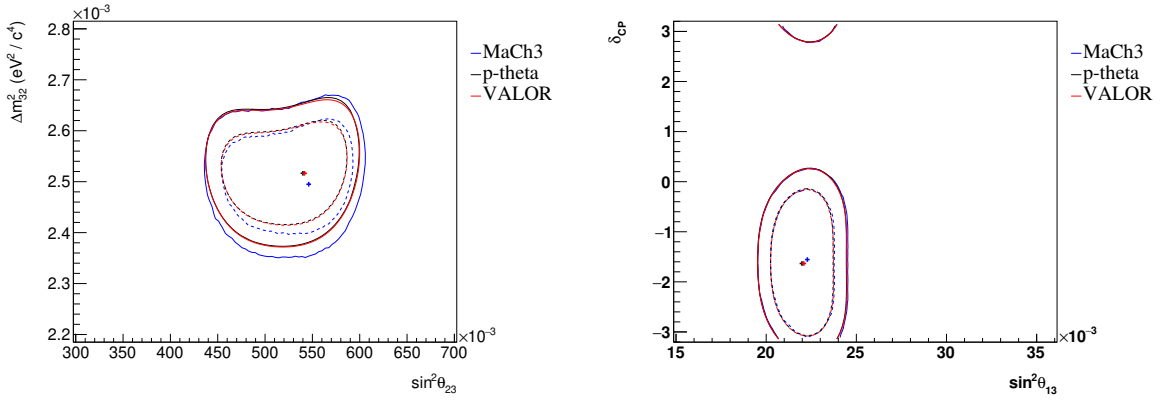
(b) Asimov A 68% and 95% confidence-level comparisons without the reactor constraint on θ_{13}: NH δ_{CP}–$\sin^2\theta_{13}$

Figure 4.12: Three-fitter comparisons of the confidence levels for the normal mass hierarchy $\sin^2\theta_{23}$–Δm^2_{32} and $\sin^2\theta_{13}$–δ_{CP}. All the intervals are from the run 1–9 Asimov A fit, without the reactor constraint on θ_{13}. There is a good agreement between the fitters, with the largest differences coming from MaCh3.

With the reactor constraint on θ_{13}

Figure 4.13 shows the comparisons of the confidence intervals between the three groups, extracted from the Asimov fits with the reactor constraint on θ_{13} – shown for normal hierarchy only. The agreement between the three fitters is better with the external constraint, with some small differences remaining in the muon neutrino disappearance channel (figure 4.13a).

It is interesting to note that we can use the fit without the reactor constraint, and reweight the posterior probability used to make figure 4.12 (where we see slightly worse agreement between MaCh3 and the rest of the fitters) from a prior flat in θ_{13} to the reactor constraint prior, and still obtain the contours the same as in the figure 4.13b (where we see a good agreement between all the three fitters). This gives an extra confidence that both contours are correct. More discussions about prior-reweighting in section 4.4.3.



(a) Asimov A 68% and 95% confidence-level comparisons with the reactor constraint on θ_{13}: NH $\sin^2\theta_{23}$ –Δm^2_{32}

(b) Asimov A 68% and 95% confidence-level comparisons with the reactor constraint on θ_{13}: NH $\sin^2\theta_{13}$ –δ_{CP}

Figure 4.13: Three-fitter comparison of the confidence levels for the normal mass hierarchy $\sin^2\theta_{23}$–Δm^2_{32} and $\sin^2\theta_{13}$–δ_{CP}. All the intervals are from the run 1–9 Asimov A fit, with the constraint on θ_{13} from the reactor neutrino experiments, taken from PDG 2018. There is a good agreement between the three fitters, with the largest differences coming from MaCh3 in the disappearance channel.

4.4 Joint detector run 1–9 oscillation analysis

4.4.1 Changes in the analysis

The run 1–9 data analysis was originally planned to have larger changes to the cross-section treatment and extra ND280 data up to run 8. Although the new cross-section changes were implemented and validated based on work described in [114], all these changes were retracted due to bugs in the ND280 code. These changes were progressed to the 2020 analysis instead (with the run 1–10 data, not described in this thesis), presented at Neutrino 2020 conference.

The only difference to the systematic treatment in the run 1–9 oscillation analysis was an addition of a post-fit smearing procedure on the contours of the Δm^2_{32} parameter, and an addition of an effective binding energy (E_{b})parameter, both based on the biases found in a fake-data studies done by the P-Theta group, described further below.

The reactor constraint was updated from the 2016 PDG [111] values to the new 2018 PDG [18], which has seen a shift in the best fit point in $\sin^2\theta_{13}$ and considerably tighter constraints. MaCh3, along with the other OA groups, has originally produced fit results 1–9d SK data with the old reactor constraint, which allowed for direct comparisons between the two reactor constraints. The posterior probability distribution was reweighted to new 2018 reactor constraint for $\sin^2\theta_{13}$. The validations and plots showing the effect of the new reactor constraint on the T2K results are described in section 4.4.3.

4.4.2 Biases in $\sin^2\theta_{23}$ and Δm^2_{32}

A study involving fake data generated from the T2K Monte Carlo with an alternative cross-section model was done by the P-Theta group, together with the near-detector fitting groups, BANFF and MaCh3, to evaluate the effect of different cross-section models on the SK fit [115]. With fifteen different models, fake data can be generated and studied with the fitters to see whether there are any biases with respect to the standard Asimov (unfluctuated MC) fit. For the 2016–2019 analyses, large biases in $\sin^2\theta_{23}$ and Δm^2_{32} were found with the fake dataset containing the nucleon binding energy needed to remove a nucleon from the nucleus, parametrized as E_b. This removal energy shifts the expected final lepton energy seen in the detector.

Biases in $\sin^2\theta_{23}$

The $\sin^2\theta_{23}$ parameter is non-Gaussian, mainly due to the degeneracy of the upper and lower octant (which will be shown later in figure 4.25), and mapping of the biases in different parts of phase-space for this parameter is non-trivial, making any fast or simple solution impossible. At the same time, the full E_b treatment, based on shifting the lepton momenta in respect to the nucleon binding energy for different nuclei (for example Oxygen and Carbon), for different values of E_b, was too difficult to implement and validate on a short time-scale. The E_b parameter was therefore derived from the differences between the E_b fake data and Asimov data, implemented in a form of a response function binned in the interaction mode and the fitting variables.

Biases in Δm^2_{32}

The 1D Δm^2_{32} distribution is very close to a Gaussian, and the bias is easier to map in different parts of the phase-space. The solution to tackle this bias was to perform a post-fit smearing of the parameter contours, assuming the parameter's Gaussianity. In case of MaCh3, the smearing procedure involves adding a random Gaussian number to the value of Δm^2_{32} for each MCMC step separately after the fit. The random Gaussian value is centered at 0 and with a width of the overall bias on Δm^2_{32} extracted from the fake data studies. This was calculated by summing in quadrature all the biases between the nominal Asimov fit and the different fake data fits produced with the alternative cross-section models. The total bias (and therefore the width of the smearing Gaussian applied in MaCh3) came to be $3.92 \times 10^{-5}\,\mathrm{eV}^2$

4.4.3 New constraint on $\sin^2\theta_{13}$ from the reactor experiments

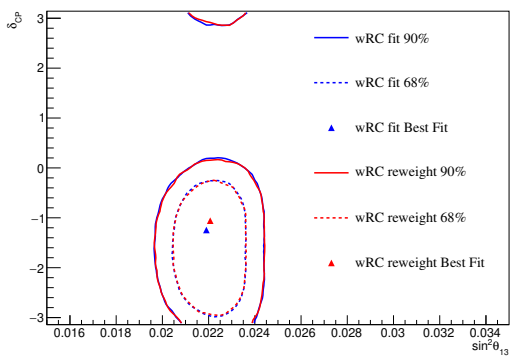
The Particle Data Group released a new constraint on $\sin^2\theta_{13}$ given by combining the results from three reactor experiments; Double Chooz, Daya Bay and RENO. It was decided that the T2K reactor constraint that was used was going to be updated to the new 2018 value, and the decision was made half-way through the already-running analysis. One of the advantages of using a Bayesian framework that returns the full posterior density, is that it enables easy manipulation of the phase-space after the sampling is finished. In this case, it was possible to reweight the already-running fits (and do extra studies in the process) to the new reactor constraint.

The Bayes theorem dictates that the posterior probability distribution is proportional to the likelihood distribution multiplied by the prior distribution. It is therefore possible to divide-out the used posterior distribution by re-weighting any binned parameter histograms, recreating the without-reactor constraint fit. We can then reweight again using the new reactor prior, giving us fully marginalized posterior distribution with a new prior on $\sin^2\theta_{13}$.

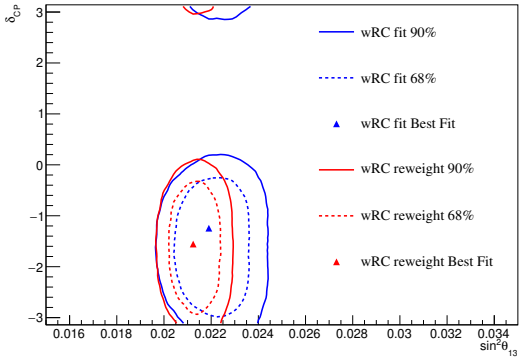
Figure 4.14 shows the plots (all normal MH) produced during the reweigting procedure that illustrate — in a more intuitive way — how the method was validated.

Plot 4.14a shows a fit without the reactor constraint on θ_{13} reweighted to the PDG 2016 reactor constraint, compared against the actual reactor-constrained fit done with PDG 2016. Given that these contours (and the intervals for all the other parameter combinations) match, reweighting from a prior flat in θ_{13} to any reactor constraint work. In this case we can, as it is done in figure 4.14b, easily reweight no-reactor-constraint fits to the new PDG 2018 and see the differences. As expected, the best fit point moved to lower values of $\sin^2\theta_{13}$ and has tighter intervals. We can use these contours to validate reweighting from non-flat reactor constraint (PDG 2016) to another (PDG 2018), which is what was needed for the already-long-running data fits with the older reactor constraint as the prior on $\sin^2\theta_{13}$. Figure 4.14c shows reactor-constrained PDG 2016 fit reweighted to the PDG 2018 constraint against the already-validated fit with a prior flat in $\sin^2\theta_{13}$ reweighted to PDG 2018. Given these match for all the parameters and mass hierarchies, it is feasible to reweight the data-fits from the PDG 2016 reactor constraint to PDG 2018. Figure 4.14d shows that the effect of the new reactor constraint on δ_{CP} (red) is relatively small with respect to the old constraint (blue).

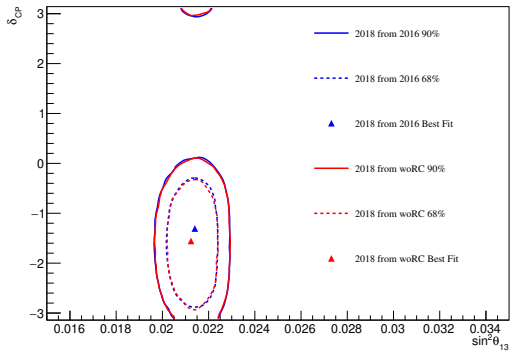
This procedure does allow for reweighing the fits if new constraints appear, or to test completely different constraints on any other parameters, but there are limitations to what can be treated this way. We need to have a substantial number of MCMC steps in the region of the new constraint that we are reweighing to, otherwise the reweighting procedure could result in contours in regions were we do not have enough statistics to make any meaningful conclusions.



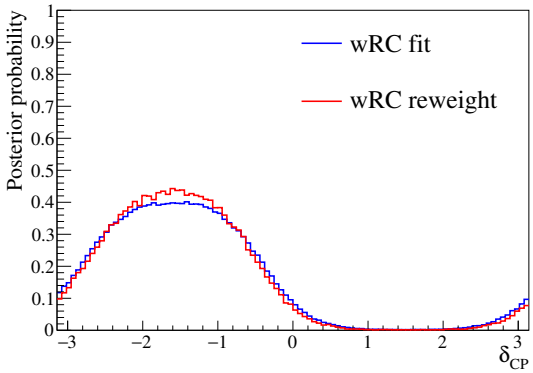
(a) Without-reactor-constraint appearance contours reweighted to the PDG 2016 values (red, woRC reweight) against the PDG 2016 fit (blue, wRC fit). Given both contours match, the reweighting scheme works.



(b) PDG 2016 reactor-constrained appearance contours reweighted to the PDG 2018 values (red, wRC reweight) against the PDG 2016 fit (blue, wRC fit). Given the previous check worked, these should be the correct intervals with the new constraint.



(c) Reactor-constrained appearance contours with PDG 2016 constraint reweighted to PDG 2018 (blue, 2018 from 2016) against the flat-prior fit reweighted to PDG 2018 (red, 2018 from woRC). Given all these match, it is now possible to use the existing MCMC chains with the old reactor constraint.

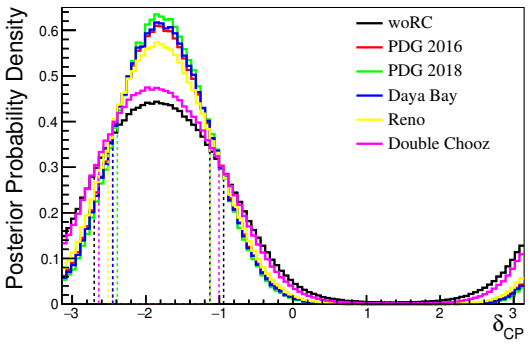


(d) 1D δ_{CP} posterior probability with the PDG 2018 reactor constraint (red, wRC reweight) against the PDG 2016 reactor constraint (blue, wRC fit). The new reactor constraint gives marginally higher sensitivity to δ_{CP}.

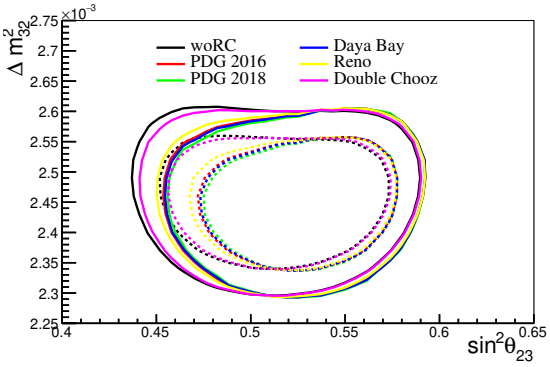
Figure 4.14: Validation plots showing posterior probabilities for appearance parameters from Asimov A run 1–9 fit reweighted to different PDG reactor constraints on $\sin^2\theta_{13}$.

The new PDF 2018 reactor constraint on $\sin\theta_{23}$ comes from the weighted least square of the results from the Daya Bay, Reno and Double Chooz experiments. The Double Chooz experiment has two separate results from different target materials; Gadolinium and Hydrogen, with the results on Hydrogen that seemed to be quite different from all the other reactor experiments. This prompted a further study, where the δ_{CP} and $\sin^2\theta_{23}$ vs Δm^2_{32} posteriors were produced with various reactor constraints to see how different reactor experiments and total constraints from the PDG compare against each other.

This study was done using the same prior-reweighting procedure described earlier in this section, with the results shown in figure 4.15a — where we see the δ_{CP} posterior probabilities against each other produced using different reactor constraints. We can see that Reno and Daya Bay alone produce contours similar to the joint PDG values, whereas Double Chooz constraints have a very small impact. Figure 4.15b provides the same conclusions from the disappearance contours. We can see that the contours in $\sin^2\theta_{23}$ prefer the higher octant (so the higher values of $\sin^2\theta_{23}$) with stronger $\sin^2\theta_{13}$ constraint, which makes sense. These two parameters can have a similar effect on the predicted number of ν_e and $\bar{\nu}_e$ events, meaning that if, for example, $\sin^2\theta_{13}$ decreases, $\sin^2\theta_{23}$ should increase (if all the other parameters are fixed). The new reactor constraint prefers lower values of $\sin^2\theta_{13}$, so our preference towards the higher octant of $\sin^2\theta_{23}$ gets stronger. Given that both RENO and Daya Bay agree with each other, whereas Double Chooz agrees with the T2K-only data, there are no discrepancies or reasons to change the reactor constraint to values different than the ones provided by the PDG group.



(a) δ_{CP} posterior probability with 1 σ contours (dashed line) produced without the reactor constraint on $\sin^2\theta_{32}$ (black line) against 2016 and 2018 PDG values (red and green, respectively) and constraints from the three reactor constraints separately



(b) $\sin^2\theta_{23}$ vs Δm^2_{32} posterior probability with 68% and 90% contours (dashed and solid lines, respectively) produced without the reactor constraint on $\sin^2\theta_{32}$ (black line) against 2016 and 2018 PDG values (red and green, respectively) and constraints from the three reactor constraints separately

Figure 4.15: Validation plots showing posterior probabilities for δ_{CP} and the disappearance contours produced with various reactor constraints against each other

4.4.4 Results with T2K-only data; without the reactor $\sin^2\theta_{13}$ constraint

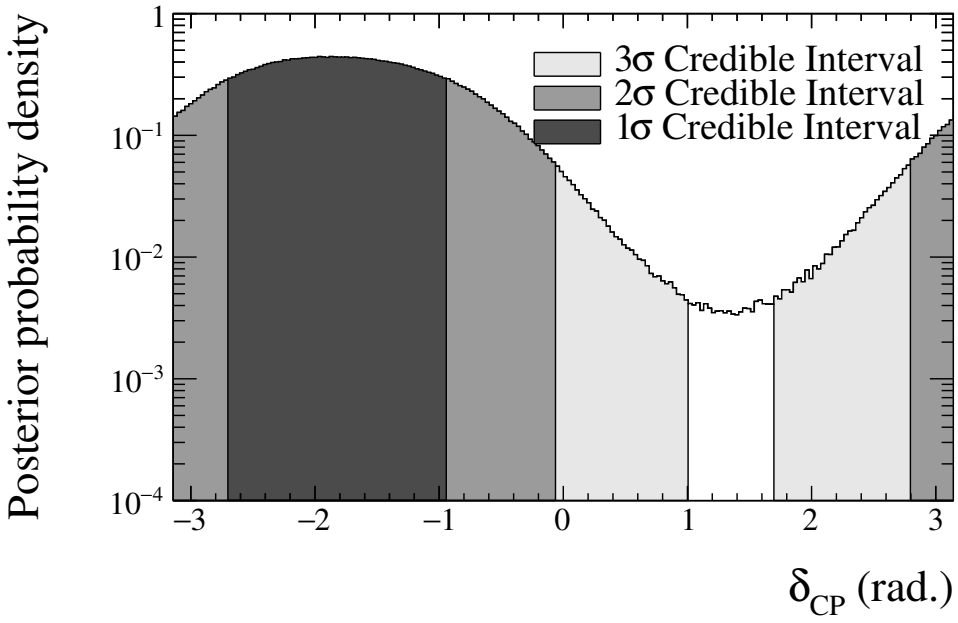
The data samples were first fit with the T2K-only data, so without the reactor constraint on $\sin^2\theta_{23}$. The Markov Chain to produce the posterior probability density contains ∼22 million MCMC steps with an autocorrelation of 20,000, providing about 1,100 independent samples.Although this is not enough for stable $3\,\sigma$ credible intervals, the expected variations are small. The best-fit PMNS-model parameters together with their 1, 2 and $3\,\sigma$ intervals, marginalized over both mass hierarchies, are shown in Table 4.7. Given these are not the main output from the T2K experiment, not all the without-reactor-constraint results are presented here.

	$\sin^2\theta_{23}$	Δm^2_{32} ($\times 10^{-3}$eV^2)	$\sin^2\theta_{13}$	δ_{CP} (rad.)
2D best fit	0.507	2.46	0.0256	-1.93
68.27% C.I. ($1\,\sigma$) range	0.473 – 0.547	2.35 – 2.58	0.0233 – 0.0334	-2.70 – -0.94
95.45% C.I. ($2\,\sigma$) range	0.447 – 0.580	2.30 – 2.63 & -2.64 – -2.34	0.0192 – 0.0399	$-\pi$ – -0.06 & 2.76 – π
99.73% C.I. ($3\,\sigma$) range	0.423 – 0.600	2.24 – 2.70 & -2.71 – -2.30	0.0159 – 0.0476	$-\pi$ – 1.01 & 1.70 – π

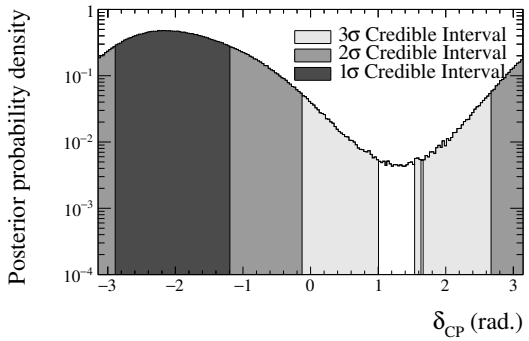
Table 4.7: Best-fit values with 1, 2 and $3\,\sigma$ 1D credible interval ranges for oscillation parameters for the data fit without reactor constraint. The 2D best-fit values are taken from the mode of the 2D marginal posterior distributions in $\sin^2\theta_{23}$ –Δm^2_{32} and $\sin^2\theta_{13}$ –δ_{CP}. All these values are marginalized over both mass hierarchies.

Figure 4.16 shows the binned posterior probability density for δ_{CP} marginalized over both MH (top), NH (left) and IH (right), with 1, 2 and $3\,\sigma$ credible intervals marked. Interestingly, the CP conserving value of $\pi/2$ is rejected with $3\,\sigma$ credible intervals when marginalized over both MH or IH but not when marginalized over NH-only. The $3\,\sigma$ contours are however closed in all the cases, which is different from all the previous analyses. It seems like T2K-data is becoming "competitive" against the fits with the reactor constraint, and has enough "strength" to start drawing strong conclusions about δ_{CP} by itself.

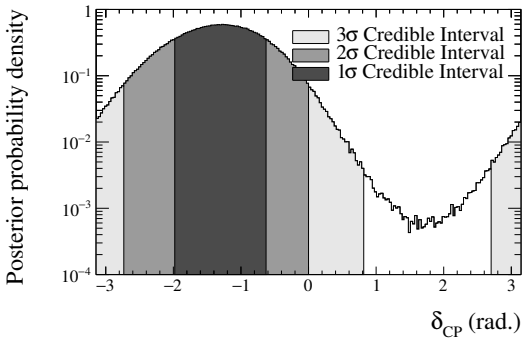
Figure 4.17 shows the 2D credible intervals for the appearance parameters; $\sin^2\theta_{13}$ against δ_{CP}. Interestingly enough, the $3\,\sigma$ contours are not closed anymore when marginalized over both MH or NH - as was the case with 1D δ_{CP} posterior. The rejected area is still far larger than in the run 1–8 analysis, which is promising.



(a) Posterior probability for δ_{CP} marginalized over both mass hierarchies. A large phase-space of the CP conversation rejected at $2\,\sigma$, although there is no full 0–π excluded yet



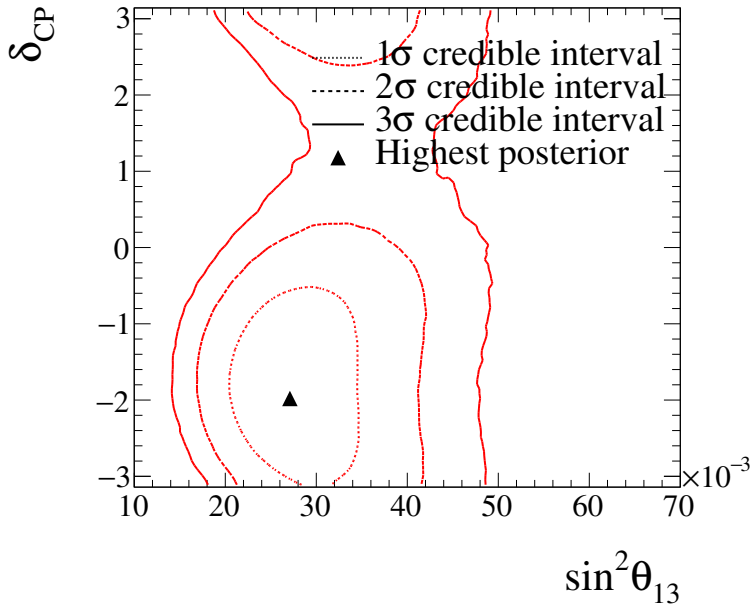
(b) Posterior probability for δ_{CP} for NH. Similar to the plot marginalized over both hierarchies due to high preference towards NH



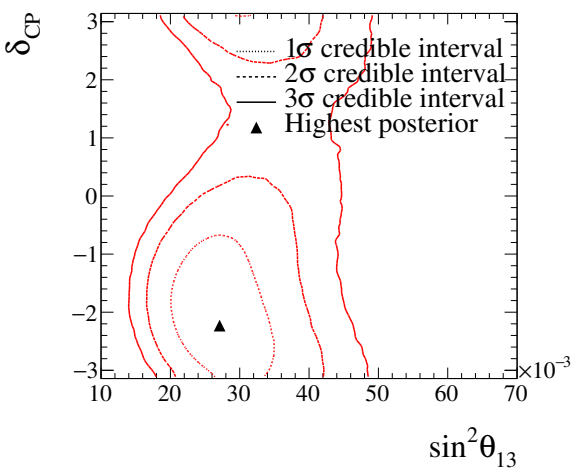
(c) Posterior probability for δ_{CP} for IH. CP conserving values for IH δ_{CP} are rejected with 2σ Credible Interval

Figure 4.16: Posterior probability on log axis for δ_{CP} marginalized over both NH and IH with 1, 2 and 3σ credible intervals from 1–9d SK data fits.

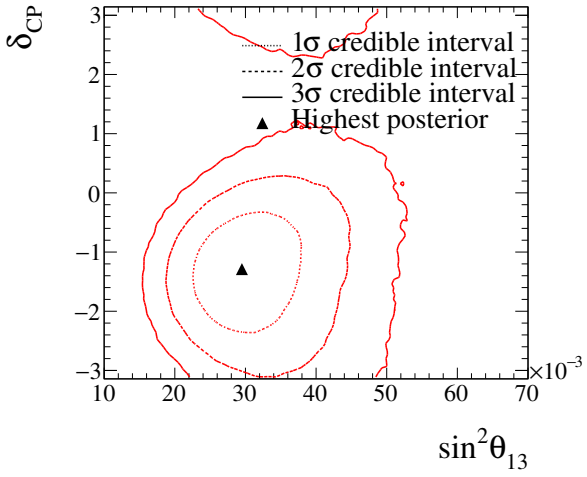
Figure 4.18 shows the remaining oscillation parameters of T2K interest against each other. The best-fit value of $\sin^2\theta_{23}$ is in the higher-octant, which now contains most of the posterior probability. This, again, is different from the run 1–8 analysis, where most of the posterior probability was in the lower octant. This is a positive change: T2K data-



(a) Both mass hierarchies



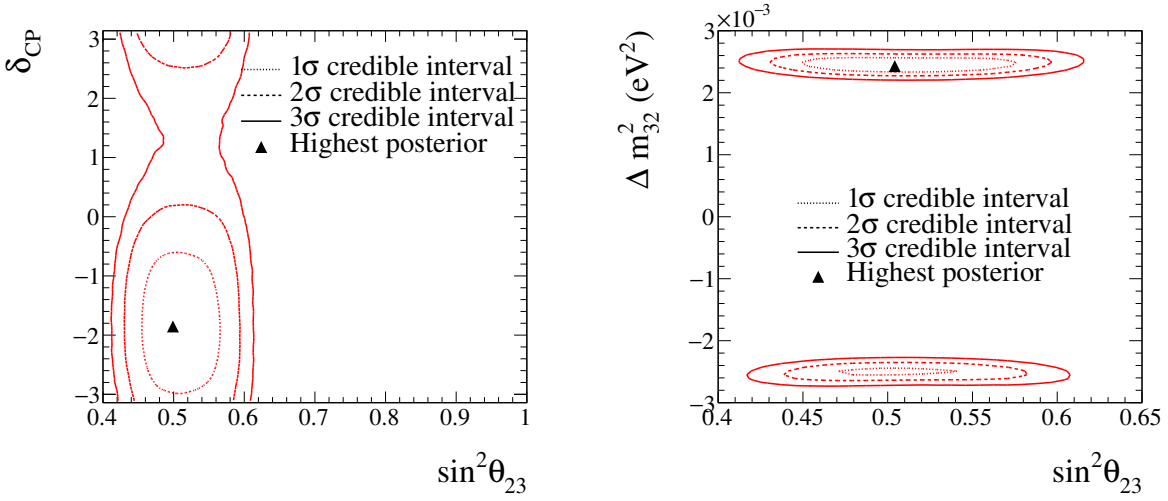
(b) Normal hierarchy



(c) Inverse hierarchy

Figure 4.17: 2D Credible Intervals for $\sin^2\theta_{13}$ against δ_{CP} marginalized over both, normal and inverse mass hierarchies, respectively.

MC agreement and with-without reactor constraint fits agreement are becoming better with more data. The $1\,\sigma$ contours of Δm^2_{32} now cover less space in the IH than before, giving us more preference towards the NH. This is more visible in the fits with the reactor constraint, described in the next section.



(a) Credible intervals for $\sin^2\theta_{23}$ against δ_{CP} marginalized over both mass hierarchies, from the run 1–9d data fit without the reactor constraint.

(b) Credible intervals for $\sin^2\theta_{23}$ against Δm^2_{32} marginalized over both mass hierarchies, from the run 1–9d data fit without the reactor constraint.

Figure 4.18: Credible intervals for $\sin^2\theta_{23}$ against δ_{CP} (left) and the disappearance (right), both marginalized over both mass hierarchies.

4.4.5 Results with the $\sin^2\theta_{13}$ constraint from the reactor experiments

The main T2K results are from the fit to the data samples with a prior constraint on $\sin^2\theta_{13}$ taken from the joint results from the reactor neutrino experiments, provided by the PDG group. The MCMC sampler was run to produce 102 million MCMC steps across the T2K parameter space. Given the autocorrelation of ∼30,000 MCMC steps, this produced 3400 independent MCMC samples; enough to produce the first $3\,\sigma$ credible intervals with enough fitting statistical certainty to be able to draw strong conclusions from them. Table 4.8 shows the highest-posterior oscillation parameters values with 1, 2 and $3\,\sigma$ credible intervals, all marginalized over both mass hierarchies of Δm^2_{32}.

	$\sin^2\theta_{23}$	Δm^2_{32} ($\times 10^{-3}$ eV^2)	$\sin^2\theta_{13}$	δ_{CP}
2D best fit	0.537	2.46	0.0214	-1.82
68.27% C.I. ($1\,\sigma$) range	0.501 – 0.564	2.37 – 2.54	0.0206 – 0.0222	-2.39 – -1.13
95.45% C.I. ($2\,\sigma$) range	0.466 – 0.587	-2.58 – -2.41 & 2.28 – 2.63	0.0199 – 0.0230	-2.95 – -0.50
99.73% C.I. ($3\,\sigma$) range	0.438 – 0.606	-2.68 – -2.32 & 2.22 – 2.70	0.0191 – 0.0238	-π – 0.13 & 2.80 – π

Table 4.8: Highest-posterior values and 1, 2 and $3\,\sigma$ 1D credible interval ranges for oscillation parameters for the run 1–9d data fit with reactor constraint. The highest-posterior values are taken from the mode of the 2D marginal posterior distributions in $\sin^2\theta_{23}$–Δm^2_{32} and $\sin^2\theta_{13}$–δ_{CP}. All these values are marginalized over both mass hierarchies.

New constraint on δ_{CP}

Figure 4.19 shows the binned log posterior probability density for all the values of δ_{CP} marginalized over both mass hierarchies, with 1, 2 and 3 σ credible intervals marked, with the normal and inverted mass hierarchies shown in 4.20 and 4.21, respectively. We can see that a large portion of the CP-conserving δ_{CP} phase-space is rejected with at least 3 σ Credible Interval, whereas the highest posterior area is still very compatible with maximal CP violating value of $\delta_{CP} = -\pi/2$. In the inverse hierarchy we are close to rejecting CP conservation at 3 σ, with π already rejected.

This is the first 3 σ result for δ_{CP} in the T2K experiment, and world first for fully

closed δ_{CP} interval. The Credible Intervals from this analysis are part of the main result in the Nature paper [53].

Since the Bayesian analysis is capable of reweighting the posterior probability given various priors, we can easily see the effect of a prior on the δ_{CP} intervals. The framework is sampling the parameter space using a prior that is flat in δ_{CP}, however, the oscillation probability takes values of $\sin\delta_{CP}$ and $\cos\delta_{CP}$, so it is interesting to see the effect of a prior flat in one of these phase-spaces.

Figure 4.22 shows another posterior-probability plot for δ_{CP} marginalized over both mass hierarchies, but this time with an extra set of credible intervals (marked with arrows) from the same posterior but reweighted to a δ_{CP} prior flat in $\sin\delta_{CP}$ phase-space. Although we are less sensitive with such a prior, the interpretation of the results for this analysis remain unchanged. The slight change of results upon changing the prior for δ_{CP} is expected; T2K data still constrains this parameter rather weakly, so changing priors on δ_{CP} does have a visible effect.

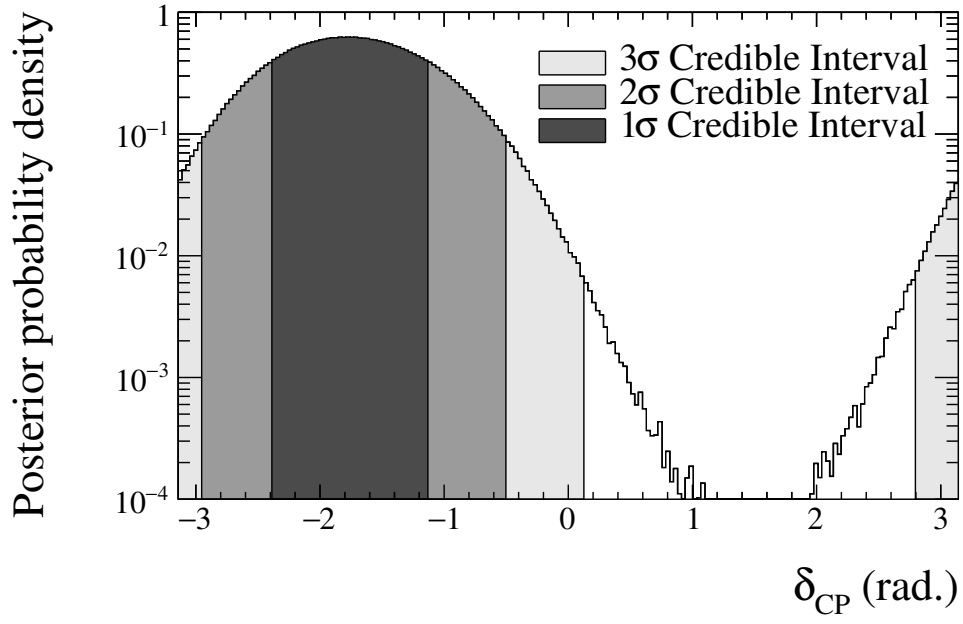


Figure 4.19: Posterior probability for δ_{CP} marginalized over both mass hierarchies, from the run 1–9d SK data fit with the reactor constraint. The CP conversation is rejected at $2\,\sigma$, with the highest posterior density area still being compatible with $\delta_{\mathrm{CP}} = -\pi/2$.

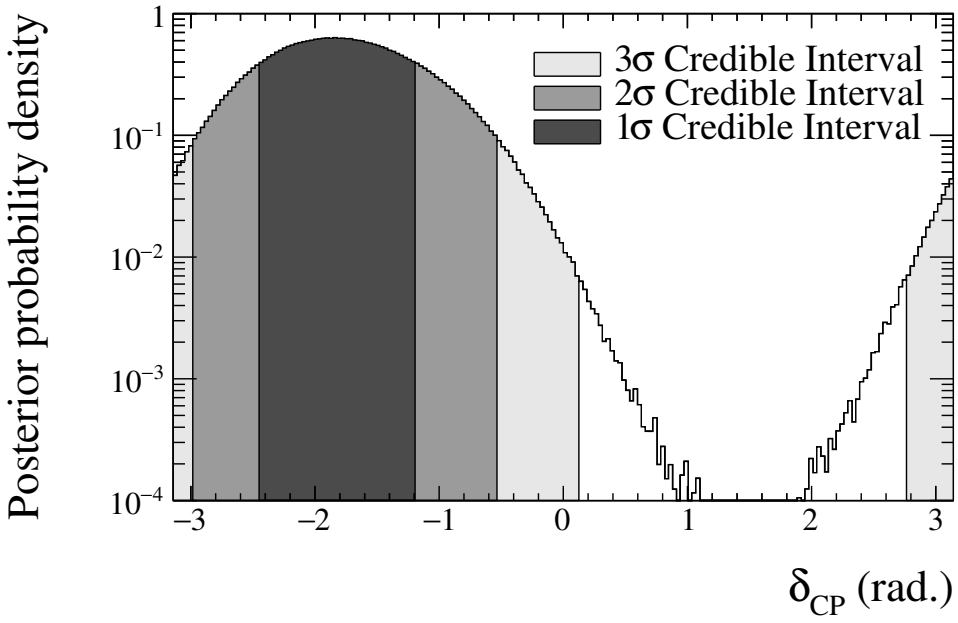


Figure 4.20: Posterior probability for δ_{CP} for NH, from the run 1–9d S data fit with the reactor constraint. Similar to the plot marginalized over both hierarchies due to high preference towards NH.

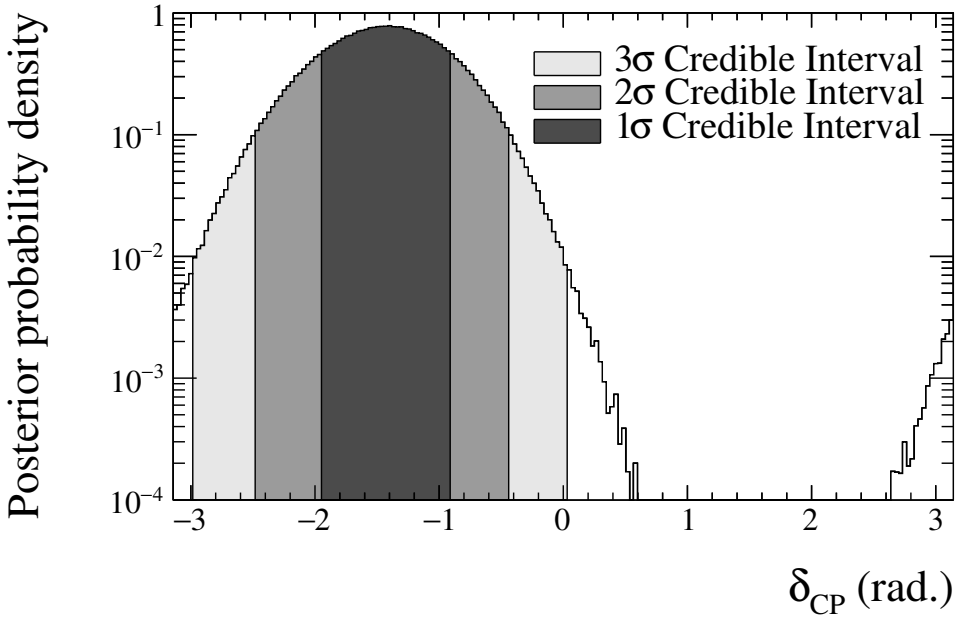


Figure 4.21: Posterior probability for δ_{CP} for IH, from the run 1–9d SK data fir with the reactor constraint. CP conserving values for IH δ_{CP} are almost rejected with 3σ Credible Interval.

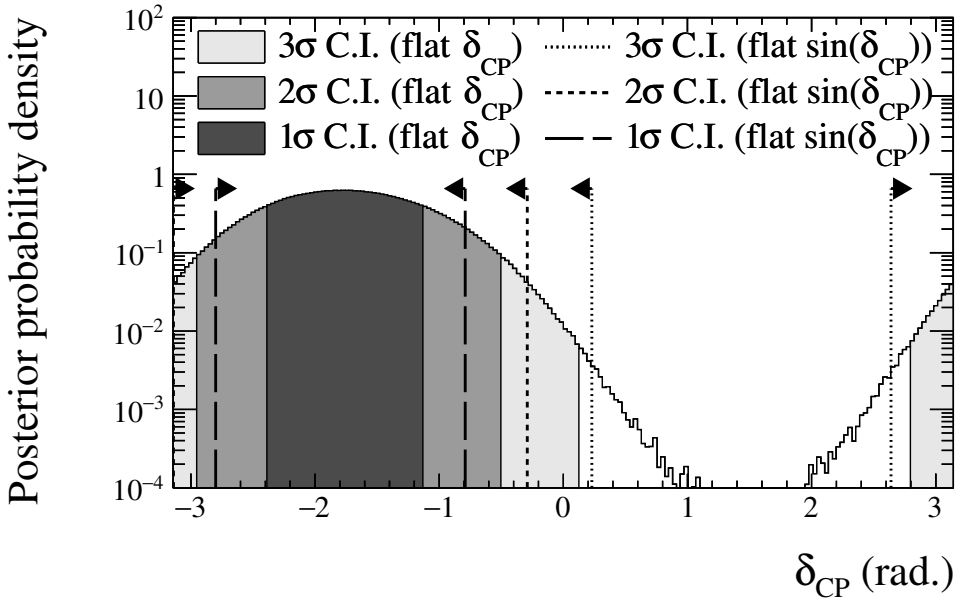


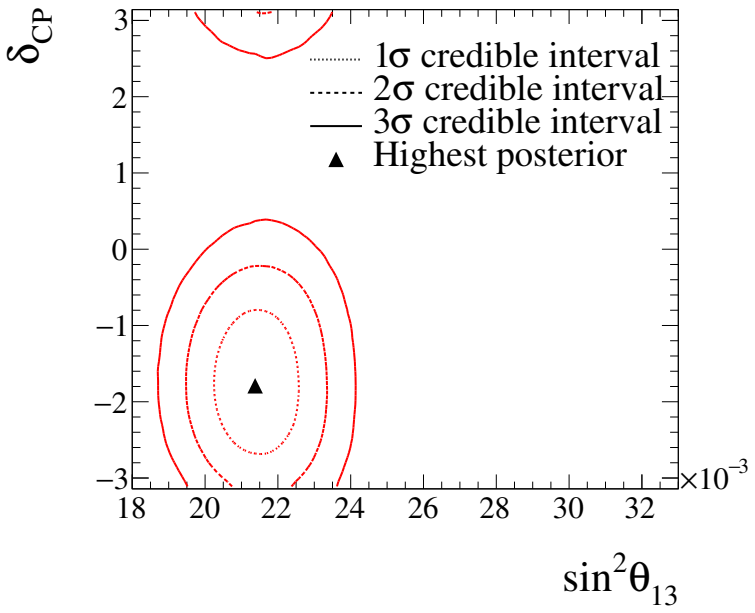
Figure 4.22: Posterior probability for δ_{CP} marginalized over both mass hierarchies, with an extra set of credible intervals from the posterior reweighted to a prior flat in $\sin\delta_{CP}$.

Electron neutrino appearance parameters

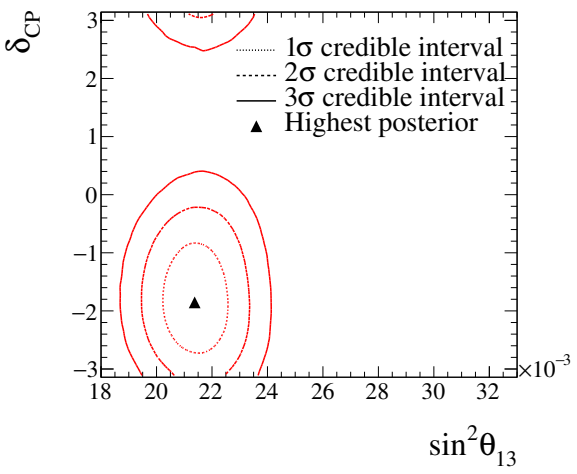
Figure 4.23 shows the credible intervals for $\sin^2\theta_{13}$ against δ_{CP}, marginalized over both mass orderings, and separately for each mass ordering. Apart from the extra $3\,\sigma$ intervals and the distribution shifting towards lower values of $\sin^2\theta_{13}$ due the new reactor constraint, these distributions did not change much since the analysis. Assuming the new data would have the same proportions as before, the expectation was that the constraint on δ_{CP} would get much stronger, especially given the new reactor constraint. Although the intervals on δ_{CP} did get tighter, the difference is relatively small. This can be explained by the new data itself; up to run 9, the data contained far fewer $\bar{\nu}_e$ events than expected, although still well within the statistical uncertainty. That however meant that the electron neutrino appearance probability, $\mathrm{P}(\bar{\nu}_\mu \to \bar{\nu}_e)$, was lowered, corresponding to low values of δ_{CP}.

With the data being slightly beyond the PMNS constraint, the MCMC steps were tightly-squeezed on the δ_{CP} space against the value of $-\pi/2$. The new data contained more $\bar{\nu}_e$ - more than expected, but also within statistical uncertainty. The overall uncertainty on $\mathrm{P}(\bar{\nu}_\mu \to \bar{\nu}_e)$ got tighter because of more available data, constraining δ_{CP} more,

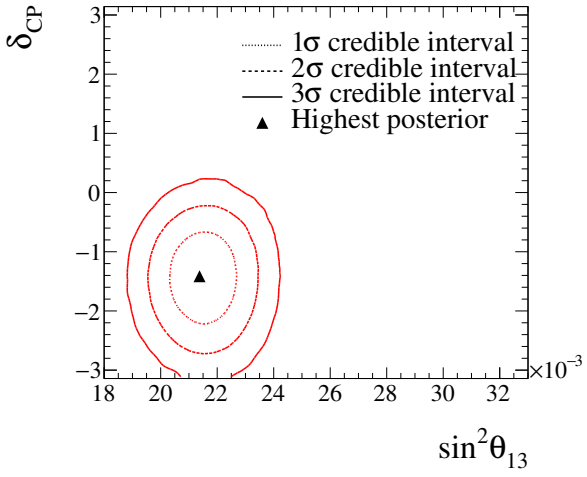
but the large statistical fluctuation on the number of ν_e somewhat disappeared, causing a weaker constraint. All these effects added together contributed to only slightly better constraint on δ_{CP} than in the previous analysis [54].



(a) Both mass hierarchies



(b) Normal hierarchy



(c) Inverse hierarchy

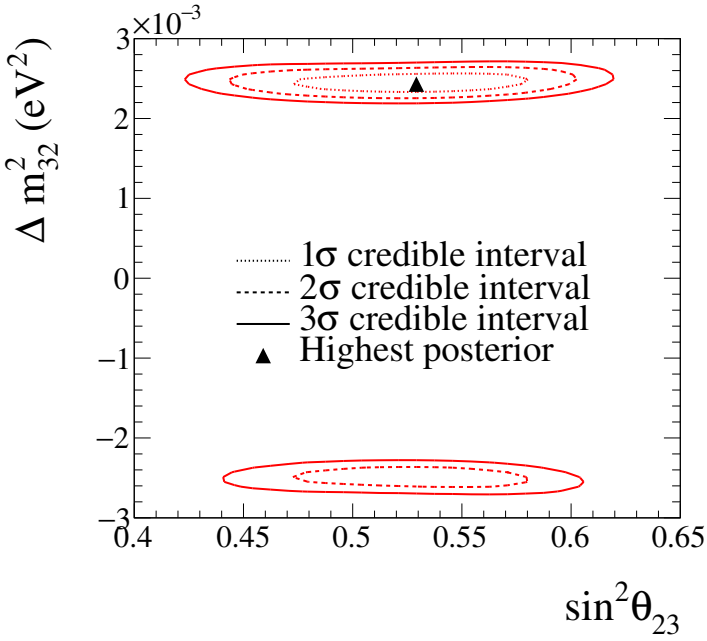
Figure 4.23: 2D Credible Intervals for $\sin^2\theta_{13}$ against δ_{CP} marginalized over both, normal and inverse mass hierarchies, respectively.

Muon neutrino disappearance parameters

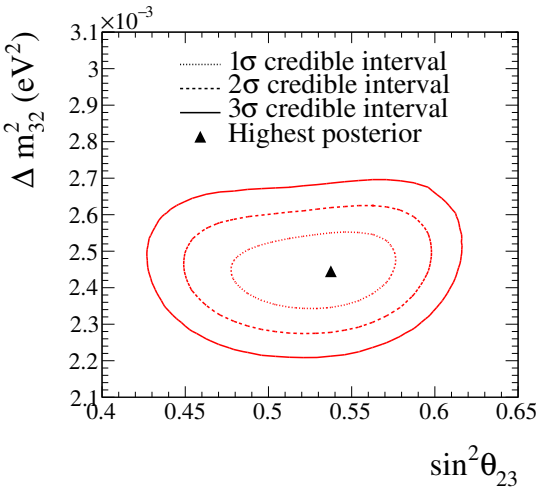
Figure 4.24 shows the disappearance contours marginalized over both mass hierarchies (4.24a), and both mass hierarchies separately (4.24b and 4.24c). Just as with the 1D plots, the $1\,\sigma$ contours are only present in the NH, and the $\sin^2\theta_{23}$ has most of its posterior probability in the upper octant.

Table 4.9 shows the posterior probabilities separated by the mass hierarchy of Δm^2_{32} and the octant of $\sin^2\theta_{23}$. The normal hierarchy contains 88.9% of posterior probability (marginalized over octant with a flat prior in $\sin^2\theta_{23}$) and the upper octant contains 79.5% of the posterior probability (marginalized over hierarchy).

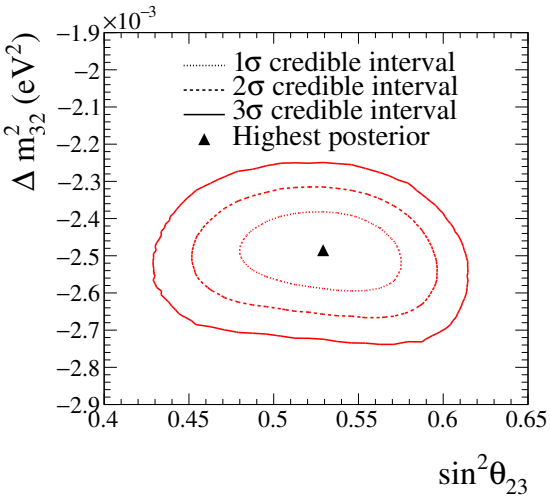
The statistical significance of these numbers is still too low to exclude either of the alternate models. The Bayes factor for the normal hierarchy over the inverted hierarchy is 8.0, where the simplest way to interpret the Bayes factor is the number of times that one model is preferred over the alternative one. In a more formal definition, a Bayes factor between $10^{1/2} = 3.16$ and 10 can be classified as "substantial" according to the Jeffreys' scale [112] or "positive" according to the Kass and Rafferty scale [113]). The significance of this result, however, has been steadily increasing from the run 1–6 analysis onwards. This can cause issues in future analyses (especially if MaCh3 will be used in, say, DUNE or Hyper-Kamiokande) where one hierarchy might be much more preferred over the other; the MCMC sampler might not be able to accept a step between the mass hierarchies, and therefore not being able to marginalize over the mass hierarchy properly. As a result, it could be more difficult to perform the mass hierarchy hypothesis testing, as there would not be enough fitting statistics (trans-hierarchy MCMC steps) to do so. This problem is further explored in section 4.4.6, with possible solutions proposed.



(a) Both mass hierarchies



(b) Normal hierarchy



(c) Inverse hierarchy

Figure 4.24: 2D Credible Intervals for $\sin^2\theta_{13}$ against δ_{CP} marginalized over both, normal and inverse mass hierarchies, respectively.

	$\sin^2\theta_{23} < 0.5$	$\sin^2\theta_{23} > 0.5$	Sum
NH ($\Delta m^2_{32} > 0$)	0.184	0.705	0.889
IH ($\Delta m^2_{32} < 0$)	0.021	0.090	0.111
Sum	0.205	0.795	1

Table 4.9: Model comparison probabilities for normal and inverted hierarchies, as well as upper and lower octants, for the run 1–9d data fit with the reactor constraint on $\sin^2\theta_{13}$.

Correlations between neutrino oscillation parameters

Figure 4.25 shows the posterior probabilities (both 2D parameter-pairs and 1D) for all the parameters of interest. Although difficult to see by eye, $\sin^2\theta_{23}$ and $\sin^2\theta_{13}$ have a small anti-correlation. Given that these parameters are multiplied by each other in the oscillation formula, an anti-correlation is not unexpected; these parameters have a similar effect on the predicted number of ν_e and $\bar{\nu}_e$. This is normally more visible without the reactor constraint, and is expected to disappear as we get more data at SK. Interestingly, marginalization effects can be seen when comparing the 1D against 2D posterior distributions and the credible intervals; the 1D intervals are tighter than the 2D intervals. Since marginalization, which is based on the probability densities, is used instead of profiling, which is based on the best-fit values, this is not unexpected, and unsurprisingly, the effect is most significant in parameters that are more strongly correlated with each other, for example, Δm^2_{32}–$\sin^2\theta_{23}$. Since the profiled contours would be based on the critical χ^2 values, they would be the same whether looking at the one parameter alone or coupled in 2D. For marginalization, however, converging the probability densities is dependent on the number of dimensions — therefore, the 1D and 2D contours should not be expected to converge around the same parameter values.

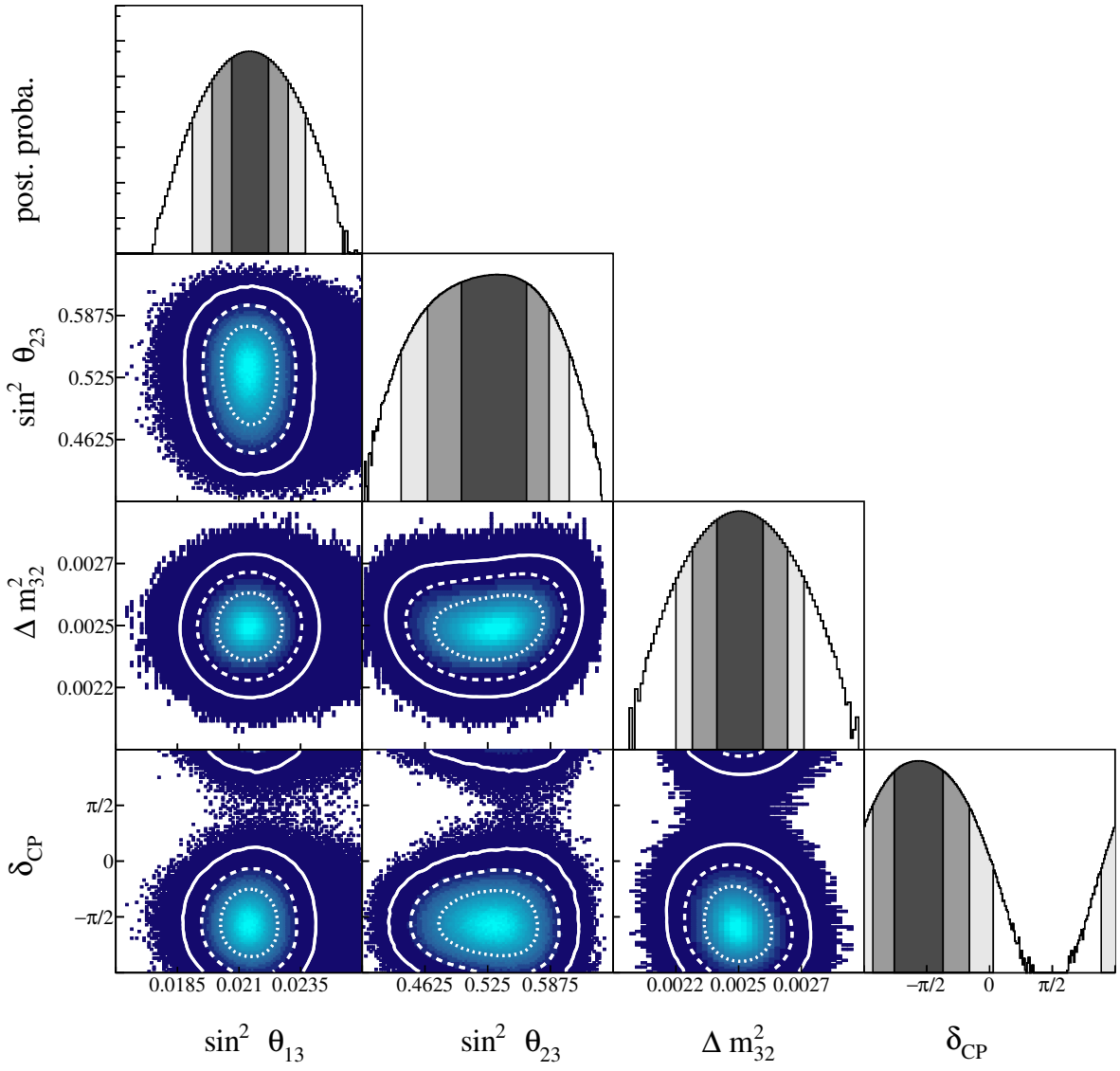


Figure 4.25: 2D posterior probabilities for all the oscillation parameters of interest plotted against each other and separately on 1D posterior plots. Each posterior has 1, 2 and 3 σ credible intervals. All the plots are marginalized over both mass hierarchies. Although the plots with Δm^2_{32} in its axis are also marginalized over both mass hierarchies, the plot boundaries are chosen to show the normal mass hierarchy only.

4.4.6 Importance Sampling: Future speed upgrade through Markov Chain reweighting

As described in 4.4.3, it is possible to run the MCMC sampler with one set of priors, and then reweight some parameters to different priors as desired. This creates an interesting opportunity for faster exploration of less-likely phase-spaces. In the simplest example that relates to T2K, we could, in principle, run the MCMC sampler with a strong prior on δ_{CP} centered around $\pi/2$, currently rejected well beyond $3\,\sigma$ CL. The likelihood surface would not change, since in the Bayesian formalism the likelihood and prior are separate from each other, and one can change independently of the other. This means we can re-weight such an MCMC sampler back to a prior flat in δ_{CP}, and we will end up with high-statistics-MCMC in the less likely regions of δ_{CP}, possibly giving us more reliable constraints in these regions.

A simple toy study shows exactly this: N random numbers generated uniformly between -10 and 10, reweighted to a Gaussian centered at 0 with a width of of 2, gave more statistics in its $3\,\sigma$ region than randomly generated N numbers directly from a Gaussian(0,2). This is illustrated in figure 4.26, where the tails of a flat distribution reweighted to a Gaussian (in red) are more stable than the distribution obtained from a real Gaussian number generator (in blue).

In this section we will describe the studies done on a reweighted chain without the reactor constraint on θ_{13}, to see if we can use it to improve the MCMC-statisticsin $3\,\sigma$ region of the with-the-reactor-constraint phase-space and reduce the number of MCMC steps needed for convergence.

We start by reweighting an MCMC chain with a prior flat in $\sin^2\theta_{13}$ to a reactor-constrained phase-space with code written and validated for the reasons described in 4.4.3. We can then look at the number of independent MCMC steps outside of $3\,\sigma$ credible intervals. Independent samples are defined here as the number of steps divided by the autocorrelation, approximated to 30,000 for chains with both and without the reactor constraint.

Table 4.10 shows the results of such a reweighting, for 1, 2 and $3\,\sigma$. We also see the

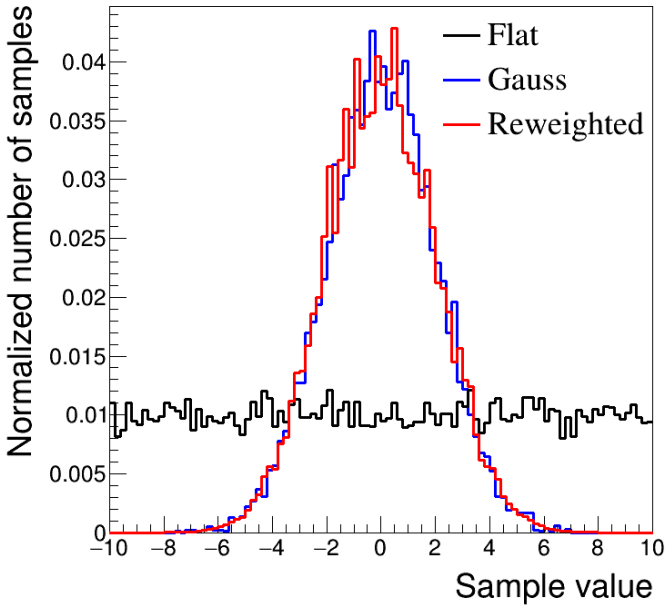


Figure 4.26: Binned (and normalized) random numbers distributed uniformly between -10 and 10 in black, reweighted to a Gaussian centered around 0 with uncertainty of 2 in red, and random numbers obtained from a Gaussian generator in blue. The tails of the reweighted Gaussian distribution have less variance bin-to-bin than the random numbers obtained from a Gaussian generator.

integrals, where our expectation is 0.6827, 0.9545 and 0.9973 for 1, 2 and 3 σ, respectively. From this it is clear that there are more MCMC-statistics outside of the 3 σ CLs. This, however, is expected, given that the sampler without the reactor constraint explores larger areas of the T2K phase-space that the reactor-constrained sampler does not.

It is difficult to find a relevant metric good enough to judge the convergence (or the "quality") of an MCMC chain in specific regions of phase-space. Perhaps a better (and in this case, final) metric would be the number of independent steps inside of the 3 σ but outside of the 2 σ intervals. This will return the MCMC-statistics measure that is only relevant to the phase-space important for the interesting regions of the reactor-constrained fit.

Table 4.11 shows that, with the new metric, we find fewer MCMC steps (about 2/3) in the relevant region with the reweighted chain than with the reactor-constrained chain. The originally-unconstrained chain, however, contains only 22 million MCMC

22 Million steps, no reactor constraint			
Interval	Integral	Samples	$\sqrt{(\text{Samples})}$
$1\,\sigma$	0.686084	351	19
$2\,\sigma$	0.954506	116	11
$3\,\sigma$	0.997335	27	5

100 Million steps, with reactor constraint			
Interval	Integral	Samples	$\sqrt{(\text{Samples})}$
$1\,\sigma$	0.684689	1083	33
$2\,\sigma$	0.957021	152	12
$3\,\sigma$	0.997422	10	4

Table 4.10: The number of independent MCMC samples in outside of 1, 2 and $3\,\sigma$ CLs with their integrals, for 100 M-step chain with the reactor constrain and 22 M-step chain with flat prior in $\sin^2\theta_{13}$, reweighted to the PDG 2018 reactor constraint. The number of independent samples in the $3\,\sigma$ interval of the 22 M-step reweighted chain is higher than that of 100 M-step chain with the reactor constraint, even when including the Poisson $\sqrt{N}$ error, despite the lower number of MCMC steps.

22 Million steps, no reactor constraint			
Interval	Integral	Samples	$\sqrt{(\text{Samples})}$
$3\,\sigma$	0.997335	89	10

100 Million steps, with reactor constraint			
Interval	Integral	Samples	$\sqrt{(\text{Samples})}$
$3\,\sigma$	0.997422	142	12

Table 4.11: The number of independent MCMC samples inside of $3\,\sigma$, but inside of $2\,\sigma$ CLs with their integrals, for 100 M-step chain with the reactor constrain and 22 M-step chain reweighted from a prior flat in $\sin^2\theta_{13}$ to the reactor constraint.

steps, whereas the constrained chain contains over 100 million steps. This means that we could still run, for example, ∼30–35 million steps without the reactor constraint MCMC, reweight it to the reactor constraint from PDG, and obtain a similar amount of MCMC statistics in the $3\,\sigma$ region than with 100 million chain ran with the reactor constraint alone.

This was presented and approved at a T2K Oscillation Analysis meeting. Instead of running ∼20 M steps without, and ∼100 M steps with the reactor constraint for the 2020 $3\,\sigma$ results, the MaCh3 group can run ∼30 M without first, then 30 M with the reactor constraint for validations (and to fill the 1σ region, which would still have less statistics with the reweighting scheme). The analysers could then continue running the chain without the reactor constraint and obtain 150-200M steps for the first-ever $4\,\sigma$ results, which are now possible in the Bayesian framework due to this study.

Although too late for the running 2020 T2K data analysis, further improvements

will be made in the future using the reweighting scheme. Instead of the standard running reactor-constrained chain, it should be possible to run it with an extra prior on δ_{CP} that biases the chain towards, for example, $\pi/2$. With the data likelihood (hopefully) still preferring $-\pi/2$ and strong enough prior around $\pi/2$, the chain could explore all the δ_{CP} phase-space more uniformly, possibly enabling $5\,\sigma$ results.

There is another method to explore, closely linked to the Parallel-Tempered MCMC described in chapter 3.4.1; through the likelihood tempering, which is also used in Simulated-Annealed MCMC methods. The likelihood function gets an additional "temperature" term (resembling the Boltzmann distribution), that effectively spreads the sampled posterior distribution in all dimensions, allowing for more samples in the less likely regions and expanding the MCMC coverage. This can be reweighted post-fit to obtain the distribution without the likelihood heating element.

This is great news not only for the T2K experiment, but even better for the future experiments like Hyper-K and DUNE. These experiments will struggle with running MCMC samplers that marginalize over both mass hierarchies at the same time; with rejection at $5\,\sigma$, the MCMC chain might not be able to easily (or at all) accept steps in the IH. Furthermore, it will take exponentially more time to achieve the MCMC convergence at $5\,\sigma$ level for all the parameters. It should, however, be possible to bias a chain with a strong prior on IH to allow the chain explore it for marginalization, and re-weight back to a flat prior after the fit, giving the analysts the confidence that both MH were explored, and the rejection is real. Tempering can be used at the same time to ensure fast MCMC stationary distribution at $5\,\sigma$ level without worrying about the correct coverage.

4.5 Summary

Table 4.8 shows the credible intervals for each oscillation parameter of interest, obtained from the T2K joint ND280-SK data. The fit was done using 5 SK and 17 ND280 data samples, both with the reactor constraint on θ_{13} and without. As expected, the T2K sensitivity to δ_{CP} without the reactor constraint is much lower, and far more data is needed for T2K to be able to produce competitive results without external constraints. The results were validated by comparing them against the two hybrid-Frequentist analysis groups: VALOR and P-Theta. All the comparison plots, for both data and sensitivity studies, are shown in appendix B.

T2K presented the world's first closed $3\,\sigma$ intervals for δ_{CP}, which is beyond the expected sensitivity, and could be explained with a statistical fluctuation in the data itself; this will be explored more in the next chapter.

A new prior-reweighting scheme was studied and used in this analysis to speed up the MCMC chain convergence to a stationary distribution. Such prior-reweighting scheme, alongside likelihood-tempering, could be used by the future T2K analyses and future long-baseline experiments to achieve stable intervals at lower computational costs, and to explore very unlikely regions of phase-space, for example, the inverse mass hierarchy.

Chapter 5

Bi-Probability Oscillation Fits

The T2K data is often shown in one plot with the number of ν_e and $\bar{\nu}_e$ candidates, against the PMNS-like prediction produced by varying the values of δ_{CP} with other oscillation parameters fixed. This creates the so-called bi-event plots, shown in 5.1, and it is a standard procedure amongst various long-baseline neutrino experiments to compare these against each other (for example, in [135]). These plots can be used not only to show the T2K data in a form of a point, but also to show where the data should be, given a set of oscillation parameters, and how far the real T2K data deviates from our prediction. Although these plots are very useful internally, theorists and other long-baseline experiments are unable to easily compare their expectations against our results with these plots alone. The number of ν_e and $\bar{\nu}_e$ events shown on these plots are very dependent on the experimental setup, i.e the full knowledge of T2K's flux, cross-sections and detector efficiencies would have to be known externally to make any independent neutrino event rate predictions for the T2K experiment.

Another way of showing the T2K sensitivities is through a bi-probability plot, where $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ predictions are shown against each other, akin to the bi-event plot. Although it is possible to show the predictions in a more mode-independent way (in other words, the knowledge of the cross-sections, detector efficiencies and so on is not needed), it is rather difficult to show the $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ predictions from the data fit with the full systematic and statistical errors, without a dedicated analysis framework.

It is, however, possible to exploit the MCMC framework described before to make it possible. This chapter will describe an analysis with MaCh3 to produce the T2K data-prediction on the bi-probability phase-space, and introduce an extra parameter, β, to break the unitary of the PMNS model and produce a model-independent prediction. The results described in this chapter are published in [136].

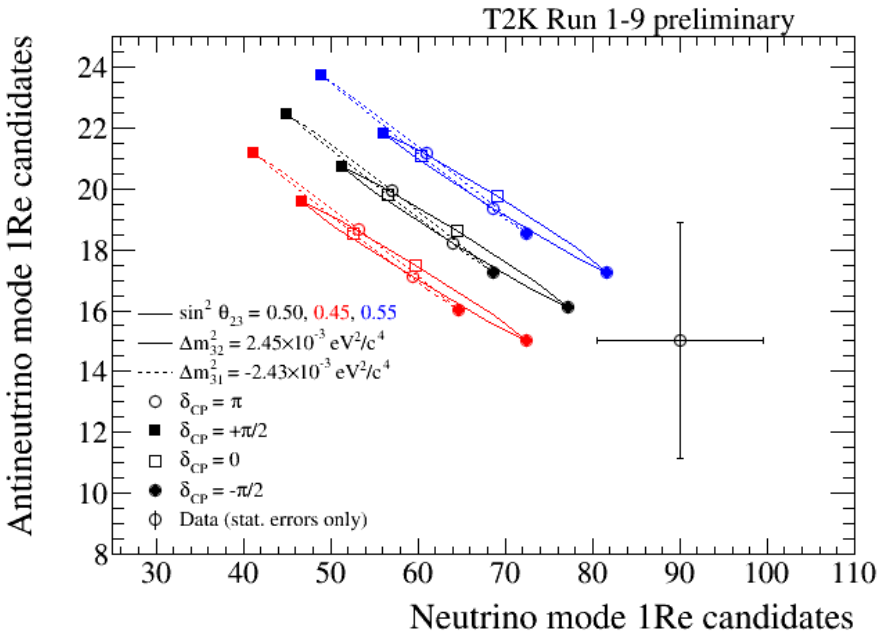


Figure 5.1: Bi-Event plot, showing the number of ν_e and $\bar{\nu}_e$ events with the statistical errors on top, together with the PMNS oval expectations. The expectations are made by fixing all the systematic and oscillation parameters, and only varying one parameter (in this case δ_{CP}) at a time.

5.1 Bi-Probability plots

As described in the chapter's introduction, instead of showing the number of neutrino and anti-neutrino events, we can show the $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ oscillation probabilities and expectations with bi-probability plots. Just like the bi-event plots, they give a clear sense of how δ_{CP} and the neutrino oscillation parameters together with mass hierarchy will affect the oscillation probabilities, and therefore how the predicted event rates for ν_e and ν_μ at Super-K should change with the extra data.

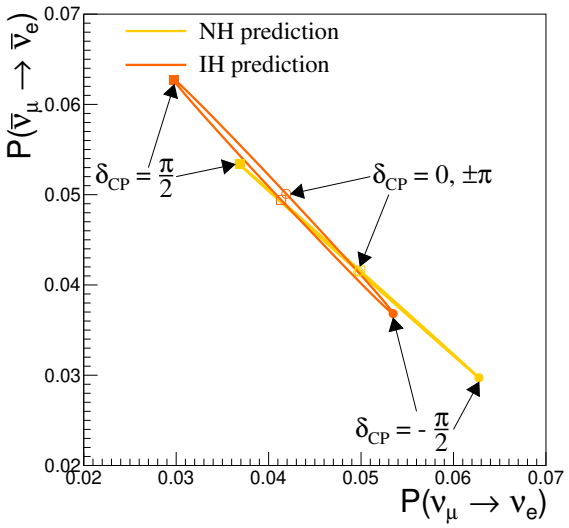
Figure 5.2 shows the predicted probabilities for the T2K experiment in fig. 5.2a and for NOvA in fig. 5.2b. Both plots were generated with an open-source neutrino oscillation calculator, Prob3++ [108], using the T2K's Asimov oscillation parameter values (listed in appendix A, table A.1) and publicly available information about both experiments' peak neutrino energy, experiment baseline length and the density of matter. The values of δ_{CP} are iterated to draw the oval-shaped predictions, separately for the inverse and normal mass hierarchies. These ovals are also affected by the values of $\sin^2\theta_{23}$ and $\sin^2\theta_{13}$; they both shift them along the line of unity. High values of $\sin^2\theta_{23}$ will also separate the mass hierarchies further, meaning the higher the $\sin^2\theta_{23}$, the easier the prediction of the mass hierarchy (and vice-versa, the lower the $\sin^2\theta_{23}$, the lower the sensitivity to the mass hierarchy).

Comparisons to external predictions, whether from other experiments or theorists, are much easier with the bi-probability plots than with the bi-event plots. Indeed, we can now easily understand the experiment's characteristics and differences against NOvA as seen in fig. 5.2. NOvA has a wider separation between the NH and IH, thanks to its longer baseline. At the same time, they would find it easier to distinguish between $\delta_{\mathrm{CP}} = 0$ and π, if nature chose these values.

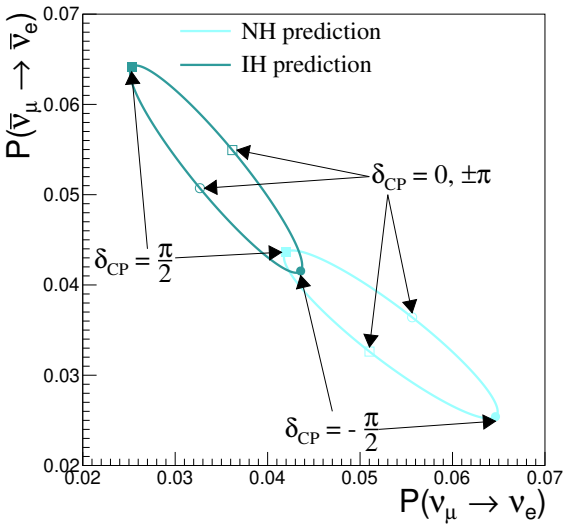
These bi-probability plots are made for a monochromatic neutrino energy; in our case it is 0.6 GeV for T2K (figure 5.2a) and 1.9 GeV for NOvA (figure 5.2b). These plots do not change much, as long as the neutrino energy is not deviated far into the tails of the experiment's neutrino beam energy distribution.

It is possible to display the T2K data fit results on this phase-space by using the output MCMC chain from one of the standard T2K Bayesian analyses, explained in Chapter 4. Since each MCMC step contains values for every oscillation and systematic parameter, it is possible to convert these back into $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ with the standard oscillation formula. This results in an MCMC chain that now consists of $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ values, marginalized over all the oscillation and systematic parameters.

Fig. 5.3 shows an example of such posterior probability densities, separately for the NH (5.3a) and for the IH (5.3b). We can see that most of the posterior density is

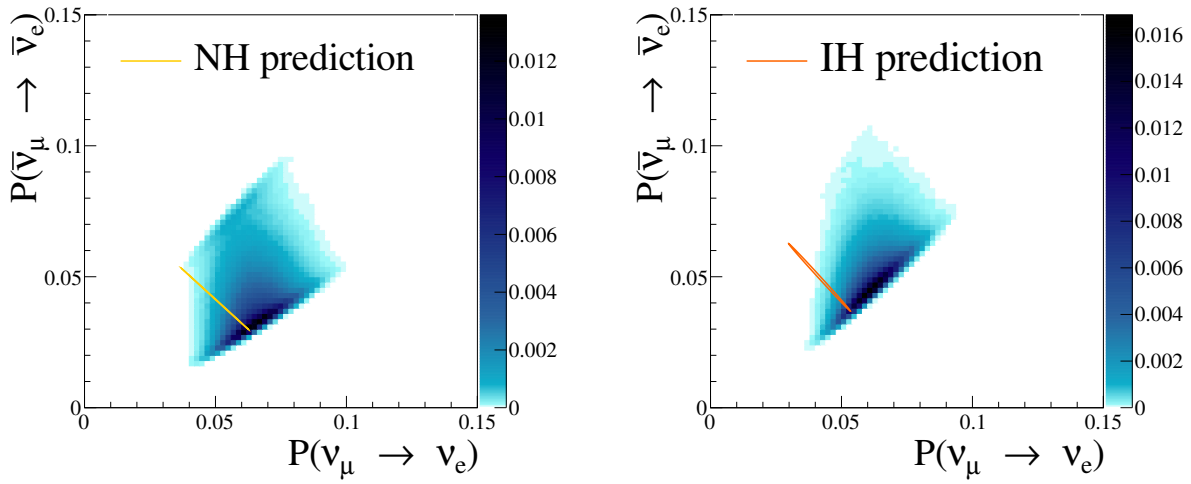


(a) T2K bi-probability ovals for NH and IH, produced by drawing over all the values of δ_{CP}. The other oscillation parameters are kept at the Asimov A values (Appendix A). Assuming monochromatic neutrino beam of 0.6 GeV.



(b) NOvA bi-probability ovals for NH and IH produced by drawing over all the values of δ_{CP}. The other oscillation parameters are kept at the Asimov A values (Appendix A). Assuming monochromatic neutrino beam of 1.9 GeV, and slightly higher earth density than T2K.

Figure 5.2: T2K and NOvA bi-probability predictions for different values of δ_{CP} and mass hierarchies. NOvA has better sensitivity for the mass hierarchy (larger separation in NH and IH bi-probabilities) due to the longer baseline, but clearly has lower sensitivity to δ_{CP} for the extreme values of $\pm\pi/2$ - it will however have better sensitivity to the δ_{CP} values of 0 and $\pm\pi$.



(a) Posterior probability from the woRC Asimov A fit with run 1–9d data on the bi-probability space with only NH steps selected. The prediction for various δ_{CP} values superimposed in yellow.

(b) Posterior probability from the woRC Asimov A fit with run 1–9d data on the bi-probability space with only IH steps selected. The prediction for various δ_{CP} values superimposed in orange.

Figure 5.3: Posterior probability from the woRC Asimov A fit with run 1–9d data on the bi-probability space marginalized over NH and IH. The dual peak at each end of the superimposed prediction is due to the prior of δ_{CP} being flat in δ_{CP}, with the oscillation formula taking values of $\sin\delta$.

concentrated around $\delta_{\text{CP}} = -\pi/2$, as expected, but we see a faint higher-density area around $\pi/2$. This is due to the choice of the prior: The T2K experiment fits with a prior flat in δ_{CP}, whereas the oscillation formula takes the values of $\cos\delta_{CP}$ and $\sin\delta_{CP}$. We can also see that the posterior density is spread roughly along the line of unity; this is due to the uncertainty on $\sin^2\theta_{23}$ and $\sin^2\theta_{13}$ and the systematic parameters.

5.2 Beyond the PMNS parametrisation

As seen in figure-5.1, the T2K data does not agree very well with the PMNS prediction (but is still well within the expected statistical fluctuations). The bi-probability plots are also constrained by the PMNS, and do not allow one to to easily overlay a data point. A new parameter was added, β, to decouple the two oscillation probabilities from each other and allow the fitter to explore the area preferred by the data, unconstrained by the unitary of the PMNS model;

$$P\left(\bar{\nu}_\mu \to \bar{\nu}_e\right) = \beta \times P_{\text{PMNS}}\left(\bar{\nu}_\mu \to \bar{\nu}_e\right), \tag{5.1}$$

$$P\left(\nu_\mu \to \nu_e\right) = 1/\beta \times P_{\text{PMNS}}\left(\nu_\mu \to \nu_e\right). \tag{5.2}$$

This new parametrisation was found to not only decouple the two probabilities from each other, but it also makes all the oscillation parameters difficult to interpret; what is the meaning of δ_{CP} when $\beta = 0.1$? Although this is non-interpretable, the resultant oscillation probabilities are still easy to understand, and were found to move in the direction preferred by the data on the bi-event plots.

The standard T2K oscillation analysis usually involves the reactor constraint on θ_{13} giving us a better sensitivity to δ_{CP} and the mass hierarchy. This constraint is driven by PMNS-like results, so it does not make any sense to include it in the bi-probability analysis which is trying to break the unitarity of the PMNS model. All the bi-probability Asimov and data fits are therefore done with a prior that is flat in $\sin^2\theta_{13}$, constrained between 0 and 1.

5.3 Validations and sensitivities

5.3.1 Asimov fits

Oscillation parameter comparisons

Standard checks described in 4.3 were done for these analyses prior to running any fits, with the predicted event rates and 1σ systematic variations agreeing with the standard output in 4.3.1.

As mentioned before, the standard oscillation parameters in the free-beta fits become effective parameters and are not to be interpreted. It is, however, still easy to compare these to the standard PMNS-like fit; we can simply constrain the MCMC chain to only use the MCMC steps with the values of β close to 1, effectively re-creating the standard PMNS-like analysis. If the framework works correctly, the constrained-β contours should match the standard analysis, whereas the unconstrained contours should explore wider area.

Figure 5.4 shows such comparisons for appearance (top) and disappearance parameters. We can see that the constrained-β plots clearly re-create the standard analysis, indicating that the new framework is working correctly. The disappearance contours do not change - unsurprisingly, since the $P(\nu_\mu \rightarrow \nu_e)$ and the $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ probabilities only really affect the appearance parameters. Such plots were generated for Asimov B fake dataset too, for both run 1–9c and run 1–9d analyses.

Fixed vs free β comparisons

Since the main goal of this analysis is to compare fixed (PMNS-like) and free (unitarity-breaking) β fits on the bi-probability phase-space, it is appropriate to generate such plots using fake MC-data tuned to the Asimov A values, just as in the standard analysis.

Figure 5.5 shows the free (top) and fixed (bottom) β fits on the left pane, with the credible interval contour comparison on the right hand side pane. We can see that the free

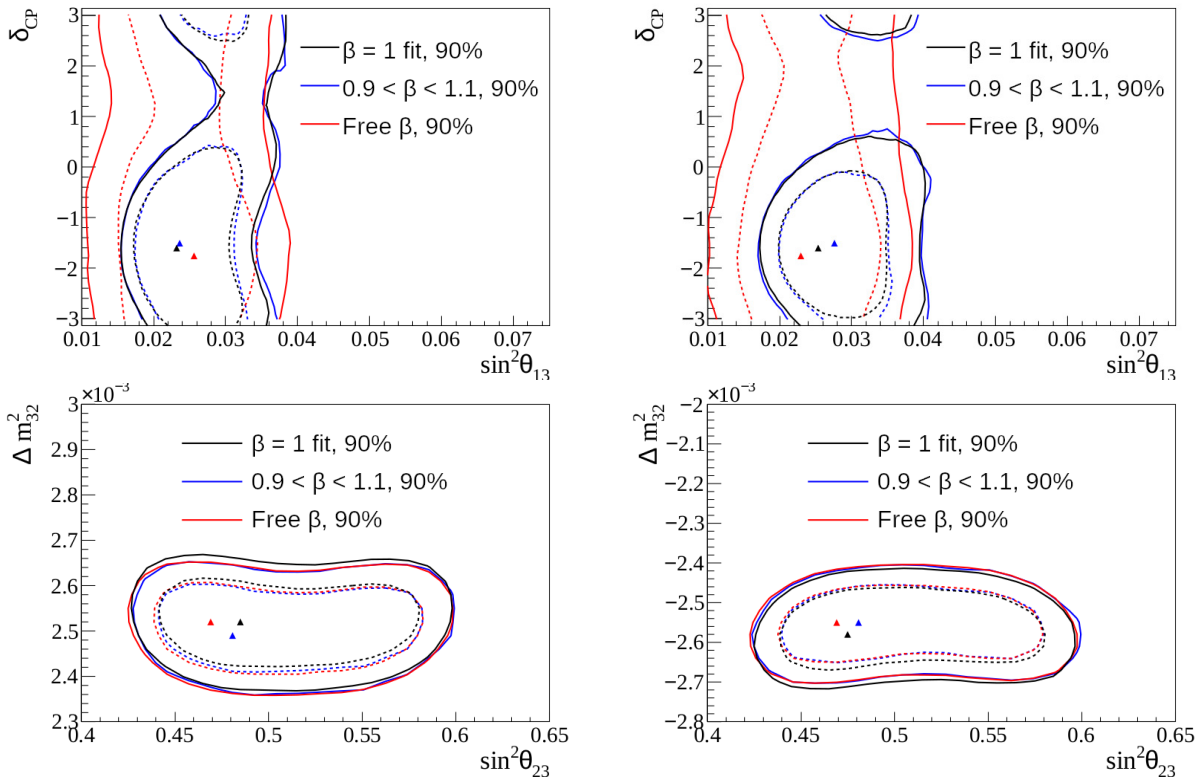


Figure 5.4: Appearance (top) and disappearance (bottom) contour comparison between the standard Asimov A fit, the free β fit and the free β with only values of β between 0.9 and 1.1 to re-create the standard fit for validation purposes. Given the relatively low number of MCMC steps, the best-fit values are not expected to converge well.

β fit explores a wider area that is not permitted by the unitarity-conserving PMNS model. We lose the sensitivity to the mass hierarchy, with the probabilities (and contours) being equally distributed between both mass hierarchies. Although very different, it is clear that the free β fit is not in any tensions with the PMNS model. Indeed, both contours overlap comfortably at all σ levels.

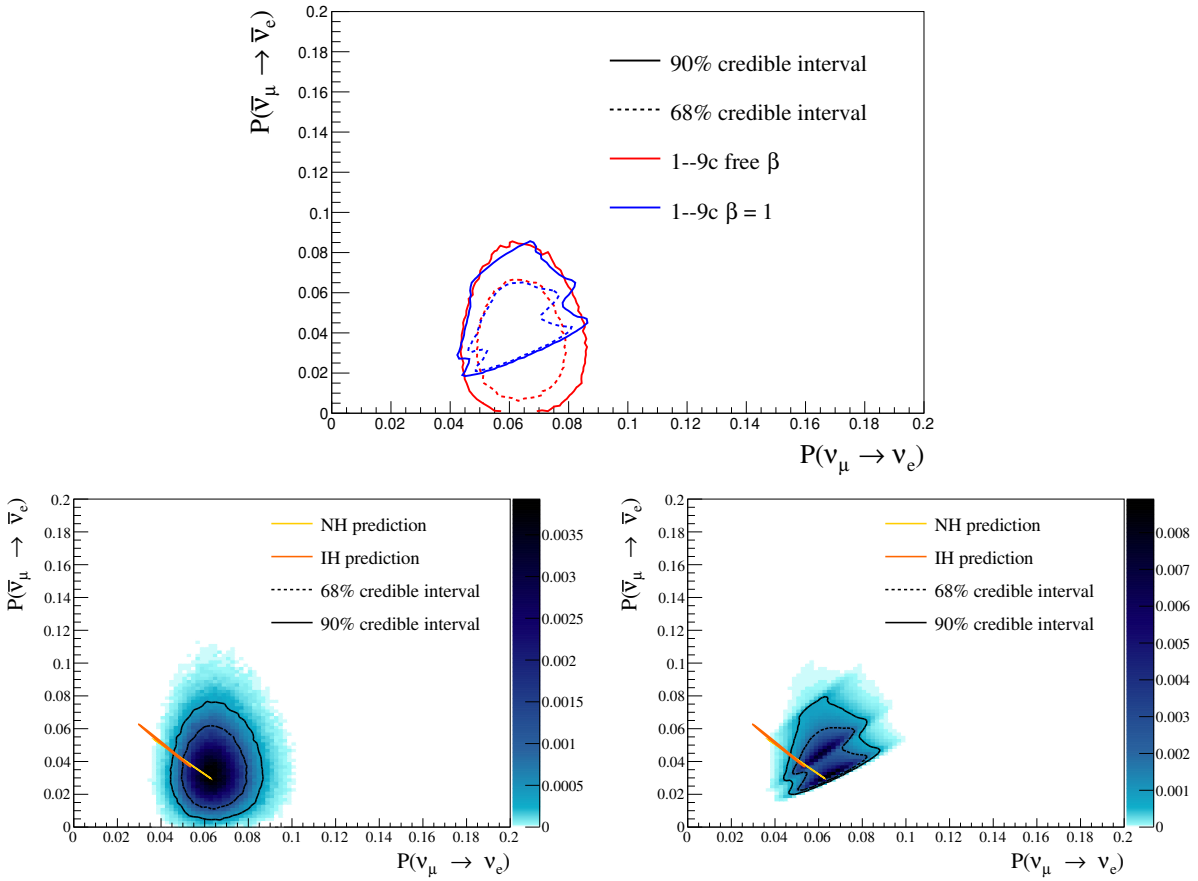


Figure 5.5: Appearance (upper) and disappearance (lower) contour comparison between the standard Asimov A fit, the free β fit and the free β with only values of β between 0.9 and 1.1 to re-create the standard fit.

5.3.2 Priors

The most important oscillation parameter prior in T2K from external experimentsis that on θ_{13}, as it has a large effect on δ_{CP}. This constraint is taken from reactor experiments that assume the PMNS model when providing their central values and errors. In this analysis we are trying to break the unitarity of the PMNS model, so it would be counter-

intuitive to use such external priors.

Although all the relevant priors are flat, it is still very important to consider the space they are flat in, for example, the prior on δ_{CP} can be flat either in δ_{CP} or $\sin\delta_{CP}$. Figure 5.6 shows the comparisons between the run 1–9 data fit results with priors for δ_{CP}, θ_{23} and θ_{13} flat in different spaces. The effect of priors on the bi-probability contours is very small. Although δ_{CP} plots do show a visible difference, this is expected given the parametrisation and is still far smaller than the statistical fluctuations from the data.

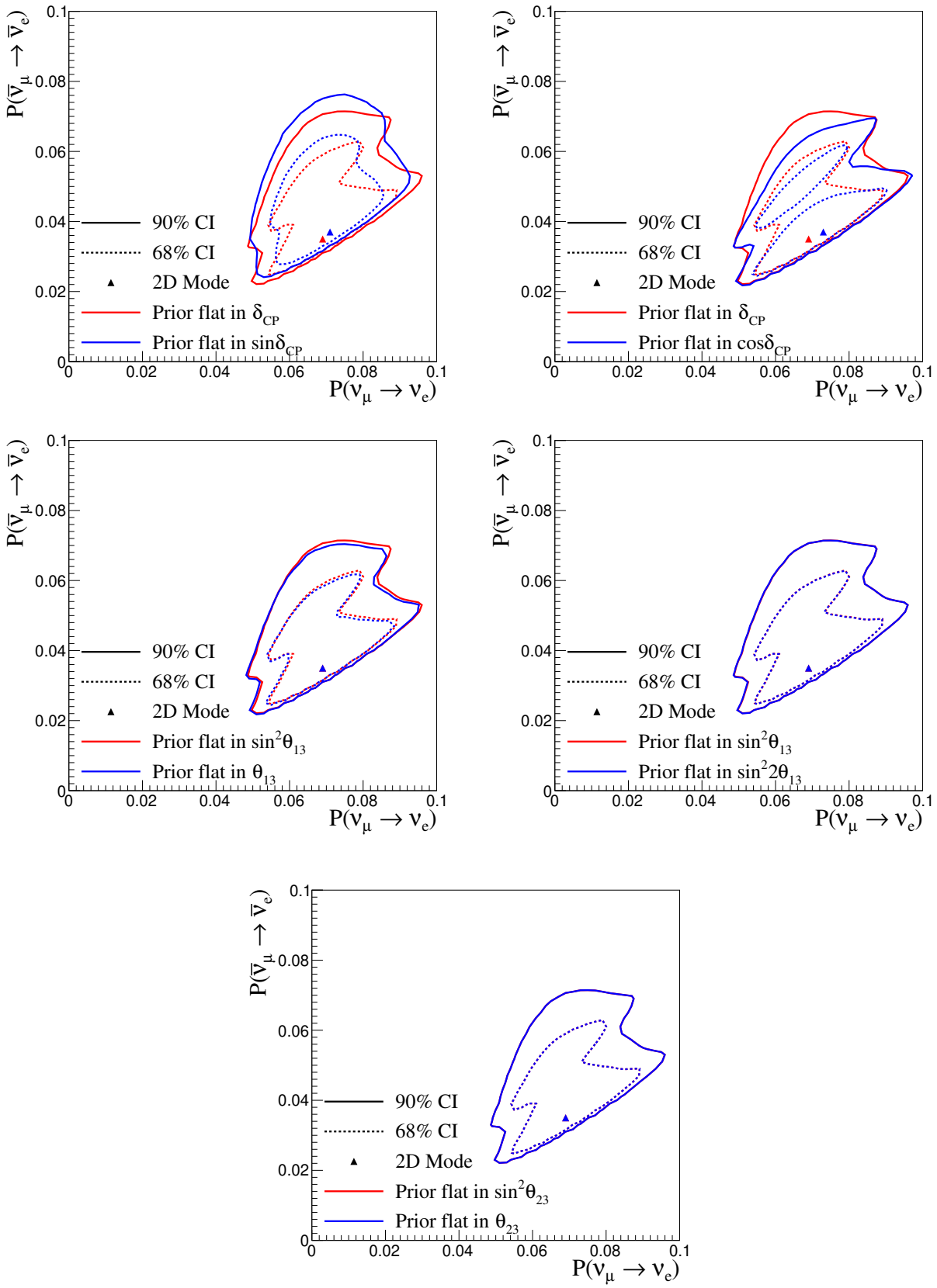


Figure 5.6: The effect of priors flat in different spaces on the fixed-β contours from the run 1–9 analysis. The largest effect is from the δ_{CP} prior, since the fit is done with a prior flat in δ_{CP}, and the oscillation probability takes the values of $\sin(\delta_{\mathrm{CP}})$ and $\cos(\delta_{\mathrm{CP}})$ (see section 4.4.5). The effect of these priors is still far smaller than even a small increase in the POT, re-confirming that the T2K experiment is still very much statistically limited.

5.4 Results for run 1–9

5.4.1 Bi-probability and bi-event plots

The bi-event plots showing the number of ν_e and $\bar{\nu}_e$ candidates are the most common way of displaying the appearance data in one plot. Given that the Bayesian analysis (including the bi-probability with free-β) outputs not only the oscillation probabilities in every MCMC step, but also all the systematic parameters, it is easy (however computationally expensive) with the currently-existing frameworks to convert these into ν_e and $\bar{\nu}_e$ event rates to plot against each other.

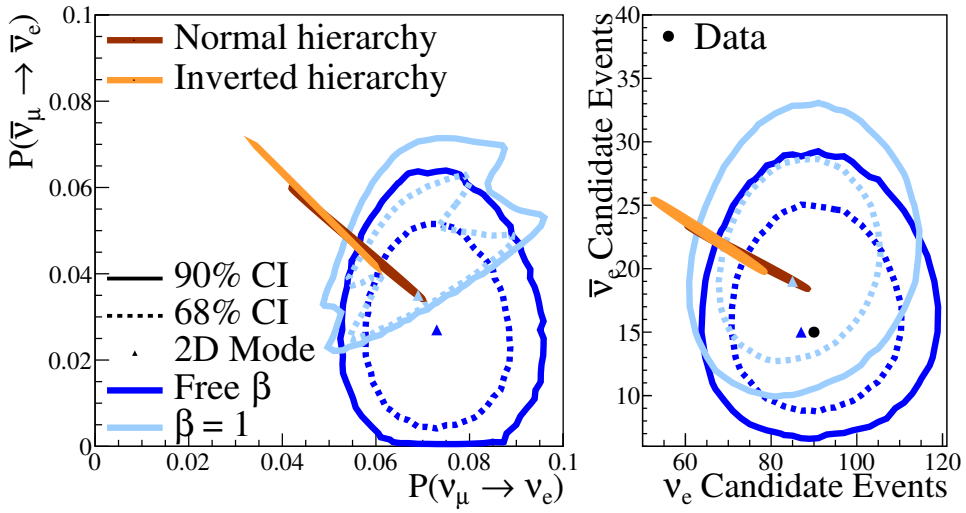


Figure 5.7: Bi-probability Bayesian credible inteval comparison (left pane) between the standard PMNS-constrained fit ($\beta = 1$) and the non-PMNS fit (free β). The right pane shows the predicted number of evnts distribution comparison, given the uncertainty in the oscillation and systematic parameters, with Poisson fluctuations applied to both PMNS-constrained and non-PMNS contours. Analysis done with the run 1–9d data and published in [136].

Figure 5.7 shows such a plot together with the data point from the run 1–9d, next to the standard bi-probability plot. We can easily see that the unitary constraint of the PMNS model does not allow for the probabilities - and therefore the event rates - to perfectly match the data point. Given the large statistical uncertainty on the data point itself,

such a deviation is still well within the statistical limits.

5.4.2 Comparisons against previous results

One of the original motivations behind the bi-probability analysis was the discrepancy between the T2K data and the model. With far fewer $\bar{\nu}_e$ events seen at SK than expected under the oscillation hypothesis, it was interesting to see whether the $\bar{\nu}_e$ appearance is real, and whether T2K actually has the sensitivity to see it if the PMNS model is not fully correct. The deviation has largely disappeared with the new data, but this analysis was also done previously with older data that still contained the apparent fluctuation.

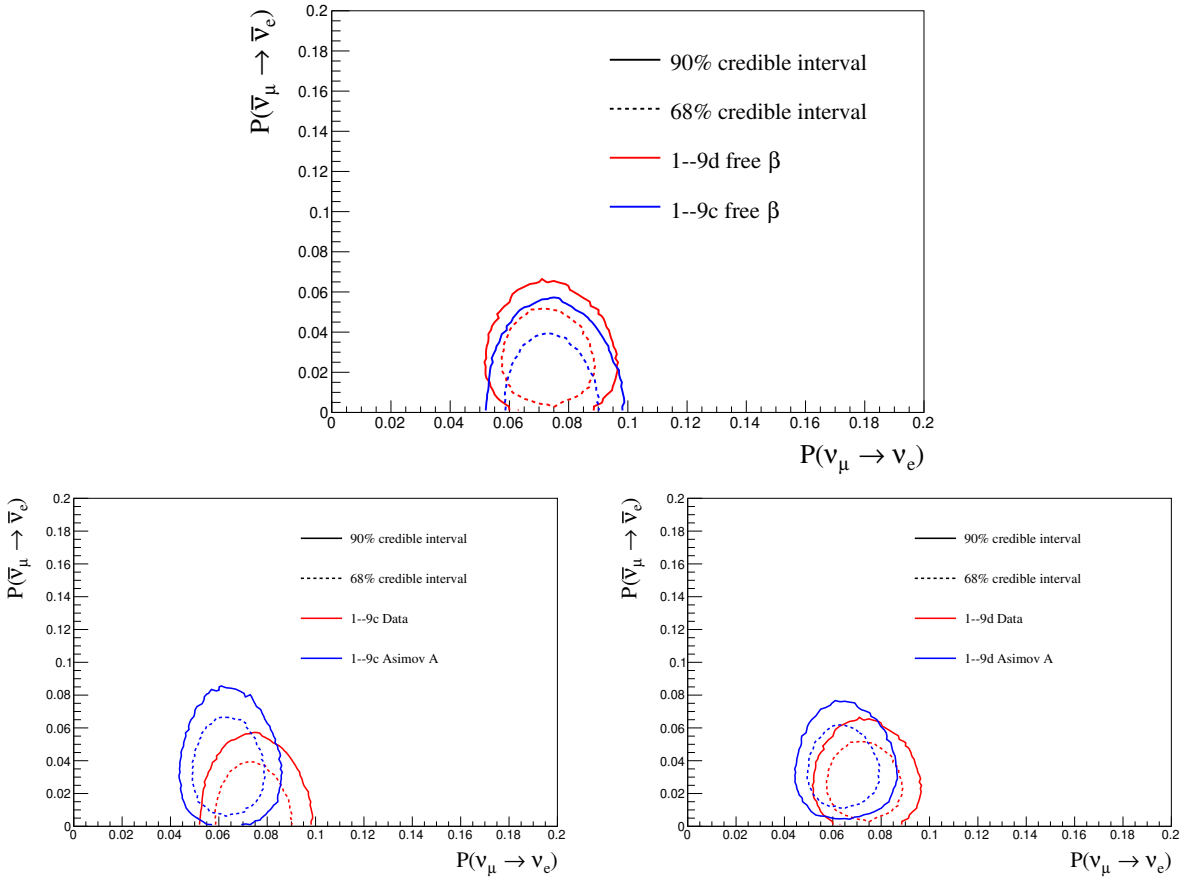


Figure 5.8: The free β bi-probability comparisons between the run 1–9d and run 1–9c data fits (top), run 1–9d Asimov vs data fits (bottom right), and run 1–9d Asimov vs data fits (bottom right).

Figure 5.8 shows various comparisons against the older fit, done with the run 1–9c data, together with the data vs Asimov comparisons. The old data clearly prefers the no

$\bar{\nu}_e$ appearance area, but with the statistical uncertainty that still covers the T2K phase-space, whereas the fits with the new data have moved away and are now closer to the Asimov expectation. This suggests the following; first, T2K currently really does not have the sensitivity to probe beyond-the-PMNS physics to a significant level. Secondly, the fluctuations in the T2K data year-to-year can have large impact on the T2K parameter space.

In many ways, this agrees with an interesting effect from the previous chapter. Although a lot more data was added, the T2K contours did not move much in the run 1–9d analysis. This can be explained by the T2K data moving closer to the expectation from the PMNS model, reducing T2K's sensitivity to δ_{CP} and the mass hierarchy. At the same time, more statistics do increase this sensitivity, and there is an interplay between these two effects that reduces the year-to-year changes in the contours.

5.5 Summary

The bi-probability plots are now a new T2K result that can be used by the theorists and other experiments to compare their predictions against the T2K data-driven results. The free β analysis breaks the PMNS unitarity and allows for oscillation parameter combinations that are not normally allowed. This formulation was used to show that the T2K data is still in a good agreement with the PMNS model and more data needs to be taken to have enough statistical significance to see any discrepancies. Furthermore, these studies have re-confirmed that the addition of new data can still significantly change the T2K phase-space, partially because of limited statistics, and partially because of the large deviation in the number of ν_e and $\bar{\nu}_e$ data events from the T2K MC. This analysis was peer reviewed and published in [136].

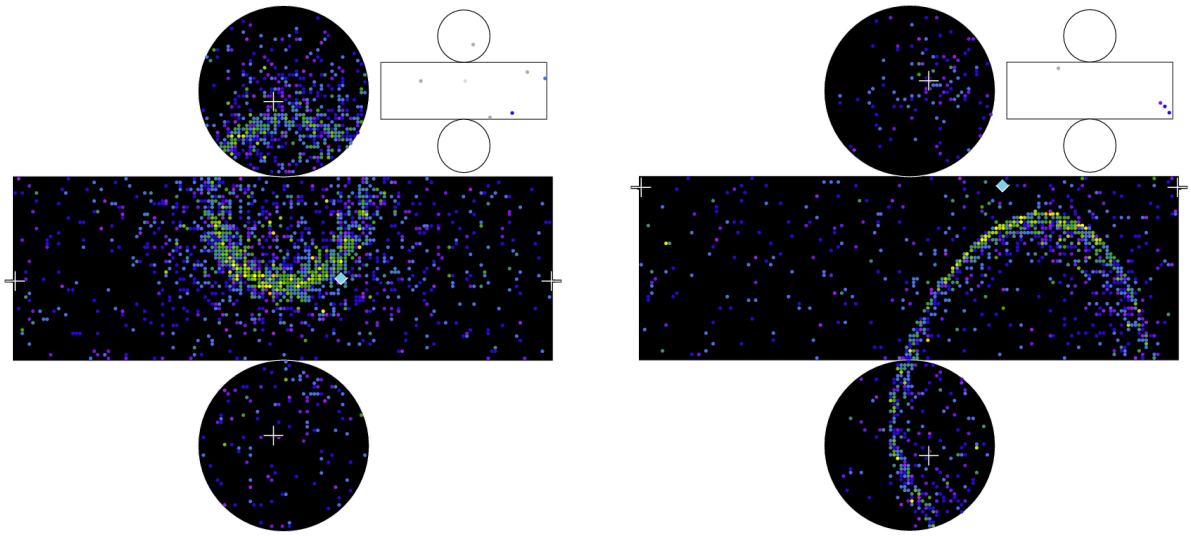
Chapter 6

Bayesian Neutrino Event Reconstruction for Water-Cherenkov detectors

This chapter describes a Water-Cherenkov event reconstruction software package used by Super-Kamiokande, fiTQun, and work to implement a new, Bayesian MCMC sampler into the package. The work described in this chapter involves predominantly studies of ideas and concepts, with more concentrated work required to develop the framework to become usable in physics analysis. Description of the standard functionality of fiTQun is described in 6.1, with the following sections describing the Reversible-Jump MCMC algorithm, and its implementation and tunings in fiTQun.

6.1 Neutrino event reconstruction at Super-Kamiokande

A neutrino interaction above the Cherenkov threshold at SK will provide two pieces of information at every PMT: the accumulated charge, and the hit time. Figure 6.1 shows example events, with their charge shown per PMT in colour. The charge and hit time per PMT are the only "signal data" recorded for each event that can be used for modelling of a neutrino inside of the detector. The reconstructed kinematic parameters are needed for any oscillation analysis, whether for atmospheric or beam neutrinos. These kinematic parameters are the charged particle's origin vertex in the 3D space of the detector $\mathbf{x}$, interaction time t, charged particle trajectory angles θ and ϕ, the visible energy based on the accumulated charge in the PMTs E_{vis}, and the particle identity, PID.



(a) Electron-like ring, with the hit color indicating the accumulated charge per PMT.

(b) Muon-like ring, with the hit color indicating the accumulated charge per PMT.

Figure 6.1: Two SK unrolled event displays with an electron-like (left) event and muon-like (right) event. The display pixel color represents the accumulated charge for each PMT separately, with the more red color indicating more accumulate charge. The Outer Detector displays are shown on top right of each plot, with very few OD hits.

6.2 fiTQun

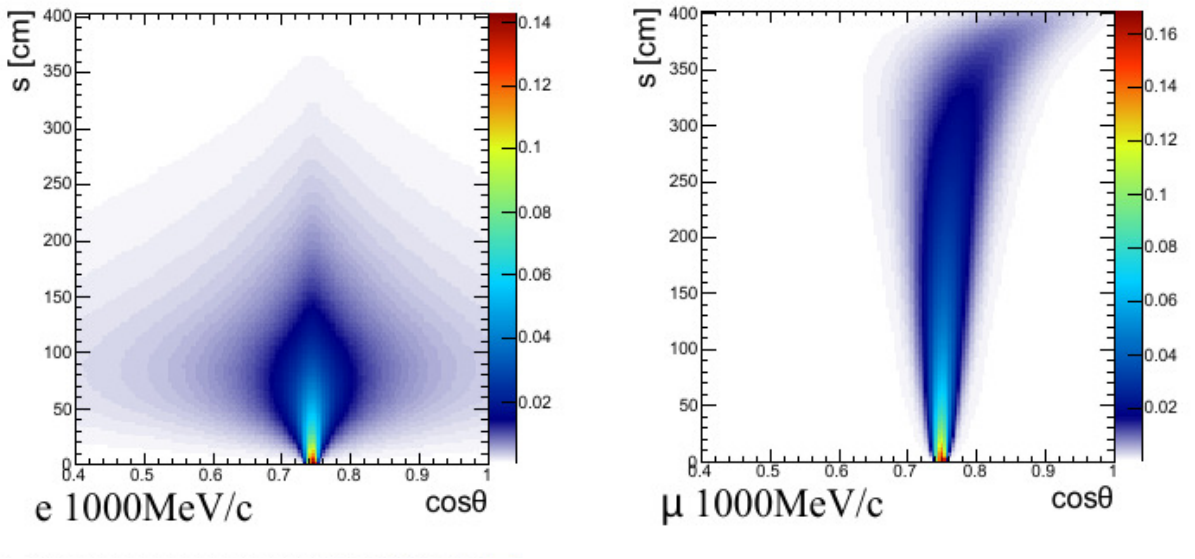
The current state-of-the-art neutrino event reconstruction software for SK, fiTQun, uses the likelihood method to extract the charged particle's kinematics, and was originally developed for the MiniBooNE experiment [118]. It constructs the joint likelihood using the available charge and timing data from each PMT, and minimizes it using MINUIT [116, 117]. The likelihood calculation at fiTQun uses information from all the PMTs, whether hit or "unhit";

$$\mathcal{L}(\mathbf{x}) = \prod_{j}^{\text{unhit}} P_j(\text{unhit}|\mu_j) \prod_{i}^{\text{hit}} \{1 - P_i(\text{unhit}|\mu_i)\} f_q(q_i|\mu_i) f_t(t_i|\mathbf{x}) \,, \tag{6.1}$$

where $f_q(q_i|\mu_i)$ is the charge likelihood given the predicted charge per PMT, μ_i, and $f_t(t_i|\mathbf{x})$ is the time likelihood. The probability of a PMT not being triggered, $P_j(\text{unhit}|\mu_j)$, given that the μ charge distribution follows the Poisson distribution, is simply proportional to $e^{-\mu}$ with additional MC-derived correction terms to account for the PMT threshold. Finally, the predicted charge per PMT, μ_i, is defined as the mean number of liberated photoelectrons.

The time likelihood, $f_t(t_i|\mathbf{x})$, has contributions from the direct light, assumed to be a Gaussian PDF fitted to an MC simulation of various PIDs at different momenta and predicted charges, and the indirect light PDF (scattered and reflected) is included in a form of a simple function that is not dependent on the momenta or predicted charges.

The $\mathbf{x}$ from the direct light can be parametrized in terms of the particle distance from the PMT, s, and the angle between the particle and PMT, θ. This parametrization is used for the charge from the direct light, μ^{dir}, which is constructed using the Cherenkov emission profiles shown in figure 6.2. These emission profiles are made in the GEANT3-based SK detector simulator, SKDETSIM, generated at different discrete momenta for each PID. We can see from the figure that as the muon travels, the cone angle changes — this creates the "cone collapse" effect, where the muon Cherenkov rings are sharp on the outside, and become more fuzzy on the inside. Furthermore, although not shown here, the Cherenkov opening angle is the same for electrons of all the momenta, and varies for muon momenta. The predicted charge from the indirect light, μ^{sec}, is made



(a) Electron predicted charge distribution for selected momentum

(b) Muon predicted charge distribution for selected momentum

Figure 6.2: Cherenkov emission profiles for electron (left) and muon (right) as seen by a PMT in terms of the particle distance from the PMT, s, and the angle between the particle's trajectory and the PMT, θ. Both shown emission profiles are for a fixed momentum of 1000 MeV/c.

by building scattering tables from simulations of 3 MeV/c electrons originating from a common vertex, with isotropic angular distribution.

6.2.1 Pre-fitting and event clustering

The $-\ln\mathcal{L}$ from equation 6.1 is minimized using gradient descent with MINUIT, which is sensitive to any local minima in the likelihood distribution. This poses a large challenge for fiTQun, as many local minima can be found even in the 1-ring hypothesis likelihood – an effect that gets worse the more Cherenkov rings we try to reconstruct. A series of pre-fitters is run before progressing to the main fiTQun ring fitter. These pre-fitters try to find the global minimum and split the event into sub-events when needed (for example, when we have a muon and a Michel electron).

Vertex pre-fitter

The vertex pre-fitter is a fast algorithm that tries to find an approximate global minimum by only looking at the PMT hit times to find the particle's $\mathbf{x}$ and time by maximizing the "goodness" metric,

$$\mathcal{G}(\mathbf{x}, t) = \sum_{j}^{\text{hit}} \exp(-(T^{i}_{\text{res}}/\sigma)^2/2), \tag{6.2}$$

where T^{i}_{res} is the residual time dependent on the PMT position, particle kinematics to be extracted and the PMT hit time. An iterative grid-search is performed, with σ being the time-window width for each grid, and MINUIT is used for the maximization of the equation 6.2. After the first grid-search is finished, the grid size with the time-window of 500 ns are decreased, and grid search repeated, until the time-window is $\sigma = 4$ ns to provide the final pre-fit vertex and timing.

Sub-event clustering

An event for reconstruction at SK is recorded inside a ∼400 ns time window. These events might contain sub-events, such as decays, scatters, and so on. There are a few sub-event clustering algorithms that try to split an event into its sub components.

A peak-finder algorithm tries to find multiple peaks in the "goodness" distribution using equation 6.2. The approximate vertex from the vertex pre-fitter is fixed, and the "goodness" function is scanned over the time. For each scan point, the threshold function for selecting a subevent (a simple function of "goodness") is calculated and compared against the "goodness". Figure 6.3 shows the "goodness" for each scan point compared against a "threshold" function, with two peaks found (red vertical lines) for an MC event with one muon and a Michel electron, with both successfully found.

A peak is chosen by comparing the scan points against a threshold, which is optimized using atmospheric samples to minimize MC-Data reconstructed differences.

There is another algorithm that tries to cluster hits by time, creating "time clusters". It first bins the event in hit-time into 10 ns bins, and finds the starting position, t_{start}, where the first 7 bins contain more than 9 hits – a number determined from the simulated

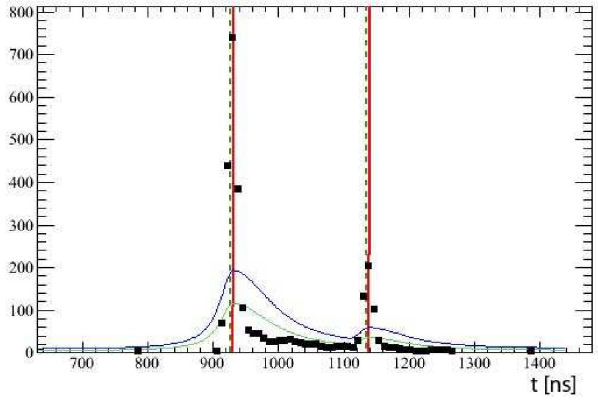


Figure 6.3: The "goodness" distribution, $\mathcal{G}(\mathbf{x}, t)$ calculated for each scan point (black squares), as a function of time for a muon-like MC event with a Michel electron. The overlaid lines are the MC-true vertices (dashed vertical lines), "threshold" functions (blue and green lines), and the final pre-fit vertices (solid vertical red lines). The "goodnes" points are above the "threshold" function in two places, indicating that both subevents were found successfully. Figure taken from [116].

MC noise – and moves the t_{start} backwards in time by 30 ns to catch early hits. It then keeps moving the 70 ns window by 10 ns at a time until the number of hits is less than 9, and the difference between t_{start} and t_{end} is larger than 70 ns. Finally, t_{end} is extended by a function of total hits in the window, to catch later hits from, say, reflections, and lower-energy events that fall below the threshold too quickly.

A third algorithm tries to match the sub-events found by the peak-finder and the time-clusters found by the time-clustering algorithm described above. The best-fit vertices found by the peak-finder are associated with every time cluster to see how many hits match in these time-windows, and the "goodness" is re-evaluated using these hits. Peaks with no time-clusters associated through the "goodness", and time-clusters with no associated peaks, are discarded to reduce the number of fake peaks and clusters that can happen when the vertex estimation is poor. For some events, there will be multiple peaks associated with the same time-window, called "in-gate" sub-events. When two time-clusters are very close to each-other they can be passed through the clustering algorithm as one. If there is one peak associated with the double-peaked time cluster, all the hits within a cluster are used by fiTQun. If there are multiple peaks associated with a multiple-peaked time cluster, the secondary sub-events are associated with each

PMT separately. This is done within that window using $-30\,\text{ns} < T^i_{res} < 60\,\text{ns}$, where T^i_{res} is the residual time difference between the expected photon arrival time (assuming it is produced from the middle of the track), and the PMT hit-time. All the hits not associated with secondary sub-events are assigned to the best-goodness event within the time-window.

6.2.2 Main ring fitter

One-ring reconstruction

After the sub-events are found and matched between the peak-finder and the time-clustering algorithms, fiTQun uses MINUIT minimizer on the likelihood function (Eq. 6.1) to find the best-fit kinematics for each of these sub-events. First, after the seeding value is set to the main pre-fit peak, fiTQun performs a likelihood grid-scan over the track direction angles, with all the other kinematics fixed. A second likelihood scan is done over different values of momentum, this time with the direction (and other kinematics) fixed. MINUIT is then finally used, with all the kinematics seeded and free in the fit, to minimize the negative log-likelihood function and obtain the best-fit. This is done for the electron hypothesis first. The second fit, for the muon hypothesis, is first seeded with the best-fit values from the electron hypothesis fit. The log-likelihood is scanned over different values of momentum once again, before MINUIT is used once again with the newly-seeded values, and repeated again, this time for a charged pion hypothesis.

It is not as simple as choosing the best-likelihood hypothesis to determine the particle PID – although this in principle should be true, it is dependent on the momentum. Figure 6.4 shows atmospheric ν_e and ν_μ 1R CCQE events, with their $\ln\mathcal{L}$ ratio against the reconstructed momentum from the electron-hypothesis fit. The electrons mis-reconstructed as muons on the left plot (above the cut line) are mainly low-energy muons close to the Cherenkov threshold.

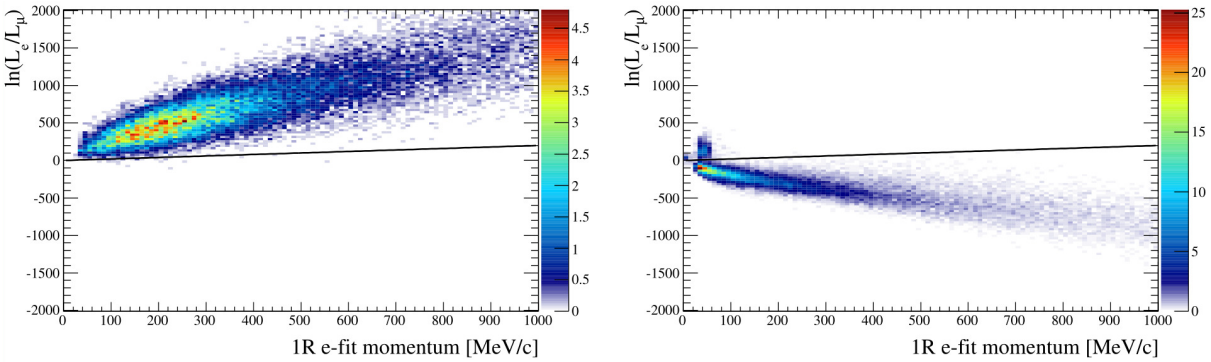


Figure 6.4: The log likelihood ratio, $\ln(\mathcal{L}_e/\mathcal{L}_\mu)$, between one-ring fits to the atmospheric ν_e (left) CCQE and ν_μ (right) CCQE samples, as a function of the electron-fit momentum. The black line represents the cut on $\ln(\mathcal{L}_e/\mathcal{L}_\mu)$ to select the PID, obtained from the MC.

Multi-ring reconstruction

The single-ring fit is relatively easily expanded into a multi-ring fit. The likelihood probability is expanded by adding an additional secondary-track term to the time PDF. This is done by making the time distribution more dependent on the distance between each track centre and the PMT for each track. The charge distribution PDF is just the integral over all the tracks per PMT. The difficulty due to extra non-linearities in the likelihood distribution causes the fit to assume a shared reconstructed 4-vertex between all the tracks in the free fit.

The best one-ring hypothesis is chosen as the first ring, with the kinematic parameters used as the initial seed for the fit to the second ring chosen by running the angle and momentum scans, similarly to the one-ring fit. The multi-ring fitter repeats the fitting procedure for all the combinations of the best one-ring-hypothesis and all the PID hypotheses for the secondary ring, with similar a log-likelihood ratio to choose the second ring. Another log-likelihood ratio is then needed to choose the best number-of-rings hypothesis. This is non-trivial, since the log-likelihoods with different numbers of parameters do not map to each other in the way different PIDs (or simply different kinematic parameter combinations) do. MC-derived cuts are needed, with figure 6.5 showing the $\ln(\mathcal{L})$ ratio between the 2-ring and 1-ring hypotheses, with an MC sample that only contains 1 electron track. We can see far more mis-reconstructed events than for the PID-hypothesis testing, contributing to the multi-ring samples being less efficient.

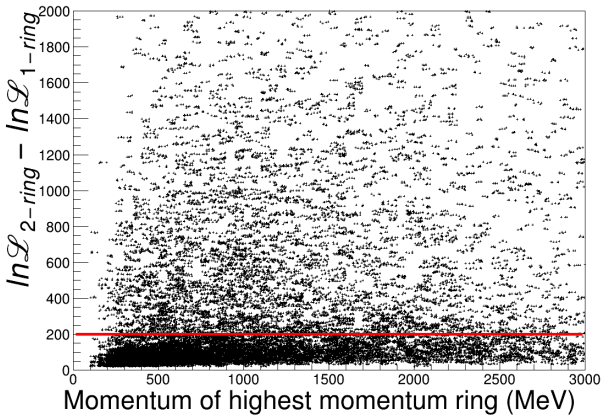


Figure 6.5: The log likelihood ratio, $\ln(\mathcal{L}_{2R}/\mathcal{L}_{1R})$, between one-ring and two-ring fits for a ν_e CCQE-only sample, with a red line showing a cut to separate the 1R (below the line) and 2R (above the line) hypotheses.

There is a dedicated fitter for the π^0 event hypothesis that results in two gammas with specific invariant mass. Just like in the general multi-ring fitter, the best-fit one-ring hypothesis (in this case must be an electron) is chosen for the first ring seed, and a log-likelihood scan is performed over the second ring's direction. The shared vertex is chosen to be 50 cm, which is the average photon-to-electron conversion length in water. The momentum is then scanned over for both electron-rings, since the 1-ring fitter usually overestimates the momentum for π^0 events. The full fit with all the parameters free (aside from the shared vertex) is performed, with similar log-likelihood ratios used to choose between 1Re and π^0 hypotheses.

6.2.3 Drawbacks

The reconstruction package works quite well for one-ring samples, although the multi-ring fitter uses approximations, together with incremental likelihood scans and fits for different hypotheses representing different number of rings, each of which possesses their own PIDs. Being a fitter, fiTQun finds the best-fit kinematic parameters, without providing uncertainties that are fundamental to the data. These uncertainties could be of use for, say, events with a low number of PMT hits or close to the wall. Having full uncertainties for each kinematic variable could potentially improve the selections and make cleaner signal samples — e.g. if we choose the *wall* cut to be a vertex at least $1\,\sigma$ away from the

fiducial volume cut.

Although the pre-fitters do a good job at trying to find the global minima and clustering the events in time, fiTQun still can mis-reconstruct events, especially for multi-ring hypotheses. The log-likelihood cuts need to be determined empirically from the MC, which can be difficult and is constantly being updated and improved. This decreases the efficiency of the multi-ring samples and even has an impact on single-ring hypotheses, as seen in figure 6.5.

Finally, the single-vertex approximation makes it difficult for fiTQun to deal with pileup-events, and does alter the best-fit reconstructed kinematics, however slightly.

6.3 Reversible-Jump MCMC

An MCMC algorithm, Reverse-Jump MCMC (RJMCMC), was proposed and partially implemented to overcome the inherent difficulties faced by fiTQun. Using MCMC means obtaining full posterior probabilities for each kinematic parameter, and being able to construct credible intervals that could then be used for "smarter" selection cuts for various data samples. Being a sampler, it could potentially also overcome local minima and reconstruct each track with its own vertex – which could make the reconstructed-MC – MC-truth differences smaller.

Reversible-Jump MCMC algorithm is a more generalized version of Metropolis-Hastings MCMC described in chapter 3, with the capability to add and subtract components from its sampled Mixture Model as a valid MCMC step [92, 122].

A Mixture Model could be used to sample from all the number-of-ring and PID hypotheses simultaneously [122]. It is defined as the total sum of all the PDF combinations, with each combination having its own assigned weight, with all the weights summing to one,

$$f_{\text{mix}} = \sum_{i=1} w_i f(\theta_i), \tag{6.3}$$

where θ_i is the parameter combination for the i'th component, in our case, a Cherenkov ring with all its seven kinematic parameters. Since all the weights have to sum to 1, the

total probability stays the same with a changing number of components. These weights are free parameters in the sampler; deviating the weights from their current value is a valid MCMC step, as long as they sum to unity.

6.3.1 Trans-dimensional step acceptance probability

The standard Metropolis-Hastings acceptance ratio must change to take the dimension-matching into account when comparing the log-likelihoods (this is analogous to the log-likelihood difference explained in 6.2.2). The new RJMCMC acceptance probability is

$$A = \min\left[1, \frac{\mathcal{L}_{new} \times P_{new} \times Q_{new \to old}}{\mathcal{L}_{old} \times P_{old} \times Q_{old \to new} \times \Delta q_{\text{old}\to\text{new}}} \times |\mathcal{J}|\right], \tag{6.4}$$

with $\mathcal{L}$ and P being the standard data-likelihoods and priors seen in the Metropolis-Hastings MCMC. The new terms are; Q, which defines the probability of proposing a new step from the previous one (or vice-versa): $\Delta q_{\text{old}\to\text{new}}$, which is the prior difference between the proposed and the current model in the MCMC proposal, and $|\mathcal{J}|$ which is the determinant of the Jacobian matrix. The last factor is not required when we do not transform the parameters during a trans-dimensional step proposal.

There are a few interesting points that can be made about the new acceptance probability 6.4. The probabilities to propose a new step with a new parametrization, $Q_{new \to old}$ and $Q_{old \to new}$, cancel each other out if the proposal is symmetric. This is not only true for RJMCMC; it is possible to implement such terms in the Metropolis-Hastings algorithm, and have an asymmetric step proposal function without violating the "subtle balance", which could have interesting applications. The $\Delta q_{old \to new}$ term matches the two different dimensions together by padding the phase-space with extra penalties drawn from the prior. The sign of the penalty is dependent on the direction in which the trans-dimensional step is taken. The $|\mathcal{J}|$ term is usually equal to 1 depending on the nature of the trans-dimensional proposal. If the new component's parameters are not chosen from the prior (i.e. are transformations of already-existing components), or, when removing one component, other component's parameters change, this term assures the subtle balance is preserved.

6.3.2 Trans-dimensional step proposal

There are various ways of proposing trans-dimensional steps in RJMCMC that are commonly used. One way, which is usually the least efficient in complex applications, is to propose a new number of components directly from a prior. This means, for example, allowing a direct jump from one component in the model to ten components in a model. This can ensure easy handling of the Jacobian term (reduces to 1), and implementation, whilst possibly sacrificing the efficiency.

A more commonly used method of proposing a new trans-dimensional step is via Birth/Death and Split/Merge proposals, first described in [122]. Both moves only change the number of model components by one, and are meant to be symmetrical in nature, with use of weak priors.

Split/Merge

The Split/Merge move usually has a probability b_{old} to split one component into two, and $d_{old} = 1 - b_{old}$ probability to merge two components together. The b_{old} probability can be set to be 0.5 at all times, except for when at a boundary (for example, when we have 1 component in the model), or set to be a function of the number of components depending on the prior. When merging, two model components are chosen at random, and their component parameters are combined using a weighted sum – with the weights corresponding to the ones in the equation 6.3.

Consequently, the "Split" proposal is performed by choosing a random component, splitting its weight, ensuring all the model weights still sum up to 1, and then using the two new weights to split the component's parameters into two.

Birth/Death

The Birth/Death proposals either add a new component to the model, or remove a random component. These proposals are simpler in principle, but can be very difficult to implement effectively in practice. The choice between Birth and Death is made prob-

abilistically; b_{old} and d_{old} are used, just like in the Split/Merge move. For a birth, a new weight w_{new} is chosen for the new component, and other weights are re-scaled with $w_{i'} = w_i(1 - w_{new})$ first. New parameters are chosen from priors, which significantly reduces the complexity; the Jacobian does not have to be calculated.

For the Death move, a component is simply removed from the Mixture Model, and the weights renormalized so that they add up to one. This is all simple in principle; in reality choosing new component parameters directly from priors often makes for a very inefficient algorithm. Indeed, it is a common issue with RJMCMC; the Birth proposal function is often the most challenging, and consequently, inefficient part of the implementation [123].

6.4 fiTQun-RJMCMC

There are six MCMC proposals implemented in fiTQun; Split/Merge, UpdateParameters, UpdatePID, UpdateEvis, UpdateEloss and BirthDeath. Every time a new MCMC step is proposed, one of these proposals is chosen at random, with equal probabilities for being picked for each. It is good to note that only Split/Merge and Birth/Death proposals are the ones that require the additional terms in the acceptance probability for dimension-matching – the remaining proposals are basically the same as in the standard Metropolis-Hastings MCMC. Figure 6.6 shows what may happen for each RJMCMC step in fiTQun.

6.4.1 Update proposals

All the proposals with "Update" in their name in figure 6.6 are from within the currently-sampled model, i.e. the number of rings does not change. The step proposal functions are a random Gaussian number centered at 0, that are added to each variable being updated, with the Gaussian width being tunable from configuration files.

The UpdateEloss proposal updates the E_{loss} variable, which is only relevant for a charged pion hypothesis. Pions, having a similar mass to muons, should deposit similar

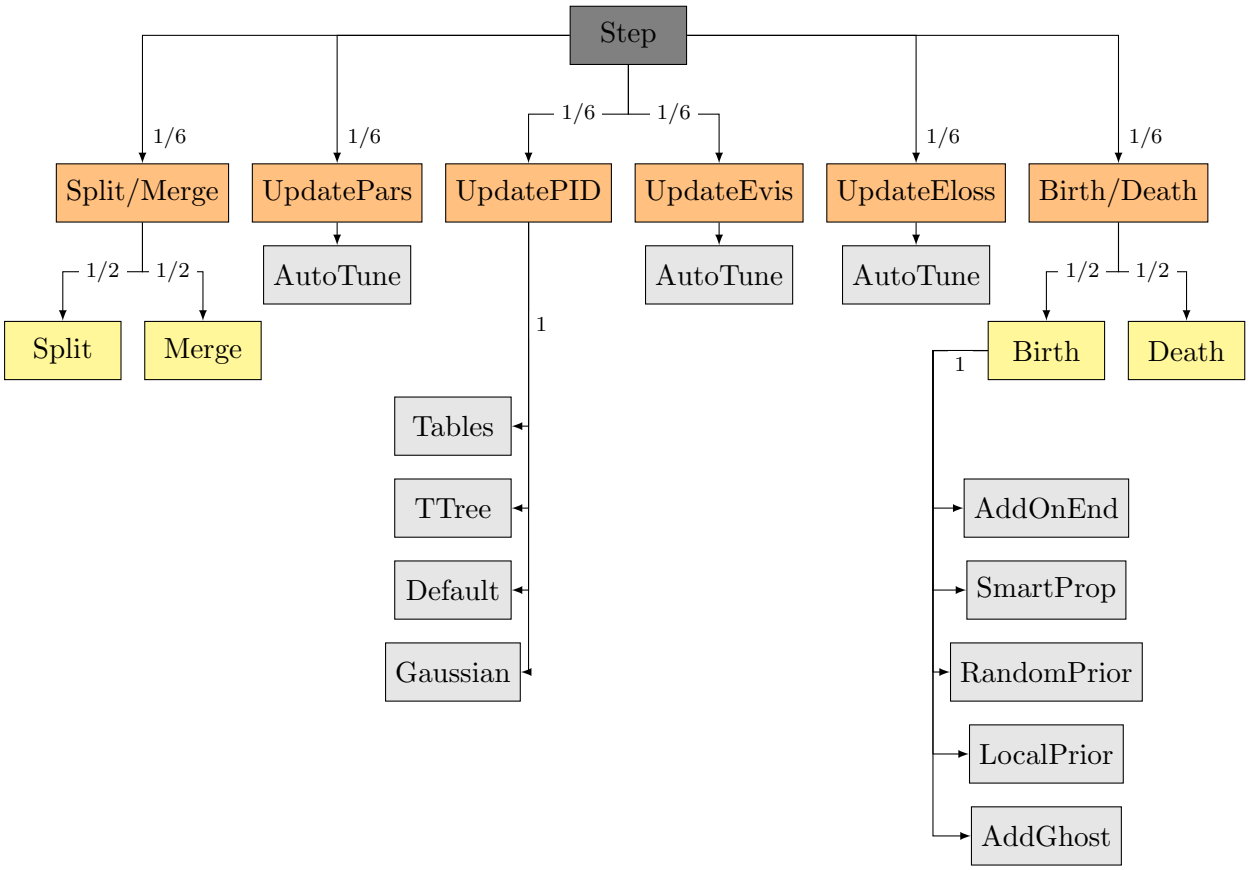


Figure 6.6: A box diagram showing how an example RJMCMC step is executed. Every step there is an equal probability of choosing one of the proposals (orange boxes). Some proposals are split into two with an equal chance of choosing either (yellow).

amounts of charge in the detector. They do, however, undergo hadronic interactions that might suddenly stop the particle. There were a few implementations the UpdateEloss proposal considered when we have a non-pion hypothesis – the final implementation was to repeat the same E_{loss} for non-pions when UpdateEloss is called, and only vary E_{loss} for rings that correspond to pion PID.

The UpdateEvis proposal updates the total reconstructed visible energy in the detector. This visible energy is being split by model-component weights between the rings, and is being used, together with the PID and its Cherenkov threshold, to calculate the particle's momentum.

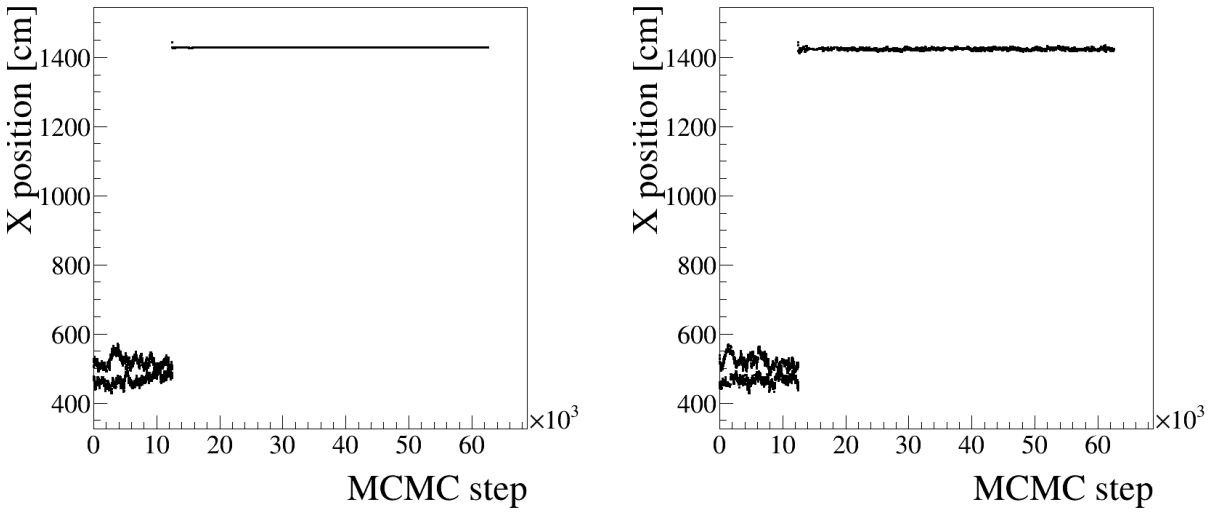
The UpdatePars proposal simply updates the remaining kinematic variables, together with the model components' weights (and therefore, momenta for each ring, without altering the total visible energy). This, it turned out, is difficult to tune; since events

are often very different from each other when trying to reconstruct, their posterior distributions can be very different, and consequently, the optimal step-sizes. In fact, switching between the number-of-rings hypotheses might already be difficult; a very-well localized muon ring might require far smaller step-sizes than a low-energy, "fuzzy" Michel electron. This created a problem where step-sizes that work for one event are completely failing for another. Furthermore, the more rings (and therefore parameters) we add to the model, the more we should reduce the step-sizes as the autocorrelations increase with the dimensionality.

A new "AutoTune" function was implemented, which tries to tune the step-sizes for each separate MC event during the MCMC burn-in (the case of fiTQun, 50–75% of the chain). For each update-style proposal (apart from UpdatePID), it calculates the acceptance ratios in a 500-step window for every number-of-rings hypothesis. If that acceptance ratio is lower than 0.2 in this hypothesis, the step-sizes are reduced, and if the ratio is above 0.3 – increased. Figure 6.7 shows the parameter traces before and after the implementation of automatic step-size tuning.

6.4.2 Update PID

The proposal updating PID, UpdatePID, went through a few implementations in fiTQun — it turned out to be one of the more challenging proposals to tune, with the full implementation still to be finished. The "Default" way of switching the PID, by simply changing it, results in a large jump in the likelihood, causing a very low acceptance ratio and almost all the PID-swap proposals being rejected. Figure 6.8 shows the log-likelihood difference when throwing an electron $\rightarrow$ muon proposal. It is clear that a simple PID switch results in a large step in the log-likelihood, which turns out to be true for most MC events that the framework was tested on. To understand this further, we need to look at the differences between how an electron and muon signal are seen by the detector. The best-fit vertex (and consequently, the reconstructed interaction time) for a muon and electron hypothesis can be very different for the same event, whether the true PID is one or another. Furthermore, the best-fit momenta should differ too, although that will be taken care of by the E_{vis}-to-momentum calculator.



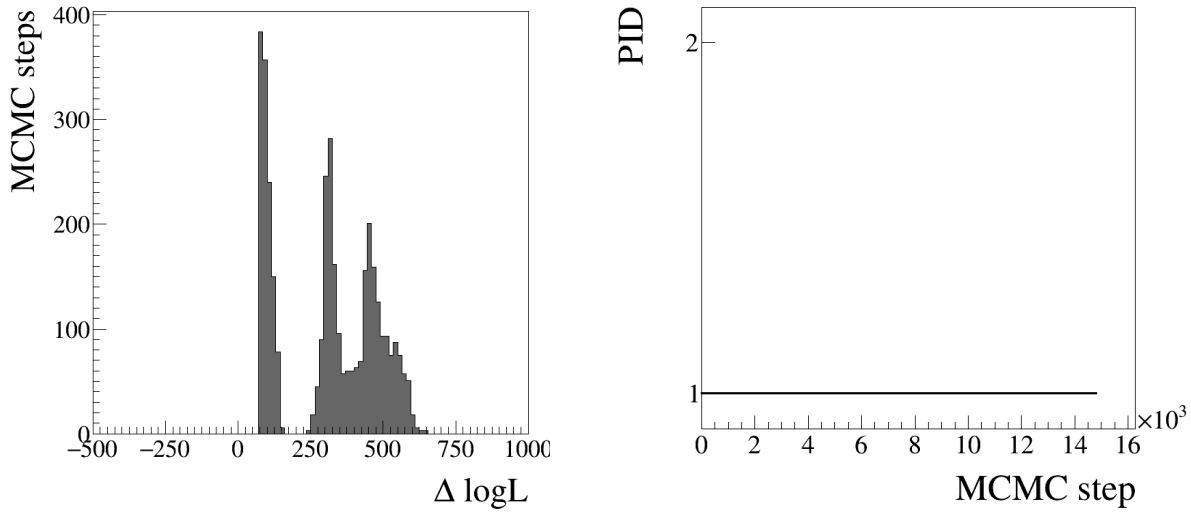
(a) Without autotuning: sampler stops moving after $\sim$ 20,000 steps

(b) With autotuning: sampler explores the one-ring hypothesis

Figure 6.7: MCMC traces showing the evolution of the horizontal X position with the number of MCMC steps for an example MC event without autotuning (left) and after autotuning (right). The first 50-75% of the chain is considered a burn-in, so the final posterior will only contain the one-ring hypothesis. We can see that without autotuning the MCMC is not changing the parameter value once converged to $\sim$ 1400 cm.

Originally there was an option to use PID transition tables of a somewhat unknown origin – they could have been derived from the MC, whether SK atmospheric or T2K. An example PID transition table for the muon to electron hypothesis is shown in figure 6.9, where a random sample from this table is taken for every $\mu \to e$ transition. The sample's time is added to the ring's reconstructed time, and the distance is added to the reconstructed vertex in the upstream direction.

There were two other PID-transition methods implemented and tested with atmospheric MC sample; one involved using a random Gaussian to vary the vertex position, and another involved an MC-derived table (of a known origin this time, using SK atmospheric MC). The random Gaussian implementation, in various tests, was centered at -200, -100, -50, 0, 50, 100 and 200 cm, with widths of 50, 100, 200 cm. For every transition a random Gaussian number was drawn and added to the reconstructed vertex in the upstream direction. The reconstructed time was re-calculated using the distance



(a) Log-likelihood difference between electron and muon hypotheses during the $e \rightarrow \mu$ MCMC proposal

(b) PID trace, with no MCMC steps accepted from the electron hypothesis (PID = 1).

Figure 6.8: The log-likelihood difference between electron and muon hypotheses when proposing an $e \rightarrow \mu$ PID step (left), and the PID trace with no accepted PID transitions, with where $1 = e$ and $2 = \mu$ (right).

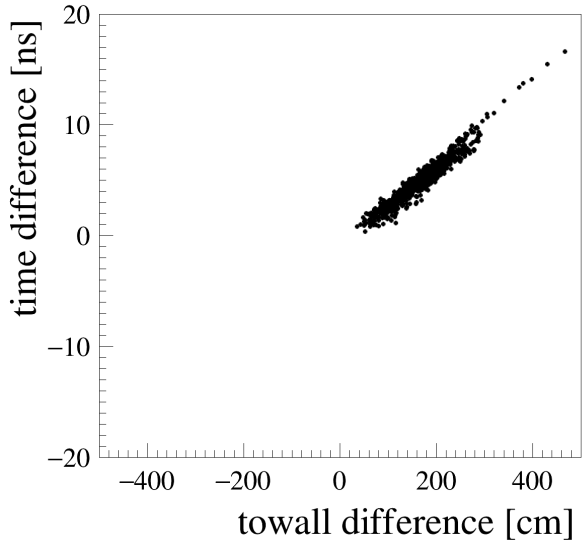


Figure 6.9: PID transition table for $\mu \rightarrow e$ proposal. A random sample from this table is taken for each proposal, and its distance added to the reconstructed vertex in the track's upstream direction, together with the time.

difference in the resultant reconstructed vertex. No tests of the Gaussian transition with the central positions and widths above resulted in any better performance or different log-likelihood difference between the proposals.

The method using MC-derived tables implementation involved producing a large n-tuple using SK atmospheric samples containing the differences between best-fit PID hypothesis – in a similar way to how the PID transition tables were made. The standard fiTQun saves all of its fits (described in 6.2.2), including all the nring and PID combinations. Fully reconstructed SK Atmospheric MC was used to find the differences in the reconstructed vertex position and time between various best-fit PID hypotheses. This was done for every event, and saved in a form of n-tuple, with an example for 1Rμ–1Re difference shown in 6.10.

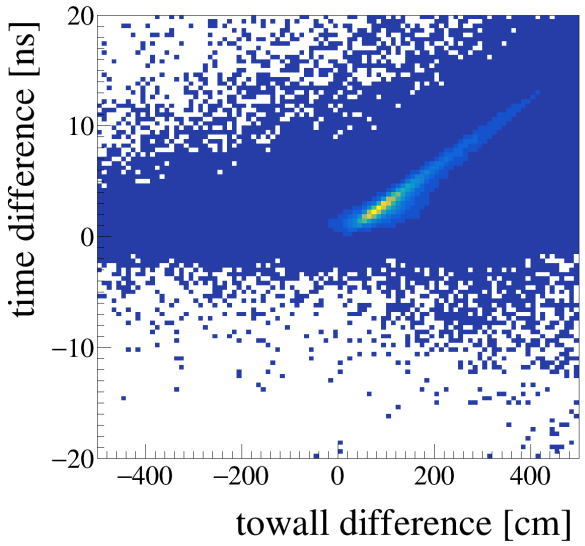


Figure 6.10: PID transition tree for $\mu \rightarrow e$ proposal, produced from the SK atmospheric MC. A sample from this distribution is drawn for each $\mu \rightarrow e$ step, and the time/towall difference added to the proposed ring parameters.

The example in figure 6.10 is very similar to the one from the PID transition table shown in figure 6.9 – with a very similar effect. The log-likelihood differences using the full transition n-tuple were still too large to have any UpdatePID proposal accepted in MCMC run tests.

More investigation is needed to make the PID transition smooth, resulting in lower

log-likelihood differences between the current and proposed MCMC steps. For example, it would interesting to look into the E_{vis}/momenta differences between the different PID hypotheses – since fiTQun uses different predicted-charge distribution tables. Another way to force a smoother transition is through the likelihood-tempering method described in the chapter 3, which involves adding a "temperature" term to the likeligood, effectively broadening the posterior distribution. It would ensure lower log-likelihood differences between the PID hypotheses, allowing for movement between the two minima. The tempering would have to be implemented for all the proposal-types, however, reducing the efficiency and requiring longer MCMC chains together with a post-fit procedure to reweight the posterior probability distributions back to normal "temperature".

6.4.3 Split/Merge

The Split/Merge proposal has a 50% chance of either splitting a ring into two, or merge two rings into one. For the splitting, a random ring is chosen first. The ring's weight is split into two random halves. When splitting the vertex, angles and the reconstructed time, the splitting magnitudes are thrown from user-defined priors. These magnitudes are then modified by both the weights, so that the newly split-ring with a higher weight (and therefore momentum) is deviated less than the one with lower weight. The PIDs for both rings are the same as for the parent ring. The log-likelihood is then weighted by evaluating the splitting-proposal PDFs for each variable – as requested by the $\Delta q_{old \to new}$ term in the acceptance ratio from equation 6.4.

The Merge move is the opposite; two random rings are chosen, and their weights used to determine the mid-point for the vertex, time, and the two angles. The PID from the ring with higher weight is used for the newly merged ring. The difference between the two rings is taken to calculate split-prior PDF satisfying $\Delta q_{\mathrm{old} \to \mathrm{new}}$ in equation 6.4.

With the splitting priors being Gaussian distributions centred at 0 and with widths of 50 cm, 20 ns and 0.02 rad for the vertex, time and angle respectively, no Split or Merge steps were accepted in any of the RJMCMC runs. This is not surprising; if we have a two-ring hypothesis that matches with the data well, and we merge them to become one ring positioned in the middle, this proposal will be rejected. The same will be true for the

opposite scenario, where we have a one-ring hypothesis that converged to where the data suggests, and we try splitting it into two rings with diverged angles. The Split/Merge move could work well for some boosted-π^0 events where there are two rings that are strongly overlapping with each other.

6.4.4 Birth/Death

As mentioned in 6.3.2, the Birth proposal is usually the most problematic to implement in RJMCMC. In many cases, high dimensionality makes the random prior throw inefficient, especially when the minima in the likelihood are very localized. With seven kinematic parameters, it is difficult to randomly create a ring that hits a high-likelihood area. Figure 6.11 shows example 2D posterior distributions and the variations of the number of rings with time for an example MC event. It is clear that the posterior distributions are very localized – the probability of throwing a new ring randomly, that is close to such a localized area, is very low.

AddOnEnd

There were a few alternative methods already implemented, with varying success. One already-existing method, "AddOnEnd", tries to exploit the physical geometry of interactions in the SK tank; secondary particles' vertices are correlated in space with the primary particle track. The method works by choosing a random ring first, and creating a new vertex in the direction of the first ring's motion, with a random direction using a uniform spherical prior. In various implementations, the new ring's time was either thrown from a prior (uniform with hard limits), randomly deviated from its parent ring, or extrapolated from its parent vertex by adding the time-of-flight difference.

Although this method seemed to work better, the chain would still converge on one number-of-rings hypothesis and be unable to either remove and add a new ring. The hypothesis it would converge on would usually not be the right one either: MCMC finds a minimum, and is unable to escape it – whether it is a global or local minimum in the likelihood. Various modifications were tested by changing the ring's weights (i.e. the

momenta) and the reconstructed time, but the results never improved.

LocalPrior

The "LocalPrior" method generates a prior for newly-proposed ring's vertex based on the positions of the already-existing rings. Very similarly to the "AddOnEnd" method, a random ring is chosen first. The ring's vertex is used as a central position for a vertex Gaussian prior, with a user-adjustable width (50–200 cm tested). Variants with a random Gaussian added in the *towall* direction and added in all three directions were tested, together with all the reconstructed-time deviations from "AddOnEnd". The results were similar those previous; each time the MCMC was run, it would converge on different local minimum, unable to remove or add a ring.

AddGhost

The "AddGhost" was an idea to add a very low-energy ring that would most likely be accepted, and allowing it converge to some minimum. A few variants were tested with different energies (starting from 5 MeV), drawn either from a uniform prior in the detector phase-space or a local prior described above. In this implementation, a new ring would be successfully added for each proposal, but not removed during the "Death" move. This resulted in the maximum allowed number of rings limit (set to 32) being hit quite early in the chain, some of them slowly converging to have similar energy to each other.

SmartProposal

The "SmartProposal" method generates distinct prior distributions for each event separately, using a low-energy vertex fitting method implemented in the BONSAI (Branch Optimization Navigating Successive Annealing Iterations) fitter in SK [124]. Four random hit PMTs are selected, with the lowest hit-time used as "event origin". Using the PMTs positions, hit times and the time-of-flight, a vertex is reconstructed using these four hits. This reconstructed vertex represents one "sample", with 10,000 samples taken with random 4-hit PMTs used for each sample.

Figure 6.13 shows a vertex distribution on the X vs Y SK space, together with the angular distribution – obtained from the BONSAI-like vertex pre-fitter. These distributions were used as priors for a new-ring proposal. Unfortunately, this method had a similar performance to using a ring drawn with random initial parameters – the chain would converge on a local minimum early in the run, and stop accepting any proposals,

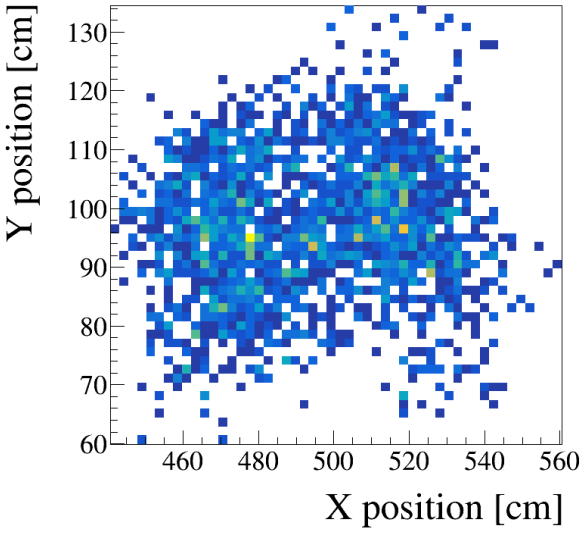
Outlook for the future

Trying to find a method that allows movement between the local minima in the likelihood, but at the same time not resulting in a negligible probability to accept a new ring, is a challenging task for fiTQun-RJMCMC.

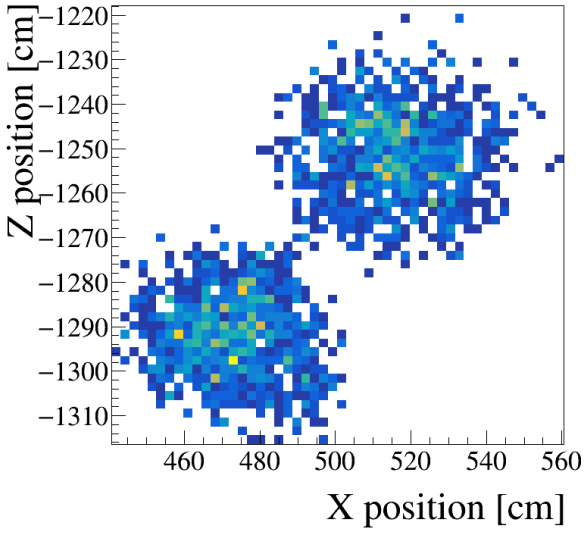
One way to ensure the phase-space is explored, would be – just like in the UpdatePID case – to use likelihood tempering. It would effectively "relax" the likelihood, merging local minima and allowing for the parameter exploration, whilst reducing the MCMC efficiency; the MCMC chains would have to be longer.

Clearly, exploiting known geometries of the particle behaviour and interactions in the SK tank should be beneficial. The AddOnEnd method (perhaps together with local prior) is worth exploring further. With its – geometrically correct – way of proposing a new vertex, it could be used together with "SmartPrior", which would propose the new ring's direction in a more informed way.

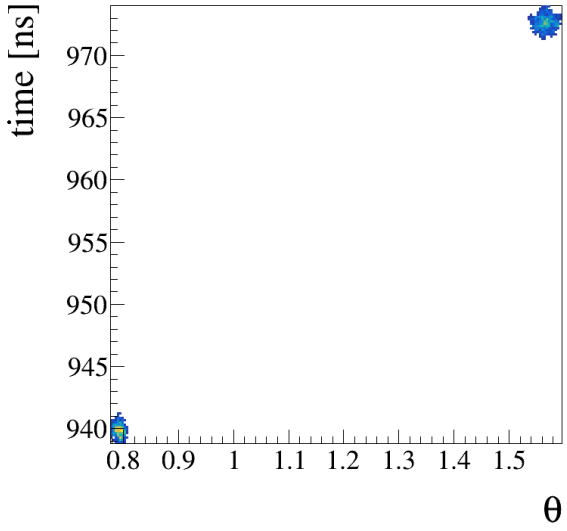
Lastly, the phase-space could be reduced to simplify the log-likelihood, which would also undermine some of the original motivations behind fiTQun-RJMCMC. It is possible that the between-model exploration could be possible with a single-vertex approximation, however, it would not reduce the issue of the θ and ϕ angles' posteriors being very localized and apart from each other.



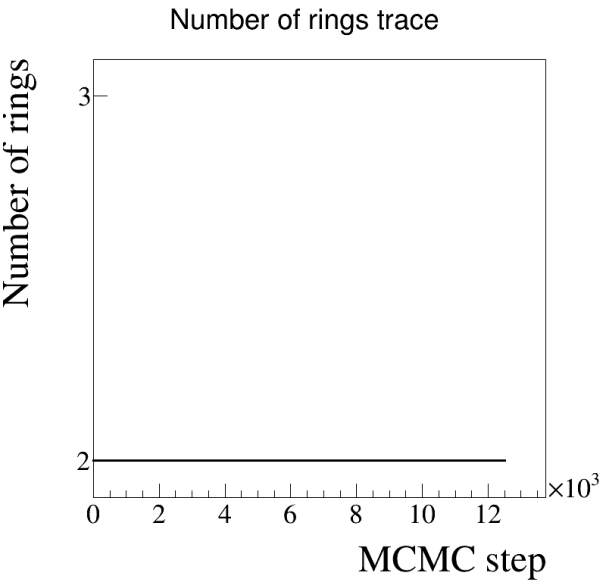
(a) 2D posterior density for X vs Y positions in the SK tank



(b) 2D posterior density for X vs Z positions in the SK tank



(c) 2D posterior density for θ angle vs reconstructed time in the SK tank



(d) Trace showing the variation of the reconstructed number of rings as the MCMC is progressing

Figure 6.11: 2D posterior densities for a number of reconstructed parameters against each other, and a trace plot showing the variation of the sampled number of rings as the MCMC chain is progressing. It is clear that the posteriors are very localized in the parameter space – if there is a third ring present that is similarly localized, the probability of randomly creating a new ring in its vicinity is very low.

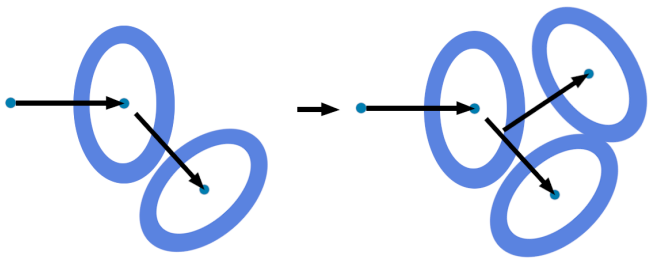
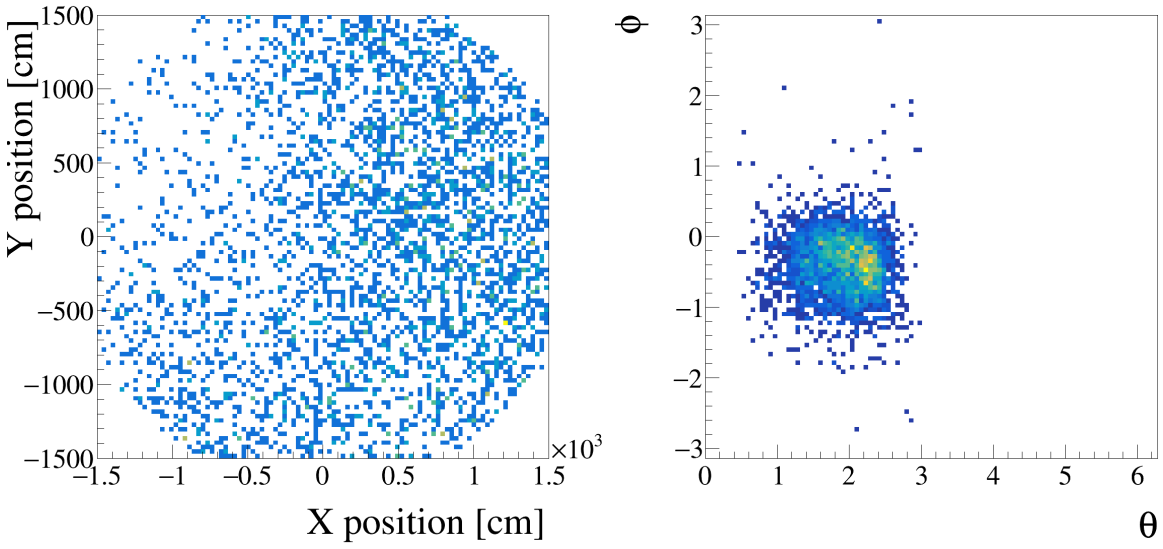


Figure 6.12: A diagram showing how "AddOnEnd" proposal works. A new vertex is thrown by extrapolating from an already-existing vertex in its *towall* direction, and a new direction is thrown from a uniform spherical prior.



(a) The X vs Y position vertex distribution obtained from a method similar to 4-hit fitter from BONSAI

(b) The θ vs ϕ angular distribution obtained from a method similar to 4-hit fitter from BONSAI

Figure 6.13: The X vs Y and θ vs ϕ distributions obtained from the 4-vertex pre-fitter similar to the one used in BONSAI [124]. These distributions can be produced before sampling each MC/Data event, and used as priors for new ring proposals.

6.5 Summary

The new mechanism for neutrino event reconstruction using Bayesian method was proposed for Water-Cherenkov detectors, using an already existing fiTQun framework for likelihood calculations. With the method implementation underway, there are still many obstacles to overcome before large validations or full analyses using this framework can take place.

The standard Metropolis-Hastings method was improved to allow within-model phase-space exploration for different MC events, using automatic step-size tuning. All explored methods for the PID update and new ring proposals result in fiTQun-RJMCMC being stuck in different local minima with each MCMC run.

More investigation is needed into various PID differences using MC, together with improving the ways to propose new rings. The PID transition must must change the kinematic parameters like E_{vis}/momenta and the vertex position to produce smaller log-likelihood difference between the current and the proposed step. For the trans-dimensional step, a mixture between AddOnEnd and SmartPrior proposals could be explored, as it joins the knowledge of the physical processes in the detector and reduces the issue with an unknown angle by pre-sampling the detector phase-space. Both the PID transition and the trans-dimensional steps could be further explored with likelihood tempering described in section 3.4.1 to further reduce the different in the log-likelihoods between the proposals.

Chapter 7

Conclusions

This thesis describes two distinct analyses of the joint ND280 and SK data, together with a continuation of an effort to implement a Bayesian method of neutrino event reconstruction at SK.

Both oscillation analysis used the same five data samples at SK and fourteen data samples at ND280, constraining all the systematic parameters simultaneously. The analyses were done with the beam exposure of 5.82×10^{20} POT in the neutrino mode (FHC) and 2.84×10^{20} POT in the anti-neutrino mode (RHC) at ND280, and with 14.94×10^{20} POT in the FHC mode and 16.35×10^{20} POT in the RHC mode at SK. This corresponds to 1 % increase in the FHC and 116 % increase in the FHC SK data since the results from the 2017 data-taking, published in [119]. With $90\,\nu_e$, $15\,\bar{\nu}_e$, $243\,\nu_\mu$ and $140\,\bar{\nu}_\mu$ beam neutrino events seen at SK, the CP violation is preferred with at least $2\,\sigma$ credible interval, with the highest posterior at $\delta_{\mathcal{CP}} = -1.82$. This oscillation analysis also marked the first closed $3\,\sigma$ intervals for δ_{CP}, published in Nature [53].

All the oscillation parameters' highest-posterior densities with their credible intervals are shown in table 7.1.

This result was produced under the PMNS-model constraint. Over the period of work, the T2K result moved closer to the "Asimov" sensitivity prediction – made by fitting to the Monte-Carlo with the same PMNS model constraint.

	$\sin^2\theta_{23}$	Δm^2_{32} ($\times 10^{-3}$ eV^2)	$\sin^2\theta_{13}$	δ_{CP}
2D best fit	0.537	2.46	0.0214	-1.82
68.27% C.I. ($1\,\sigma$) range	0.501 – 0.564	2.37 – 2.54	0.0206 – 0.0222	-2.39 – -1.13
95.45% C.I. ($2\,\sigma$) range	0.466 – 0.587	-2.58 – -2.41 & 2.28 – 2.63	0.0199 – 0.0230	-2.95 – -0.50
99.73% C.I. ($3\,\sigma$) range	0.438 – 0.606	-2.68 – -2.32 & 2.22 – 2.70	0.0191 – 0.0238	$-\pi$ – 0.13 & 2.80 – π

Table 7.1: Best-fit values with 1, 2 and $3\,\sigma$ credible interval ranges for oscillation parameters for the run 1–9 data fit with constraint on θ_{13} from the reactor neutrino experiments.

In a separate analysis, a fit was done with an extra degree of freedom, decoupling the $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ oscillation probabilities from each other and allowing them to take any non-PMNS values preferred by the data. A fit without the (PMNS-like) constraint on θ_{13} from the reactor experiment has shown no tensions between the T2K data and the PMNS model, a result that was published in the PRL journal [136].

Finally, the implementation of a novel Bayesian method for neutrino event reconstruction in Super-Kamiokande was continued, using the proficiency with MCMC methods gained from the oscillation analysis, with many obstacles found towards completing it. More work on the framework needs to be done for it to be used in the analysis, with possible solutions found and explained.

The T2K experiment will continue to take more data, with planned ND280 upgrades [126] for the T2K-II phase. The planned T2K-SK and T2K-NOvA analyses will provide interesting ways of testing the cross-section model, and show interesting neutrino oscillation results before the next-generation experiments start. The PMNS model will be sampled to far higher significance by the Hyper-Kamiokande and DUNE experiments, possibly obtaining the world's first $5\,\sigma$ intervals for the δ_{CP}, and with the JUNO [127] experiment resolving the neutrino mass hierarchy problem.

Bibliography

[1] Wolfgang Pauli. "Liebe Radioaktive Damen und Herren". In: *Congrès des physiciens de Tübingen* (1930).

[2] A.D Sakharov. "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe". In: *JETPL* 5.1 (Jan. 1967), pp. 24–27. URL: http://www.jetpletters.ac.ru/ps/1643/article_25089.shtml.

[3] H. Georgi and S.L. Glashow. "Unity of All Elementary Particle Forces". In: *Phys. Rev. Lett.* 32 (1974), pp. 438–441. DOI: 10.1103/PhysRevLett.32.438.

[4] Rabindra N. Mohapatra and Goran Senjanovi ć. "Neutrino masses and mixings in gauge models with spontaneous parity violation". In: *Phys. Rev. D* 23 (1 Jan. 1981), pp. 165–180. DOI: 10.1103/PhysRevD.23.165. URL: https://link.aps.org/doi/10.1103/PhysRevD.23.165.

[5] B. Pontecorvo. "Mesonium and anti-mesonium". In: *Zh.Eksp.Teor.Fiz.* 33 (1957), p. 429.

[6] B. Pontecorvo. "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". In: *Sov. Phys. JETP* 26 (1968), pp. 984–988.

[7] J Chadwick. "Intensitätsverteilung im magnetischen Spectrum der \β-Strahlen von radium B + C". In: *Verhandl. Dtsc. Phys. Ges.* 16 (1914), p. 383.

[8] Charles Drummond Ellis and W. A. Wooster. "The average energy of disintegration of radium E". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.776 (Dec. 1927), pp. 109–123. ISSN: 0950-1207. DOI: 10.1098/rspa.1927.0168. URL: https://royalsocietypublishing.org/doi/10.1098/rspa.1927.0168.

[9] E. Fermi. "Versuch einer Theorie der β-Strahlen. I". In: *Zeitschrift für Physik* 88.3-4 (Mar. 1934), pp. 161–177. ISSN: 14346001. DOI: 10.1007/BF01351864.

[10] F. Reines and C. L. Cowan. *Detection of the free neutrino*. Nov. 1953. DOI: 10.1103/PhysRev.92.830.

[11] J. B. Brown and K. Mendelssohn. *A new technique for studying the helium film [1]*. 1947. DOI: 10.1038/160670a0.

[12] C. L. Cowan et al. "Large Liquid Scintillation Detectors". In: *Phys. Rev.* 90 (3 May 1953), pp. 493–494. DOI: 10.1103/PhysRev.90.493. URL: https://link.aps.org/doi/10.1103/PhysRev.90.493.

[13] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9 (1 July 1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[14] M.M. Block et al. "Neutrino interactions in the CERN heavy liquid bubble chamber". In: *Physics Letters* 12 (Oct. 1964), pp. 281–285. DOI: 10.1016/0031-9163(64)91104-7.

[15] M. L. Perl et al. "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation". In: *Phys. Rev. Lett.* 35 (22 Dec. 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[16] K. Kodama et al. "Observation of tau neutrino interactions". In: *Physics Letters B* 504.3 (2001), pp. 218–224. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(01)00307-0. URL: http://www.sciencedirect.com/science/article/pii/S0370269301003070.

[17] K. Abe et al. "Solar Neutrino Measurements in Super-Kamiokande-IV". In: *Phys. Rev. D* 94.5 (2016), p. 052010. DOI: 10.1103/PhysRevD.94.052010. arXiv: 1606.07538 [hep-ex].

[18] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[19] A. Gando et al. "Reactor On-Off Antineutrino Measurement with KamLAND". In: *Phys. Rev. D* 88.3 (2013), p. 033001. DOI: 10.1103/PhysRevD.88.033001. arXiv: 1303.4667 [hep-ex].

[20] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. "Search for Neutrinos from the Sun". In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[21] K. S. Hirata et al. "Observation of ^{8}B solar neutrinos in the Kamiokande-II detector". In: *Phys. Rev. Lett.* 63 (1 July 1989), pp. 16–19. DOI: 10.1103/PhysRevLett.63.16. URL: https://link.aps.org/doi/10.1103/PhysRevLett.63.16.

[22] Q. R. Ahmad et al. "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (1 June 2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.011301.

[23] T. J. Haines et al. "Calculation of Atmospheric Neutrino-Induced Backgrounds in a Nucleon-Decay Search". In: *Phys. Rev. Lett.* 57 (16 Oct. 1986), pp. 1986–1989. DOI: 10.1103/PhysRevLett.57.1986. URL: https://link.aps.org/doi/10.1103/PhysRevLett.57.1986.

[24] K.S. Hirata et al. "Experimental study of the atmospheric neutrino flux". In: *Physics Letters B* 205.2 (1988), pp. 416–420. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(88)91690-5. URL: http://www.sciencedirect.com/science/article/pii/0370269388916905.

[25] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[26] K. Abe et al. "Indication of Electron Neutrino Appearance from an Accelerator-Produced Off-Axis Muon Neutrino Beam". In: *Phys. Rev. Lett.* 107 (4 July 2011), p. 041801. DOI: 10.1103/PhysRevLett.107.041801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.041801.

[27] P. Adamson et al. “Measurement of the Neutrino Mass Splitting and Flavor Mixing by MINOS”. In: *Phys. Rev. Lett.* 106 (18 May 2011), p. 181801. DOI: 10.1103/PhysRevLett.106.181801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.181801.

[28] Y. Abe et al. “Indication for the disappearance of reactor electron antineutrinos in the Double Chooz experiment”. In: *Physical Review Letters* 108.13 (Dec. 2011). DOI: 10.1103/PhysRevLett.108.131801. arXiv: 1112.6353. URL: http://arxiv.org/abs/1112.6353%20http://dx.doi.org/10.1103/PhysRevLett.108.131801.

[29] F. P. An et al. “Observation of electron-antineutrino disappearance at Daya Bay”. In: *Physical Review Letters* 108.17 (Mar. 2012). DOI: 10.1103/PhysRevLett.108.171803. arXiv: 1203.1669. URL: http://arxiv.org/abs/1203.1669%20http://dx.doi.org/10.1103/PhysRevLett.108.171803.

[30] Soo-Bong Kim. “Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment”. In: *Physical Review Letters* 108.19 (Apr. 2012). DOI: 10.1103/PhysRevLett.108.191802. arXiv: 1204.0626. URL: http://arxiv.org/abs/1204.0626%20http://dx.doi.org/10.1103/PhysRevLett.108.191802.

[31] K. Abe et al. “Observation of Electron Neutrino Appearance in a Muon Neutrino Beam”. In: *Phys. Rev. Lett.* 112 (6 Feb. 2014), p. 061802. DOI: 10.1103/PhysRevLett.112.061802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.112.061802.

[32] K. Eguchi et al. “First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance”. In: *Phys. Rev. Lett.* 90 (2 Jan. 2003), p. 021802. DOI: 10.1103/PhysRevLett.90.021802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.90.021802.

[33] T. Araki et al. “Measurement of Neutrino Oscillation with KamLAND: Evidence of Spectral Distortion”. In: *Phys. Rev. Lett.* 94 (8 Mar. 2005), p. 081801. DOI: 10.1103/PhysRevLett.94.081801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.94.081801.

[34] M. H. Ahn et al. “Measurement of neutrino oscillation by the K2K experiment”. In: *Phys. Rev. D* 74 (7 Oct. 2006), p. 072003. DOI: 10.1103/PhysRevD.74.072003. URL: https://link.aps.org/doi/10.1103/PhysRevD.74.072003.

[35] P. Adamson et al. “Measurement of Neutrino Oscillations with the MINOS Detectors in the NuMI Beam”. In: *Phys. Rev. Lett.* 101 (13 Sept. 2008), p. 131802. DOI: 10.1103/PhysRevLett.101.131802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.101.131802.

[36] K. Hirata et al. “Observation of a neutrino burst from the supernova SN1987A”. In: *Phys. Rev. Lett.* 58 (14 Apr. 1987), pp. 1490–1493. DOI: 10.1103/PhysRevLett.58.1490. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.1490.

[37] R. Abbasi et al. “Extending the Search for Neutrino Point Sources with IceCube above the Horizon”. In: *Phys. Rev. Lett.* 103 (22 Nov. 2009), p. 221102. DOI: 10.1103/PhysRevLett.103.221102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.103.221102.

[38] Katarzyna Frankiewicz. “Searching for Dark Matter Annihilation into Neutrinos with Super-Kamiokande”. In: *ArXiv pre-print* (Oct. 2015). arXiv: 1510.07999. URL: http://arxiv.org/abs/1510.07999.

[39] T Araki et al. “Experimental investigation of geologically produced antineutrinos with KamLAND”. In: *Nature* 436.7050 (2005), pp. 499–503. ISSN: 1476-4687. DOI: 10.1038/nature03980. URL: https://doi.org/10.1038/nature03980.

[40] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. “Remarks on the Unified Model of Elementary Particles”. In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870. eprint: https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf. URL: https://doi.org/10.1143/PTP.28.870.

[41] 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: *Phys. Rev. D* 96 (9 Nov. 2017), p. 092006. DOI: 10.1103/PhysRevD.96.092006. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.092006.

[42] K. Abe et al. "Measurements of the T2K neutrino beam properties using the INGRID on-axis near detector". In: *Nucl. Instrum. Meth. A* 694 (2012), pp. 211–223. DOI: 10.1016/j.nima.2012.03.023. arXiv: 1111.3119 [physics.ins-det].

[43] S. Assylbekov et al. "The T2K ND280 off-axis pi-zero detector". In: *Nucl. Instrum. Meth. A* 686 (2012), pp. 48–63. DOI: 10.1016/j.nima.2012.05.028.

[44] T2K Collaboration. "T2K ND280 Conceptual Design Report". In: *Internal Report* (2005).

[45] P.A. Amaudruz et al. "The T2K Fine-Grained Detectors". In: *Nucl. Instrum. Meth. A* 696 (2012), pp. 1–31. DOI: 10.1016/j.nima.2012.08.020. arXiv: 1204.3666 [physics.ins-det].

[46] D Allan et al. "The electromagnetic calorimeter for the T2K near detector ND280". In: *Journal of Instrumentation* 8.10 (Oct. 2013), P10019–P10019. DOI: 10.1088/1748-0221/8/10/p10019. URL: https://doi.org/10.1088%2F1748-0221%2F8%2F10%2Fp10019.

[47] Y. Fukuda et al. "The Super-Kamiokande detector". In: *Nucl. Instrum. Meth. A* 501 (2003), pp. 418–462. DOI: 10.1016/S0168-9002(03)00425-X.

[48] T2K collaboration internal website. *Five Sample Joint Oscillation Analysis with T2K Run1-9 Data*. Tech. rep. TN-367. T2K, 2019. eprint: https://www.t2k.org/docs/plots/055/oa-plots/P-theta.

[49] M. Friend et al. *Flux Prediction and Uncertainty Updates with NA61 2009 Thin Target Data and Negative Focussing Mode Predictions*. Tech. rep. TN-217. T2K, 2015. eprint: www.t2k.org/docs/technotes/217.

[50] *K2K collabration website*. URL: https://neutrino.kek.jp/intro/k2k.html (visited on 2020).

[51] *J-PARC website*. URL: https://j-parc.jp/public/Acc/en/index.html (visited on 2020).

[52] *T2K collabration internal website*. URL: https://t2k.org (visited on 2020).

[53] K Abe et al. "Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations". In: *Nature* 580.7803 (2020), pp. 339–344. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2177-0. URL: https://doi.org/10.1038/s41586-020-2177-0.

[54] Morgan Wascko. *T2K Status, Results, and Plans*. June 2018. DOI: 10.5281/zenodo.1286752. URL: https://doi.org/10.5281/zenodo.1286752.

[55] K. Matsuoka et al. "Design and performance of the muon monitor for the T2K neutrino oscillation experiment". In: *Nucl. Instrum. Meth. A* 624.3 (2010), pp. 591–600. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2010.09.074. URL: http://www.sciencedirect.com/science/article/pii/S016890021002098X.

[56] D. Beavis et al. *Long baseline neutrino oscillation experiment at the AGS. Physics design report*. Tech. rep. Upton, NY: Brookhaven National Laboratory (BNL), Apr. 1995. DOI: 10.2172/52878. URL: http://www.osti.gov/servlets/purl/52878-fzmjLc/webviewable/.

[57] Yuichi Oyama. "Results from K2K and status of T2K". In: (Dec. 2005). DOI: 10.1007/978-1-4020-4965-1{_}9. URL: http://arxiv.org/abs/hep-ex/0512041%20http://dx.doi.org/10.1007/978-1-4020-4965-1%7B%5C_%7D9.

[58] K. Abe et al. "T2K neutrino flux prediction". In: *Phys. Rev. D* 87.1 (2013), p. 012001. DOI: 10.1103/PhysRevD.87.012001. arXiv: 1211.0469 [hep-ex].

[59] K. Abe et al. "The T2K Experiment". In: *Nucl. Instrum. Meth. A* 659 (2011), pp. 106–135. DOI: 10.1016/j.nima.2011.06.067. arXiv: 1106.1238 [physics.ins-det].

[60] Y. Itow et al. *The JHF-Kamioka neutrino project*. 2001. arXiv: hep-ex/0106019 [hep-ex].

[61] G. Cowan. *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. ISBN: 9780198501558. URL: https://books.google.co.uk/books?id=ff8ZyW0nlJAC.

[62] K. E. Duffy et al. *Joint $\nu_\mu + \nu_e + \bar{\nu}_\mu + \bar{\nu}_e$+ND280 analysis with run 1-9c data*. Tech. rep. TN-320. T2K internal, 2017. eprint: http://www.t2k.org/docs/technotes/320.

[63] G. O. Roberts, A. Gelman, and W. R. Gilks. "Weak convergence and optimal scaling of random walk Metropolis algorithms". In: *Annals of Applied Probability* 7.1 (Feb. 1997), pp. 110–120. DOI: 10.1214/aoap/1034625254.

[64] Steve Brooks et al. *Handbook of Markov Chain Monte Carlo*. CRC press, 2011. ISBN: 9780429138508. DOI: 10.1201/b10905.

[65] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (Apr. 1970), pp. 97–109. ISSN: 0006-3444. DOI: 10.1093/biomet/57.1.97. eprint: https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf. URL: https://doi.org/10.1093/biomet/57.1.97.

[66] Charles J. Geyer. "Markov Chain Monte Carlo Maximum Likelihood". In: (1991). URL: http://conservancy.umn.edu/handle/11299/58440.

[67] G. Altekar et al. "Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference". In: *Bioinformatics* 20.3 (Feb. 2004), pp. 407–415. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btg427. URL: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btg427.

[68] Jonathan Goodman and Jonathan Weare. "Ensemble samplers with affine invariance". In: *Communications in applied mathematics and computational science* 5.1 (2010), pp. 65–80. DOI: 10.2140/camcos.2010.5.65.

[69] Daniel Foreman-Mackey et al. "emcee: The MCMC Hammer". In: *Publications of the Astronomical Society of the Pacific* 125.925 (Mar. 2013), pp. 306–312. DOI: 10.1086/670067. URL: https://doi.org/10.1086%2F670067.

[70] Radford M. Neal. *MCMC using Hamiltonian dynamics*. 2012. arXiv: 1206.1901 [stat.CO].

[71] A. Habig. "The NOvA Experiment". In: *Nuclear Physics B - Proceedings Supplements* 229-232 (2012). Neutrino 2010, p. 460. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2012.09.097. URL: http://www.sciencedirect.com/science/article/pii/S0920563212003143.

[72] Simon Duane et al. "Hybrid Monte Carlo". In: *Physics Letters B* 195.2 (1987), pp. 216–222. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(87)91197-X. URL: http://www.sciencedirect.com/science/article/pii/037026938791197X.

[73] S. Zsoldos et al. *Updated P-theta run 1–9 disappearance analysis.* Tech. rep. TN-387. T2K, 2019. eprint: www.t2k.org/docs/technotes/387.

[74] N. Abgrall et al. "Pion emission from the T2K replica target: Method, results and application". In: *Nucl. Instrum. Meth. A* 701 (2013), pp. 99–114. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2012.10.079.

[75] N. Abgrall et al. "Measurements of cross sections and charged pion spectra in proton-carbon interactions at 31 GeV/c". In: *Phys. Rev. C* 84 (3 Sept. 2011), p. 034604. DOI: 10.1103/PhysRevC.84.034604. URL: https://link.aps.org/doi/10.1103/PhysRevC.84.034604.

[76] N. Abgrall et al. "Measurements of $\pi^{\pm}$, $K^{\pm}$ and proton double differential yields from the surface of the T2K replica target for incoming 31 GeV/c protons with the NA61/SHINE spectrometer at the CERN SPS". In: *Eur. Phys. J. C* 79.2 (2019), p. 100. DOI: 10.1140/epjc/s10052-019-6583-0. arXiv: 1808.04927 [hep-ex].

[77] A. Ferrari et al. *FLUKA: a multi-particle transport code (program version 2005).* Tech. rep. LAC-R-773. 2005.

[78] René Brun et al. "GEANT Detector Description and Simulation Tool". In: (Oct. 1994). DOI: 10.17181/CERN.MUHF.DMJ1.

[79] C. Zeitnitz and T.A. Gabriel. "The GEANT-CALOR interface and benchmark calculations of ZEUS test calorimeters". In: *Nucl. Instrum. Meth. A* 349.1 (1994), pp. 106–111. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(94)90613-0. URL: http://www.sciencedirect.com/science/article/pii/0168900294906130.

[80] Yoshinari Hayato. "A neutrino interaction simulation program library NEUT". In: *Acta Phys. Polon. B* 40 (2009). Ed. by Arthur Ankowski and Jan Sobczyk, pp. 2477–2489.

[81] Y. Hayato. "NEUT". In: *Nucl. Phys. B Proc. Suppl.* 112 (2002). Ed. by J.G. Morfin, M. Sakuda, and Y. Suzuki, pp. 171–176. DOI: 10.1016/S0920-5632(02)01759-0.

[82] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas. "Inclusive charged-current neutrino-nucleus reactions". In: *Phys. Rev. C* 83 (4 Apr. 2011), p. 045501. DOI: 10.1103/PhysRevC.83.045501. URL: https://link.aps.org/doi/10.1103/PhysRevC.83.045501.

[83] J. Nieves, J. E. Amaro, and M. Valverde. "Inclusive quasielastic charged-current neutrino-nucleus reactions". In: *Phys. Rev. C* 70 (5 Nov. 2004), p. 055503. DOI: 10.1103/PhysRevC.70.055503. URL: https://link.aps.org/doi/10.1103/PhysRevC.70.055503.

[84] R. Gran et al. "Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 GeV". In: *Phys. Rev. D* 88 (11 Dec. 2013), p. 113007. DOI: 10.1103/PhysRevD.88.113007. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.113007.

[85] A. A. Aguilar-Arevalo et al. "First measurement of the muon neutrino charged current quasielastic double differential cross section". In: *Phys. Rev. D* 81 (9 May 2010), p. 092005. DOI: 10.1103/PhysRevD.81.092005. URL: https://link.aps.org/doi/10.1103/PhysRevD.81.092005.

[86] A. A. Aguilar-Arevalo et al. "First measurement of the muon antineutrino double-differential charged-current quasielastic cross section". In: *Phys. Rev. D* 88 (3 Aug. 2013), p. 032001. DOI: 10.1103/PhysRevD.88.032001. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.032001.

[87] L. Fields et al. "Measurement of Muon Antineutrino Quasielastic Scattering on a Hydrocarbon Target at $E_\nu \sim 3.5$ GeV". In: *Phys. Rev. Lett.* 111 (2 July 2013), p. 022501. DOI: 10.1103/PhysRevLett.111.022501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.111.022501.

[88] C. Wilkinson et al. "Testing charged current quasi-elastic and multinucleon interaction models in the NEUT neutrino interaction generator with published datasets from the MiniBooNE and MINERνA experiments". In: *Phys. Rev. D* 93 (7 Apr. 2016), p. 072010. DOI: 10.1103/PhysRevD.93.072010. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.072010.

[89] Callum Wilkinson et al. "Reanalysis of bubble chamber measurements of muon-neutrino induced single pion production". In: *Phys. Rev. D* 90 (11 Dec. 2014), p. 112017. DOI: 10.1103/PhysRevD.90.112017. URL: https://link.aps.org/doi/10.1103/PhysRevD.90.112017.

[90] E. J. Moniz et al. "Nuclear Fermi Momenta from Quasielastic Electron Scattering". In: *Phys. Rev. Lett.* 26 (8 Feb. 1971), pp. 445–448. DOI: 10.1103/PhysRevLett.26.445. URL: https://link.aps.org/doi/10.1103/PhysRevLett.26.445.

[91] Richard Gran. "Model uncertainties for Valencia RPA effect for MINERvA". In: *arXiv preprint arXiv:1705.02932* (2017).

[92] PETER J. GREEN. "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination". In: *Biometrika* 82.4 (Dec. 1995), pp. 711–732. ISSN: 0006-3444. DOI: 10.1093/biomet/82.4.711. eprint: https://academic.oup.com/biomet/article-pdf/82/4/711/699533/82-4-711.pdf. URL: https://doi.org/10.1093/biomet/82.4.711.

[93] S. Bienstock et al. *Constraining the Flux and Cross Section Models with Data from the ND280 Detector using FGD1 and FGD2 for the 2017 Joint Oscillation Analysis*. Tech. rep. TN-324. T2K internal, 2017. eprint: https://www.t2k.org/docs/technotes/324.

[94] Clarence Wret. "Minimising Systematic Uncertaintiesin the T2K Experiment Using Near-Detector and External Data". PhD thesis. Imperial College London, 2018.

[95] Xiaoyue Li. "A Joint Analysis of T2K Beam Neutrino andSuper-Kamiokande Sub-GeV Atmospheric Neutrino Data". PhD thesis. Stony Brook University, 2018.

[96] J. Imber et al. *T2K-SK Systematic Error Summary for 2017 OA*. Tech. rep. TN-326. T2K internal, 2017. eprint: https://www.t2k.org/docs/technotes/326.

[97] Th.A Mueller and S Mine. *SK pi0 systematic errors for nue analysis with T2K 6.4 x 10e20POT (RUN1-4) data*. Tech. rep. TN-156. T2K internal, 2013. eprint: https://www.t2k.org/docs/technotes/156.

[98] A Missert. *SK Atmospheric Neutrino Fit and Fiducial Volume Optimization*. Tech. rep. TN-318. T2K internal, 2017. eprint: https://www.t2k.org/docs/technotes/318.

[99] R. Akutsu et al. *Super-Kamiokande Data Quality In T2K Runs 1-9*. Tech. rep. TN-355. T2K internal, 2019. eprint: https://www.t2k.org/docs/technotes/355.

[100] P. Bartet et al. *ν_μ CC event selections in the ND280 tracker using Run 2+3+4 data*. Tech. rep. TN-212. T2K internal, 2015. eprint: https://www.t2k.org/docs/technotes/212.

[101] V. Berardi et al. *CC $\bar{\nu}_\mu$ event selection in the ND280 tracker using Run 5c and Run 6 anti-neutrino beam data*. Tech. rep. TN-246. T2K internal, 2015. eprint: https://www.t2k.org/docs/technotes/246.

[102] V. Berardi et al. *CC ν_μ background event selection in the ND280 tracker using Run 5c+Run 6 anti-neutrino beam data*. Tech. rep. TN-248. T2K internal, 2015. eprint: https://www.t2k.org/docs/technotes/248.

[103] K. Iwamoto et al. *Inelastic Single Pion Signal Study in ν_e Appearance using Modified Decay Electron Cut*. Tech. rep. TN-233. T2K internal, 2016. eprint: https://www.t2k.org/docs/technotes/233.

[104] A. Himmel et al. *Super-Kamiokande events and data quality studies for1T2K Runs 5 and 6*. Tech. rep. TN-219. T2K internal, 2015. eprint: https://www.t2k.org/docs/technotes/219.

[105] X. Li and M. Wilking. *FiTQun Event Selection Optimization*. Tech. rep. TN-317. T2K internal, 2017. eprint: https://www.t2k.org/docs/technotes/319.

[106] A. Kaboth, R. Calland, and D. Payne. *A Joint ND80-SK 1R_μ-SK 1R_e Fit using MCMC*. Tech. rep. TN-171. T2K internal, 2014. eprint: \url{www.t2k.org/docs/technotes/171}.

[107] Richard G. Calland, Asher Kaboth, and Doris L. Payne. "Accelerated event-by-event neutrino oscillation reweighting with matter effects on a GPU". In: *Journal of Instrumentation* 9 (2014), p. 04016.

[108] R. Wendell and et. al et. "Prob3++ software for computing three flavorneutrino oscillation probabilities." In: (2012). eprint: http://www.phy.duke.edu/~raw22/public/Prob3++/.

[109] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[110] C. Andreopoulos et al. *T2K Neutrino and Anti-Neutrino 3-Flavour Joint Analysis of 1Run 1-9 (*1.4938×10^{21}*-POT*ν 1.6346×10^{21}*-POT*$\bar{\nu}$*) data sets - Statistical Update*. Tech. rep. TN-360. T2K internal, 2018. eprint: https://www.t2k.org/docs/technotes/360.

[111] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[112] H. Jeffreys. *The Theory of Probability*. OUP Oxford, 1998. ISBN: 9780191589676. URL: https://books.google.ca/books?id=vh9Act9rtzQC.

[113] Robert E. Kass and Adrian E. Raftery. "Bayes Factors". In: *Journal of the American Statistical Association* 90.430 (1995), pp. 773–795. DOI: 10.1080/01621459.1995.10476572. eprint: http://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572.

[114] Elder Sebastian Pinzon Guerra. "Measurement of Pion-Carbon Cross Sections at DUET and Measurement of Neutrino Oscillation Parameters at the T2K Experiment". PhD thesis. York University, 2017.

[115] S. Bienstock and Bolognesi, S. *Assessing the cross-section model uncertainties on the T2K oscillation analyses with fake data studies using the BANFF, MaCh3, P-Theta and VALOR frameworks*. Tech. rep. TN-331. T2K internal, 2017. eprint: \url{www.t2k.org/docs/technotes/331}.

[116] S. Berkman, P. de Piero, and A. Konaka. *fiTQun: A New Reconstruction Algorithm for Super-K*. Tech. rep. TN-146. T2K internal, 2013. eprint: https://www.t2k.org/docs/technotes/146.

[117] S. Berkman, H. A. Tanaka, and S. Tobayama. *Data/Monte Carlo Simulation Comparisons with fiTQun v3r1*. Tech. rep. TN-143. T2K internal, 2013. eprint: https://www.t2k.org/docs/technotes/153.

[118] R.B. Patterson et al. "The extended-track event reconstruction for MiniBooNE". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 608.1 (2009), pp. 206–224. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2009.06.064. URL: http://www.sciencedirect.com/science/article/pii/S0168900209012480.

[119] K. Abe et al. "Search for CP Violation in Neutrino and Antineutrino Oscillations by the T2K Experiment with 2.2×1021 Protons on Target". In: *Physical Review Letters* 121.17 (Oct. 2018). ISSN: 1079-7114. DOI: 10.1103/physrevlett.121.171802. URL: http://dx.doi.org/10.1103/PhysRevLett.121.171802.

[120] M Jiang et al. "Atmospheric neutrino oscillation analysis with improved event reconstruction in Super-Kamiokande IV". In: *Progress of Theoretical and Experimental Physics* 2019.5 (May 2019). ISSN: 2050-3911. DOI: 10.1093/ptep/ptz015. URL: http://dx.doi.org/10.1093/ptep/ptz015.

[121] Peter J. Green. "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination". In: *Biometrika* 82.4 (Dec. 1995), pp. 711–732. ISSN: 0006-3444. DOI: 10.1093/biomet/82.4.711. URL: https://doi.org/10.1093/biomet/82.4.711.

[122] Sylvia. Richardson and Peter J. Green. "On Bayesian Analysis of Mixtures with an Unknown Number of Components (with discussion)". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 59.4 (1997), pp. 731–792. DOI: 10.1111/1467-9868.00095. eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9868.00095. URL: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00095.

[123] Fan Y. and Sisson S.A. *Reversible jump Markov chain Monte Carlo*. 2010. arXiv: 1001.2055.

[124] Smy Michael. *Proceedings, 30th International Cosmic Ray Conference (ICRC 2007)*. Merida, Mexico, July 2007.

[125] Evgeny K. Akhmedov, V.A. Rubakov, and A.Yu. Smirnov. "Baryogenesis via neutrino oscillations". In: *Phys. Rev. Lett.* 81 (1998), pp. 1359–1362. DOI: 10.1103/PhysRevLett.81.1359. arXiv: hep-ph/9803255.

[126] K. Abe et al. "T2K ND280 Upgrade - Technical Design Report". In: (Jan. 2019). arXiv: 1901.03750 [physics.ins-det].

[127] T. Adam et al. *JUNO Conceptual Design Report*. 2015. arXiv: 1508.07166 [physics.ins-det].

[128] L. Wolfenstein. "Neutrino oscillations in matter". In: *Phys. Rev. D* 17 (9 May 1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369. URL: https://link.aps.org/doi/10.1103/PhysRevD.17.2369.

[129] S.P. Mikheyev and A.Yu. Smirnov. "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos". In: *Sov. J. Nucl. Phys.* 42 (1985), pp. 913–917.

[130] Y. Yamanoi et al. "Large horn magnets at the KEK neutrino beam line. II". In: *IEEE Transactions on Applied Superconductivity* 10.1 (2000), pp. 252–255.

[131] Eric Christensen et al. "Antineutrino Monitoring for Heavy Water Reactors". In: *Phys. Rev. Lett.* 113 (4 June 2014), p. 042503. DOI: 10.1103/PhysRevLett.113.042503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.042503.

[132] André de Gouvêa and Ivan Martinez-Soler. *Reactor Antineutrino Oscillations at Super-Kamiokande*. 2020. arXiv: 2006.01155 [hep-ph].

[133] Hisakazu Minakata and Shinji Watanabe. "Solar neutrinos and leptonic CP violation". In: *Physics Letters B* 468.3-4 (Dec. 1999), pp. 256–260. ISSN: 0370-2693. DOI: 10.1016/s0370-2693(99)01224-1. URL: http://dx.doi.org/10.1016/S0370-2693(99)01224-1.

[134] Tadashi Koseki et al. "Beam commissioning and operation of the J-PARC main ring synchrotron". In: *Progress of Theoretical and Experimental Physics* 2012.1 (Dec. 2012). 02B004. ISSN: 2050-3911. DOI: 10.1093/ptep/pts071. eprint: https://academic.oup.com/ptep/article-pdf/2012/1/02B004/11595767/pts071.pdf. URL: https://doi.org/10.1093/ptep/pts071.

[135] Gary J. Feldman. "Long-baseline accelerator neutrino experiments". In: *International Conference on History of the Neutrino: 1930-2018*. 2019. arXiv: 1901.09431 [hep-ex].

[136] K. Abe et al. "Search for Electron Antineutrino Appearance in a Long-Baseline Muon Antineutrino Beam". In: *Physical Review Letters* 124.16 (Apr. 2020). ISSN: 1079-7114. DOI: 10.1103/physrevlett.124.161802. URL: http://dx.doi.org/10.1103/PhysRevLett.124.161802.

[137] S Bolognesi et al. *NIWG model and uncertainties for 2017 oscillation analysis*. Tech. rep. TN-315. T2K, 2017. eprint: www.t2k.org/docs/technotes/315.

[138] S Bolognesi et al. *Assessing the effect of cross-section model uncertainties on the T2K oscillation analyses with fake data studies using the BANFF, MaCh3 and VALOR frameworks*. Tech. rep. TN-285. T2K, 2016. eprint: www.t2k.org/docs/technotes/285.

Appendix A

Predicted SK event rates for run 1–9d

The predicted SK event rates are produced prior to the far detector (or joint) fits, partially to initially assess the sensitivities and changes with respect to the previous analyses, and partially to ensure that each of the analysis groups is in agreement before starting the Asimov (and later, data) fits. The event rates are produced by selecting some oscillation and systematic parameter values, reweighing the MC to these values, and looking at the predicted event rates split by interaction mode. This is done for three sets of oscillation parameters, unoscillated, Asimov A and Asimov B, listed in table A.1, and using the systematic parameter values before the ND280 fit (so taken from prior) and after the ND280 fit (where the best-fit values are used). The tables following table A.1 show the nominal event rates at SK, produced using the nominal systematic parameters' values taken from the BAFF ND280 fit.

Table A.1: Three sets of the neutrino oscillation parameter values used when producing the predicted event rates at SK

	Unoscillated	Asimov A	Asimov B
$\sin^2\theta_{12}$	0.0	0.304	
$\sin^2\theta_{13}$	0.00	0.0212	
$\sin^2\theta_{23}$	0.0	0.528	0.45
$\Delta m^2_{21}(\times 10^{-5}$ eV$^2)$	7.53		
$\Delta m^2_{32}(\times 10^{-3}$ eV$^2)$	2.509		
δ_{CP}	0	-1.601	0

Table A.2: Nominal post-ND280 fit event rate prediction for FHC 1R$_\mu$ sample: 14.9380×10^{20} POT, unoscillated

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	965.66282	0.00436	23.40636	0.00020	0.00000	0.00000	989.07374
CC1π	77.79879	0.00215	4.18909	0.00009	0.00000	0.00000	81.99012
CC coherent	0.76614	0.00000	0.17503	0.00000	0.00000	0.00000	0.94117
CCnπ	6.18317	0.00047	0.45082	0.00008	0.00000	0.00000	6.63454
CC other	0.82883	0.00043	0.05326	0.00000	0.00000	0.00000	0.88253
NCπ^0	0.62495	0.01869	0.02138	0.00143	0.00000	0.00000	0.66646
NC$\pi^{+/-}$	4.99423	0.09960	0.17726	0.00980	0.00000	0.00000	5.28090
NC coherent	0.00000	0.00000	0.00044	0.00003	0.00000	0.00000	0.00048
NC other	2.10279	0.07529	0.13268	0.00923	0.00000	0.00000	2.31998
2p-2h	134.78903	0.00163	2.57646	0.00008	0.00000	0.00000	137.36720
NC 1γ	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Sample totals	1193.75076	0.20262	31.18278	0.02095	0.00000	0.00000	
Total	1225.15712						

Table A.3: Nominal post-ND280 fit event rate prediction for FHC $1R_e$ sample: 14.9380×10^{20} POT, unoscillated

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	1.01072	6.89531	0.02101	0.27029	0.00000	0.00000	8.19734
CC1π	0.18445	1.01319	0.00522	0.07092	0.00000	0.00000	1.27378
CC coherent	0.00077	0.00896	0.00031	0.00416	0.00000	0.00000	0.01421
CCnπ	0.01783	0.11418	0.00057	0.00865	0.00000	0.00000	0.14123
CC other	0.01042	0.01870	0.00000	0.00093	0.00000	0.00000	0.03005
NCπ^0	1.72930	0.03761	0.06069	0.00357	0.00000	0.00000	1.83117
NC$\pi^{+/-}$	0.17178	0.00480	0.01033	0.00057	0.00000	0.00000	0.18749
NC coherent	0.53883	0.00729	0.04868	0.00296	0.00000	0.00000	0.59776
NC other	0.32454	0.01730	0.02153	0.00129	0.00000	0.00000	0.36467
2p-2h	0.13769	1.71864	0.00241	0.04543	0.00000	0.00000	1.90416
NC 1γ	0.96236	0.01722	0.05061	0.00235	0.00000	0.00000	1.03254
Sample totals	5.08870	9.85319	0.22136	0.41113	0.00000	0.00000	
Total	15.57439						

Table A.4: Nominal post-ND280 fit event rate prediction for RHC $1R_\mu$ sample : 16.3456×10^{20} POT, unoscillated

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	78.48710	0.00191	289.84935	0.00207	0.00000	0.00000	368.34042
CC1π	14.17658	0.00076	25.18321	0.00056	0.00000	0.00000	39.36111
CC coherent	0.15038	0.00000	1.63684	0.00011	0.00000	0.00000	1.78733
CCnπ	2.26618	0.00012	1.47460	0.00000	0.00000	0.00000	3.74091
CC other	0.29672	0.00013	0.16638	0.00000	0.00000	0.00000	0.46323
NCπ^0	0.13032	0.00479	0.13243	0.00358	0.00000	0.00000	0.27112
NC$\pi^{+/-}$	0.78284	0.03295	0.99469	0.02624	0.00000	0.00000	1.83671
NC coherent	0.00000	0.00000	0.00439	0.00000	0.00000	0.00000	0.00439
NC other	0.69330	0.03560	0.43187	0.01681	0.00000	0.00000	1.17759
2p-2h	17.81507	0.00061	23.59162	0.00061	0.00000	0.00000	41.40790
NC 1γ	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Sample totals	114.79849	0.07686	343.46537	0.04998	0.00000	0.00000	
Total	458.39070						

Table A.5: Nominal post-ND280 fit event rate prediction for RHC $1R_e$ sample: 16.3456×10^{20} POT, unoscillated

nue no osc

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.08897	1.19030	0.30234	1.93411	0.00000	0.00000	3.51572
CC1π	0.03616	0.23231	0.03905	0.32695	0.00000	0.00000	0.63447
CC coherent	0.00014	0.00226	0.00310	0.02415	0.00000	0.00000	0.02965
CCnπ	0.00842	0.04172	0.00227	0.02338	0.00000	0.00000	0.07579
CC other	0.00372	0.00477	0.00262	0.00162	0.00000	0.00000	0.01273
NCπ^0	0.30059	0.01164	0.38057	0.00988	0.00000	0.00000	0.70269
NC$\pi^{+/-}$	0.05384	0.00275	0.04314	0.00171	0.00000	0.00000	0.10143
NC coherent	0.10435	0.00481	0.44089	0.00678	0.00000	0.00000	0.55682
NC other	0.12374	0.00477	0.05756	0.00212	0.00000	0.00000	0.18819
2p-2h	0.01398	0.32814	0.01537	0.25961	0.00000	0.00000	0.61710
NC 1γ	0.17384	0.00959	0.38085	0.00822	0.00000	0.00000	0.57250
Sample totals	0.90775	1.83306	1.66775	2.59852	0.00000	0.00000	
Total	7.00709						

Table A.6: Nominal post-ND280 fit event rate prediction for FHC CC-1π^+ sample: 14.9380×10^{20} POT, unoscillated

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.35021	0.02687	0.00619	0.00185	0.00000	0.00000	0.38512
CC1π	0.55706	0.83485	0.00721	0.00403	0.00000	0.00000	1.40315
CC coherent	0.00368	0.01928	0.00021	0.00008	0.00000	0.00000	0.02325
CCnπ	0.04411	0.10623	0.00210	0.00264	0.00000	0.00000	0.15509
CC other	0.13342	0.01095	0.00074	0.00021	0.00000	0.00000	0.14532
NCπ^0	0.02080	0.00077	0.00084	0.00008	0.00000	0.00000	0.02249
NC$\pi^{+/-}$	0.06592	0.00167	0.00387	0.00033	0.00000	0.00000	0.07179
NC coherent	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NC other	0.21033	0.00847	0.01326	0.00090	0.00000	0.00000	0.23296
2p-2h	0.08352	0.02367	0.00179	0.00101	0.00000	0.00000	0.10998
NC 1γ	0.02908	0.00000	0.00068	0.00007	0.00000	0.00000	0.02983
Sample totals	1.49812	1.03277	0.03689	0.01120	0.00000	0.00000	
Total	2.57898						

Table A.7: Nominal post-ND280 fit event rate prediction for FHC $1R_\mu$ sample: 14.9380×10^{20} POT, oscillated with oscillation parameters set A

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	178.42434	0.00419	11.15508	0.00019	0.02257	0.00015	189.60652
CC1π	28.22707	0.00203	2.65499	0.00009	0.02377	0.00007	30.90802
CC coherent	0.29208	0.00000	0.09287	0.00000	0.00000	0.00000	0.38495
CCnπ	4.81427	0.00044	0.36286	0.00008	0.00000	0.00001	5.17765
CC other	0.70680	0.00042	0.04543	0.00000	0.00011	0.00000	0.75276
NCπ^0	0.62495	0.01869	0.02138	0.00143	0.00000	0.00000	0.66646
NC$\pi^{+/-}$	4.99423	0.09960	0.17726	0.00980	0.00000	0.00000	5.28090
NC coherent	0.00000	0.00000	0.00044	0.00003	0.00000	0.00000	0.00048
NC other	2.10279	0.07529	0.13268	0.00923	0.00000	0.00000	2.31998
2p-2h	35.79919	0.00154	1.44610	0.00008	0.02472	0.00004	37.27167
NC 1γ	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Sample totals	255.98572	0.20220	16.08909	0.02094	0.07116	0.00027	
Total	272.36939						

Table A.8: Nominal post-ND280 fit event rate prediction for FHC $1R_e$ sample: 14.9380×10^{20} POT, oscillated with oscillation parameters set A

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.16326	6.41268	0.00638	0.25474	45.63934	0.27065	52.74705
CC1π	0.04800	0.95056	0.00286	0.06769	4.44467	0.04638	5.56016
CC coherent	0.00011	0.00842	0.00018	0.00394	0.03386	0.00317	0.04969
CCnπ	0.01267	0.11032	0.00034	0.00839	0.07763	0.00292	0.21227
CC other	0.00068	0.01817	0.00000	0.00091	0.00839	0.00022	0.02836
NCπ^0	1.72930	0.03761	0.06069	0.00357	0.00000	0.00000	1.83117
NC$\pi^{+/-}$	0.17178	0.00480	0.01033	0.00057	0.00000	0.00000	0.18749
NC coherent	0.53883	0.00729	0.04868	0.00296	0.00000	0.00000	0.59776
NC other	0.32454	0.01730	0.02153	0.00129	0.00000	0.00000	0.36467
2p-2h	0.03028	1.60663	0.00119	0.04313	8.46347	0.03828	10.18298
NC 1γ	0.96236	0.01722	0.05061	0.00235	0.00000	0.00000	1.03254
Sample totals	3.98182	9.19099	0.20280	0.38955	58.66736	0.36161	
Total	72.79413						

Table A.9: Nominal post-ND280 fit event rate prediction for RHC $1R_\mu$ sample: 16.3456×10^{20} POT, oscillated with oscillation parameters set A

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	33.19935	0.00183	64.81563	0.00199	0.00183	0.00286	98.02348
CC1π	8.61403	0.00072	9.82402	0.00054	0.00127	0.00129	18.44187
CC coherent	0.10168	0.00000	0.37866	0.00010	0.00000	0.00014	0.48058
CCnπ	1.78891	0.00012	1.14781	0.00000	0.00000	0.00000	2.93684
CC other	0.24670	0.00013	0.13957	0.00000	0.00000	0.00000	0.38639
NCπ^0	0.13032	0.00479	0.13243	0.00358	0.00000	0.00000	0.27112
NC$\pi^{+/-}$	0.78284	0.03295	0.99469	0.02624	0.00000	0.00000	1.83671
NC coherent	0.00000	0.00000	0.00439	0.00000	0.00000	0.00000	0.00439
NC other	0.69330	0.03560	0.43187	0.01681	0.00000	0.00000	1.17759
2p-2h	9.62226	0.00056	6.24008	0.00058	0.00069	0.00043	15.86460
NC 1γ	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Sample totals	55.17938	0.07670	84.10914	0.04984	0.00379	0.00471	
Total	139.42356						

Table A.10: Nominal post-ND280 fit event rate prediction for RHC $1R_e$ sample: 16.3456×10^{20} POT, oscillated with oscillation parameters set A

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.02799	1.10913	0.05225	1.82035	2.15891	5.95579	11.12442
CC1π	0.01655	0.21958	0.01605	0.30983	0.30267	0.70350	1.56818
CC coherent	0.00005	0.00214	0.00090	0.02278	0.00241	0.06511	0.09339
CCnπ	0.00623	0.04044	0.00151	0.02263	0.01914	0.01292	0.10285
CC other	0.00212	0.00465	0.00100	0.00158	0.00253	0.00091	0.01279
NCπ^0	0.30059	0.01164	0.38057	0.00988	0.00000	0.00000	0.70269
NC$\pi^{+/-}$	0.05384	0.00275	0.04314	0.00171	0.00000	0.00000	0.10143
NC coherent	0.10435	0.00481	0.44089	0.00678	0.00000	0.00000	0.55682
NC other	0.12374	0.00477	0.05756	0.00212	0.00000	0.00000	0.18819
2p-2h	0.00630	0.30851	0.00491	0.24506	0.48497	0.69685	1.74660
NC 1γ	0.17384	0.00959	0.38085	0.00822	0.00000	0.00000	0.57250
Sample totals	0.81559	1.71799	1.37963	2.45093	2.97063	7.43509	
Total	16.76987						

Table A.11: Nominal post-ND280 fit event rate prediction for FHC CC-1π^+ sample: 14.9380×10^{20} POT, oscillated with oscillation parameters set A

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.02911	0.02487	0.00118	0.00173	0.20772	0.00259	0.26720
CC1π	0.07759	0.77591	0.00260	0.00385	4.79591	0.00222	5.65807
CC coherent	0.00029	0.01791	0.00008	0.00007	0.10924	0.00010	0.12770
CCnπ	0.01937	0.10193	0.00108	0.00255	0.11477	0.00137	0.24107
CC other	0.01868	0.01062	0.00007	0.00020	0.00206	0.00007	0.03170
NCπ^0	0.02080	0.00077	0.00084	0.00008	0.00000	0.00000	0.02249
NC$\pi^{+/-}$	0.06592	0.00167	0.00387	0.00033	0.00000	0.00000	0.07179
NC coherent	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NC other	0.21033	0.00847	0.01326	0.00090	0.00000	0.00000	0.23296
2p-2h	0.00994	0.02210	0.00050	0.00096	0.15555	0.00103	0.19007
NC 1γ	0.02908	0.00000	0.00068	0.00007	0.00000	0.00000	0.02983
Sample totals	0.48109	0.96424	0.02416	0.01074	5.38527	0.00739	
Total	6.87289						

Table A.12: Nominal post-ND280 fit event rate prediction for FHC 1R$_\mu$ sample: 14.9380×10^{20} POT, oscillated with oscillation parameters set B

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	189.17456	0.00419	11.27825	0.00019	0.01524	0.00022	200.47266
CC1π	28.73398	0.00203	2.66702	0.00009	0.01627	0.00008	31.41947
CC coherent	0.29753	0.00000	0.09361	0.00000	0.00000	0.00001	0.39115
CCnπ	4.82414	0.00044	0.36345	0.00008	0.00000	0.00001	5.18811
CC other	0.70763	0.00042	0.04548	0.00000	0.00011	0.00000	0.75364
NCπ^0	0.62495	0.01869	0.02138	0.00143	0.00000	0.00000	0.66646
NC$\pi^{+/-}$	4.99423	0.09960	0.17726	0.00980	0.00000	0.00000	5.28090
NC coherent	0.00000	0.00000	0.00044	0.00003	0.00000	0.00000	0.00048
NC other	2.10279	0.07529	0.13268	0.00923	0.00000	0.00000	2.31998
2p-2h	36.93442	0.00154	1.45567	0.00008	0.01631	0.00005	38.40807
NC 1γ	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Sample totals	268.39423	0.20220	16.23526	0.02094	0.04793	0.00036	
Total	284.90092						

Table A.13: Nominal post-ND280 fit event rate prediction for FHC $1R_e$ sample: 14.9380×10^{20} POT, oscillated with oscillation parameters set B

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.17385	6.41268	0.00651	0.25474	30.69151	0.33017	37.86945
CC1π	0.04945	0.95056	0.00288	0.06769	3.24038	0.05466	4.36561
CC coherent	0.00012	0.00842	0.00018	0.00394	0.02469	0.00371	0.04107
CCnπ	0.01271	0.11032	0.00035	0.00839	0.06532	0.00336	0.20045
CC other	0.00081	0.01817	0.00000	0.00091	0.00740	0.00025	0.02754
NCπ^0	1.72930	0.03761	0.06069	0.00357	0.00000	0.00000	1.83117
NC$\pi^{+/-}$	0.17178	0.00480	0.01033	0.00057	0.00000	0.00000	0.18749
NC coherent	0.53883	0.00729	0.04868	0.00296	0.00000	0.00000	0.59776
NC other	0.32454	0.01730	0.02153	0.00129	0.00000	0.00000	0.36467
2p-2h	0.03138	1.60663	0.00120	0.04313	5.91439	0.04587	7.64260
NC 1γ	0.96236	0.01722	0.05061	0.00235	0.00000	0.00000	1.03254
Sample totals	3.99513	9.19099	0.20296	0.38955	39.94369	0.43802	
Total	54.16034						

Table A.14: Nominal post-ND280 fit event rate prediction for RHC $1R_\mu$ sample: 16.3456×10^{20} POT, oscillated with oscillation parameters set B

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	33.72265	0.00183	67.51569	0.00199	0.00143	0.00378	101.24737
CC1π	8.66063	0.00072	9.96922	0.00054	0.00092	0.00153	18.63356
CC coherent	0.10209	0.00000	0.39261	0.00010	0.00000	0.00015	0.49495
CCnπ	1.79220	0.00012	1.15008	0.00000	0.00000	0.00000	2.94240
CC other	0.24704	0.00013	0.13974	0.00000	0.00000	0.00000	0.38691
NCπ^0	0.13032	0.00479	0.13243	0.00358	0.00000	0.00000	0.27112
NC$\pi^{+/-}$	0.78284	0.03295	0.99469	0.02624	0.00000	0.00000	1.83671
NC coherent	0.00000	0.00000	0.00439	0.00000	0.00000	0.00000	0.00439
NC other	0.69330	0.03560	0.43187	0.01681	0.00000	0.00000	1.17759
2p-2h	9.69609	0.00056	6.42560	0.00058	0.00051	0.00052	16.12385
NC 1γ	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Sample totals	55.82716	0.07670	87.15632	0.04984	0.00286	0.00598	
Total	143.11885						

Table A.15: Nominal post-ND280 fit event rate prediction for RHC $1R_e$ sample: 16.3456×10^{20} POT, oscillated with oscillation parameters set B

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.02862	1.10913	0.05492	1.82035	1.52181	7.11449	11.64931
CC1π	0.01672	0.21958	0.01626	0.30983	0.23935	0.83862	1.64036
CC coherent	0.00005	0.00214	0.00092	0.02278	0.00189	0.07699	0.10477
CCnπ	0.00624	0.04044	0.00151	0.02263	0.01687	0.01511	0.10280
CC other	0.00214	0.00465	0.00102	0.00158	0.00226	0.00105	0.01270
NCπ^0	0.30059	0.01164	0.38057	0.00988	0.00000	0.00000	0.70269
NC$\pi^{+/-}$	0.05384	0.00275	0.04314	0.00171	0.00000	0.00000	0.10143
NC coherent	0.10435	0.00481	0.44089	0.00678	0.00000	0.00000	0.55682
NC other	0.12374	0.00477	0.05756	0.00212	0.00000	0.00000	0.18819
2p-2h	0.00636	0.30851	0.00501	0.24506	0.36272	0.82948	1.75714
NC 1γ	0.17384	0.00959	0.38085	0.00822	0.00000	0.00000	0.57250
Sample totals	0.81649	1.71799	1.38266	2.45093	2.14490	8.87574	
Total	17.38872						

Table A.16: Nominal post-ND280 fit event rate prediction for FHC CC-1π^+ sample: 14.9380×10^{20} POT, oscillated with oscillation parameters set B

	ν_μ	ν_e	$\bar{\nu}_\mu$	$\bar{\nu}_e$	ν_e signal	$\bar{\nu}_e$ signal	Total
CCQE	0.03352	0.02487	0.00124	0.00173	0.14230	0.00335	0.20700
CC1π	0.08327	0.77591	0.00264	0.00385	3.42831	0.00262	4.29659
CC coherent	0.00033	0.01791	0.00008	0.00007	0.07474	0.00012	0.09325
CCnπ	0.01957	0.10193	0.00109	0.00255	0.09283	0.00160	0.21957
CC other	0.02019	0.01062	0.00008	0.00020	0.00185	0.00008	0.03302
NCπ^0	0.02080	0.00077	0.00084	0.00008	0.00000	0.00000	0.02249
NC$\pi^{+/-}$	0.06592	0.00167	0.00387	0.00033	0.00000	0.00000	0.07179
NC coherent	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NC other	0.21033	0.00847	0.01326	0.00090	0.00000	0.00000	0.23296
2p-2h	0.01081	0.02210	0.00051	0.00096	0.11297	0.00124	0.14858
NC 1γ	0.02908	0.00000	0.00068	0.00007	0.00000	0.00000	0.02983
Sample totals	0.49380	0.96424	0.02430	0.01074	3.85300	0.00901	
Total	5.35509						

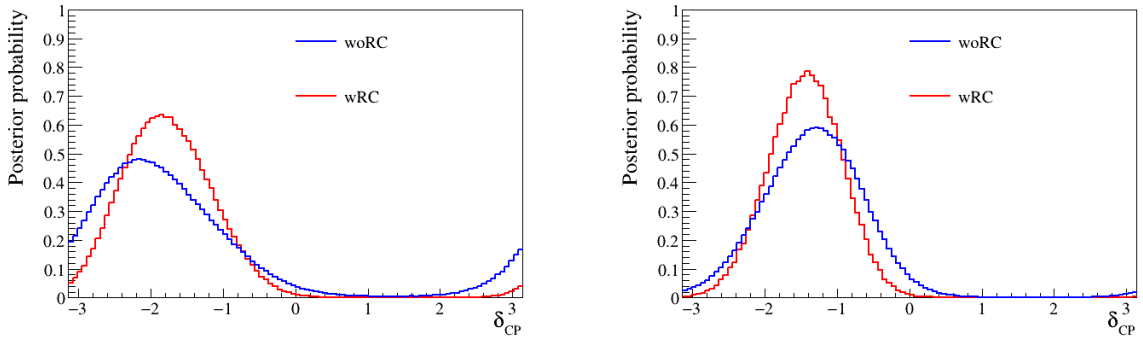
Appendix B

Comparisons involving run 1–9 analysis

This appendix will list various oscillation parameter interval comparisons made that involve the run 1–9 analysis. The comparisons are made for the with against without the reactor constraint on $\sin^2\theta_{13}$, the full run 1–9 (here called 1–9d) against the older 1–9c, and, finally, the Asimov A against the data fits.

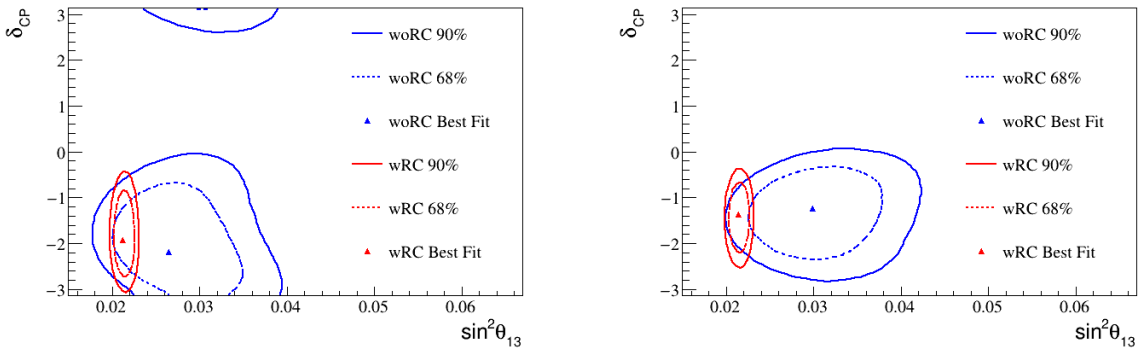
B.1 With vs without the reactor constraint on $\sin^2\theta_{13}$

B.1.1 Data



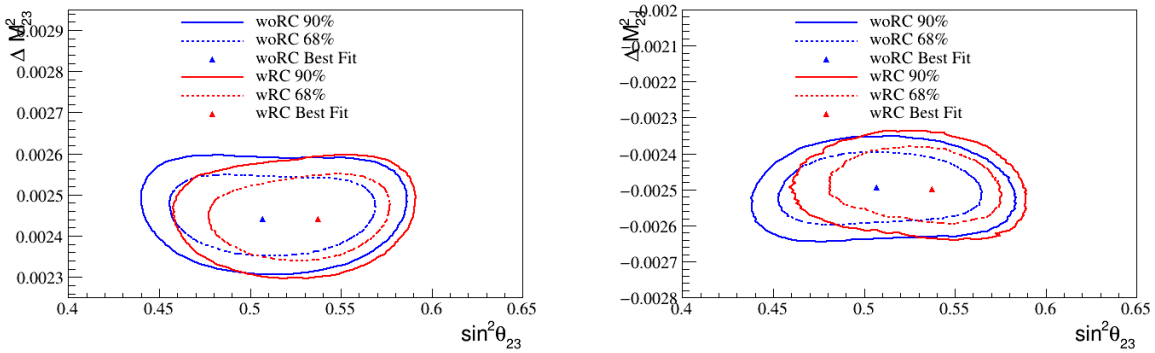
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.1: Comparisons of the δ_{CP} results from the run 1–9 data fit with a prior on $\sin^2\theta_{13}$ (woRC), against fit with a prior from the reactor neutrino experiments (wRC).



(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.2: Comparisons of the appearance results from the run 1–9 data fit with a prior on $\sin^2\theta_{13}$ (woRC), against fit with a prior from the reactor neutrino experiments (wRC).

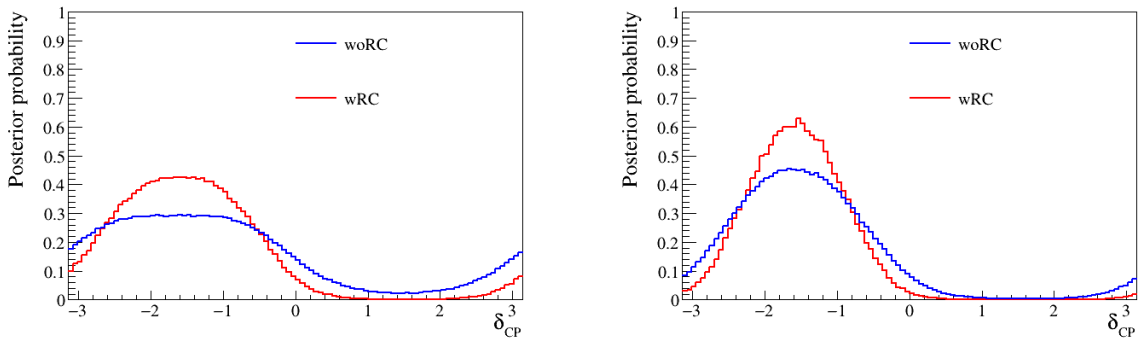


(a) Normal mass hierarchy

(b) Inverse mass hierarchy

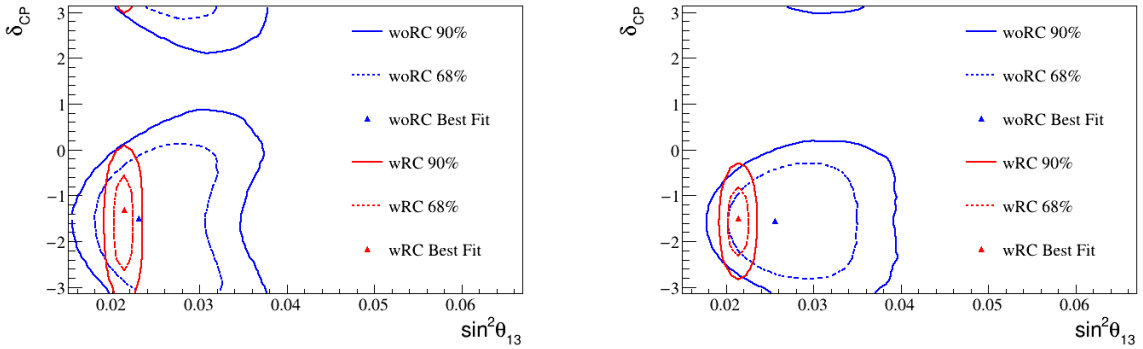
Figure B.3: Comparisons of the disappearance results from the run 1–9 data fit with a prior on $\sin^2\theta_{13}$ (woRC), against fit with a prior from the reactor neutrino experiments (wRC).

B.1.2 Asimov A



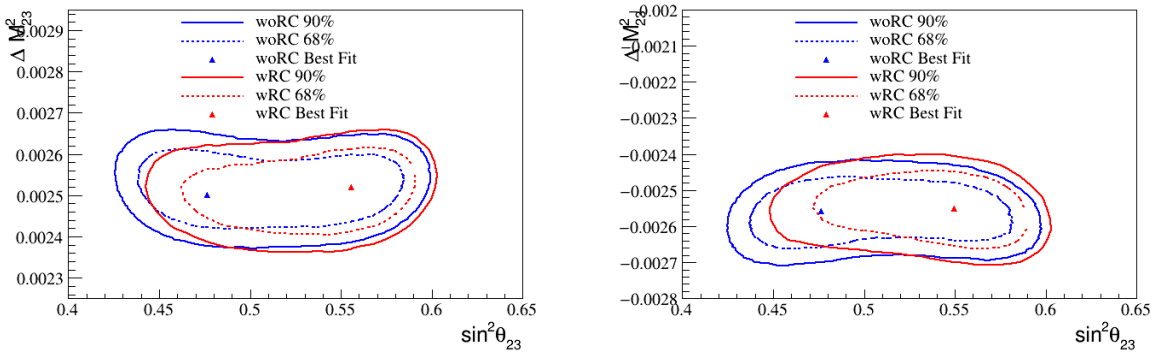
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.4: Comparisons of the δ_{CP} results from the run 1–9 Asimov A fit with a prior on $\sin^2\theta_{13}$ (woRC), against fit with a prior from the reactor neutrino experiments (wRC).



(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.5: Comparisons of the appearance results from the run 1–9 Asimov A fit with a prior on $\sin^2\theta_{13}$ (woRC), against fit with a prior from the reactor neutrino experiments (wRC).



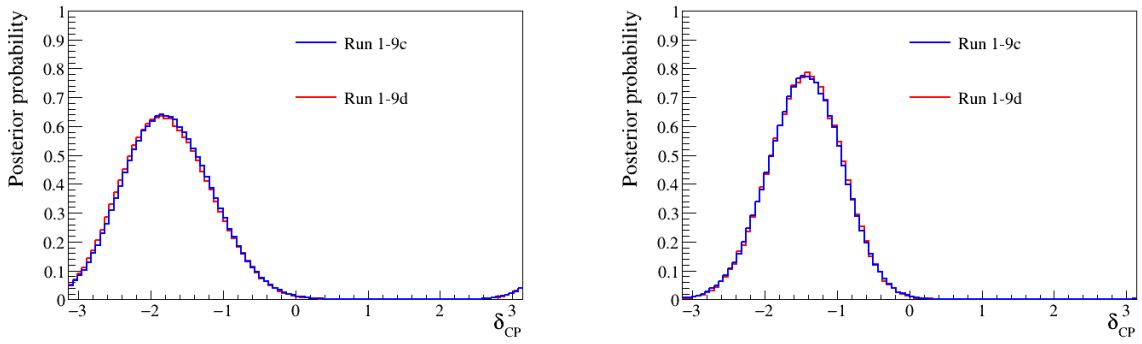
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.6: Comparisons of the disappearance results from the run 1–9 Asimov A fit with a prior on $\sin^2\theta_{13}$ (woRC), against fit with a prior from the reactor neutrino experiments (wRC).

B.2 Run 1–9d vs run 1–9c

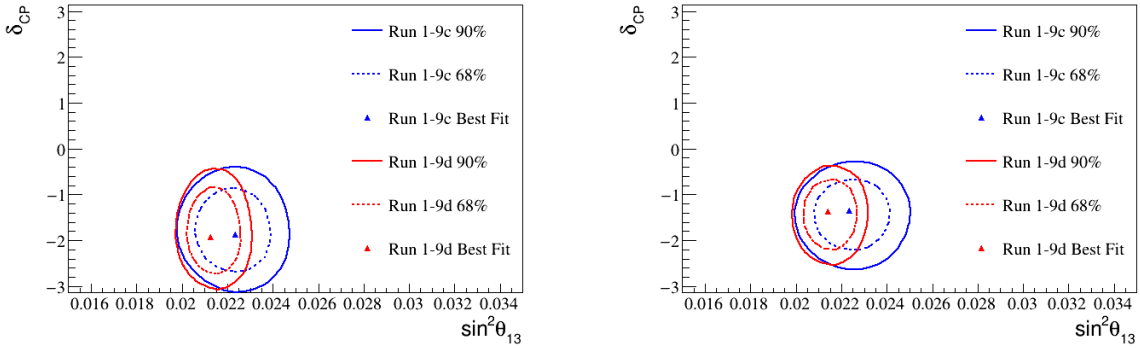
Run 1–9c analysis was presented during the Neutrino 2018 conference [54], with 1.49×10^{21} FHC POT and 1.12×10^{21} RHC POT, with run 1–9d representing the main analysis described in this thesis.

B.2.1 Data, with the $\sin^2\theta_{13}$ constraint



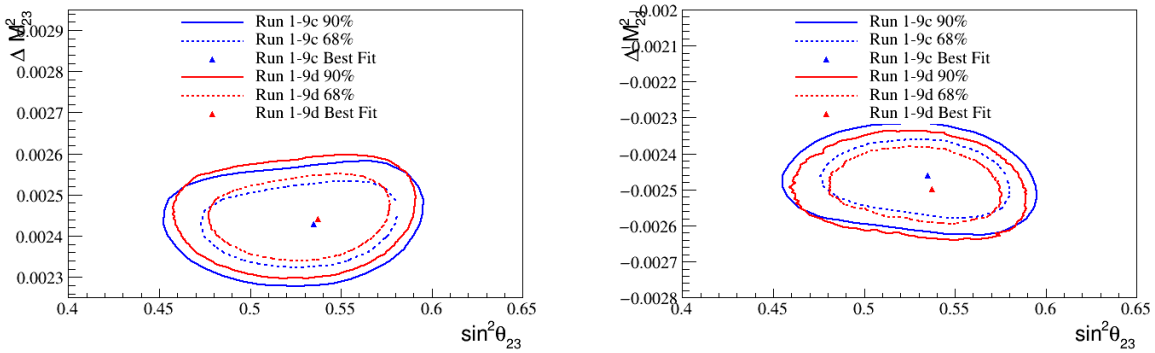
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.7: Comparisons of the δ_{CP} results from the run 1–9d data fit against fit from the run 1–9c, with the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy (b) Inverse mass hierarchy

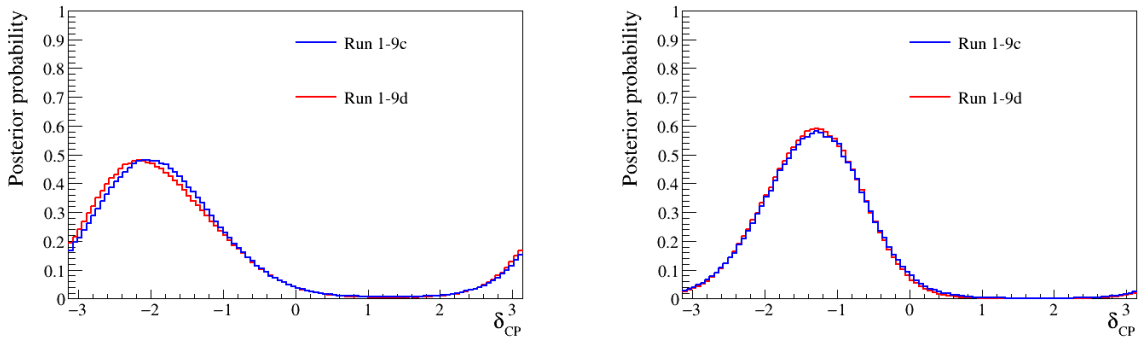
Figure B.8: Comparisons of the appearance results from the run 1–9d data fit against fit from the run 1–9c, with the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy (b) Inverse mass hierarchy

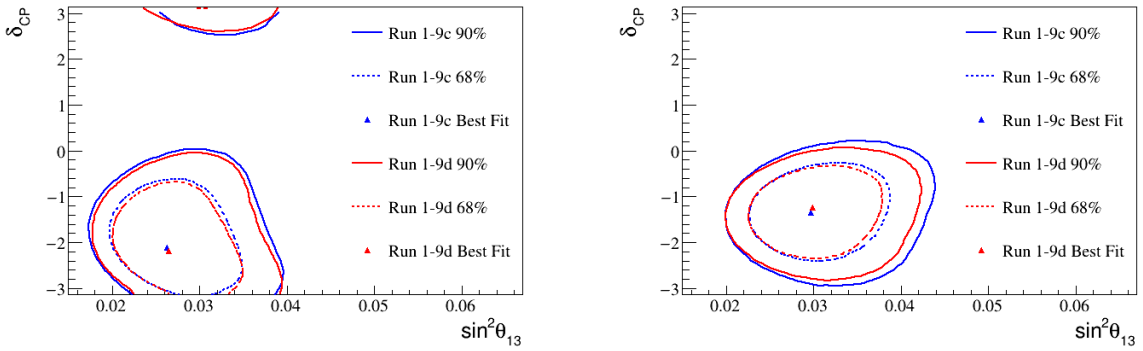
Figure B.9: Comparisons of the disappearance results from the run 1–9d data fit against fit from the run 1–9c, with the reactor constraint on $\sin^2\theta_{13}$.

B.2.2 Data, with a flat prior on $\sin^2\theta_{13}$



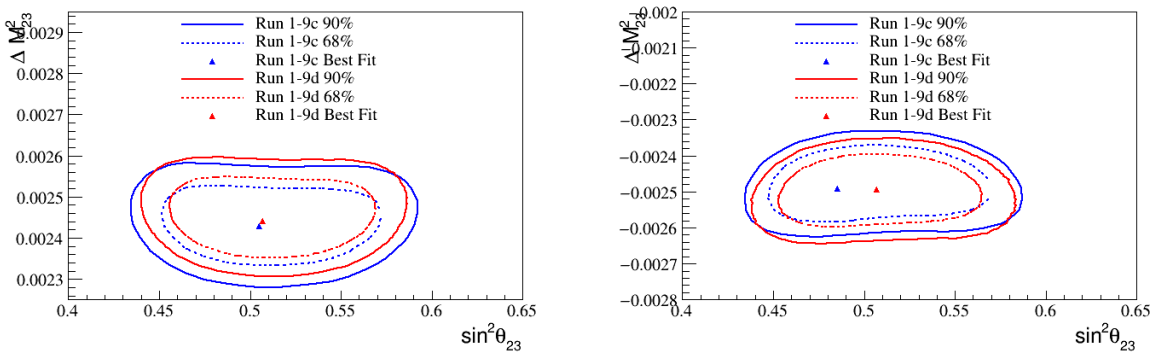
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.10: Comparisons of the δ_{CP} results from the run 1–9d data fit against fit from the run 1–9c, without the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.11: Comparisons of the appearance results from the run 1–9d data fit against fit from the run 1–9c, without the reactor constraint on $\sin^2\theta_{13}$.

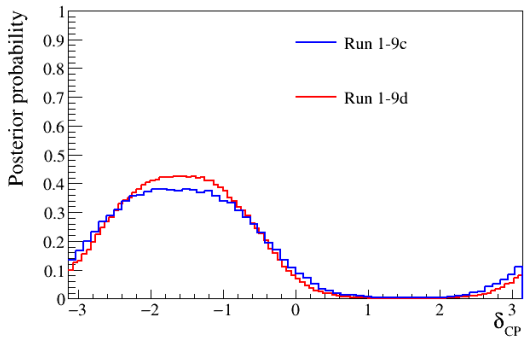


(a) Normal mass hierarchy

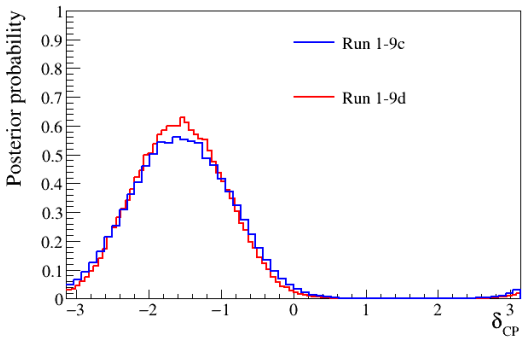
(b) Inverse mass hierarchy

Figure B.12: Comparisons of the disappearance results from the run 1–9d data fit against fit from the run 1–9c, without the reactor constraint on $\sin^2\theta_{13}$.

B.2.3 Asimov A, with the $\sin^2\theta_{13}$ constraint

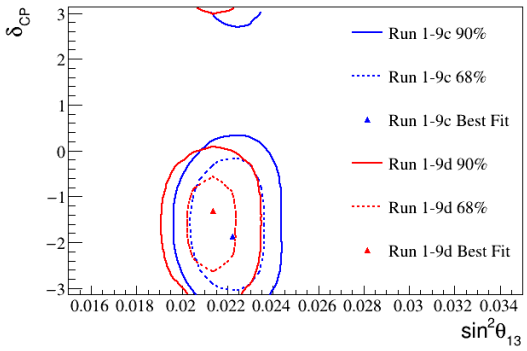


(a) Normal mass hierarchy

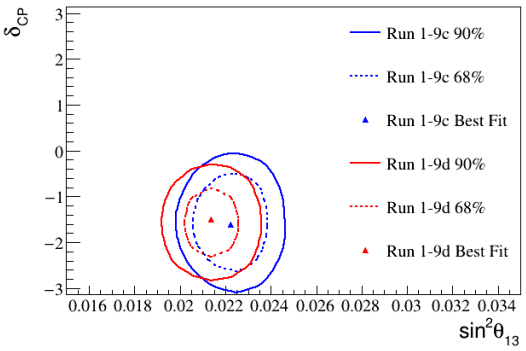


(b) Inverse mass hierarchy

Figure B.13: Comparisons of the δ_{CP} results from the run 1–9d Asimov A fit against fit from the run 1–9c, with the reactor constraint on $\sin^2\theta_{13}$.

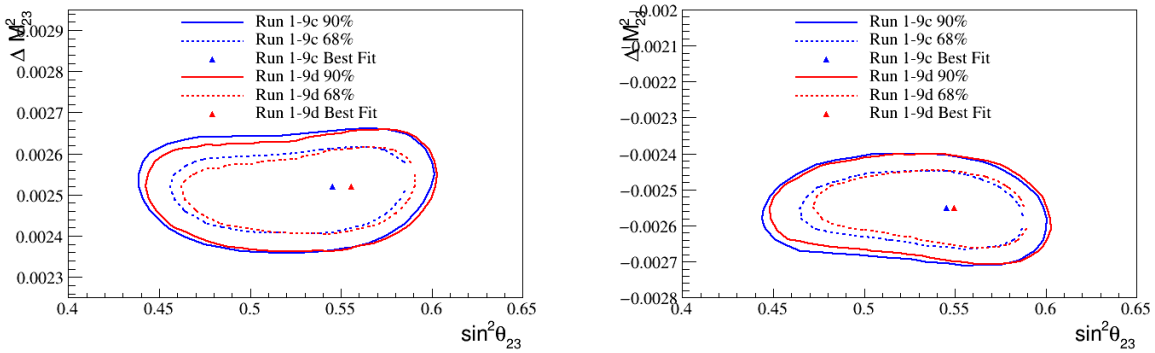


(a) Normal mass hierarchy



(b) Inverse mass hierarchy

Figure B.14: Comparisons of the appearance results from the run 1–9d Asimov A fit against fit from the run 1–9c, with the reactor constraint on $\sin^2\theta_{13}$.

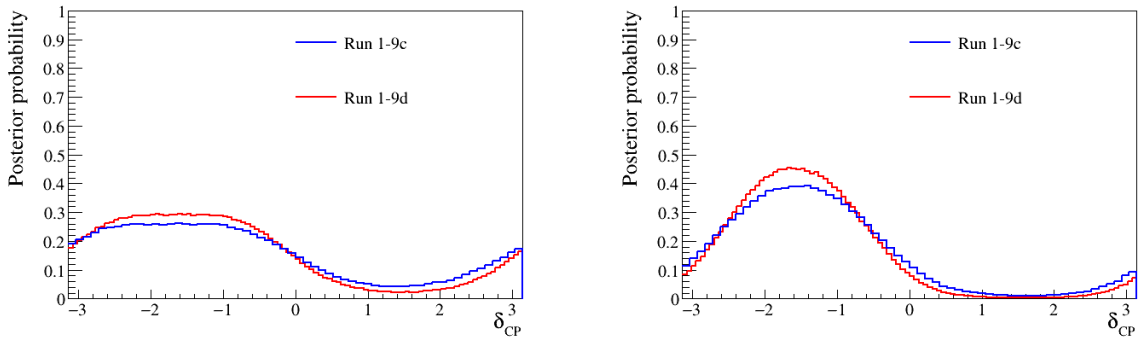


(a) Normal mass hierarchy

(b) Inverse mass hierarchy

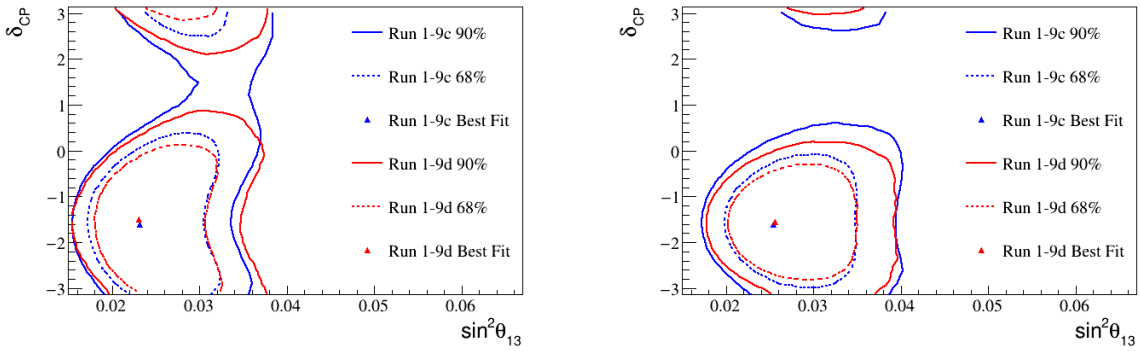
Figure B.15: Comparisons of the disappearance results from the run 1–9d Asimov A fit against fit from the run 1–9c, with the reactor constraint on $\sin^2\theta_{13}$.

B.2.4 Asimov A, with a flat prior on $\sin^2\theta_{13}$



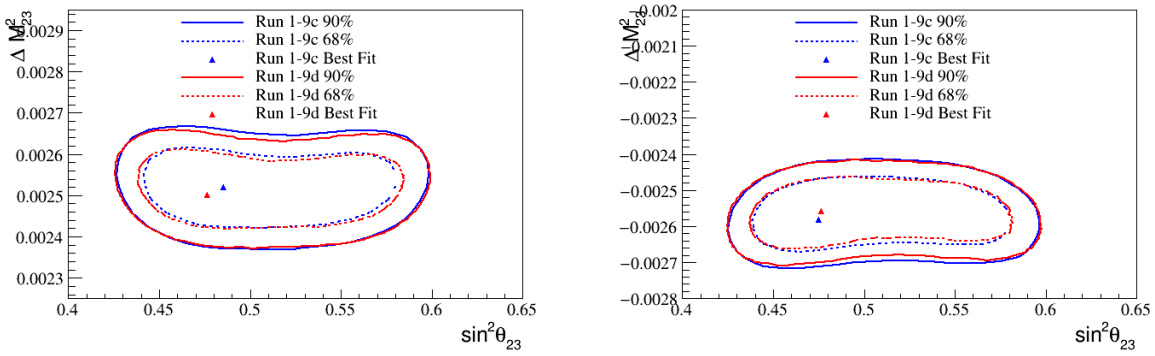
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.16: Comparisons of the δ_{CP} results from the run 1–9d Asimov A fit against fit from the run 1–9c, without the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.17: Comparisons of the appearance results from the run 1–9d Asimov A fit against fit from the run 1–9c, without the reactor constraint on $\sin^2\theta_{13}$.

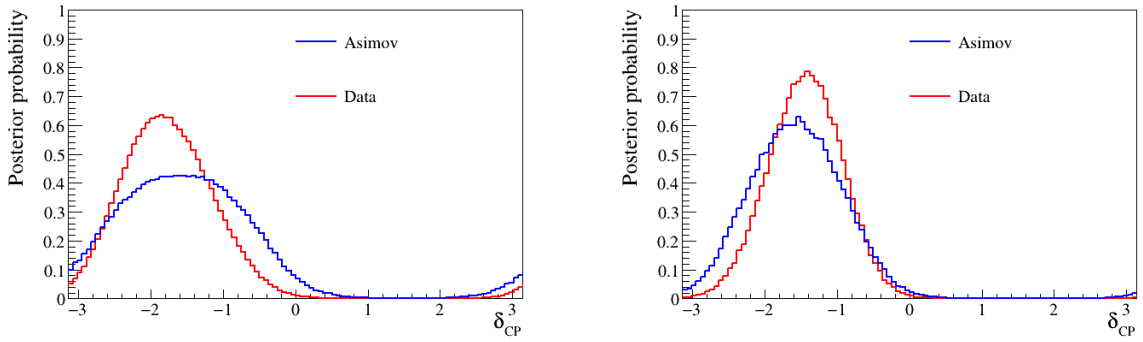


(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.18: Comparisons of the disappearance results from the run 1–9d Asimov A fit against fit from the run 1–9c, without the reactor constraint on $\sin^2\theta_{13}$.

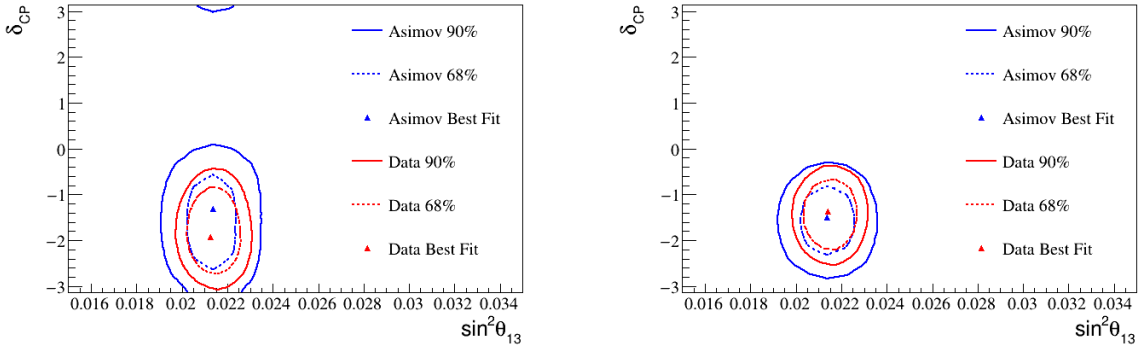
B.3 Data vs Asimov A

B.3.1 With the $\sin^2\theta_{13}$ constraint



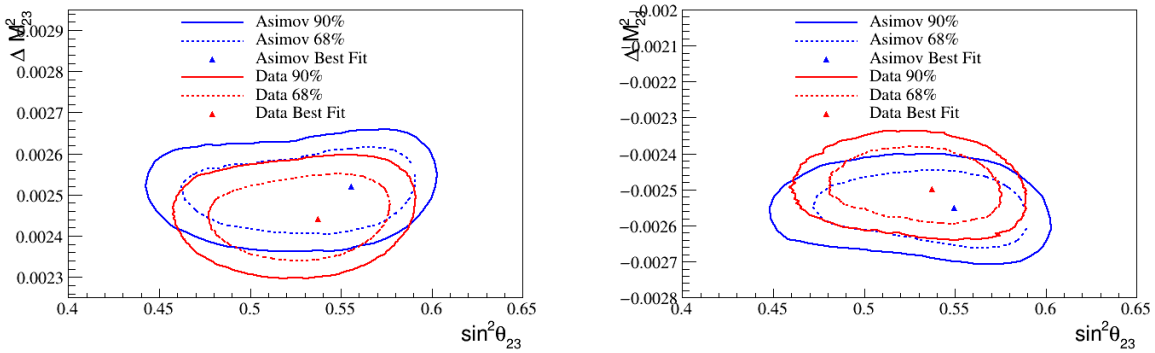
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.19: Comparisons of the δ_{CP} results from the run 1–9 Asimov A fit against the data fit, with the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy (b) Inverse mass hierarchy

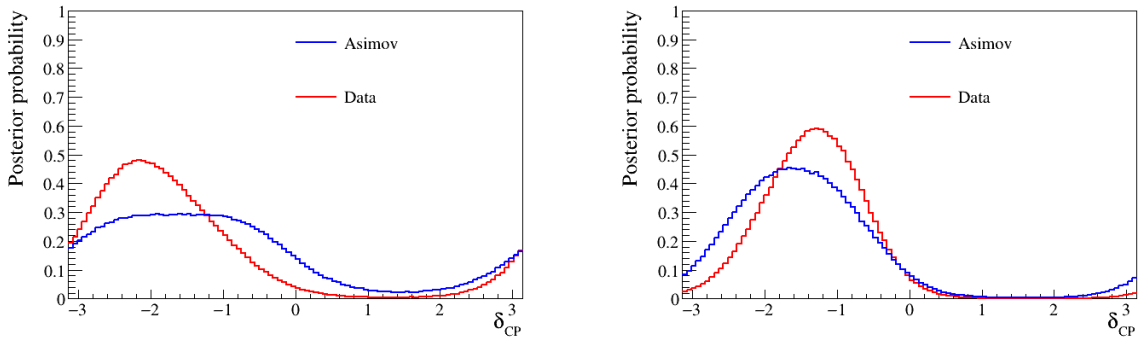
Figure B.20: Comparisons of the appearance results from the run 1–9 Asimov A fit against the data fit, with the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy (b) Inverse mass hierarchy

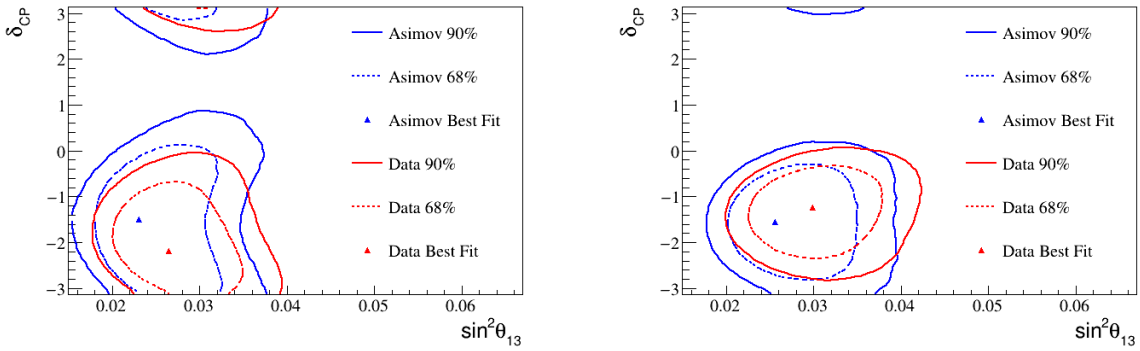
Figure B.21: Comparisons of the disappearance results from the run 1–9 Asimov A fit against the data fit, with the reactor constraint on $\sin^2\theta_{13}$.

B.3.2 With a flat prior on $\sin^2\theta_{13}$



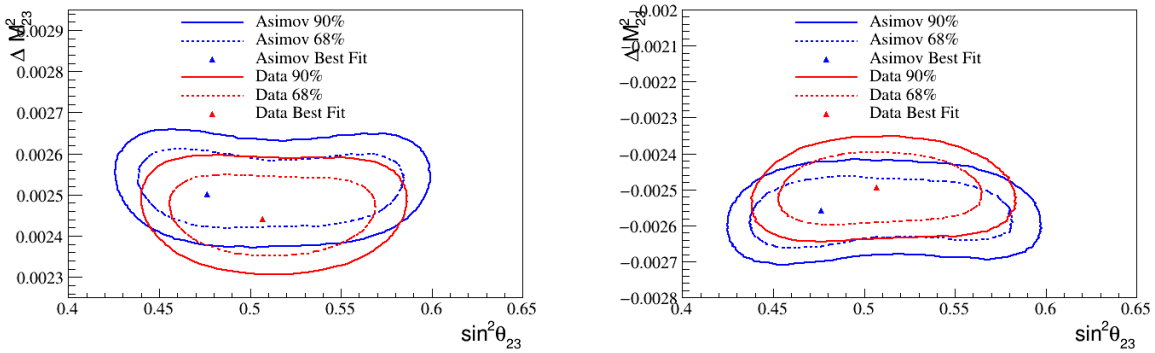
(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.22: Comparisons of the δ_{CP} results from the run 1–9 Asimov A fit against the data fit, without the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy (b) Inverse mass hierarchy

Figure B.23: Comparisons of the appearance results from the run 1–9 Asimov A fit against the data fit, without the reactor constraint on $\sin^2\theta_{13}$.



(a) Normal mass hierarchy

(b) Inverse mass hierarchy

Figure B.24: Comparisons of the disappearance results from the run 1–9 Asimov A fit against the data fit, without the reactor constraint on $\sin^2\theta_{13}$.