Simulatenous measurement of neutrino mixing parameters in beam and atmospheric oscillations

Daniel Robert Clement Barrow

Magdalen College,
Oxford University

Simulatenous measurement of neutrino mixing parameters in beam and atmospheric oscillations

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Nulla aliquet porttitor lacus luctus accumsan tortor posuere. Pulvinar neque laoreet suspendisse interdum. Sem viverra aliquet eget sit. Nunc sed velit dignissim sodales ut eu sem integer vitae. At erat pellentesque adipiscing commodo elit at imperdiet dui accumsan. Fames ac turpis egestas integer eget aliquet nibh. Scelerisque eu ultrices vitae auctor eu augue. Purus non enim praesent elementum facilisis leo vel. Sollicitudin nibh sit amet commodo. Vitae auctor eu augue ut. Vel quam elementum pulvinar etiam. A condimentum vitae sapien pellentesque habitant morbi tristique senectus. Viverra accumsan in nisl nisi scelerisque eu ultrices. Sed viverra ipsum nunc aliquet bibendum enim. Sit amet purus gravida quis. In vitae turpis massa sed elementum tempus egestas sed sed. Vivamus arcu felis bibendum ut tristique et egestas. Senectus et netus et malesuada fames ac turpis. Ac auctor augue mauris augue.

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. This dissertation does not exceed the word limit for the respective Degree Committee.

Daniel Robert Clement Barrow

Acknowledgements

at pretium nibh ipsum. Eget nunc scelerisque viverra mauris in aliquam. Arcu vitae elementum curabitur vitae nunc sed velit dignissim. Sed arcu non odio euismod lacinia at quis risus sed. Vitae tempus quam pellentesque nec nam aliquam sem et tortor. Viverra aliquet eget sit amet tellus cras adipiscing. Purus sit amet luctus venenatis lectus magna. In aliquam sem fringilla ut morbi tincidunt augue. Fermentum dui faucibus in ornare. Aliquam malesuada bibendum arcu vitae elementum curabitur vitae nunc sed.

Ultricies leo integer malesuada nunc vel risus commodo. Tellus cras adipiscing enim eu turpis egestas pretium. Dictumst quisque sagittis purus sit amet volutpat consequat mauris nunc. Vitae congue mauris rhoncus aenean vel elit scelerisque mauris pellentesque. Vel facilisis volutpat est velit egestas dui id ornare. Suscipit adipiscing bibendum est ultricies. At in tellus integer feugiat scelerisque varius morbi enim. Cras semper auctor neque vitae tempus. Commodo sed egestas egestas fringilla phasellus faucibus. Cras pulvinar mattis nunc sed blandit. Pretium viverra suspendisse potenti nullam ac tortor vitae. Purus sit amet volutpat consequat. Orci sagittis eu volutpat odio facilisis mauris. Sit amet massa vitae tortor condimentum lacinia quis. Commodo sed egestas egestas fringilla phasellus. Sed libero enim sed faucibus turpis. Vitae tempus quam pellentesque nec.

Blandit massa enim nec dui. Viverra tellus in hac habitasse platea dictumst vestibulum. Bibendum enim facilisis gravida neque convallis. Sagittis nisl rhoncus mattis rhoncus urna neque. Nisl rhoncus mattis rhoncus urna neque. Ac tortor vitae purus faucibus ornare. Aenean sed adipiscing diam donec adipiscing tristique risus. Sapien nec sagittis aliquam malesuada bibendum. Et leo duis ut diam quam nulla. Tellus rutrum tellus pellentesque eu tincidunt tortor aliquam nulla facilisi. Fermentum leo vel orci porta non pulvinar neque. Eget sit amet tellus cras adipiscing enim eu. Sed viverra tellus in hac habitasse platea dictumst vestibulum.

Sit amet consectetur adipiscing elit. At varius vel pharetra vel turpis nunc eget lorem. Elit scelerisque mauris pellentesque pulvinar pellentesque habitant morbi tristique senectus. Odio euismod lacinia at quis risus sed vulputate odio. Vitae suscipit tellus mauris a diam maecenas sed enim ut. Dignissim convallis aenean et tortor at risus viverra. Diam sollicitudin tempor id eu. Erat velit scelerisque in dictum. Blandit cursus risus at ultrices. Ac tortor vitae purus faucibus ornare suspendisse sed.

Contents

Chapter 1

Introduction

Chapter 2

Neutrino Oscillation Physics

Neutrino Oscillation Physics Chapter

2.1. Discovery of Neutrinos

At the start of the 20th century, the electrons emitted from β-decay of the nucleus were found to have a continuous energy spectrum [1,2]. This observation seemingly broke the conservation of energy invoked within the nuclear models of that period. Postulated in 1930 by Pauli as the solution to this problem, the neutrino (originally termed "neutron") was theorised to be an electrically neutral spin-1/2 fermion with a mass on the same order of magnitude as the electron [3]. This particle was to be emitted with the electron in β-decay to ellivate the apparent breaking of energy conservation. As a predecessor of weak interaction model, Fermi's theory of β-decay developed the understanding by coupleing the four consistuent particles; electron, proton, neutron and neutrino, into a consistent model [4].

Whilst Pauli was not convinced of the ability to detect neutrinos, the first observations of the particle were made in mid-1950s by using inverse β-decay (IBD) process, $\bar{\nu}_e + p \rightarrow n + e^+$, from neutrinos generated by nuclear reactors [5,6]. The detector consisted of two parts; a neutrino interaction medium and a liquid scintialltor. The interaction medium is built from two water tanks, each loaded with cadium chloride to allow increased efficiency of neutron capture. The positron emitted from IBD annihaltes, $e^+ + e^- \rightarrow 2\gamma$, generating a prompt signal, whereas the neutron is captured on the cadmium via $n + ^{108}Cd \rightarrow ^{109m}Cd \rightarrow ^{109}Cd + \gamma$, producing a delayed signal. The experiment also observed an increase in the neutrino event rate when the reactor was operating compared to when it was switched off, in much the same way as modern reactor neutrino experiments operate.

After the discovery of the ν_e, the natural question of how many flavours of neutrino exist and how that aligns with the number of charged leptons was asked. In 1962, a measurement of the ν_μ was conducted at the Brookhaven National Laboratory [7].

A proton beam was directed at a beryllium target, generating a π dominated beam which then decayed via $\pi^+ \rightarrow \mu^+ + \nu_\mu$ and the subsequent interactions of the ν_μ were observed. The final observation to be made was that of the ν_τ from the DONUT experiment [8].

2.2. Theory of Neutrino Oscillation

As direct evidence of beyond Standard Model physics, a neutrino generated with lepton flavour α can mutate into a different lepton flavour β after propagating some distance. This phenomena is called neutrino oscillation and requires that neutrinos must have a non-zero mass. This behaviour has been characterised by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) [9–11] mixing matrix which describes how the flavour and mass eigenstates are associated. This is analogous to the Cabbibo-Kobayashi-Maskawa (CKM) [12] matrix found in quark physics.

2.2.1. Three Flavour Oscillations

The PMNS parameterisation defines three flavour eigenstates, ν_e, ν_μ and ν_τ (denoted ν_α), which are assigned based upon the weak interaction flavour states, and three mass eigenstates, ν_1, ν_2 and ν_3 (denoted ν_i). Each mass eigenstate is the superposition of all three flavour states,

$$|\nu_i\rangle = \sum_\alpha U_{\alpha i} |\nu_\alpha\rangle \,. \tag{2.1}$$

U is the PMNS matrix which correlates the mass and flavour eigenstates

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \,. \tag{2.2}$$

The propogation of a neutrino through space is determined by the superposition of the three mass eigenstates. Therefore, the propogation of a neutrino flavour eigenstate can be re-written as a plane-wave solution to the time-dependent Schrödinger equation,

$$|\nu_\alpha(t)\rangle = \sum_i U^*_{\alpha i} |\nu_i\rangle e^{-i\phi_i}. \tag{2.3}$$

The probability of observing a neutrino of flavour eigenstate β from one which originated as flavour α can be calculated as,

$$P(\nu_\alpha \to \nu_\beta) = \left|\langle \nu_\beta | \nu_\alpha(t)\rangle\right|^2 = \sum_{i,j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} e^{-i(\phi_j - \phi_i)} \tag{2.4}$$

The ϕ_i term can be expressed in terms of the energy, E_i, and magnitude of the three momentum, p_i, of the neutrino, $\phi_i = E_i t - p_i x$. Therefore,

$$\phi_j - \phi_i = E_j t - E_i t - p_j x + p_i x. \tag{2.5}$$

For a relativistic particle, $E_i \gg m_i$,

$$p_i = \sqrt{E_i^2 - m_i^2} \approx E_i - \frac{m_i^2}{2E_i}. \tag{2.6}$$

Making the approximation that the neutrino mass eigenstates were created with the same energy and that $x = t = L$, where L is the distance travelled by the neutrino, Equation 2.5 then becomes

$$\phi_j - \phi_i = \frac{\Delta m_{ij}^2 L}{2E}, \tag{2.7}$$

where $\Delta m_{ij}^2 = m_i^2 - m_j^2$. This, teamed with further use of unitarity relations results in Equation 2.4 becoming

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4 \sum_{i>j} \mathbb{R} \left(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^* \right) \sin^2 \left(\frac{\Delta m_{ij}^2 L}{4E} \right) \tag{2.8}$$
$$+ (-) 2 \sum_{i>j} \mathbb{I} \left(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^* \right) \sin \left(\frac{\Delta m_{ij}^2 L}{2E} \right).$$

Where the negative sign is included for the oscillation probability of antineutrinos.

Typically, the PMNS matrix is parameterised into three mixing angles, a charge parity (CP) violating phase δ_{CP}, and two Majorana phases $\alpha_{1,2}$,

$$U = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}}_{\text{Atmospheric, Accelerator}} \underbrace{\begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta_{CP}} & 0 & c_{13} \end{pmatrix}}_{\text{Reactor, Accelerator}} \underbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Reactor, Solar}} \underbrace{\begin{pmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Majorana}}.$$
$$\tag{2.9}$$

Where $s_{ij} = \sin(\theta_{ij})$ and $c_{ij} = \cos(\theta_{ij})$. The mixing angles are often referred to by the style of experiment which best constrains these parameters; $(1,2)$ as "solar", $(2,3)$ as "atmospheric" and $(1,3)$ as "reactor". Many neutrino experiments aim to measure the PMNS parameters from a wide array of origins, as is the purpose of this thesis.

The Majorana phase containing matrix included within Equation 2.9 is only included for completeness. For the purposes of an oscillation analysis experiment, any term in this oscillation probability calculation containing this phase disappears due to taking the expectation value of the PMNS matrix.

A two flavour approximation can be attained when one assumes the third mass eigenstate is degenerate with another. As discussed in section 2.3, it is found that $\Delta m_{21}^2 \ll |\Delta m_{31}^2|$ such that this two flavour approximation is reasonable for understanding the features of the oscillation. In this two flavour case, the mixing matrix becomes

$$U_{2\text{ Flav.}} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \tag{2.10}$$

which culminates in the oscillation proabability,

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - (+)\sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right). \tag{2.11}$$

The oscillation probability is a sinusodial function depending upon the distance over which the neutrino propagates. The frequency and amplitude of oscillation is dependent upon the ratio of the $\Delta m^2/4E$ and $\sin^2 2\theta$, respectively. For more human-readable units, the maximum oscillation probability for a fixed value of θ is given at $L[km]/E[GeV] \sim 1.27/\Delta m^2$. It is this calculation that determines the best L/E value for a given experiment to be designed around for measurements of a specific value of Δm^2.

2.2.2. The MSW Effect

subsection 2.2.1 describes the theory of neutrino oscillation in a vacuum but beam neutrinos and atmospheric neutrinos originating from below the horizon propagate through matter in the Earth. The coherent scattering of neutrinos from a material target modifies the energy of the mass eigenstates resulting in a change to the oscillation probability. Notably, charged current scattering ($\nu_e + e^- \to \nu_e + e^-$, propagated by a W boson) only effects electron neutrinos compared to the neutral current scattering ($\nu_l + l^- \to \nu_l + l^-$, propagated by a Z^0 boson) which does not favour any neutrino flavour. In the two-flavour limit, the effective mixing parameter becomes

$$\sin^2(2\theta) \to \sin^2(2\theta_m) = \frac{\sin^2(2\theta)}{(A/\Delta m^2 - \cos(2\theta))^2 + \sin^2(2\theta)}, \tag{2.12}$$

where $A = 2\sqrt{2}G_F N_e E$ with N_e is the electron density of the medium and G_F is Fermi's constant. It is clear to see that there exists a value of $A = \Delta m^2 \cos(2\theta)$ for $\Delta m^2 > 0$ which results in a divergent mixing parameter. This is known as the Mikheyev-Smirnov-Wolfenstein (MSW) effect (or more clloquially, the matter resonance) which regenerates the electron neutrino component of the neutrino flux [13–15]. The density at which the resonance occurs is given by

$$N_e = \frac{\Delta m^2 \cos(2\theta)}{2\sqrt{2}G_F E} \tag{2.13}$$

At densities lower than this critical value, the oscillation probability will be be much closer to that from vacuum oscillation. As seen, the resonance occuring from the MSW effect depends on the sign of Δm^2. Therefore, any neutrino oscillation experiment which obsreves neutrinos which have propagated through matter can have some sensitivity to the ordering of the neutrino mass eigenstates.

For an experiment observing atmospheric neutrinos propagating through the Earth such as the studies presented in this thesis, a model of the Earth's density and layering is required. The model used within this analysis is the Preliminary Reference Earth Model (PREM) [16]. This model approximates the Earth as four layers of constant density as described in Table 2.1. The density measurements provided in this model are provided in terms of masss density, whereas neutrino oscillations are sensitive to the electron number density. Consequently, the chemical compositon of each layer multiplied by the mass density gives the relevent value for oscillation probabilities.

Layer	Outer Radius [km]	Density [g/cm^3]	Chemical composition (Z/A)
Inner Core	1220	13	0.468 ± 0.029
Outer Core	3480	11.3	0.468 ± 0.029
Lower Mantle	5701	5.0	0.497
Transition Zone	6371	3.3	0.497

Table 2.1.: Description of the four layers of the Earth invoked within the PREM model [16].

The beam oscillation probability in this thesis uses a baseline of 295km and density 2.6g/cm^3.

2.3. Neutrino Oscillation Measurements

As evidence of beyond standard model physics, the 2015 Nobel Prize in Physics was awarded to the Super-Kamiokande (SK) and Sudbury Neutrino Observatory (SNO) collaborations for the first definitive observation of solar and atmospheric neutrino oscillation. Since then, the field has seen a wide array of oscillation measurements from a variety of neutrino sources. As seen in subsection 2.2.1, the neutrino oscillation probability is dependent on the ratio of the propagation baseline, L, to the neutrino energy, E, which determines the flavour of neutrino oscillation a particular experiment is sensitive to.

As illustrated in Figure 2.1, neutrino energies span a wide range of energies. The least energetic neutrinos are from diffuse supernovae and terrestrial neutrinos at $O(1)$eV whereas the most energetic neutrinos originate from atmospheric and galactic neutrinos of $> O(1)$TeV.
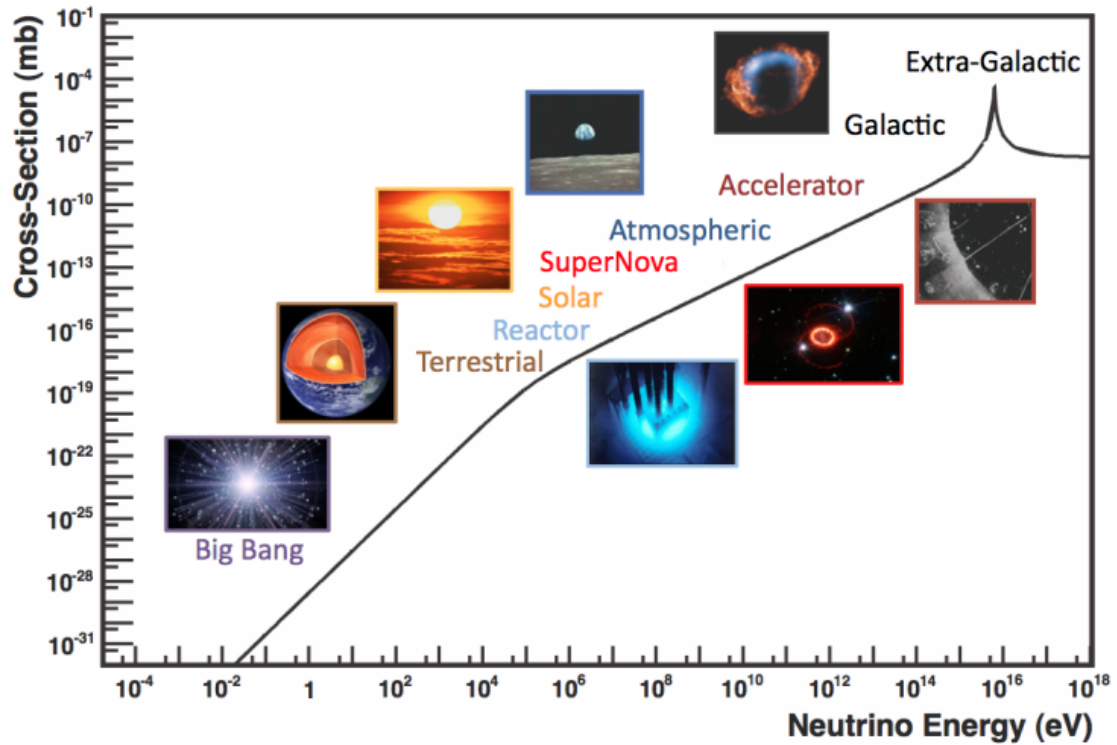

Figure 2.1.: The energy spectrum of neutrinos from various natural and man-made sources. Taken from [17]

2.3.1. Solar Neutrinos

Solar neutrinos are emitted from interaction chains contained within the fusion reaction at the center of the Sun. The solar neutrino flux given as a function of neutrino energy for different fusion and decay chains is illustrated in Figure 2.2 (From [18]). Whilst proton-proton fusion generates the largest flux of neutrinos, the neutrinos are of low energy and are difficult to reconstruct. Consequently most experiments focus on the neutrinos from the decay of ^8B (via $^8B \rightarrow^8 Be^* + e^+ + \nu_e$), which are higher energy.
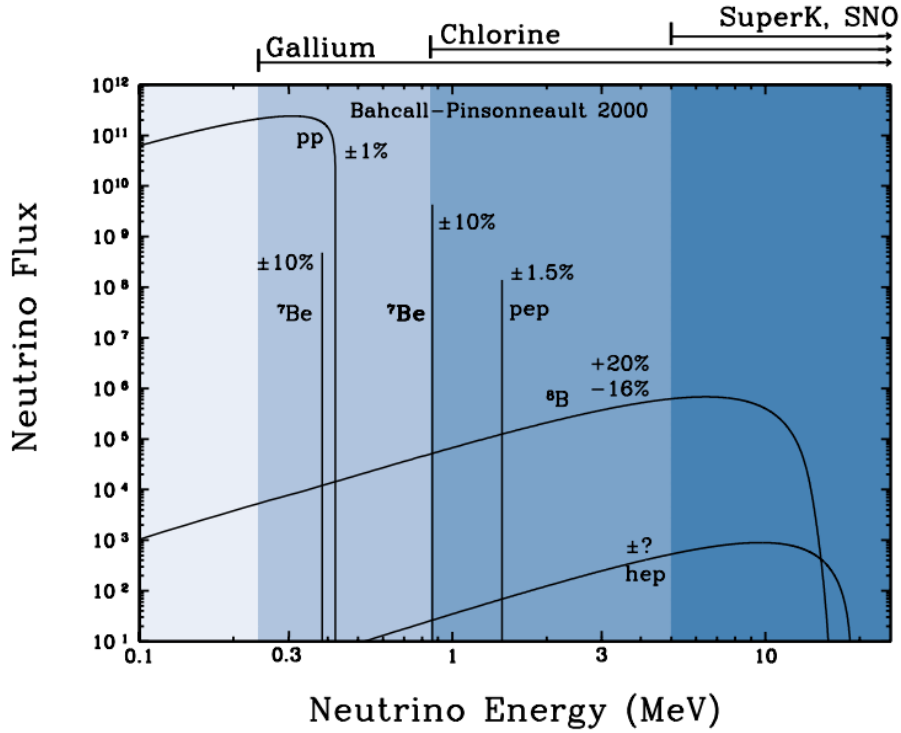
Figure 2.2.: The solar neutrino flux as a function of neutrino energy for various fusion reaction and decay chains as predicted by the Standard Solar Model. Taken from [18].

The first measurements of solar neutrinos observed a significant reduction in the event rate compared to predictions from the Standard Solar Model [19, 20]. The proposed solution to this "solar neutrino problem" was $\nu_e \leftrightarrow \nu_\mu$ oscillations in a precursory version of the PMNS model [21]. The Kamiokande [22], Gallex [23] and Sage [24] experiments confirmed the ~ 0.5 deficit of solar neutrinos.

The conclusive solution to the "solar neutrino problem" was determined by the SNO collaboration [25]. Using a deutrium water target to observe ^8B neutrinos, the event rate of charged current (CC), neutral current (NC) and elastic scattering (ES)

interactions (Given in Equation 2.14) measured. CC events can only occur for electron neutrinos, whereas the other interaction channels are agnostic to neutrino flavour (Although the ES reaction is more sensitive to electron neutrino interactions). This meant that there was a direct measurement of the ν_e and total neutrino flux. It was concluded that the CC and ES reaction rates were consistent with the deficit previouly observed but notably that the NC reaction rate was only consistent with the other interaction rates under the hypothesis of flavour transformation.

$$
\begin{aligned}
\nu_e + d &\to p + p + e^- & (CC) \\
\nu_x + d &\to p + n + \nu_x & (NC) \\
\nu_x + e^- &\to \nu_x + e^- & (ES)
\end{aligned}
\tag{2.14}
$$

Since then, many experiments have precisely measured the neutrino flux of different interaction chains within the sun [26–28]. The most recent measurement was that of CNO neutrinos which were recently observed with 5σ certainty by the Borexino collaboration. Future neutrino experiments aim to further the spectroscopic measurements of different fusion chains within the Sun [29–31]. Whilst not of direct focus, dark mattter experiments like DARWIN [32] will also be sensitive to the solar neutrinos emitted by the Sun.

2.3.2. Atmospheric Neutrinos

The interactions of primary cosmic ray protons (and heavier nuclei) in Earth's upper atmospheric generate showers of energetic hadrons, mostly pions and kaons, which when they decay produce a natural source of neutrinos spanning energies of MeV to TeV [33]. This decay is via

$$
\begin{aligned}
\pi^\pm &\to \mu^\pm + \overset{(-)}{\nu_\mu} \\
\mu^\pm &\to e^\pm + \overset{(-)}{\nu_\mu} + \overset{(-)}{\nu_e}
\end{aligned}
\tag{2.15}
$$

such that for a single pion decay, three neutrinos are produced. The atmopsheric neutrino flux energy spectra as predicted by the Bartol [34], Honda [35,36] and FLUKA

[37] models is illustrated in Figure 2.3. The flux distribution peaks at an energy of around 10GeV. The uncertainties associated with these models are dominated by the interaction modelling of kaon and pions as well as the primary cosmic flux.
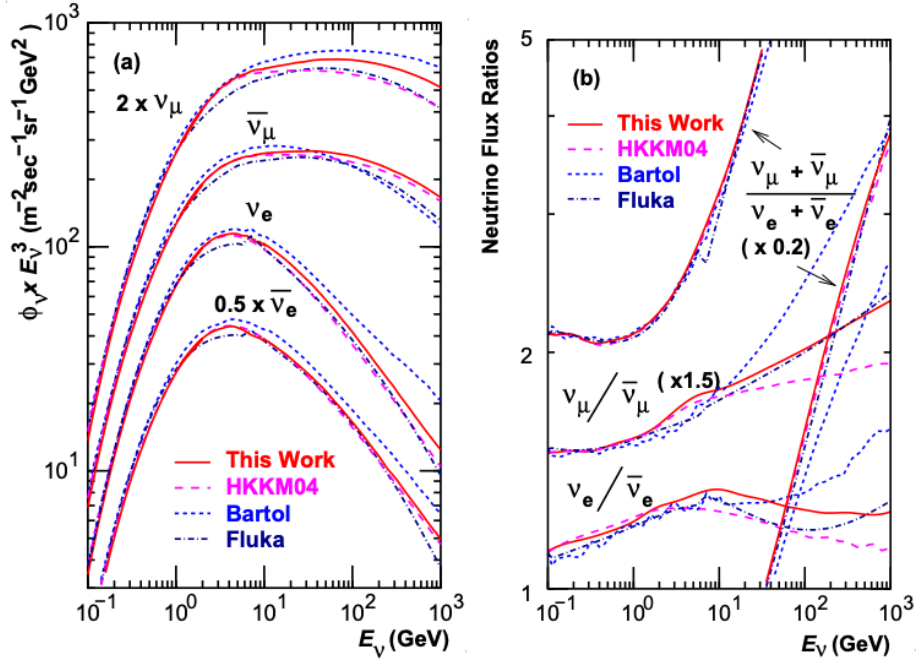

Figure 2.3.: Left panel: The atmospheric neutrino flux for different neutrino flavours as a function of neutrino energy which compares the 2007 Honda model ("This work") [35], the 2004 Honda model ("HKKM04") [36], the Bartol model [34] and the FLUKA model. Right panel: indicates the ratio of the flux models between each model. Taken from [35].

The oscillations present in atmospheric neutrinos still follow the same L/E probability presented in subsection 2.2.1, but unlike a fixed beam experiment, the baseline for each neutrino is dependent upon the zenith angle with respect to Z-axis of the detector as illustrated by Figure 2.4. Neutrinos coming from interactions taking place in the atmospheric above the detector ($\cos(\theta) = 1.0$) only travel the height of the atmosphere before being observed within the detector; a distance of $O(20)$km. However, neutrinos which are observed as coming from directly below the detector ($\cos(\theta) = -1.0$) have travelled $O(6 \times 10^3)$km from interactions in the atmosphere on the opposite side of the Earth. As discussed in subsection 2.2.2, any neutrino passing through the Earth is subject to matter effects. These effects are most notable for any neutrino which passes through the Earth's core ($\cos(\theta) < -0.45$).

Figure 2.5 highlights the neutrino flux as a function of zenith angle for different slices of neutrino energy. For medium to high energy neutrinos (and to a lesser degree
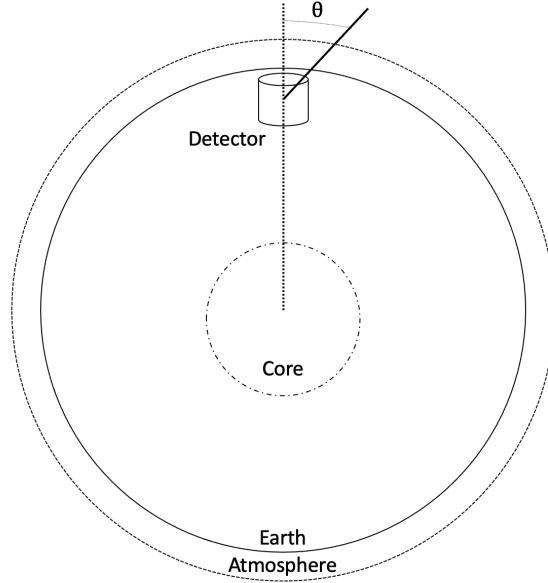
Figure 2.4.: A diagram illustrating the definition of zenith angle as used in the Super Kamiokande experiment [38].

for low energy neutrinos), the flux is approximately symmetric around $\cos(\theta) = 0$. To the accuracy of this approximation, the systematic uncertainties associated with atmospheric flux for comparing upward-going and down-going neutrino cancels. This allows the down-going events, which are mostly insensitive to oscillation probabilities due to the small L/E, to act as a unoscillated prediction (similar to a near detector in an accelerator neutrino experiment).

Precursory hints of atmospheric neutrinos were observed in the mid-1960s searching for $\overset{(-)}{\nu_\mu} + X \to X^* + \mu^\pm$ [39]. These experiments were located in deep underground laboratories to reduce the cosmic muon background. These were replaced with the IMB-3 [40] and Kamiokande [41] experiments which measured a deficit of muon neutrinos compared to electron neutrinos $R(\nu_\mu/\nu_e)$. Both experiments were found to have a consistent measurement, with $R(\nu_\mu/\nu_e) = 0.67 \pm 0.17$ and $R(\nu_\mu/\nu_e) = 0.60^{+0.07}_{-0.06} \pm 0.05$. Soudan-2 [42] determined similar measurements. Super-Kamiokande (SK) [38] extended this analysis by fitting oscillation parameters in $\nu_\mu \to \nu_\tau$ which found best fit parameters $\sin^2(2\theta) > 0.92$ and $1.5 \times 10^{-3} < \Delta m^2 < 3.4 \times 10^{-3} \text{eV}^2$.

Since then, atmospheric neutrino exeeriments have begun making precision measurements of the $\sin^2(\theta_{23})$ and $|\Delta m^2_{32}|$ oscillation parameters, and to a lesser extent the sign of Δm^2_{23} through the matter resonance present for any neutrinos passing through the Earth. Atmospheric neutrino oscillation is dominated by $P(\nu_\mu \to \nu_\tau)$,
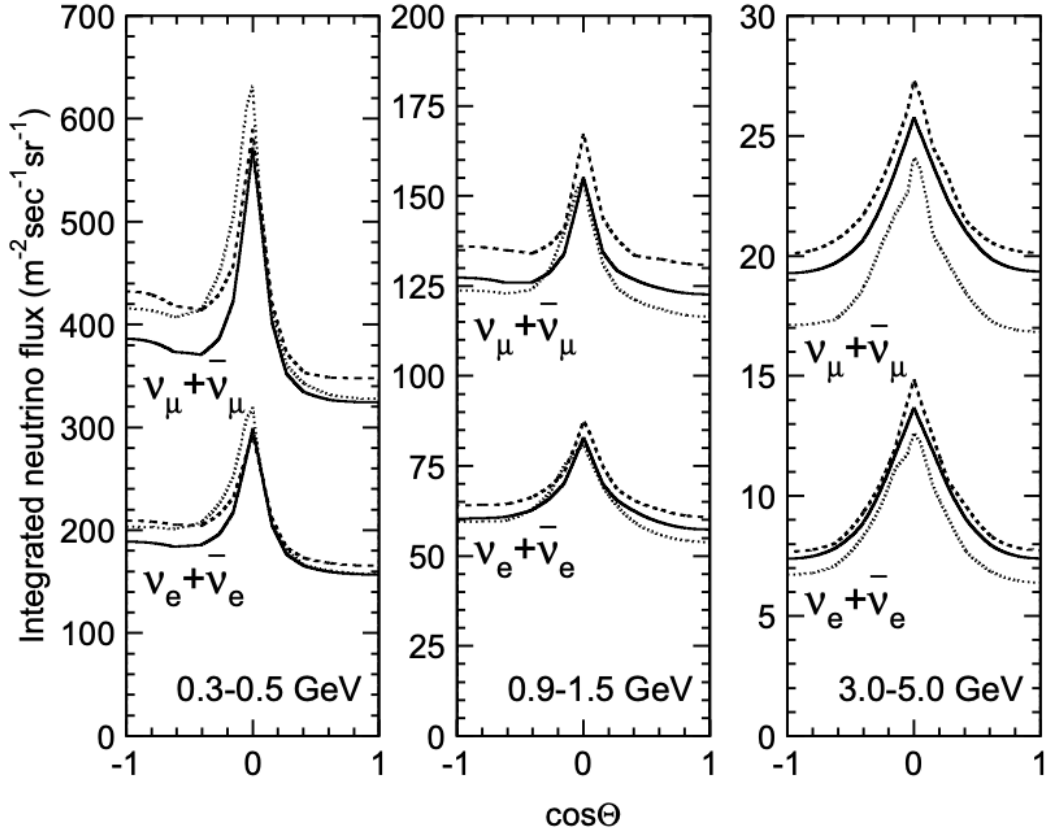
Figure 2.5.: Predictions of the summed neutrino and antineutrino flux for electron and muon neutrinos from the Bartol [34], Honda [35] and FLUKA [37] models as a function of zenith angle with respect to the detector. Left panel: $0.3 < E_\nu < 0.5$. Middle panel: $0.9 < E_\nu < 1.5$. Right panel: $3.0 < E_\nu < 5.0$. Taken from [38].

where SK observed a 4.6σ discovery of ν_τ appearance [43]. Figure 2.6 illustrates the current estimates on the atmospheric mixing parameters from a wide range of atmospheric and accelerator neutrino observatories.

2.3.3. Accelerator Neutrinos

The concept of using a man-made source of neutrinos, a "neutrino beam", was first realised in 1962 [50] and led to the first discovery that ν_e and ν_μ were in fact different particles. Since then, many large scale experiments have followed which all using the same fundamental concepts. A proton beam is aimed at a target, producing charged mesons which then decay to neutrinos and the associated leptons. The mesons can be sign-selected by the use of magenetic focussing horns to generate a neutrino or antineutrino beam. Absorbing material and the rock between the target and detector

Figure 2.6.: Constraints on the atmospheric oscillation parameters, $\sin^2(\theta_{23})$ and Δm_{23}^2, from atmospheric and long baseline experiments SK [44], T2K [45], NOνA [46], IceCube [47] and MINOS [48]. Figure taken from [49].
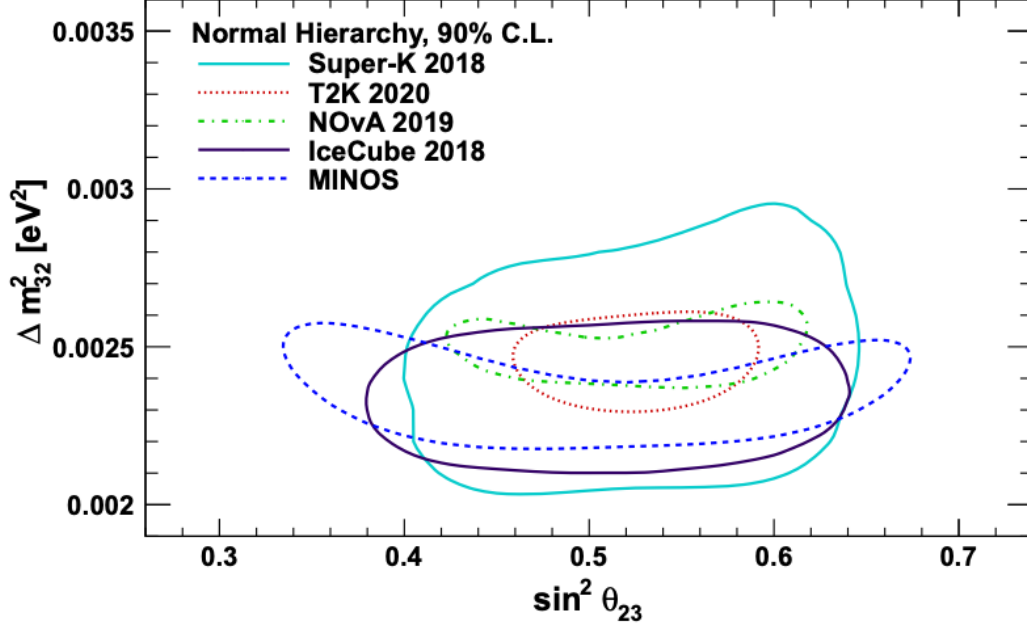
absorb all particles barring the neutrinos. Pions are the primary meson which decay and depending on the orientation of the focussing, a predominately muon neutrino beam is generated via $\pi^+ \to \mu^+ + \nu_\mu$ or $\pi^- \to \mu^- + \bar{\nu}_\mu$. Despite this, an approximately 10% "wrong-sign" neutrino background of ν_e is generated via the decay of muons or kaons.

As the inital proton beam can be tuned resulting in a tunable neutrino energy spectra, the advantage of these type of experiments is that they can be focused in on the oscillation dip presented by the L/E term in Equation 2.11 using the two flavour approximation. This tuning means that accelerator experiments are typically more sensitive to the atmospheric mixing parameters. However, the disadvantage compared to atmospheric neutrino experiments is that the baseline (or length of the neutrino beam) has to be shorter due to the lower flux. Consequently, there is typically less sensitivity to matter effects and the ordering of the neutrino mass eigenstates.

A neutrino experiment measures

$$R(\vec{x}) = \Phi(E_\nu) \times \sigma(E_\nu) \times \epsilon(\vec{x}) \times P(\nu_\alpha \to \nu_\beta), \qquad (2.16)$$

where $R(\vec{x})$ is the event rate of neutrinos at position $\vec{x}$, $\Phi(E_\nu)$ is the flux of neutrinos with energy E_ν, $\sigma(E_\nu)$ is the cross section of the interaction and $\epsilon(\vec{x})$ is the efficiency of the detector. Thus, in order to leverage the most out of an accelerator neutrino experiment, the flux and cross section systematics need to be constrained. This is typically done via the use of a "near detector", situated at a baseline of $O(1)$km which observes the unoscillated neutrino flux and constrains the parameters used within the flux and cross section model.

Long baseline experiments became the forerunners of precision measurements of oscillation parameters. MINOS [51] and K2K [52] confirmed the $\nu_\mu \rightarrow \nu_\mu$ oscillations seen in atmospsheric neutrino experiments by finding consistent mixing parameter values for $\sin^2(\theta_{23})$ and Δm^2_{23}. The current generation of accelerator neutrino experiments, T2K and NOνA extended this field by observing $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ and lead the sensitivity to atmospheric mixing parameters as seen in Figure 2.6 [53]. The two experiments differ in their peak neutrino energy, baseline and detection technique. The NOνA experiment is situated at a baseline of 810km from the NuMI beamline which delivers 2GeV neutrinos whereas T2K neutrinos are peaked around 0.6GeV and propagate 295km. The NOνA experiment also uses functionally identical detectors (near and far) which allows the approximately cancellation of detector systematics whereas T2K uses a plastic scintillator technique at the near detector and a water Cherenkov far detector. The future generation experiments DUNE [54] and Hyper-Kamiokande [55] will succeed these experiments as the high precision-era of neutrino oscillation parameter measurements develops.

Several anomalous results have been observed in the LSND [56] and MiniBooNE [57] detectors which were designed with purposefully short baselines. Parts of the neutrino community attributed these results to be oscillations induced by a fourth "sterile" neutrino [58] but several searches in other experiments, MicroBooNE [59] and KARMEN [60] found no hints of additional neutrino species.

2.3.4. Reactor Neutrinos

As illustrated in the first discovery of neutrinos (section 2.1), nuclear reactors are a very powerful man-made source of electron antineutrinos. For reactors which use low-enriched uranium ^{235}U as fuel, the antineutrino flux is dominated by the β-decay fisson of ^{235}U, ^{238}U, ^{239}Pu and ^{241}Pu [61] as illustrated in Figure 2.7.
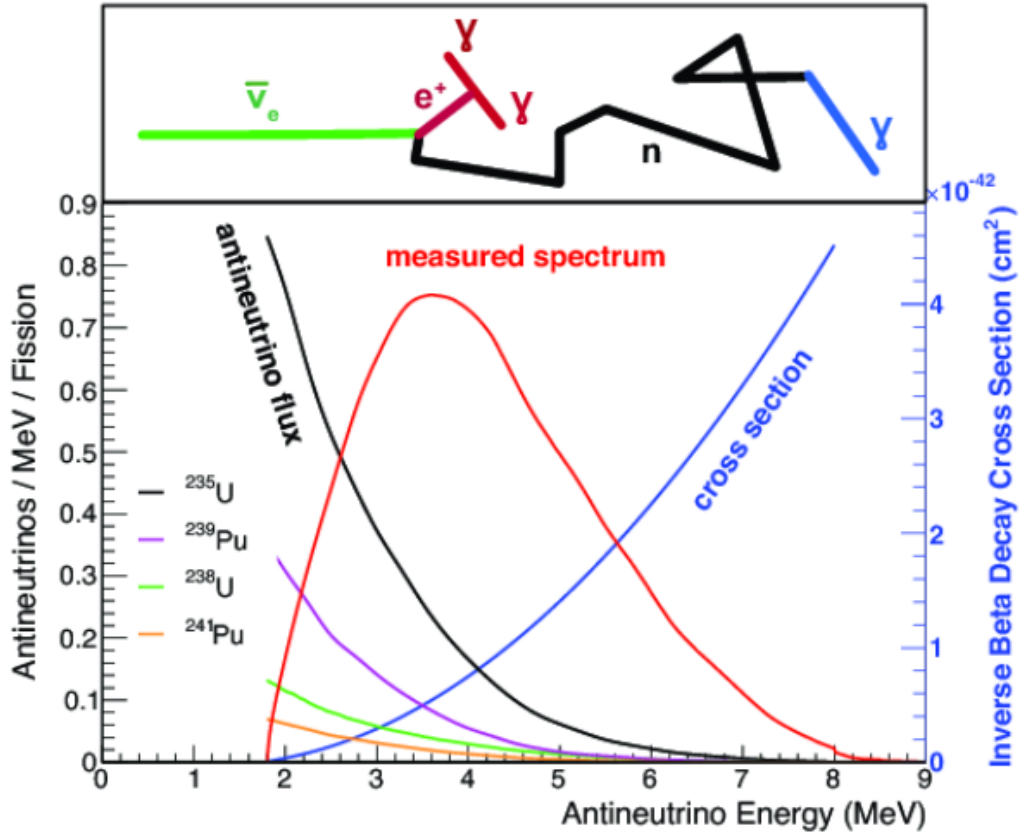
Figure 2.7.: Reactor electron antineutrino fluxes for ^{235}U (Black), ^{238}U (Green), ^{239}Pu (Purple) and ^{241}Pu (Orange) isotopes. The inverse β-decay cross section is given in Blue and corresponding measurable neutrino spectrum is highlighted in Red. Top panel: Schematic of Inverse β-decay interaction. Taken from [62].

Due to their low energy, reactor electron antineutrinos interact via the inverse β-decay (IBD) interaction. The typical signature contains two signals delayed by $O(200)\mu$s; firstly the prompt photons from positron annihilation, and secondly the photons emitted ($E_{tot}^{\gamma} = 2.2$MeV) from de-excitation after neutron capture on hydrogen. Recently, SK included gadolinium dopanants into the ultra-pure water to increase the energy released from the photon cascade to ~ 8MeV and reduce the time of the delayed signal to $\sim 28\mu$s. This improves the detector's ability to distinguish between background and signal events [63].

There are many short baseline experiments (L $\sim O(1)$km) which have measured the $\sin^2(\theta_{13})$ and Δm_{23}^2 oscillation parameters. Daya Bay [64], RENO [65] and Double Chooz [66] have all provided precise measurements, with the first discovery of a non-zero θ_{13} made by Daya Bay and RENO. The constraints on $\sin^2(\theta_{13})$ by the reactor experiments lead the field and are often used as external inputs to accelerator neutrino

experiments to improve their sensitivity to δ_{CP} and mass hierarchy determination. One curiosity of these short baseline reactor experiments is the '5MeV excess' [67]. First observed in 2014 [68, 69], all three experiments listed observed a shape excess in events around $E_\nu \sim 5\text{MeV}$. The reason behind this excess is speculated to be either oscillations to sterile neutrinos or a fault in the Huber-Mueller model [70]. At this time, the latter is favoured as Daya Bay [71] observed substantial evidence (4.0σ) of correlation between the excess and the ^{235}U electron antineutrino flux. JUNO-TAO [72], a small collaboration within the larger JUNO experiment, is a next generation reactor experiment which aims to precisely measure the isotopic antineutrino yields from the different fission chains. Alongside this, it aims to explain the '5MeV excess' by conducting a search for sterile neutrinos with a mass scale of around 1eV.

Kamland [73] is the only experiment to have observed reactor neutrinos using a long baseline (flux weighted averaged baseline of $L \sim 180\text{km}$) which allows it to have sensitivity to Δm^2_{12}. Combined with the SK solar neutrino experiment, the combined analysis puts the most stringent constraint on Δm^2_{12} [74] which is used as a prior uncertainty with accelerator neutrino experiments.

2.4. Measurement Summary

Chapter 3

T2K and SK Experiment Overview

This needs some references!

3.1. The Super-Kamiokande Experiment

The SK experiment began taking data in 1996 and has had many modifications throughout its lifespan. There has been six defined periods of data taking as noted in Table 3.1. Between the SK-I and SK-II periods, a significant proportion of the PMTs were damaged during maintainence. Those that survived were equally distributed throughout the detector in the SK-II era, which resulted in a reduced photo-coverage. From SK-III onwards, significant repair efforts were made such that the full suite of PMTs were operational. Before the start of SK-IV, the data acquisition and electronic systems were upgraded. Between SK-IV and SK-V, a significant effort was placed into tank open maintainence and repair/replacement of defective PMTs, a task in which the author of this thesis was required for. Consequently, the detector conditions were significantly different between the two operational periods. SK-VI saw the start of the 0.01% gadonlium doped water being introduced into the tank and the detector continues to operate to this day. Efforts are currently underway to increase the gadolinium concentrate in the likely start of the next SK period.

Period	Start Date	End Date	Live-time (days)
I	April 1996	July 2001	1489.19
II	October 2002	October 2005	798.59
III	July 2006	September 2008	518.08
IV	September 2008	May 2018	3244.4
V	January 2019	July 2020	461.02
VI	July 2020	Ongoing	583.3

Table 3.1.: The various SK periods and respective live-time. The SK-VI live-time is calculated until 1st April 2022.

3.1.1. The SK Detector

The basic structure of the Super-Kamiokande (SK) detector is a cylindrical tank with diameter 39.3m and hieght 41.1m filled with ultrapure water. A diagram of the signficant parts of the SK detector is illustrated in Figure 3.1. The SK detector is situated in the Kamioka mine in Gifu, Japan. The cavern in which the detector is located is underground with roughly 1km rock overburden (2.7km water equivalent overburden). At this depth, the rate of cosmic ray muons is significantly decreased to a value of $\approx$ 2Hz, compared to a surface-level detector. The top of the tank is covered with stainless steel which is designed as a working platform for maintainence, calibration and location for high voltage and data acquisition electronics.
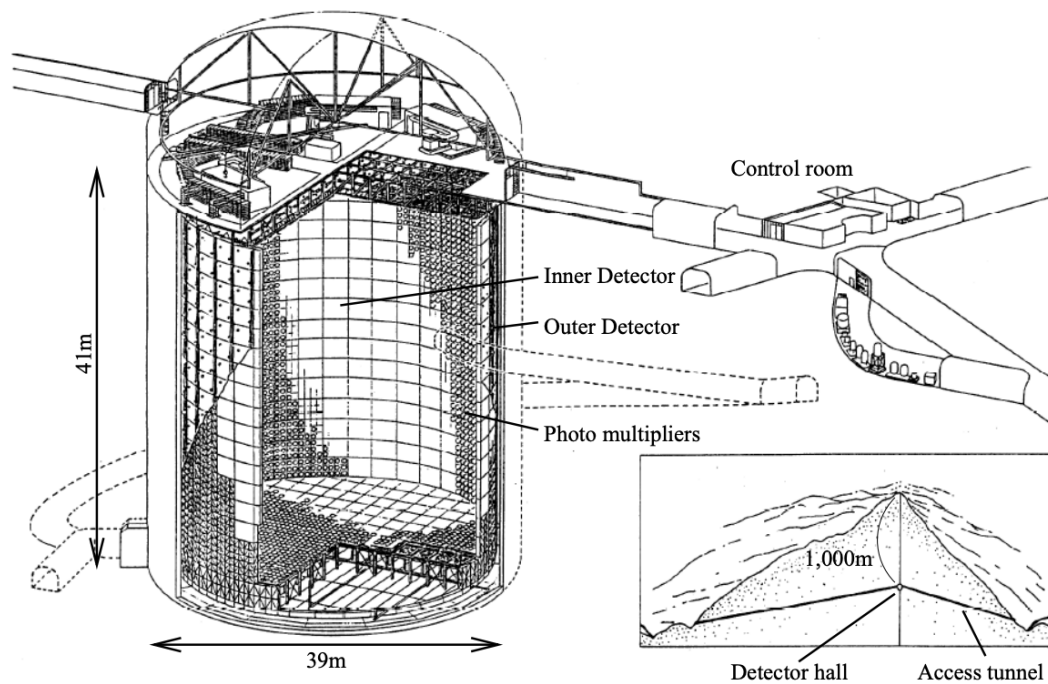
Figure 3.1.: A schematic diagram of the Super-Kamiokande Detector. Taken from [75].

A smaller cylindrical structure (36.2m diameter, 33.8m hieght) is situated inside the tank, with an approximate 2m gap between this structure and the outer tank wall. The purpose of this structure is to support the photomultiplier tubes (PMTs). The volume inside and outside the support structure are referred to as the inner detector (ID) and outer detector (OD), respectively. The ID and OD are instrumented by 11129 50cm and 1885 20cm PMTs respectively. The inner detector contains a 32kton volume of water, although many analyses performed at SK use a "fiducial volume" defined by

the volume of water inside the ID excluding some distance to the ID wall. This reduces the volume of detector which is sensitive to neutrino events but reduces radioactive backgrounds and allows for better constraints on the reconstruction systematics. The nominal fiducial volume is defined as the area contained inside 2m from the ID wall for a total of 22.5kton water.

The two regions of the detector (ID and OD) are optically separated with opaque black plastic. The purpose of this is to deteremine whether a track entered or exited the ID. This allows cosmic muon rays and partially contained events to be tagged and separated from neutrino events entirely contained within the detector's sensitive region. This black plastic is also used to cover the area between the ID PMTs to reduce photon reflection from the ID walls. Contrary to this, the OD is lined with a reflective material (called Tyvek) to allow photons to reflect around inside the OD until collected by one of the PMTs. Furthermore, each OD PMT is backed with 50×50cm plates of wavelength shifting acrylic which increases the efficiency of light collection.

In the SK-IV data taking period, the photocathode coverage of the detector, or the fraction of the ID wall instrumented with PMTs, is $\sim 40\%$. The PMTs have a quantum efficiency (the ratio of detected electrons to incident photons) of $\sim 21\%$. The proportion of photoelectrons which produce a signal in the dynode of a PMT, termed the collection efficiency, is $> 70\%$. One disadvantage of using PMTs as the detection media is that the Earth's geomagnetic field can modify their response. Therefore, a set of compensation coils is built around the inner surface of the detector to mitigate this effect.

PMT Specifics from Takenaka-san's thesis

As mentioned, the SK detector is filled with ultrapure water, which in a perfect world would contain no impurities. However bacteria and organic compounds can significantly degrade the water quality, thus decreasing the attenuation length thus reducing the total number of photons which could be detected by the PMTs. To combat this, a sophisticated water treatment system has been developed for SK. UV lights, mechanical filters and membrane degasifiers are used to reduce the bacteria, suspended particulates and radioactive materials from the water. The flow of water within the tank is also critical as it can remove stagnant bacterial growth or build up of dust on the surfaces within the tank. Gravity drifts impurities in the water towards the bottom of the tank which if left uncontrolled, can create asymetric water conditions between the top and bottom of the tank. Typically, the water entering the tank is

cooled below the ambient temperature of the tank to control convection and inhibit bacteria growth. Furthermore, the dark noise hits within PMTs is sensitive to the PMT temperature so controlling the temperature gradients within the tank is beneficial for stable measurements.

SK-VI is the first phase of the SK experiment to use gadonlium dopants within the ultrapure water. As such, the SK water system had to be replaced to avoid removing the gadolium concentrate from the ultrapure water. For an inverse β-decay (IBD) interaction in a water target, the emitted neutron is thermally captured on hydrogen. This process releases 2.2MeV γ rays which are difficult to detect due to Compton scattered elecrtons from a γ ray of this energy is very close to the Cherenkov threshold, limiting the number of photons produced. Thermal capture of neutrons on gadolinium generates γ rays with higher energy meaning they are more easily detected. SK-VI has 0.01% Gd loading (0.02% gadolinium sulphate by mass) which causes $\approx 50\%$ of neutrons emitted by IBD to be captured on gadolinium. Whilst predominantly useful for low energy analyses, Gd loading allows $\nu/\bar{\nu}$ separation for atmopsheric neutrino event selections [76]. Efforts are currently in place to increase the gadolinium concentrate to 0.03% for $\approx 75\%$ neutron capture efficiency on gadonlium. The final stage of loading targets 0.1% concentrate.

More details of water purification and air purification details

3.1.2. Calibration

The calibration of the SK detector is documented in [77] and summarised below. The analysis presented within this thesis is dependent upon 'high energy events' (Charged particles with $O(> 100)$MeV momenta). These are events which are expected to generate a larger number photons such that each PMT will be hit with multiple photons. The reconstruction of these events depends upon the charge deposited within each PMT and the timing response of each individual PMT. Therefore, the most relevent calibration techniques to this thesis are outlined.

Before installation, 420 PMTs were calibrated to have identical charge response and then distributed throughout the tank in a cross-shape pattern (As illustrated by Figure 3.2). These are used as a standardised measure for the rest of the PMTs installed at similar geometric positions within SK to be calibrated against. This allows each PMT to have it's high voltage set so that all PMTs give the same signal for an identical

light collection. To perform this calibration, a xenon lamp is located at the center of the SK tank which flashes uniform light at 1Hz. This allows for geometrical effects, water quality variation and timing effects to be measured in-situ throughout normal data taking periods.
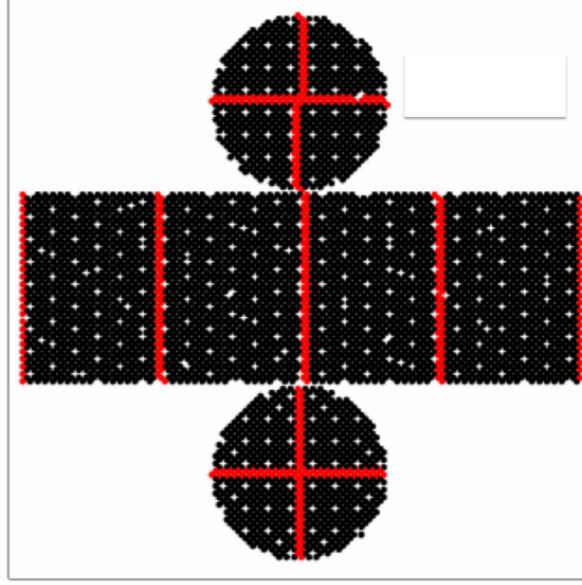

Figure 3.2.: The location of "standard PMTs" (red) inside the SK detector. Taken from [77].

When specifically performing calibration of the detector (in out-of-data taking mode), the water in the tank was circulated to avoid top/bottom asymmetric water quality. Any non-uniformity within the tank significantly effects the PMT hit probability through scattering or absorption. This becomes a dominant effect for the very low intensity light sources discussed later which are designed such that only one photon is incident upon a given PMT.

The "gain" of a PMT is defined as the ratio of total charge of the signal produced compared to the charge of photoelectrons emitted by the photocathodes within the PMT. To calibrate the signal of each PMT, the "relative" and "absolute" gain values are measured. The relative gain is the variation of gain among each of the PMTs whereas the absolute gain is the average gain of all PMTs.

The relative gain is calibrated as follows. A laser light is used to generate measurements under two conditions; a high-intensity flash which illuminates every PMT with a sufficient number of photons, and a low-intensity flash in which only a small number of PMTs collect light. The first measurement creates an average charge, $Q_{obs}(i)$ on PMT i, whereas the second measurement ensures that each hit PMT only generates

a single photoelectron. For the low intensity measurement, the number of times each PMT records a charge larger than some threshold value, $N_{obs}(i)$ is counted. [77] indicates that the specified threshold value does not induce any significant biases in the measurement of gain. The values measured can be expressed as

$$
\begin{aligned}
Q_{obs}(i) &\propto I_H \times f(i) \times \epsilon(i) \times G(i), \\
N_{obs}(i) &\propto I_L \times f(i) \times \epsilon(i),
\end{aligned}
\tag{3.1}
$$

Where I_H and I_L is the intensity of the high and low flashes, $f(i)$ is the acceptance efficiency of the ith PMT, $\epsilon(i)$ is the product of the quantum and collection efficiency of the ith PMT and $G(i)$ is the gain of the ith PMT. The relative gain for each PMT can determined by taking the ratio of these quantities.

The absolute gain calibration is performed by observing fixed energy γ-rays of $E_\gamma \approx 9\text{MeV}$ emitted isotropically from neutron capture on a NiCf source situated at the centre of the detector. This generates a photon yeild of about 0.004 photoelectrons/PMT/event, meaning that $> 99\%$ of PMT signals are generated from a single photoelectron. A charge distribution is generated by performing this calibration over all PMTs, and the average value of this distribution is taken to be the absolute gain value. It has been found that the absolute gain increases as a function of time despite there being no known reason for this behaviour.

As mentioned in subsection 3.1.1, the average quantum and collection efficiency for the SK detector is $\sim 21\%$ and $> 70\%$ respectively. However, these values do differ between each PMT. Consequently, the NiCf source is also used to calibrate the "quantum $\times$ collection" efficiency (denoted "QE") value of each PMT. This value is calculated as calibrating the signal generated by one incident photon is of more importance than the individual efficiency of photoelectron emission or detection on the dynode. The NiCf low intensity source is used as the PMT hit probability is proportional to the QE ($N_{obs}(i) \propto \epsilon(i)$ in Equation 3.1). A Monte Carlo prediction which includes photon absorption, scattering and reflection is made to estimate the number of photons incident on each PMT and the ratio of the number of predicted to observed hits is calculated. The difference is attributed to the QE efficiency of that PMT. This technique is extended to calculate the relative QE efficiency by normalizing the average of all PMTs which removes the dependence of the light intensity.

Due to differering cable lengths and readout electronics, the timing response between a photon hitting the PMT and the signal being captured by the data acquisition can be different between each PMT. Due to threshold triggers (Described in subsection 3.1.3), the time at which a pulse reaches a threshold is dependent upon the size of the pulse. This is known as the 'time-walk' effect and also needs to be accounted for in each PMT. To calibrate the timing response, a pulse of light with width 0.2nsec is emitted into the detector through a diffuser and observed by an external PMT which is used to measure the pulse start time. Two dimensional distributions of time and pulse height (or charge) are made for each PMT and are used to calibrate the timing response. This is performed in-situ whilst data taking with the light source pulsing at 0.03Hz.

The top/bottom water quality asymmetry is measured using the NiCf calibration data and cross-referencing these results to the "standard PMTs". The water attenuation length is continously measured by the rate of vertically-downgoing cosmic-ray muons which enter via the top of the tank.

Light absorption and scattering details

3.1.3. Data Acquisition and Triggering

Dark noise is the phenomenon where a PMT registers a pulse that is consistent with a single photoelectron emitted from photon detection despite being the PMT being in complete darkness. This is predominately caused by two processes. Firstly there is intrinsic dark noise which is where photoelectrons gain enough thermal energy to be emitted from the photocathode, and secondly, the radioative decay of contaminants inside the structure of the PMT which mimics the same effect. Typical dark noise rate for PMTs used within SK are $O(3)$kHz [78] which equates to about 12 dark noise hits per 220ns [79]. Link to Frank-Tamm equation. Which wavelength photons are PMTs adapted for

The maximum distance that a photon can travel within the SK detector is ~ 200ns [80]. The earlier data acquisition and triggering are documented in [80, 81]. The analysis presented in this thesis uses the SK-IV period of the SK experiment so only the relevant points of the data acquisition and triggering systems will be discussed.

Before the SK-IV period started, the existing front-end electronics were replaced with "QTC-Based Electrons with Ethernet, QBEE" systems [82]. When the QBEE

observes a signal above a 1/4 photoelectron threshold, the charge-to-time (QTC) converter generates a rectangular pulse. The start of the rectangular pulse indicates the time at which the analogue photoelectron signal was recieved and the width of the pulse indicates the total charge intergrated throughout the signal. This is then digitized by time-to-digital converters and sent to the "front-end" PCs. The digitized signal from every QBEE is then chronologically order and sent to the "merger" PCs. It is the merger PCs which apply the software trigger and send any 'interesting' events the "organizer" PC. This sorts the data stream of multiple merger PCs into chrolonolgically ordered events in time which is then saved to disk. The schematic of data flow from PMTs to disk is illustrated in Figure 3.3.
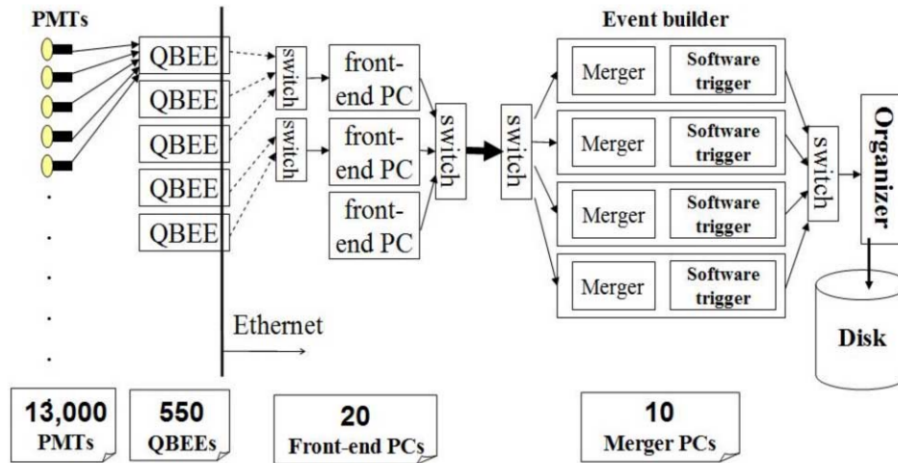

Figure 3.3.: Taken from [83].

The software trigger applied in SK-IV replaces the hardware trigger applied in SK-I to SK-III. This operates by determining the number of PMT hits within a 200ns sliding window, N_{200}, which coincides with the maximum time that a Cherenkov photon would take to traverse the length of the SK tank. For lower energy events which generate fewer photons, this technique is usful for eliminating background processes like dark noise and radioactive decay which would be expected to separated in time. When the value of N_{200} exceeds some threshold, a software trigger is issued. There are several trigger thresholds used within the SK-IV period which are detailed in Table 3.2. If one of these thresholds is met, the PMT hits within an extended time window is also readout and saved to disk. The length of the time window is also included in the tabulated threshold conditions. In the special case of an event which exceeds the SHE trigger but does not exceed the OD trigger, the AFT trigger looks for delayed coincidences of 2.2MeV gamma rays emitted from neutron capture in a $535\mu s$

window after the SHE trigger. A similar but more complex "Wideband Intelligent Trigger (WIT)" has been deployed and is described in [79].

Trigger	Acronym	Condition	Extended time window (μs)
Super Low Energy	SLE	>34/31 hits	1.3
Low Energy	LE	>47 hits	40
High Energy	HE	>50 hits	40
Super High Energy	SHE	>70/58 hits	40
Outer Detector	OD	>22 hits in OD	N/A

Table 3.2.: The trigger thresholds and extended time windows saved around an event which were utilised throughout the SK-IV period. The exact thresholds can change and the values listed here represent the thresholds at the start and end of the SK-IV period.

3.1.4. Cherenkov Radiation

Cherenkov light is emitted from any highly energetic charged particle travelling with relativistic velocity β greater than the local speed of light in a medium with refractive index $n > 1.0$. This occurs due to the charged particle exciting the polarised media, which de-excites via photon emission. From Huygen's principal, the emitted waves propagate outwards but only generate coherent wavefronts when the charged particle moves faster than the phase velocity of that media. Consequently, Cherenkov light is formed at the surface of a cone with characteristic pitch angle,

$$\cos(\theta) = \frac{1}{\beta n}. \tag{3.2}$$

Consequently, the Cherenkov momentum threshold (where the requirement is $\beta = 1/n$) is dependent upon the mass, m, of the charged particle moving through the media,

$$P_{thres} = \frac{m}{\sqrt{n^2 - 1}} \tag{3.3}$$

For water, where $n = 1.33$, the Cherenkov threshold momentum and energy for various particles is given in Table 3.3. For completeness, γ-rays are directed indirectly via the combination of photons generated by compton scattering and pair production. The threshold for detection in the SK detector is higher than the threshold for photon production. This is due to the fact that the attenuation of photons in the water means that typically $\sim 75\%$ of Cherenkov photons reach the ID PMTs. Then the collection and quantum efficiencies described in subsection 3.1.1 results in the number of detected photons being lower than the number of photons which reach the PMTs.

Particle	Threshold Momentum (MeV)	Threshold Energy (MeV)
Electron	0.5828	0.7751
Muon	120.5	160.3
Pion	159.2	211.7
Proton	1070.0	1423.1

Table 3.3.: The threshold momentum and energy for a particle to generate Cherenkov light in ultrapure water, as calculated in Equation 3.2 in ultrapure water where $n = 1.33$.

Typical 43 cone Frank-Tamm equation - Number of photons: (2*pi/137)*(1-(1+(105/444)**2)/1.33**2 1e7/700) approx equal to 352 photons per cm for a muon of momentum 444MeV Plot of muon momentum showing difference between Cherenkov threshold and detected threshold

3.2. The Tokai to Kamiokande Experiment

3.2.1. The Neutrino Beam

3.2.2. The Near Detector at 280m

3.2.3. The Interactive Neutrino GRID

Chapter 4

Bayesian Statistics Implemented Through Markov Chain Monte Carlo Techniques

The analysis throughout this thesis is based upon a Bayesian oscillation analysis. To extract the oscillation parameters, a Markov Chain Monte Carlo (MCMC) method is used. This chapter explains the theory of how parameter estimates can be determined using this technique. Referenece MCMC Literature.

The oscillation parameter determination presented within this thesis is built upon a simultaneous fit to near detector, far detector beam and atmospheric neutrino data. In total, there are four oscillation parameters of interest ($\sin^2(\theta_{23})$, $\sin^2(\theta_{13})$, Δm_{23}^2 and δ_{CP}), two oscillation parameters which this study will not be sensitive to ($\sin^2(\theta_{12})$, Δm_{12}^2) and many nuisance parameters that control the systematic uncertainty models invoked within this study. The systematic uncertainties can be grouped into cateogries depending on how they are defined; 574 bin-normalisations due to the near detector response, 45 bin-normalisations to describe the far detector response to neutrino beam events, 27 parameters to describe the detector response to atmospheric neutrino events, 100 to model the bin-normalisation due to beam flux uncertanties, 18 which model the atmospheric flux uncertainties, and 87 to describe the correlated cross section model. An alternative parameterisation, where the far detector response is correlated between the beam and atmospheric samples, replaces the bin-normalisation parameters with 224 shift and smear systematics. Section SystChap describes the systematic model in more depth.

The MCMC technique generates a multi-dimensional probability distribution across all of the model parameters used in the fit. To determine the parameter estimate of a single parameter, this multi-dimensional object is integrated over all other parameters. This process is called Marginalisation and is further described in subsection 4.3.1. Monte Carlo techniques approximate the probability distribution of each parameter in the limit of generating infinite samples. As ever, generating a large number of samples is time and resource dependent and as such the MCMC technique utilised within this analysis relies upon the powerful Metropolois-Hastings algorithm which

reduces the required number of samples to sufficiently sample the parameter space. This technique is described in further detail in subsection 4.2.1.

4.1. Bayesian Statistics

According to Bayesian Inference, observables and parameters of a statistical model are treated on an equal footing. To estimate model parameters $\vec{\theta}$ and data D, one needs to define the joint probability distribution $P(D|\vec{\theta})$ which can be described as the prior distribution for model parameters $P(\vec{\theta})$ and the likelihood of the data given the model parameters $P(D|\vec{\theta})$,

$$P(D,\vec{\theta}) = P(D|\vec{\theta})P(\vec{\theta}). \tag{4.1}$$

The prior distribution, $P(\vec{\theta})$, describes all previous knowledge about the parameters within the model. For example, if the risk of developing health problems is known to increase with age, the prior distribution would describe the increase. For the purpose of this analysis, the prior distribution are typically the best-fit values taken from external data measurements with a Gaussian uncertainty. The prior distribution can also contain correlations between model parameters. In an analysis using Monte Carlo techniques, the likelihood of measuring some data assuming some set of model parameters is calculated by comparing the Monte Carlo prediction generated at that particular set of model parameters to the data.

It is parameter estimation that is important for this analysis and as such, we apply Bayes' theorem Ref to Bayes' theorem to calculate the probability for each parameter to have a certain value given the observed data $P(\vec{\theta}|D)$, known as the posterior distribution (often just the posterior). This can be expressed as

$$P(\vec{\theta}|D) = \frac{P(D|\vec{\theta})P(\vec{\theta})}{\int P(D|\vec{\theta})P(\vec{\theta})d\vec{\theta}}. \tag{4.2}$$

The demoninator in Equation 4.2 is the integral of the joint probability distribution over all values of all parameters used within the fit. For brevity, we say that the posterior distribution is

$$P(\vec{\theta}|D)\alpha P(D|\vec{\theta})P(\vec{\theta}). \tag{4.3}$$

In subsection 4.3.1, we see that for the cases used within this analysis, it is reasonable to know the posterior to some normalisation constant.

4.2. Monte Carlo Simulation

Monte Carlo techniques are used to numerically solve a complex problem. These techniques all rely on generating samples from a desired distribution and in the limit of high number of samples, approximate the properties of these samples as the properties of the distribution. A typical example of a simple use of Monte Carlo techniques is to calculate the area underneath a curve. For example, take the problem of calculating the area under a straight line with gradient $M = 0.4$ and constant $C = 1.0$. Analytically, one can calculate the area under the line is equal to 3 units for $0 \leq x \leq 10$. Using Monte Carlo techniques, one can calculate the area under this line by throwing many random values for the x and y components of each sample and then calculating whether that point falls below the line. The area can then be calculated by the ratio of points below the line to the total number of samples thrown multiplied by the total area of which samples were scattered. The study shown in Figure 4.1 highlights this technqiue and finds the area under the curve to be 29.9 compared to an analytical solution of 30.0. The deviation of the numerical to analytical solution can be attributed to the number of samples used within the study. The accuracy of the approximation in which the properties of the Monte Carlo samples replicate those of the desired distribution is dependent on the number of samples used. Replicating this study with differing number of Monte Carlo samples used in each study (As shown in Figure 4.2) highlights how the Monte Carlo techniques are only accurate in the limit of high number of samples. Whilst this example has an analytical solution, these techniques are just as applicable to complex solutions. Clearly though, any numerical solution is only as useful as its efficiency. As discussed, the accuracy of the Monte

Carlo technique is dependent upon the number of samples the problem is evaluated at. Furthermore if the positions at which the samples are evaluated are not cleverly picked, the efficiency of the Monte Carlo technique significantly drops. Given the example in Figure 4.1, if the region in which the samples are scattered significantly extends passed the region of interest, many calculatations will be calculated but do not add to the ability of the Monte Carlo technique to achieve the correct result. For instance, any sample evaluated at a $y \geq 5$ could be removed without affecting the final result. This does bring in an aspect of the 'chicken and egg' problem in that to achieve efficient sampling, one needs to know the distribution beforehand.
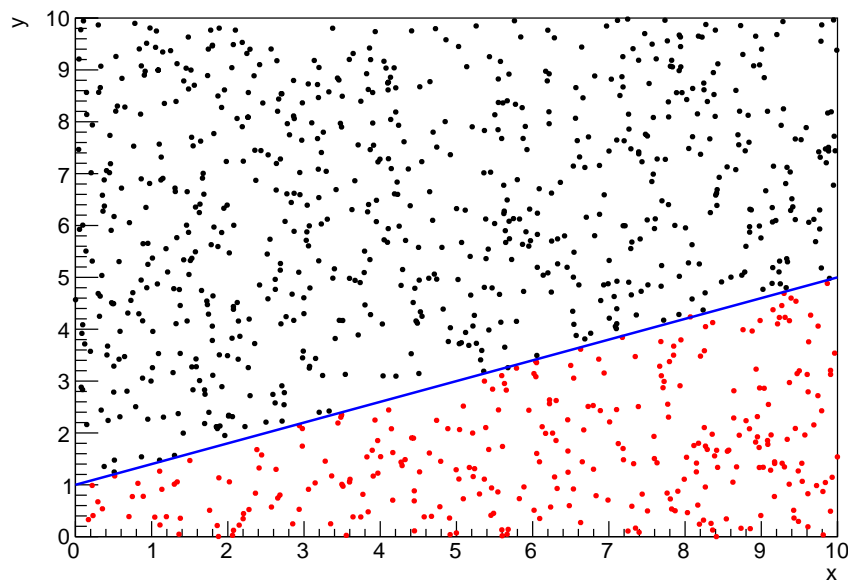

Figure 4.1.: Example of using Monte Carlo techniques to find the area under the blue line. The gradient and intercept of the line are 0.4 and 1.0 respectively. The area found to be under the curve using one thousand samples is 29.9.

4.2.1. Markov Chain Monte Carlo

This analysis utilises a multi-dimensional probability distribution, with some dimensions being significantly more constrained than others (whether that be from prior knowledge of parameter distributions from external data or un-physical regions in which parameters can not exist). Consequently, the Monte-Carlo techniques used need to be as efficient as possible. For this analysis, the Markov Chain Monte Carlo (MCMC) technique is choosen. An MCMC technique is a Monte Carlo technqiue
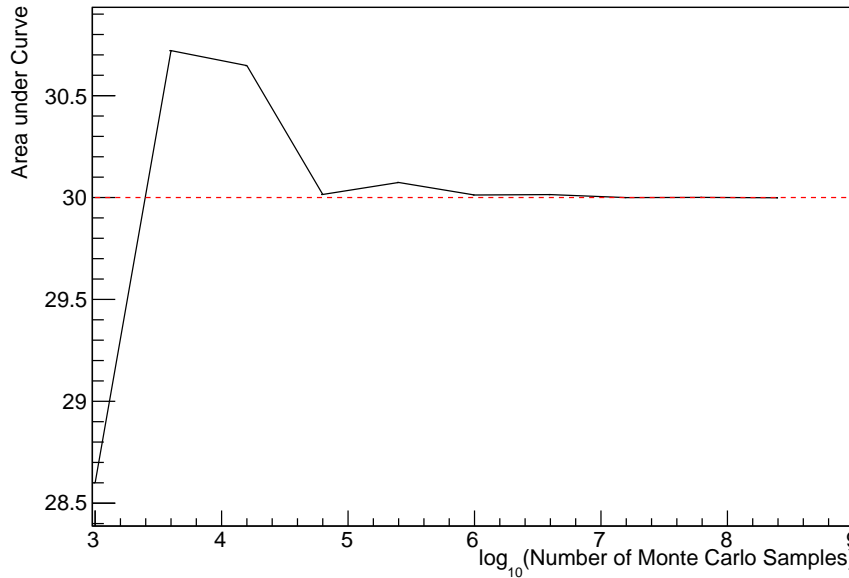
Figure 4.2.: The area under a line of gradient 0.4 and intercept 1.0 for the range $0 \leq x \leq 10$ as calculated using Monte Carlo tecniques as a function of the number of samples used in each repetition. The analytical solution to the area is 30 units as given by the red line.

which uses a Markov chain to select which points at which to sample tthe parameter distribution. This techniques performs a semi-random stochastic walk though the allowable parameter space, which builds the posterior distribution which has the property that the density of sampled points is proportional to the probability density of that parameter. This does mean that the samples produced by this technique are not statistically independent, but will cover the space of the distribution.

A Markov chain functions by selecting the position of step $\vec{x}_{i+1}$ based on the position of $\vec{x}_i$. The space in which the Markov chain selects samples is dependent upon the total number of parameters utilised within the fit, where a discrete point in this space is descibed by the N-dimensional space $\vec{x}$. In a perfectly operating Markov chain, the position of next step depends solely on the previous step and not on the further history of the chain ($\vec{x}_0$, $\vec{x}_1$, etc.). However, in solving the multi-dimesionality of the fit used within this analysis, each step becomes correlated with several of the steps procceding itself. This behaviour is further explained in subsection 4.2.3. Providing the MCMC chain is well optimised, it will begin to converge towards a unique stationary distribution. The period between the chain's initial starting point and the convergence to the unique stationary distribution is colliqually known as burn-in period and is discussed further in subsection 4.2.3. Once the chain reaches the

stationary distribution, all points sampled after that pointt will looked like samples from that distribution.

Further details of the theories underpinning MCMC techniques are discussed in Link to MCMC techniques but can be summarised by the requirement that the chain satifies the three 'regularity conditions':

- Irreducibility: From every position in the parameter space $\vec{x}$, there must exist a non-zero probability for every other position in the parameter space to be reached.

- Recurrence: Once the chain arrives at the stationary distribution, every step following from that position must be samples from the same stationary distribution.

- Aperiodicity: The chain must not repeat the same sequence of steps at any point throughtout the sampling period.

To then utilise MCMC techniques for parameter estimation, one needs to construct a Markov chain which has the property that the unique stationary distribution is the posterior distribution described in section 4.1. Consequently, the output of the chain after burn-in (ie. the sampled points after the chain has reached the stationary distribution) can then be used to approximate the posterior distribution and model parameters $\vec{\theta}$. To achieve the requirement that the unique stationary distribution found by the chain be the posterior distribution, one can use the Metropolis Hastings algorithm which guides the stochastic process depending on the likelihood of the current proposed step compared to that of the previous step. Implementation and other details of this technique are discussed in subsection 4.2.2.

4.2.2. Metropolis Hastings Algorithm

As a requirement for MCMC, the Markov chain implemented in this technique must have a unique stationary distribution which is equivalent to the posterior distribution as highlighted in subsection 4.2.2. To ensure this requirement and that the regulairty conditions are met, this analysis utilises the Metropolis-Hastings algorithm M.H algorithm citations. For the ith step in the chainn, this method determines the position in the allowable parameter space to which the chain steps to based on the current step, $\vec{x}_i$, and the proposed step, $\vec{y}_{i+1}$. The proposed step is randomomly selected from some proposal function $f(\vec{x}_{i+1}|\vec{x}_i)$, which depends solely on the current step (ie. not the further history of the chain). The next step in the chain $\vec{x}_{i+1}$ can be either the current

step or the proposed step determined by whether the proposed step is accepted or rejected. To decide if the proposed step is selected, the acceptance probability, $\alpha(\vec{x}_i, \vec{y}_i)$, is calculated as

$$\alpha(\vec{x}_i, \vec{y}_{i+1}) = min\left(1, \frac{P(\vec{y}_{i+1}|D)f(\vec{x}_i|\vec{y}_{i+1})}{P(\vec{x}_i|D)f(\vec{y}_{i+1}|\vec{x}_i)}\right). \tag{4.4}$$

Where $P(\vec{y}_{i+1}|D)$ is the posterior distribution as introduced in section 4.1. To simplify this calculation, the proposal function is required to be symmetric such that $f(\vec{x}_i|\vec{y}_{i+1}) = f(\vec{y}_{i+1}|\vec{x}_i)$. In practice, a multi-variate Gaussian distribution is used to throw parameter proposals from. This reduces Equation 4.4 to

$$\alpha(\vec{x}_i, \vec{y}_{i+1}) = min\left(1, \frac{P(\vec{y}_{i+1}|D)}{P(\vec{x}_i|D)}\right). \tag{4.5}$$

After calculating this quantity, a random number, β, is generated uniformly between 0 and 1. If $\beta \leq \alpha(\vec{x}_i, \vec{y}_{i+1})$, the proposed step is accepted. Otherwise, the chain sets the next step equal to the current step and this procedure is repeated. Consequently, one can notice that if the posterior probability of the proposed step is greater than that of the current step, $(P(\vec{y}_{i+1}|D) \geq P(\vec{x}_i|D))$, the proposed step will always be accepted. If the opposite is true, $(P(\vec{y}_{i+1}|D) \leq P(\vec{x}_i|D))$, the proposed step will be accepted with probability $P(\vec{x}_i|D)/P(\vec{y}_{i+1}|D)$. This ensures that the Markov chain does not get trapped in any local minima present in the potentially non-Gaussian posterior distribution. The outcome of this technique is that the density of steps taken in a discrete region is directly proportional to the probability density in that region.

4.2.3. MCMC Optimisation

As discussed in subsection 4.2.2, the proposal function invoked within the Metropolis Hastings algorithm can take any form and the chain will still converge to the stationary distribution; for this analysis a multi-variate Gaussian distribution was selected. As discussed in Link to Analysis Strategy, this analysis performs the Monte Carlo reweighting on an event-by-event basis resulting in a large amount of computational resources required to perform a parameter fit. Consequently, the amount of steps taken

before the unique stationary distribution is found should be minimised as only steps after convergence add information into the fit. Furthermore, the chain should entirely cover the allowable parameter space to ensure that all values have been considered. As a solution to both of these problems, tuning the distance that the proposal function jumps between steps on a parameter-by-parameter basis can both minimise the length of the burn-in period and ensure that the correlation between step $\vec{x}_i$ and $\vec{x}_{i-n}$, where n is some choosen number of steps away from the current step, is sufficiently small.

The effect of changing the width of the proposal function is highlighted in Figure 4.3. Three scenarios, each with the same underlying stationary distribution (A Gaussian of width 1.0 and mean 0.), are presented. The only difference between the three scenarios is the width of the proposal function, colliqually known as the 'step size', denoted $f(\sigma)$, where σ is the width of the Gaussian proposal function. Each scenario also starts at an initial parameter value of 10.0 which would be considered an extreme variation, For the case where $\sigma = 0.1$, it is clear to see that the chain takes a long time to reach the expected region of the parameter. This indicates that this chain would have a large burn-in period and would not converge to the stationary distribution until approximately step 500. Futhermore, whilst the chain does move towards the expected region, each step is signficantly correlated with the previous. Now considering the case where $\sigma = 5.0$, the chain approachs the expected parameter region almost instantly meaning that the burnin period is not significant. However, there are clearly large regions of steps where the chain does not move as the chain proposes steps in the tails of the stationary distribution, with low probability of being accepted. Consequently, this chain would take a signficant number of steps to fully span the allowable parameter region. For the final scenario, where $\sigma = 0.5$, you can see a relatively small burn-in period of approximately 100 steps. Once the chain reaches the stationary distribution, it moves throughout the expected region of parameter values many times, sufficiently sampling the full parameter region. This example is a single parameter varying in across a continuous distribution and does not fully reflect the difficulties in the many-hundred multi-variate parameter distribution used within this analysis but it does give an conceptual idea of the importance of selecting the proposal function and associated step size.

As discussed, step size tuning directly correlates to the number of steps the chain needs to have sufficient coverage of the full parameter space. It also directly ties into the average step acceptance rate. If the step size is too small, many steps will be accepted but the chain moves slowly. If the opposite is true, many steps will be
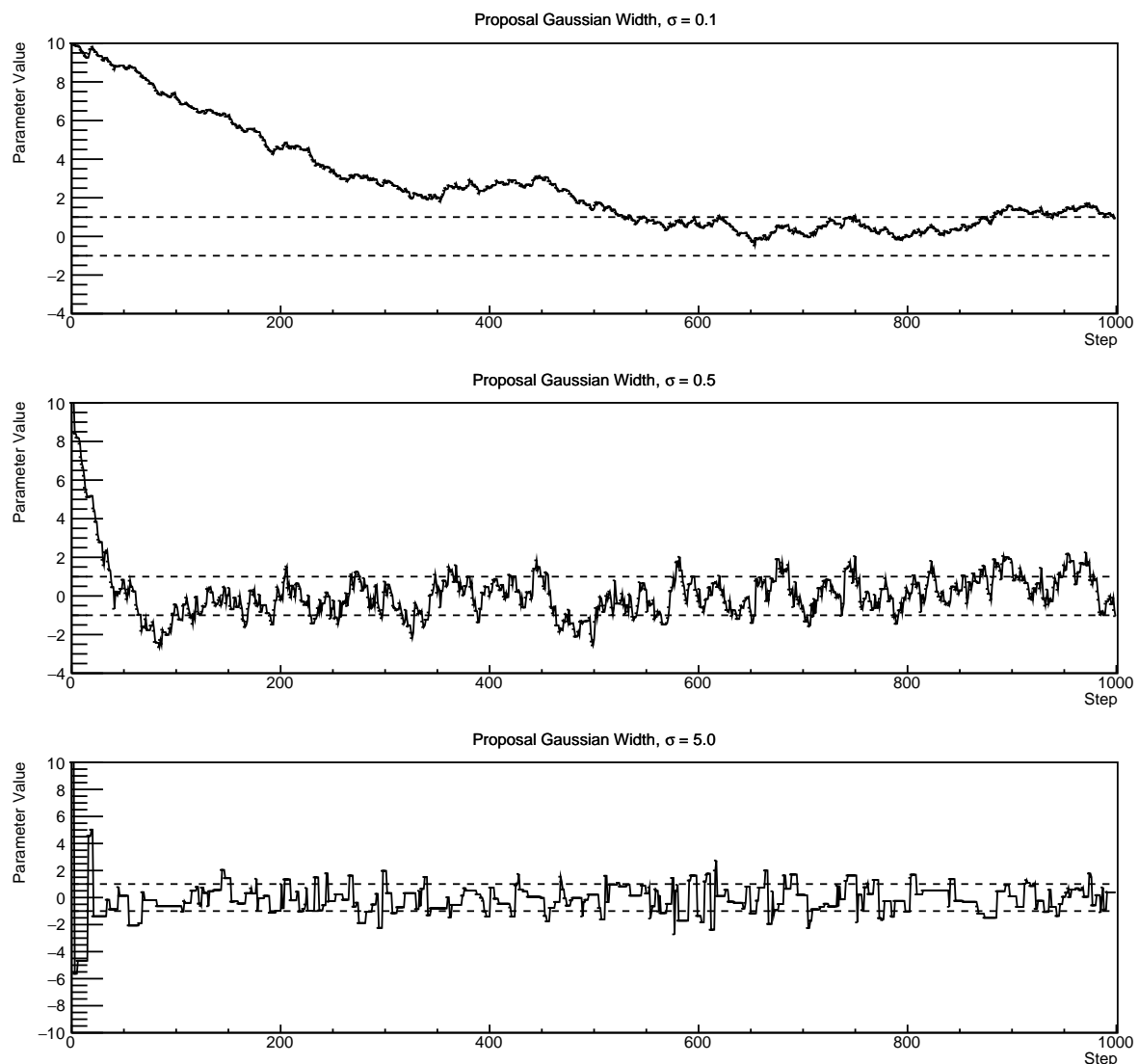
Figure 4.3.: Three MCMC chains, each with a stationary distribution equal to a Gaussian centered at 0 and width 1 (As indicated by the black dotted lines). All of the chains use a Gaussian proposal function but have different widths (or 'step size σ'). The top panel has $\sigma = 0.1$, middle panel has $\sigma = 0.5$ and the bottom panel has $\sigma = 5.0$.

rejected as the chain proposes steps in the tails of the distribution. Discussion in the literature Some juicy MCMC links suggest that the 'ideal' acceptance rate of a high dimension MCMC chain should be approximately $\sim 25\%$. Many efforts have been taken to hand select a set of parameter-by-parameter step sizes to approximately reach the ideal acceptance rate.

Figure 4.3 highlights the likelihood as calculated by the fit in Link to AsimovA official inputs section as a function of the number of steps in each chain. In pratice, many independent MCMC chains are run simultaneously to parallelise the task of

performing the fit. This figure overlays the distribution found from each chain. As seen, the likleihood decreases from it's initial value and converges towards a stationary distribution after 1×10^5 steps. As each fit (whether it be different asimov fits or data fit) will have a different set perfered parameter values which results in a different stationary distribution. For each fit in presented in this thesis, a burn-in period of 1×10^5 steps was found to be sufficient.

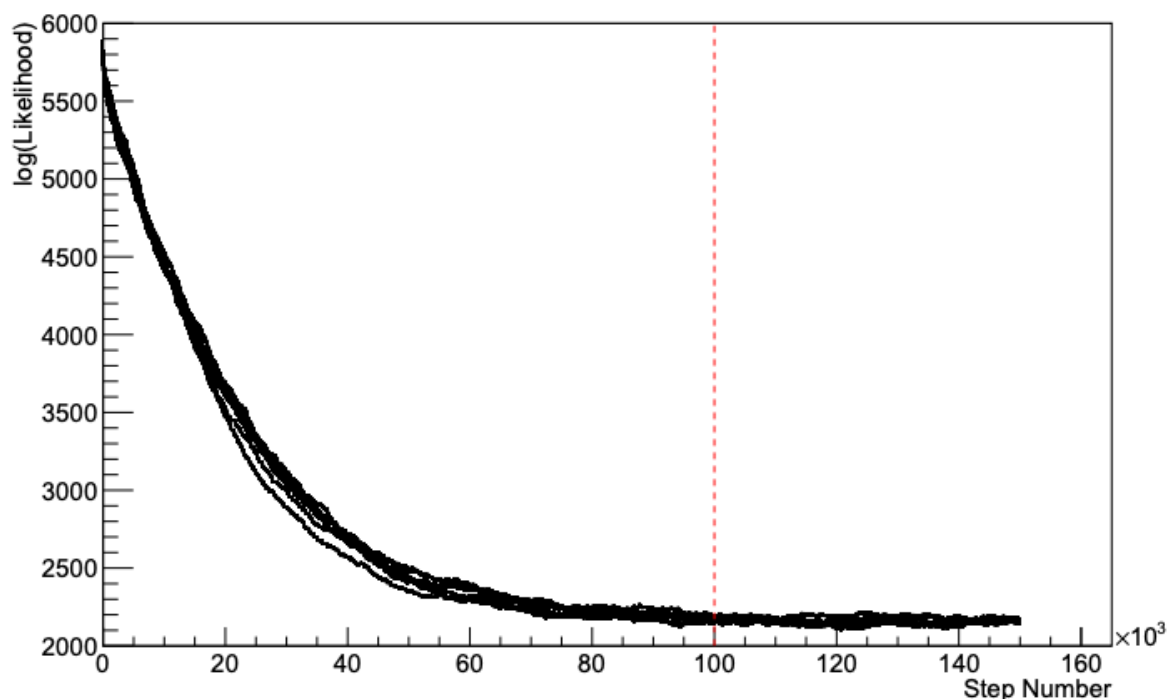**Figure 4.4.:** The log-likelihood from the fit detailed in Link to AsimovA official inputs as a function of the number of steps accumulated in each fit. Many independent MCMC chains were run in parallel and overlayed on this plot. The red line indicates the 1×10^5 step burn-in period after which the log-likelihood becomes stable.

4.3. Understanding the MCMC Results

Whilst section 4.1 and section 4.2 describe how to interpret Bayesian statistics and explains the MCMC techniques used within this analysis, there is no mention of how to intrepret the output of chain. The posterior distribution output from the chain is a high dimension object, with as many dimensions as there are parameters included in the fit. However, we typically aren't too interested in this multi-dimension object as it is difficult to conceptualise and instead opt for one or two dimensional projections of

this probability distribution, most commonally the oscillation parameters. To do this, we invoke the marginalisation technique highlighted in subsection 4.3.1.

4.3.1. Maginalisation

The output of the MCMC chain is a highly dimensional probability distribution which is very difficult to interpret. From the standpoint of an oscillation analysis experiment, the one or two dimensional 'projections' of the oscillation parameters of interest are more relevant to analysis conclusions. Despite this, the best fit values and uncertainties on the oscillation parameters of interest should correctly encapsulate the correlations to the other systematic uncertainties (colloquially called 'nuisance' parameters). For this joint beam and atmospheric analysis, the oscillation parameters of interest are $\sin^2(\theta_{23})$, $\sin^2(\theta_{13})$, Δm^2_{23} and δ_{CP}. All other parameters (Including the oscillation parameter this fit is insensitive to) are deemed nuisance parameters. To generate these projections, we rely upon integrating the posterior distribution over all nuiance parameters. A simple example of this technique is to imagine the scenario where two coins are flipped. To determine the probability that the first coin returned a 'head', the exact result of the second coin flip is disregared and simply integrated over. For the parameters of interest, $\vec{\theta}_i$, we can calculate the marginalised posterior by integrating over the nuisance parameters, $\vec{\theta}_n$. In this case, Equation 4.2 becomes

$$P(\vec{\theta}_i|D) = \frac{\int P(D|\vec{\theta}_i, \vec{\theta}_n) P(\vec{\theta}_i, \vec{\theta}_n) d\vec{\theta}_n}{\int P(D|\vec{\theta}) P(\vec{\theta}) d\vec{\theta}} \tag{4.6}$$

Where $P(\vec{\theta}_i, \vec{\theta}_n)$ encodes the prior knowledge about the uncertainty and correlations between the parameters of interest and the nuisance parameters. In practice, this is simply taking the one or two dimesional projection of the multi-dimensional probability distribution.

Whilst in principle an easy solution to a complex problem, correlations between the interesting and nuisance parameters can bias the marginalised results. A similar effect is found when the parameters being marginalised over have non-Gaussian probability distributions. For example, Figure 4.5 highlights the marginalisation bias in the probability distribution found for a parameter when requiring a correlated parameter to have a positive parameter value. Due to the complex nature of this oscillation

parameter fit presented in this thesis, there are certainly correlations occuring between the oscillation parameters of interest and the other nuisance parameters included in the fit.
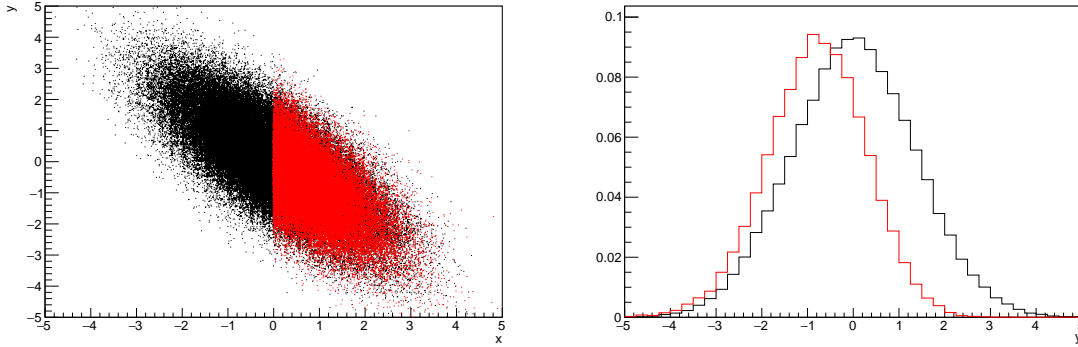

Figure 4.5.: Left: The two dimensional probability distribution for two correlated parameters x and y. The red distribution shows the two dimensional probability distribution when $0 \leq x \leq 5$. Right: The marginalised probability distribution for the y parameter found when requiring the x to be bound between $-5 \leq x \leq 5$ and $0 \leq x \leq 5$ for the black and red distribution, respectively.

4.3.2. Parameter Estimation and Credible Intervals

The purpose of this analysis is to determine the best fit values for the oscillation parameters that the beam and atmospheric samples are sensitive to; $\sin^2(\theta_{23})$, $\sin^2(\theta_{13})$, Δm^2_{23} and δ_{CP}. Typically, the results presented take the form of one or two dimension marginalised probability distributions for the appearance ($\sin^2(\theta_{13})$ and δ_{CP}) and disappearance ($\sin^2(\theta_{23})$ and Δm^2_{23}) parameters. The posterior probability density taken from the output MCMC chain is binned in the parameters of interest and the parameter best fit point is taken to be the position of highest posterior probability. This is performed in both the one dimension and two dimensions.

However, the single best fit point in a given parameter is not much use on its own. We would also like to determine the uncertainty, or credible interval, on that best fit point. The definition of the 1σ credible interval is that we have 68% belief that the parameter is within those bounds. For a more generalised definition, the credible interval is the region of the posterior distribution which contains a specific fraction of the total probability, such that

$$\int P(\theta|D)d\theta = \alpha \qquad (4.7)$$

Where θ is the parameter on which we calculate the credible interval. This technique then calculates the $\alpha \times 100\%$ credible interval.

In practice, this analysis uses the highest posterior density (HPD) credible intervals which is calculate through the following method. First, the probability distribution is area-normalised such that it has an integrated area equal to unity. The bins of probability are then summed from the highest to lowest until the sum exceeds the 1σ level (0.68 in this example). This process is repeated for a range of credible intervals, notably the 1σ, 2σ and 3σ along with other levels where the critical values for each level can be found in Table 40.1 of Link. This process can be repeated for the two dimensional probability distributions by creating two dimensional contours of credible intervals rather than a one dimensional result.

4.3.3. Application of Bayes' Theorem

Due to the matter resonance induced by oscillations occuring in the Earth in both beam and atmospheric neutrinos, this analysis has some sensitivity to the mass hierarchy of neutrino states (whether Δm_{23}^2 is positive or negative) and the octant of $\sin^2(\theta_{23})$. The Bayesian approach utilised within this analysis gives an intuitive method of model comparison by determining which hypothisis is most favourable. Taking the ratio of Equation 4.3 for the two hypotheses of normal hierarchy, NH, and inverted hierarchy, IH, gives

$$\frac{P(\vec{\theta}_{NH}|D)}{P(\vec{\theta}_{IH}|D)} = \frac{P(D|\vec{\theta}_{NH})}{P(D|\vec{\theta}_{IH})} \times \frac{P(\vec{\theta}_{NH})}{P(\vec{\theta}_{IH})}. \qquad (4.8)$$

Where the middle term defines the Bayes factor which is a data driven interpretation of how strong the data prefers one hierarchy to the other. For this analysis, equal priors on both mass hierarchy hypotheses are choosen ($P(\vec{\theta}_{NH}) = P(\vec{\theta}_{IH}) = 0.5$). In practice, the MCMC chain proposes a value of $|\Delta m_{23}^2|$ and then applies a 50% probability that the value is sign flipped. Consequently, the Bayes factor can be calculated from

the ratio of the probability density in either hypothesis which equates to counting the number of steps taken in the normal and inverted hierarchy and taking the ratio. The same approach can be taken to compare the upper octant (UO) compared to lower octant (LO) hypothesis of $\sin^2(\theta_{23})$.

Whilst the value of Bayes factor is important to a Bayesian and should always be shown for readers to determine their own conclusions, the Jeffreys scale Cite Link (highlighted in Table 4.1) gives an indication of the strength of preference for one model compared to the other. Other interpretations of the strength of preference of a model exist in Kass and Raferty Scale Link.

$\log_{10}(B_{AB})$	B_{AB}	Strength of Preference
< 0.0	< 1	No preference for Hypothesis A (Supports Hypothesis B)
$0.0 - 0.5$	$1.0 - 3.16$	Preference for Hypothesis A is weak
$0.5 - 1.0$	$3.16 - 10.0$	Preference for Hypothesis A is substantial
$1.0 - 1.5$	$10.0 - 31.6$	Preference for Hypothesis A is strong
$1.5 - 2.0$	$31.6 - 100.0$	Preference for Hypothesis A is very strong
> 2.0	> 100.0	Decisive preference for Hypothesis A

Table 4.1.: Jeffreys scale for strength of preference for two models A and B as a function of the calculated Bayes factor ($B_{AB} = B(A/B)$) between the two models Link to Jeffs scale. The original scale is given in terms of $\log_{10}(B(A/B))$ but converted to linear scale for easy of comparison throughout this thesis.

4.3.4. Comparison of MCMC Output to Expectation

Whilst not important for the extraction of oscillation parameters, understanding how the data constrains the model parameters is of importance to the understanding of this analysis. A simple method of doding this is to perform a comparison in the fitting parameters (For instance, the reconstructed neutrino energy and lepton direction for T2K far detector beam samples) of the spectra generated by the MCMC chain to 'data' which is being fit, whether that be true data or some variation of Monte Carlo prediction. This allows an easy comparison of the MCMC probability distribution to the data. To perform this, N steps from the post burn-in MCMC chain are randomly selected (Where for all plots of this style in this thesis, $N = 3000$) and the Monte Carlo prediction at each step is generated by reweighting the model parameters to the values specified at that step. Due to the probability density being directly correlated with

the density of steps in a certain region, parameter values close to the best fit value are most likely to be selected.

In practice, for each bin of the fitting parameters has a probability distribution of event rates, with one entry per sampled MCMC step. This distribution is binned and the bin with highest probability is selected as the mean and an error on the width of this probability distribution is calculated using the approach highlighted in subsection 4.3.2. Consequently the best fit distribution in the fit parameter is not necessairly that which would be attained by reweighting the Monte Carlo prediction to the most probable parameter values.

A similar study can be performed to illustrate the freedom of the model parameter space prior to the fit by instead taking parameter values that are thrown from the prior uncertainty of each parameter. This becomes troublesome for parameters with no prior uncertainty as the range is technically infinite range, but where applicable solutions to remove these have been addressed.

Chapter 5

Simulation

Simulation Chapter

5.1. Beamline

Beamline Simulation Section

5.2. Atmospheric Flux

Atmospheric Flux Simulation Section

5.3. Neutrino Interaction

Neutrino Interaction Simulation Section

5.4. Near Detector

Near Detector Simulation Section

5.5. Far Detector

Far Detector Section

5.6. Event Reconstruction

Event Reconstruction Section

Appendix A

AppendixA

This is an exmaple AppendixA