Imperial College London
Department of Physics

Study of ν_μ disappearance in the T2K experiment

Patrick Masliah

June 4, 2012

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Physics of Imperial College London
and the Diploma of Imperial College London

Declaration of originality

I herewith certify that all material in this dissertation which is not my own work has been properly acknowledged.

Patrick Masliah

Abstract

A ν_μ disappearance analysis performed on the dataset collected during the first physics run of the T2K experiment is presented. This dataset represents an accumulated exposure of 3.23×10^{19} protons on target (POT). A total of eight events passed all the selection criteria at the far detector, Super-Kamiokande, with an expectation of $22.8 \pm 4.8(\text{stat})^{+3.7}_{-3.7}(\text{syst})$.

A two flavor neutrino oscillation fit was performed with systematic uncertainties using a Markov Chain Monte Carlo method and a Bayesian approach. The $\sin^2 2\theta_{23}$ oscillation parameter was constrained to the physical region. Under these conditions, we found best fit values for the oscillation parameters of $\sin^2 2\theta_{23} = 0.83^{+0.16}_{-0.18}$ and $\Delta m^2_{32} = 2.9^{+0.8}_{-1.0} \times 10^{-3}$ eV$^2/c^4$.

Contents

List of Figures

List of Tables

Acknowledgements

I wish to first express my gratitude to Morgan Wascko, my supervisor, for his patience, encouragements, and positivity at all times. For allowing so much intellectual freedom whilst always guiding me down the right path.

I also wish to offer special thanks to Antonin Vacheret for his mentoring and friendship. I have truly learnt a lot from you. To Yoshi Uchida for his support over many years and commitment to excellence that transpired in the regular feedback provided. To Dave Wark for his inspiring leadership and fascinating anecdotes. And finally to Matthew Malek and Per Johnson for always being helpful and ready to discuss any topics.

I thank all the T2K Imperial College graduate students for creating such a great environment to work in and for the help and counsel they provided: Jim Dobson, Pawel Guzowski, Sarah Ives, Gil Kogan, Mark Scott, Samantha Short, Peter Sinclair and Benjamin Smith. It was my honour and great pleasure to work with such bright and talented people.

The T2K experiment as a whole was full of amazing people. Smart, hard-working, funny and from such varied backgrounds; I could not have wished for better experiences.

I thank my parents for always believing in me and for the sacrifices they made. I also thank all the friends I met at Imperial College for making these years so memorable. Finally I offer my sincere gratitude to Androula Alekou for the support and the great times spent together.

Introduction

Billions of neutrinos pass through our human body every single second of every single day leaving us both completely unharmed and unaware. These ghostlike particles only interact via the weakest force of nature and travel mostly unhindered through space, matter and time. In fact, given their cross section, it would take approximately one light year of lead to block 50% of the neutrinos produced in the sun.

Originally created in the big bang, they are also produced naturally on various scales. On a cosmological scale, they are produced in massive bursts in exploding stars or more constantly in nuclear reactions in the sun. On the scale of our planet earth they are produced naturally in the interactions of cosmic rays in the atmosphere or in radioactive rocks and also artificially in nuclear reactors, particle accelerators or nuclear bombs. Neutrinos are small, nearly massless (they travel close to the speed of light) and plentiful.

Their existence only started creeping into the world of science at the turn of the 20th century when physicists began to study nuclear decays. Physicists such as Chadwick studying nuclear Beta decay detected two outgoing particles: an electron (beta particle) and a proton [1]. From conservation of energy principles, they were expecting that the energy spectrum of two body decay particles must be discrete but the physicists observed that the energy spectrum of electrons was instead continuous. This led Wolfgang Pauli to postulate the existence of a new particle in his poignant "Open Letter to the Radioactivity Group at the Regional Meeting in Tubingen" [2]:

> "I have hit upon a desperate remedy to save the exchange theorem of statistics and the law of conservation of energy. This is the possibility that electrically neutral particles exist which I will call neutrons"

Enrico Fermi renamed this new particle the neutrino.

1 Neutrino history, phenomenology and experimental developments

This introductory section will cover the history of neutrino physics starting from the initial discovery by Cowan and Reines in 1956 to the current and future generation of experiments. The experimental apparatus of each neutrino experiment mentioned will be briefly described in order to illustrate the evolving techniques and technologies that led to the present generation of detectors.

1.1 The discovery

The elusive neutrino haunted physicists for more than two decades from the time it was first postulated by Wolfgang Pauli to the time that it was finally detected experimentally by Cowan and Reines in 1956 using the Savannah River nuclear reactor [3, 4].

The Savannah River nuclear reactor experiment:

The experiment at the Savannah River nuclear reactor was designed to measure the reaction

$$\overline{\nu}_e + p^+ \rightarrow e^+ + n \tag{1.1}$$

where positron annihilation produces a gamma ray pair and neutron capture a time delayed gamma ray. Three tanks of liquid scintillator with 110 5-inch photomultiplier tubes were layered with two target tanks filled with a solution of cadmium chloride. The detector was then shielded with lead and placed underground. The liquid scintillator tanks were used to detect the gamma rays and the tanks of cadmium chloride were used to absorb the neutron.

Cowan and Reines measured a four to one signal to background ratio of events with the expected signature of a neutrino event and also measured the dependence of the signal rate on the reactor power.

1.2 Flavours of neutrinos

The neutrino observed at the Savannah River nuclear reactor experiment was of course an electron anti-neutrino but the concept of flavour had not yet been introduced. Progress on that front was made from observations of disparities between the expected and measured branching ratio of rare muon decays.

1.2.1 Electron and muon neutrinos

The muon is an unstable particle that decays via the weak interaction into an electron, a neutrino and an anti-neutrino:

$$\mu \rightarrow e + \nu + \overline{\nu} \tag{1.2}$$

The concept of conservation of lepton number was postulated by Schwinger [5] in order to explain why the muon would not also decay to an electron and a photon or into three electrons:

$$\mu \rightarrow e + \gamma \tag{1.3}$$

$$\mu^{+} \rightarrow e^{+}e^{+}e^{-} \tag{1.4}$$

With this concept came the idea that there exists distinct flavours of neutrino and that each flavour is paired up with the corresponding lepton: the electron with an electron neutrino and the muon with a muon neutrino.

The two neutrino experiment in 1962 was designed specifically to test the neutrino flavour assumption by using pion decay to look for muon neutrino production [6].

The Two Neutrino Experiment:

The two neutrino experiment was based on the precept that the following two reactions should occur at the same rate if there exists only one type of neutrino:

$$\nu + p \rightarrow n + \mu \tag{1.5}$$

$$\nu + p \rightarrow n + e \tag{1.6}$$

The Alternating Gradient Synchrotron (AGS) in Brookhaven was used to create a beam of neutrinos from pion decays by striking a beryllium target with 15 GeV protons. The primary pion decay mode with probability of 99.99% is:

$$\pi^{+} \rightarrow \mu^{+} + \nu \tag{1.7}$$

$$\pi^{-} \rightarrow \mu^{-} + \overline{\nu} \tag{1.8}$$

From lepton number conservation, the neutrinos produced along side the muon have the same flavour and will produce a muon when interacting according to:

$$\overline{\nu}_{\mu} + p \rightarrow \mu^{+} + n \tag{1.9}$$

A spark chamber made up of 90 aluminium plates with neon gas filling the space in between was used to identify the charged particles (electron or muon) from neutrino interaction based on the topology of the tracks: the heavier muon results in clean tracks where as electrons showers in the detector.

Large amounts of shielding were used in order to filter out cosmic rays and remove muons from the pion decays before they themselves have the chance to decay. This included, notably, 13.5 meters of steel from a dismantled warship used as a beam dump. A timing window coincident with the beam was also used as an additional filter against cosmic rays. The detector measured an excess of muon events over electron events which confirmed the existence of a second distinct type of neutrino and validated the lepton number conservation law.

Schwartz, Lederman and Steinberger earned the Nobel prize in 1988 for their work on the design of the first neutrino beam and on the detector for measuring the neutrino reactions [7].

1.2.2 Tau neutrino

The tau neutrino is the third and heaviest neutrino and is paired with the Tau lepton. The Tau lepton was first suggested to have been observed in 1975 by Perl et al in order to account for missing energy in e^{+} e^{-} annihilation [8] in events of the form

$$e^{+} + e^{-} \rightarrow e^{\pm} + \mu^{\mp} + \geq 2 \text{ undetected particles.} \tag{1.10}$$

The experiment used the SPEAR (Stanford Positron Electron Asymmetric Rings) electron positron collider at SLAC (Stanford Linear Accelerator Center) and the LBL (Lawrence Berkeley Laboratory) magnetic detector.

The discovery was confirmed with the demonstration of the existence of other predicted decay channels: the anomalous μ - hadron, the anomalous e - hadron and semileptonic decay modes from various independent experiments including PLUTO [9] and Mark I [10].

The Tau neutrino was naturally assumed to exist once the Tau lepton had been detected but credit for the actual observation goes to the DONUT (Direct Observation of the NU Tau) collaboration at Fermilab in 2000 [11].

The DONUT Experiment:

The idea underlying the DONuT experiment was simple: a beam of ν_τ is created and ν_τ charged current events are measured:

$$\nu_\tau + N \rightarrow \tau + X, \tag{1.11}$$

Although simple in principle, the task was very complex because of the difficulty of creating an intense beam of τ neutrinos and in identifying the τ lepton from the charged current interactions.

The neutrino beam was generated primarily by the leptonic decays of charmed particles, pions and kaons form the hadronic shower created by 800 GeV protons hitting a beam dump of tungsten alloy. Measurements of the composition of the neutrino flux gave a breakdown of 97% of ν_e and ν_μ and 3% of ν_τ with the majority of the tau neutrinos coming from the decay of D_s mesons:

$$D_s \rightarrow \tau + \nu_\tau, \tag{1.12}$$

Sweeping magnets and a large amount of concrete shielding were used to reduce the two main backgrounds from the dump: muons and neutrons respectively

The most common signature for a τ lepton is a short track with a kink caused by the decay of the τ into a charged particle (this happens with a branching ratio of 86%). This characteristic topology was recognised by using a high resolution emulsion tracking system. It was necessary for the emulsion system to achieve high resolution because of the short 2 mm decay length of the τ lepton.

An electronic spectrometer was placed behind the emulsion target in order to pre-select neutrino interaction candidates and estimate the position of their interaction vertex in the emulsion.

The DONuT experiment successfully observed 9 ν_τ interactions with an estimated background of 1.5 events.

1.2.3 The Z boson constraint

Three neutrino flavours have been observed so far corresponding to the three lepton flavours: ν_e, ν_μ and ν_τ but could there be more?

A partial answer to this question is given by electroweak measurements on the Z resonance which provides a constraint on the number of "light" neutrino species (neutrinos that have a mass of less than half the Z mass).

The total width Γ_{tot} of the Z boson is the sum of the visible and invisible partial widths Γ_{vis} and Γ_{inv}. The visible width corresponds to Z decays to quarks and charged leptons whereas the invisible width is assumed to be due

to the number of light neutrino species N_ν. N_ν is given by:

$$N_\nu = \frac{\Gamma_{\text{inv}}}{\Gamma_l}\left(\frac{\Gamma_l}{\Gamma_\nu}\right)_{SM}, \tag{1.13}$$

where Γ_l and Γ_ν are respectively the lepton and neutrino partial widths and $\left(\frac{\Gamma_l}{\Gamma_\nu}\right)_{SM}$, the Standard Model value for the ratio of the charged leptonic to neutrino partial widths, is used to reduce the model dependence.

In what is one of the most precise measurement in the field of high energy physics, the number of "light" neutrino species has been determined to be 2.9840 ± 0.0082 [12].

This does not discard the possibility of the existence of "heavy" neutrinos or sterile light neutrinos.

1.3 Neutrino mass and the standard model

When the standard model was first constructed, neutrinos were assumed to be massless particles. Indeed there was no evidence at the time to the contrary. This is illustrated in the next sections on attempts at direct measurements of the neutrino mass and the measurement of neutrino helicity.

The mechanism by which mass is introduced in the standard model is reviewed in the section on Dirac mass.

1.3.1 Direct measurements of mass

Experiments on direct measurements of the electron neutrino mass look at the energy spectrum of the electron in tritium β-decay:

$$^3H \rightarrow ^3 He + e^- + \overline{\nu}_e, \tag{1.14}$$

A non-zero neutrino mass will change both the slope at the end point and the maximum allowable energy in the electron energy spectrum. No searches indicated any significant effect attributable to the neutrino which implied that the neutrino mass was either zero or very small.

1.3.2 Neutrino helicity

Helicity is the projection of a particle's spin along its direction of motion. By definition, the helicity of a massive particle can be reversed from the point of view of an observer by changing to a reference frame where the particle moves in the opposite direction. This is not possible if a particle is massless because it travels at the speed of light.

Maurice Goldhaber made a direct measurement of neutrino helicity in a famously ingenious experiment using a combined analysis of circular polari-

sation and resonant scattering of γ rays following electron capture in EU^{152m} [13]. He found that 100% of the neutrinos had negative helicity which implied once again that they are massless particles.

1.3.3 Chirality

An object is said to be chiral if it cannot be super-imposed on its mirror image. Chirality is an intrinsic quantum mechanical property of particles. Physically, it is a phase shift in the particle's wave-function. Just like helicity, a particle can have left or right chirality and the two concepts are actually the same in the case of massless particles. However, unlike helicity, chirality is invariant under Lorentz transformation.

1.3.4 Dirac mass

The mathematical formalism of the standard model of particle physics describes particles as quantum fields. Their equations of motion are solutions to the Euler-Lagrange equation. Since we know that the equation of motion of a spin 1/2 fermion is given by the Dirac equation, the Lagrangian of a free fermion is constructed such that the solution of the Euler-Lagrange equation yields the Dirac equation. Mass is included in the Lagrangian by the addition of the Dirac mass term:

$$m\overline{\Psi}\Psi = m\overline{\Psi}_L\Psi_R + m\overline{\Psi}_R\Psi_L, \tag{1.15}$$

where Ψ_L and Ψ_R are the left and right chiral states. Therefore a non-zero Dirac mass requires a particle to have both a left and a right handed chiral state and the Dirac mass is the coupling constant between the two components. For example, an electron is actually a superposition of states of a left and right chiral electron.

One particularity of the standard model is that it is a chiral theory: left and right chiral particles behave differently. For instance, the W boson which mediates the charged-current weak interactions only couples to the negative-chirality component of fermions (left handed neutrinos and right handed anti-neutrinos). As a note, the mixing between the left and right chiral fermions should be prohibited by "Gauge invariance" given that only the left handed fermion can interact with the W. The process by which this happens in the standard model is called electroweak symmetry breaking. Basically, the vacuum expectations value of the Higgs gives particles their mass but at the same time breaks the conservation of weak charge and therefore allows the different chiral states to mix without breaking Gauge invariance.

1.4 Neutrino flavour change and oscillation

Neutrino physics took a new turn in the 1970s when evidence in favour of neutrino flavour change began to emerge. By the end of the century, the evidence collected was unequivocal: neutrinos change flavours, and, by implication, leptons mix and neutrinos have non-zero masses.

1.4.1 Neutrino Oscillation formalism

The theoretical framework designed to account for neutrino oscillation rests upon the existence of neutrino mass and lepton mixing.

To demonstrate this, we postulate that neutrinos have masses and leptons mix. Consequently, from the first postulate, there is a spectrum of neutrino mass eigenstates ν_i, $i = 1, 2, \cdots$, each with mass m_i. From the second postulate, the neutrino of flavour α that associates with the lepton of same flavour is in fact a superposition of mass eigenstates according to the unitary transformation

$$|\nu_\alpha\rangle = \sum_i U^*_{\alpha i}|\nu_i\rangle, \tag{1.16}$$

where $U^*_{\alpha i}$ gives the amplitude for the mixing between the lepton of flavour α and the neutrino of mass eigenstate m_i.

Equation 1.17 can also be inverted to express each mass eigenstate ν_i as a superposition of flavours:

$$|\nu_i\rangle = \sum_\alpha U_{\alpha i}|\nu_\alpha\rangle. \tag{1.17}$$

In effect, a neutrino flavour change is an inherently quantum-mechanical effect that can be visualised in three stages enumerated below and depicted in figure 1.1 from [14].

1. A neutrino of flavour α is created at a source,
2. it travels a distance L to a detector,
3. it interacts in the detector and produces a charged lepton of flavour β.

Since ν_α is actually a coherent superposition of mass eigenstates ν_i, the particle that propagates from the neutrino source to the detector is one of the ν_i and all contributions must be added coherently.

The probability for the neutrino of flavour α to oscillate to a flavour β is given by:

$$P(\nu_\alpha \to \nu_\beta) = |\mathrm{Amp}(\nu_\alpha \to \nu_\beta)|^2, \tag{1.18}$$

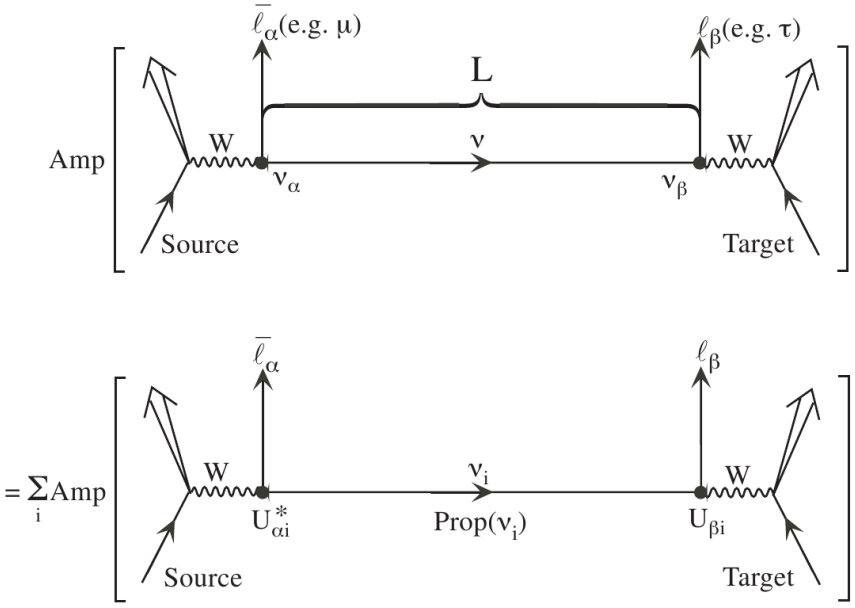


Figure 1.1: Neutrino flavour change in vacuum. "Amp" denotes amplitude [15].

where the amplitude $\mathrm{Amp}(\nu_\alpha \to \nu_\beta)$ is:

$$\mathrm{Amp}(\nu_\alpha \to \nu_\beta) = \sum_i U^*_{\alpha i}\mathrm{Prop}(\nu_i)U_{\beta i}. \tag{1.19}$$

The amplitude for the initial neutrino to propagate from the source to the detector $\mathrm{Prop}(\nu_i)$ can be shown to approximate to [14]:

$$\mathrm{Prop}(\nu_i) = \exp\left(-im_i^2\frac{L}{2E}\right), \tag{1.20}$$

in the limit that neutrinos have very small masses, and, using the fact that different mass eigenstate components of a beam that contribute coherently to the oscillation signal must have the same energy, E [16]. L stands for the neutrino travel distance between the source and the detector.

Assuming that CPT invariance holds, the probability $P(\overline{\nu}_\alpha \to \overline{\nu}_\beta)$ for the corresponding anti-neutrino oscillation is given by:

$$P(\overline{\nu}_\alpha \to \overline{\nu}_\beta) = P(\nu_\beta \to \nu_\alpha). \tag{1.21}$$

Finally, combining equations 1.18, 1.19, 1.20 and 1.21, the following rela-

tionship for neutrino oscillation can be derived:

$$P(\overset{(-)}{\nu}_\alpha \to \overset{(-)}{\nu}_\beta) = \delta_{\alpha\beta} \quad - \quad 4\sum_{i>j} \Re\left(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta i}\right)\sin^2\left(\Delta m^2_{ij}\frac{L}{4E}\right)$$
$$\overset{+}{(-)} \quad 2\sum_{i>j} \Im\left(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta i}\right)\sin\left(\Delta m^2_{ij}\frac{L}{4E}\right), \quad (1.22)$$

where $\Delta m^2_{ij} \equiv m^2_i - m^2_j$.

From equation 1.22, we see that if the mixing matrix U is complex, $P(\overline{\nu}_\alpha \to \overline{\nu}_\beta)$ and $P(\nu_\alpha \to \nu_\beta)$ will in general differ. Since $\overline{\nu}_\alpha \to \overline{\nu}_\beta$ is the CP-mirror image of $\nu_\alpha \to \nu_\beta$, $P(\overline{\nu}_\alpha \to \overline{\nu}_\beta) \neq P(\nu_\alpha \to \nu_\beta)$ would be a violation of the CP invariance [15].

For three neutrino flavours, the collection of $U^*_{\alpha i}$'s give the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix [17, 18]:

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \quad (1.23)$$

where c_{ij} is short for $\cos\theta_{\mathrm{ij}}$ and s_{ij} for $\sin\theta_{\mathrm{ij}}$.

The leptonic mixing matrix is then often factorised into three components to separate the dominant terms that arise from atmospheric, solar and reactor and beam experiments (equation 1.24). By substituting α and β with the relevant neutrino flavours in equation 1.22, we find that:

- the dominant atmospheric oscillation term $P(\nu_\mu \to \nu_\mu)$ depends on the θ_{23} mixing angle,
- the reactor/accelerator sub-dominant oscillation term $P(\nu_\mu \to \nu_e)$ depends on the θ_{13} mixing angle and the CP-violating phase δ,
- the solar oscillation term $P(\nu_e \to \nu_e)$ depends on the θ_{12} mixing angle.

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.24)$$

The minimal neutrino oscillation model therefore adds seven new parameters to the Standard Model: the masses of the three neutrino mass eigenstates, the three mixing angles θ_{ij} and the CP violating phase δ of the PMNS mixing matrix. The absolute masses have no effect on oscillations and the actual observables are the differences of the squares of the masses Δm^2_{ij}.

In the two neutrino approximation, the "survival probability" for the observation of a ν_μ of energy E (GeV) having travelled a distance L (Km) is:

$$P(\nu_\mu \rightarrow \nu_\mu) \simeq 1 - \sin^2 2\theta_{23} \sin^2 \left[\frac{1.27 \Delta m^2_{23} L}{E_\nu} \right], \tag{1.25}$$

and the probability for the ν_μ to oscillate to a ν_e is:

$$P(\nu_\mu \rightarrow \nu_e) \simeq \sin^2 2\theta_{13} \sin^2 \left[\frac{1.27 \Delta m^2_{23} L}{E_\nu} \right] \tag{1.26}$$

where Δm^2_{ij} the mass difference m^2_i-m^2_j. The mass difference Δm^2, in combination with L and E, therefore dictate the period of oscillation where as the angles θ_{ij} give the magnitude.

1.4.2 Experimental evidence for neutrino flavour change and oscillation

Neutrinos are constantly emitted in very large quantities by the sun, by hadronic decays from hadrons produces in the collisions of cosmic rays with nuclei in the upper atmosphere and by operating nuclear reactors. Physicists made good use of this fact over three decades to study the properties of neutrinos.

Solar neutrinos

The first hint for flavour change came with the observation of an anomaly in the solar electron neutrino flux. In 1970, the Homestake [19] experiment observed a deficit of electron neutrinos from predictions made by the reliable standard solar model (SSM). This anomaly was confirmed by the SAGE [20] and GALLEX [21] experiments.

The Sudbury Neutrino Observatory was the first experiment to show direct evidence for neutrino flavour transformation [22]. Using the charged current (CC) reaction $\nu_e + d \rightarrow p + p + e^-$, they confirmed the electron neutrino solar flux deficit for ^8B neutrinos. Simultaneously, using the neutral current (NC) reaction $\nu_x + d \rightarrow p + n + \nu_x$ which is equally sensitive to all active neutrino flavours (ν_e, ν_μ, ν_τ), they showed conservation of the total ^8B solar neutrino flux in agreement with the SSM prediction, thus showing that some electron neutrinos had changed to one of the other two flavours.

Reactor experiments

The KamLAND reactor experiment demonstrated that this flavour change was the consequence of neutrino oscillation by observing distortion of reactor $\overline{\nu}_e$ energy spectrum [23] as a function of the distance from the reactor. Figure

1.2 plots the ratio of the background-subtracted $\overline{\nu}_e$ candidate events, including the subtraction of geoneutrinos, to no oscillation expectation [24].

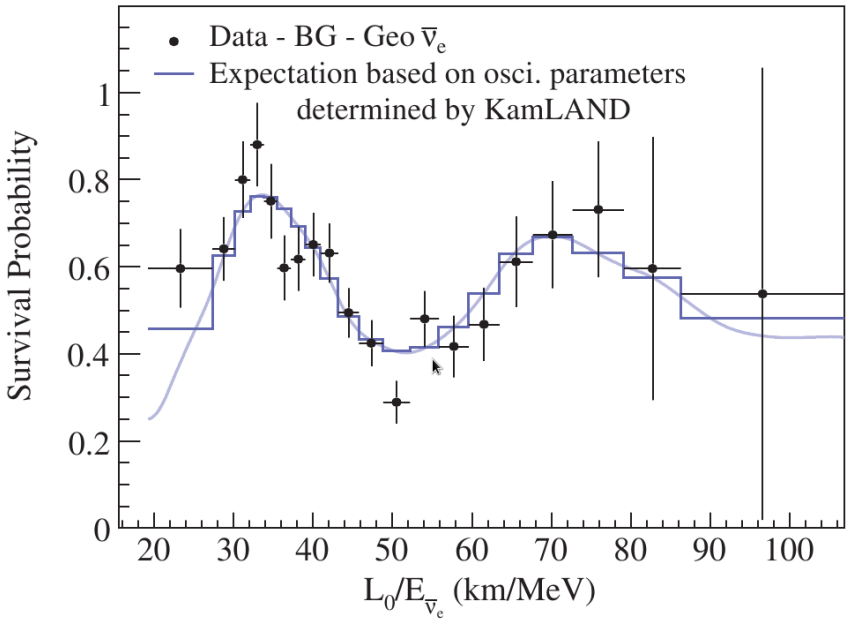


Figure 1.2: Ratio of the background and geoneutrino-subtracted $\overline{\nu}_e$ spectrum to the expectation for no-oscillation as a function of $\frac{L_0}{E}$. L_0 is the effective baseline taken as a flux-weighted average ($L_0 = 180$ km) [24].

Atmospheric neutrinos

Similarly, atmospheric neutrinos, produced as decay products in hadronic showers resulting from collisions of cosmic rays with nuclei in the upper atmosphere, were studied by the Super-Kamiokande water Cherenkov detector [25]. The data presented a zenith angle dependent deficit of muon neutrinos consistent with $\nu_\mu \leftrightarrow \nu_\tau$ oscillation. Long baseline neutrino experiments K2K [26] and MINOS [27] observed ν_μ disappearance consistent with the Super-K results. Figure 1.3 presents 68% and 90% contours for the MINOS oscillation fit and 90% contours for the Super-K and K2K experiments.

1.4.3 Current values and limits of neutrino oscillation parameters

The current values and limits of the neutrino oscillation parameters are given in table 1.1.

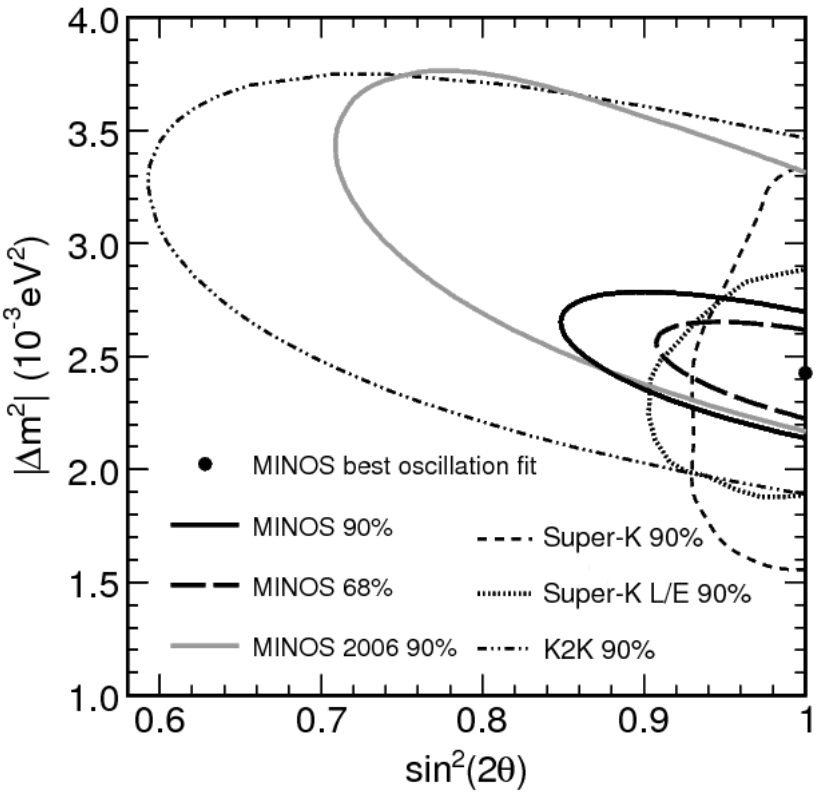


Figure 1.3: Contours for the MINOS 68% and 90% oscillation fit including systematic errors. Also shown are results from Super-K [28] and K2K [26] and the earlier MINOS result [27, 29]. Figure obtained from [30]

Parameter	Value/Limit (MeV)
Δm^2_{21}	$(7.59 \pm 0.21) \times 10^{-5} eV^2$
Δm^2_{32}	$2.32^{+0.12}_{-0.08}$
$\sin^2 2\theta_{12}$	$0.861^{+0.026}_{-0.022}$
$\sin^2 2\theta_{23}$	> 0.92, CL=90%
$\sin^2 2\theta_{13}$	0.092±0.016(stat.)±0.005(syst.)

Table 1.1: Current values and limits of the neutrino oscillation parameters [31].

2 T2K Experiment

The T2K (Tokai-to-Kamioka) experiment is a long baseline neutrino oscillation experiment.

Physics goals

The primary physics goal of the T2K experiment is the measurement of the last unknown lepton sector mixing angle θ_{13} from observations of ν_e appearance in a ν_μ beam. The aim is to improve the sensitivity to θ_{13} by an order of 20 compared to the current best limit set by the CHOOZ experiment [32]

A secondary goal is to do a precision measurement of the known atmospheric oscillation parameters $\sin^2 2\theta_{23}$ and Δm^2_{23} with precisions of respectively 0.01 and 10^{-4} eV^2 using ν_μ disappearance studies.

The T2K neutrino beam will also be used for neutrino interaction cross section studies.

Experiment Overview

A focused beam of muon neutrinos is produced at the J-PARC accelerator in Tokai, Japan using an intense proton beam and neutrino beamline. The neutrinos with an energy spectrum peaked on axis at approximately 2 GeV are sent 2.5° off-axis towards the Super-Kamiokande [33] detector 295 km. There, the neutrino rates and energies are measured.

A near detector complex composed of an on-axis detector, INGRID, and an off-axis detector, the ND280, was constructed 280 m downstream of the neutrino target to sample the beam immediately after production [34]. The INGRID detector measures the neutrino beam direction and profile. The ND280 measures the muon neutrino flux and energy spectrum in the direction of Super-K. The intrinsic electron neutrino contamination of the beam and rates for exclusive neutrino reactions are also measured by the ND280 in order to characterise signals and backgrounds at the far detector.

Figure 2.1 presents a diagram of the T2K beam.

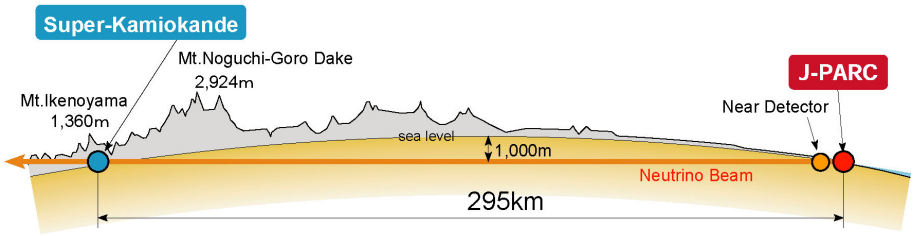


Figure 2.1: Diagram of the path of the T2K beam and location of the detectors

Off-Axis experiment

T2K is the first long baseline neutrino experiment to use an off-axis neutrino beam. By pointing the beam 2.5 degrees off-axis from the Super Kamiokande detector, a narrower neutrino energy spectrum (in the no oscillation hypothesis) is obtained centred around 600 MeV rather than 2 GeV (see fig 2.2). The peak of the spectrum at 600 MeV corresponds approximately to the expected oscillation maximum in the ν_μ survival probability formula (equation 1.25) given the best fit values for the oscillation parameters from the MINOS results (figure 1.3) and the length of the T2K baseline. This coincidence results in maximised neutrino oscillation effects and therefore an improved sensitivity. The ν_e background due to the contamination of the beam and neutral current background events that arise predominantly from higher energy neutrinos are also reduced at the off-axis position [35].

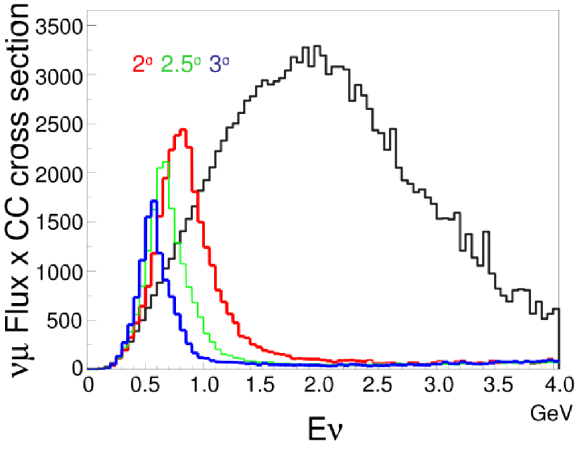


Figure 2.2: Predicted neutrino energy spectra comparison for on axis beam (black) and 2, 2.5 and 3 degree off axis beams. The beam off axis has a much narrower energy band and cuts out NC background events that come predominantly from higher energy neutrinos.

2.1 J-PARC complex and T2K beam

The J-PARC complex consists of three accelerators: a linear accelerator (LINAC), a rapid-cycling synchrotron (RCS) and a main ring (MR) synchrotron. A proton beam with 6 or 8 bunches (since June 2010) is accelerated up to 30 GeV in the main ring and extracted into the neutrino beamline by a set of 5 kicker magnets. The neutrino beamline is composed of a primary and a secondary beamline:

Primary neutrino beamline

The primary beamline redirects the proton beam to point towards Kamioka using normal and superconducting magnets. The intensity, position, profile, and loss of the beam in the primary section are monitored to ensure that it is operational and adequately tuned using respectively:

- current transformers (CTs),
- electrostatic monitors (ESMs),
- segmented secondary emission monitors (SSEMs),
- beam loss monitors (BLMs).

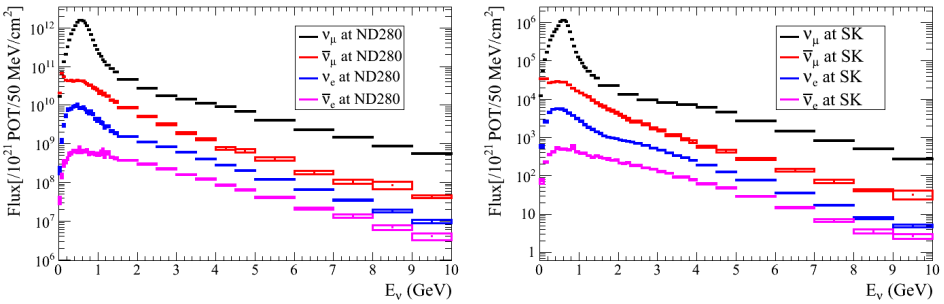
Descriptions of the CTs, ESMs, SSEMs and BLMs can be found in [34].

Secondary neutrino beamline

The secondary beamline consists of a target station, a decay volume and a beam dump. The proton beam collides in the target station on a fixed graphite rod, cooled with helium gas, to produce pions and kaons. The charged hadrons are focused into a 100 m long decay volume with powerful magnetic fields. obtained by using three sets of two concentric aluminium cylinders, called magnetic horns, with 250 kA current pulses. Pions then decay in the decay volume into muons and muon neutrinos: $\pi^+ \rightarrow \mu^+\nu_\mu$. The beam dump placed at the end of the decay volume stops the remaining pions and kaons whilst the neutrinos and high energetic muons carry on in the direction of the the near detector complex and the Super-K far detector.

T2K beam

The off-axis neutrino beam is composed of approximately 95% ν_μ, 4% $\overline{\nu}_\mu$ and less than 1% ν_e. Figure 2.3 presents the predicted flux by flavor at the ND280 and Super-K detectors simulated using FLUKA2008 [36] and GEANT3 [37].



(a) ND280 (b) Super-K

Figure 2.3: T2K beam flux prediction by flavor [38].

The ν_μ component of the beam arises from primary charged pion and kaon decays and secondary anti-muon decays. The relevant decay channels and branching ratios are presented in table 2.1.

Decay	Branching Ratio (%)
$\pi^+ \rightarrow \mu^+\nu_\mu$	99.987
$K^+ \rightarrow \mu^+\nu_\mu$	63.4
$\mu^- \rightarrow e^-\nu_\mu\overline{\nu}_e$	~ 100

Table 2.1: Beam ν_μ signal decay channels and branching ratios [39].

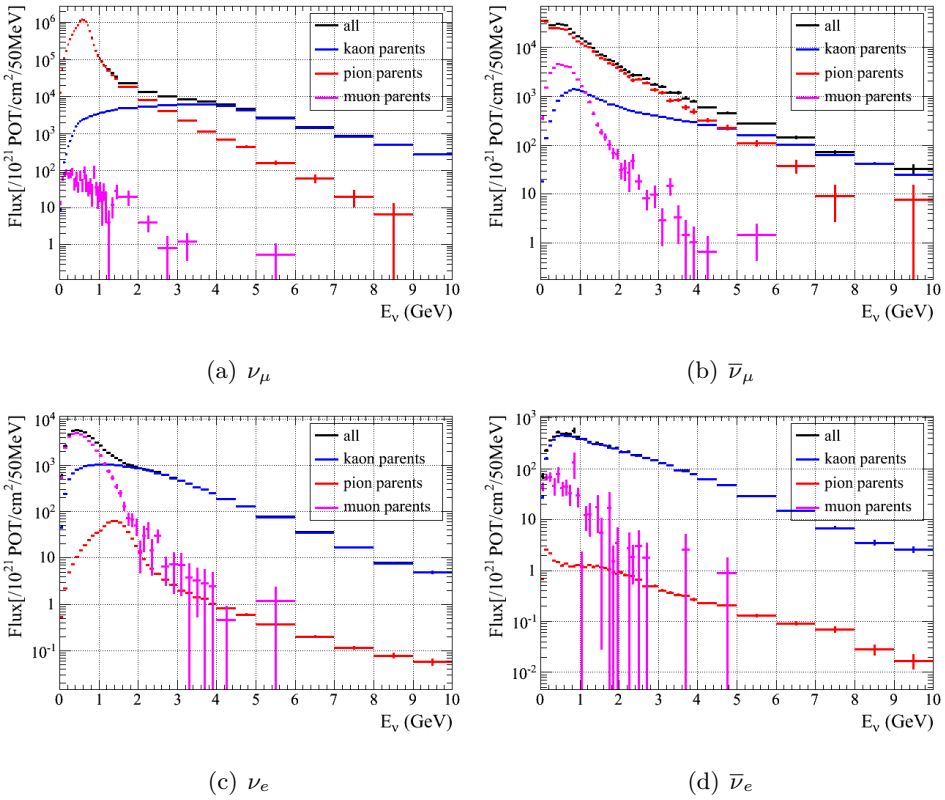
Electron neutrino beam contamination also arises from primary charged pion and kaon decays and secondary muon decays. The relevant decay channels and branching ratios are presented in table 2.2.

Decay	Branching Ratio (%)
$\pi^+ \rightarrow e^+\nu_e$	0.012
$K^+ \rightarrow \pi^0 e^+\nu_e$	4.9
$\mu^+ \rightarrow e\overline{\nu}_\mu\nu_e$	~ 100

Table 2.2: Beam ν_e contamination signal decay channels and branching ratios [39].

The high purity of the beam is possible because the $\pi^+ \rightarrow e^+\nu_e$ decay channel is heavily suppressed due to helicity effects. Muon neutrinos originate primarily from pion parents below 2.5 to 3 GeV and primarily from kaon parents above 3 GeV for both the near and far detectors. Muon anti-neutrinos arise mainly from pion decays below 5 GeV and mainly from kaon decays above. A significant fraction of lower energy events are from muon parents.

Electron neutrino contamination below 1.2 GeV originates mainly from muon decays and almost exclusively from kaon parents above. Figure 2.4 and 2.5 present respectively the predicted Super-K and ND280 parent contribution for each neutrino flavor component of the off-axis neutrino beam. Each spectrum was predicted using FLUKA2008 and GEANT3.



(a) ν_μ (b) $\overline{\nu}_\mu$

(c) ν_e (d) $\overline{\nu}_e$

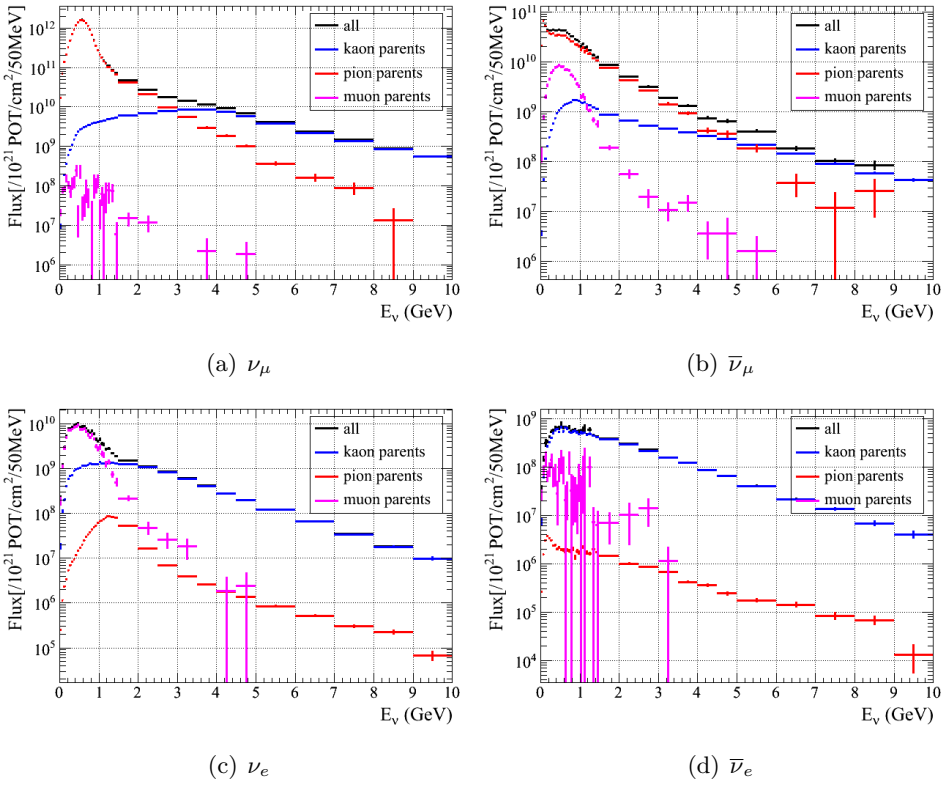
Figure 2.4: T2K beam flux parent contribution prediction at Super-K with statistical uncertainties [38].

2.2 Near detectors

As previously mentioned, the near detector complex, located 280 m from the production target, is composed of an on-axis and an off-axis detector: INGRID and ND280. It was designed to measure the neutrino beam direction, energy spectrum, flavor content and interaction rates of the beam before oscillation.

2.2.1 ND280

The ND280, shown in an exploded diagram in figure 2.6, is an aggregate of several complementary sub-detectors placed inside the UA1/NOMAD 0.2 T magnet recycled from the famous CERN experiment. The physical size of the



(a) ν_μ (b) $\overline{\nu}_\mu$

(c) ν_e (d) $\overline{\nu}_e$

Figure 2.5: ND280 beam flux parent contribution prediction at Super-K with statistical uncertainties [38].

detectors are therefore limited by the dimensions of the UA1 magnet cavity (7.0 m long, 3.6 m high, 3.5 m wide).

A Pi-zero detector (P0D) is placed upstream and a tracker downstream. The tracker is composed of three time projection chambers (TPCs) alternated with two fine grain detectors (FGDs). These innermost detectors are all surrounded by electromagnetic calorimeters (ECals). The magnet return yoke is instrumented with scintillator to measure the ranges of muons that exit the side of the detector. The scintillator bars used in the near detector were developed and produced at Fermilab [40].

UA1 Magnet

The UA1 magnet at 0.2T surrounds the ECAL allowing accurate measurement of charged-particle momenta and sign in the tracker and is instrumented with scintillator bars to detect sideways going muons (the Side Muon Range Detector). It can also act as a veto or trigger for cosmic events.

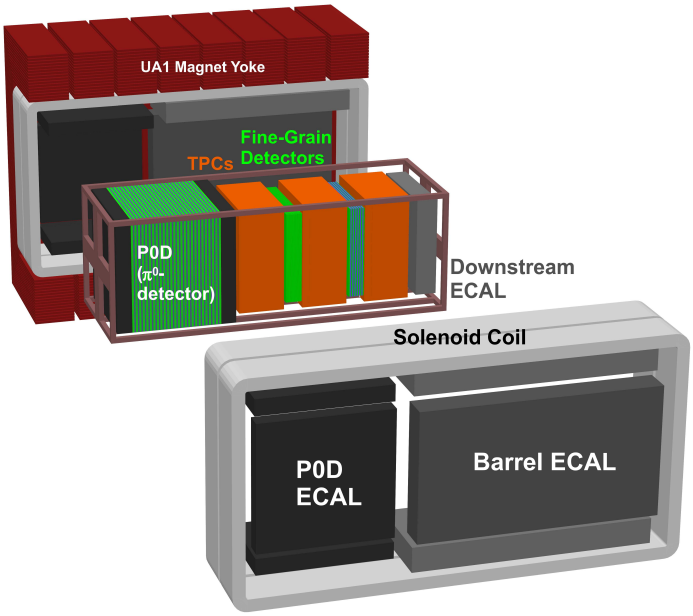


Figure 2.6: Exploded diagram of the ND280m off-axis detector. Left in the figure is upstream of the beam, right is downstream. Dimensions of the UA1 magnet cavity (7.0m long, 3.6m high, 3.5m wide) limit the sizes of the detectors

Pi-zero detector (P0D)

The P0D is composed of scintillator bars alternated with layers of lead foil, brass sheets and water. Each scintillator bar is mounted with a wavelength shifting fibre (WLS) inserted through a hole in the centre. One end of the fibre is mirrored, the other is instrumented using a multi-pixel photon-counter (MPPC) from Hamamatsu [41, 42]. Each photo-detector is read out with a TRIP-t front end board (TFB) [43]. The P0D layout is presented in figure 2.7.

The primary role of the P0D is to measure the rate of neutral current π^0 production on water, which is the main background to the ν_e appearance at Super-K:

$$\nu_\mu + N \rightarrow \nu_\mu + N + \pi^0 + X. \tag{2.1}$$

The P0D can be operated with the water targets either filled or empty which enables determination of the water target cross sections with a subtraction method. The scintillator bars provide sufficiently fine segmentation to reconstruct charged particle tracks (muons and pions) and electromagnetic showers (electrons or photons from π^0's).

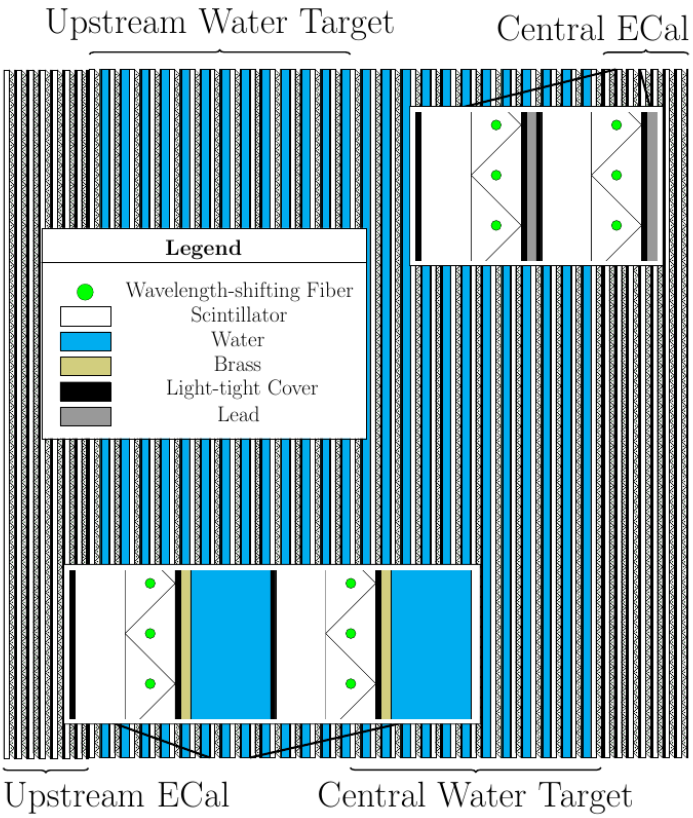


Figure 2.7: A schematic of the P0D. The beam is coming from the left and going right. Insets show details of the Water Target and ECal layers.

Time projection chamber (TPC)

The TPCs are argon-based drift gas detectors with readout planes on each end that act as anodes and a central cathode panel. A uniform electric drift field is produced as a result in the active drift volume of the TPC. Charged particles passing through the TPCs produce ionisation electrons in the gas. These electrons drift away from the central cathode and towards one of the readout electrodes where they are multiplied using an avalanche process and sampled by a charge amplifier in a Micro-Mesh Gaseous Structure (MICROMEGAS) with 7.0 mm $\times$ 9.8 mm anode pad segmentation [44]. The electronics are read out with AFTER "ASIC" front end boards [45]. Three dimensional imaging of charged particle tracks is obtained. Figure 2.8 presents a simplified diagram of the TPC design.

The TPCs operate in the magnetic field of the UA1 magnet and offer excellent imaging capabilities in three dimensions. These characteristics allow measurement of the 3-momenta and dE/dx of charged particles for particle identification and selection of high purity samples of different types of neu-

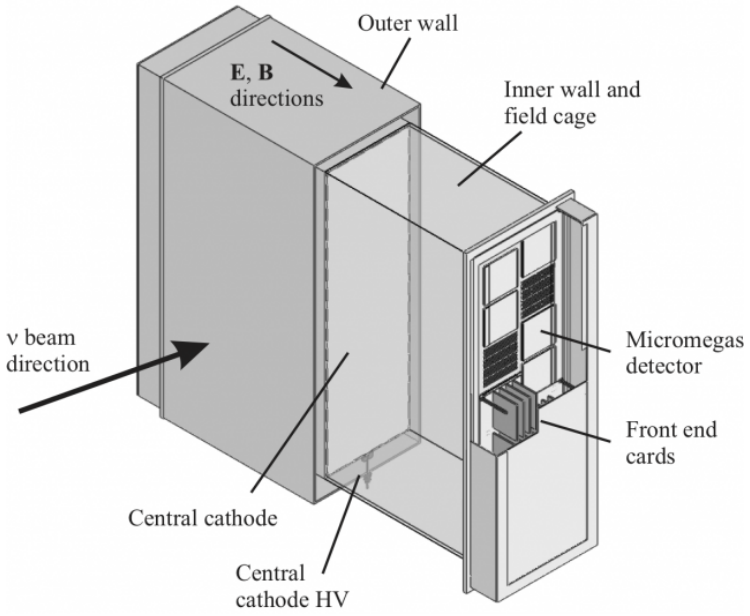


Figure 2.8: Diagram of the TPC design.

trino interactions. Deposited energy resolution of 7.8% was obtained for minimum ionising particles compared to a 10% requirement for a 3σ separation of electrons and muons in the T2K energy region [46].

Fine grained detector (FGD)

The two FGDs are placed respectively downstream of the first and of the second TPC and are composed of finely segmented plastic scintillators oriented perpendicular to the beam in either the x or y direction. The first FGD consists exclusively of plastic scintillator bars arranged in 15 XY modules. The scintillator provides the target mass for neutrino interactions and allows for tracking of charged particles produced in these interactions. The second FGD consists of seven XY modules of plastic scintillator alternating with six 2.5 cm thick layers of water to provide separate determination of exclusive neutrino cross-sections on carbon and on water. Just like with the P0D, each scintillator bar is mounted with a Y-11 fibre mirrored on one side and instrumented on the other side with an MPPC. The electronics are read out with AFTER "ASIC" front end boards rather than TRIP-t.

The aim of the FGDs is to provide target mass for neutrinos in the tracker and track charged particles from interaction vertices.

Electromagnetic calorimeter (ECal)

The ND280 calorimeters are sampling calorimeters made up of alternating layers of plastic scintillator bars for the active material and lead absorber sheets. Thirteen different modules are used to surround the inner detectors (P0D, TPCs, FGDs) on all sides so as to provide a near hermetic coverage for all particles exiting the inner detector volume. Six Barrel-ECal modules surround the tracker volume on its four sides parallel to the direction of the beam and another six surround the P0D. An additional module is placed downstream of the Tracker. The ECals use WLS fibres (single or double ended with the single ended ones being mirrored at the other end), MPPCs and TFBs to digitise the light signal produced by scintillation inside the bars.

The ECals are used to complement the inner detectors in full event reconstruction. The energy and direction of photons are measured for the reconstruction of π^0 decays. Charged particles (electrons, muons and pions) are also detected and particle identification information is extracted.

Side muon range detector (SMRD)

One half of the UA1 magnet yoke consists of 8 C-shaped flux return yokes with each yoke of 16 steel plates arranged radially (like an onion). The yokes are labelled 1 through 8 from upstream to downstream.

The SMRD consists of three layers of scintillator modules on the top and bottom for all yokes. Both sides of each yoke are instrumented with three layers for yokes 1 through 5, four layers for yoke 6 and six layers for yokes 7 and 8. Figure 2.9 presents a diagram of the UA1 magnet and the instrumentation of each yoke. Like the P0D and ECal sub-systems, the SMRD digitises the light signal from the scintillator bars using the combination of WLS fibre, MPPC and TFB.

The purpose of the SMRD is threefold. Muons escaping with high angles with respect to the beam direction are recorded and their momenta measured. The SMRD can provide a trigger on cosmic ray muons that cross the ND280. Finally, the SMRD also helps identify beam-related event interactions in the surrounding cavity walls and the iron of the magnet [47].

2.2.2 Interactive neutrino grid (INGRID)

The INGRID on-axis detector is designed to measure the on-axis beam intensity and direction. Because the energy spectrum at Super-Kamiokande is highly sensitive to the off-axis angle, an accuracy of 1mrad on the beam direction is required.

The detector consists of 14 identical modules arranged as a cross with 7 modules along both the horizontal and vertical axis. An additional two sepa-

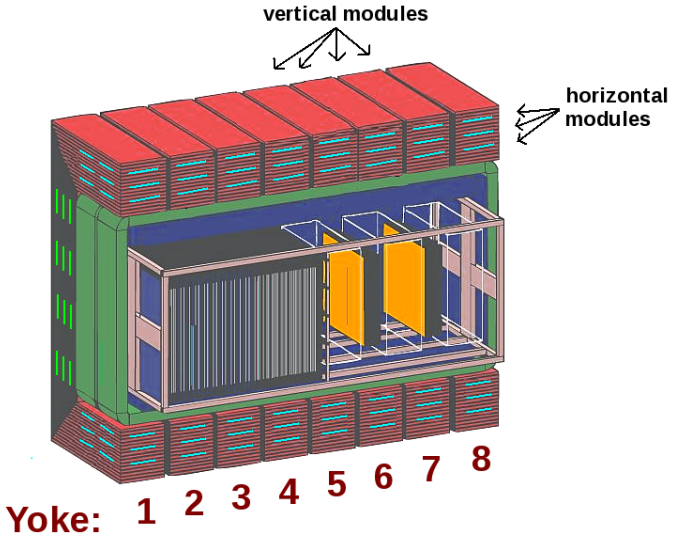


Figure 2.9: Diagram of the UA1 magnet with emphasis on SMRD scintillator modules. The light blue lines represent the 3 layers of horizontal modules for the top and bottom of each yoke. The green line represents the vertical modules on each side of each yoke with 3 layers for yokes 1 to 5, 4 layers for yoke 6 and 6 layers for yokes 7 and 8 (hidden for all yokes except the upstream side of yoke 1).

rate modules are also placed at off-axis directions in order to check the axial symmetry of the neutrino beam. Using the number of observed neutrinos in each module, the beam centre is measured to a precision better than 10 cm or 0.4 mrad.

Each module consists of nine iron plates alternated with 11 tracking scintillator planes surrounded by veto scintillator planes to reject interactions outside the module. The light signal from the scintillator bars is digitised using WLS fibres with mirroring at one end, MPPCs and TFBs.

Figure 2.10 presents the layout of the 16 INGRID modules. Figure 2.11 presents the x and y beam profiles from the horizontal and vertical axis modules fitted with Gaussian functions.

2.3 Far detector: Super Kamiokande

The Super-Kamiokande detector, famous for past measurements of flavor oscillations in atmospheric, solar and accelerator produced neutrinos [25, 49, 26, 50], is used as the far detector in the T2K experiment.

Super-K is a cylindrical water Cherenkov detector of diameter 39.3 m and height 41.4 m built 1 km underground in the Kamioka mine and filled with 50 KTon of ultra pure water (see figure 2.12). Geometrically, it consists of

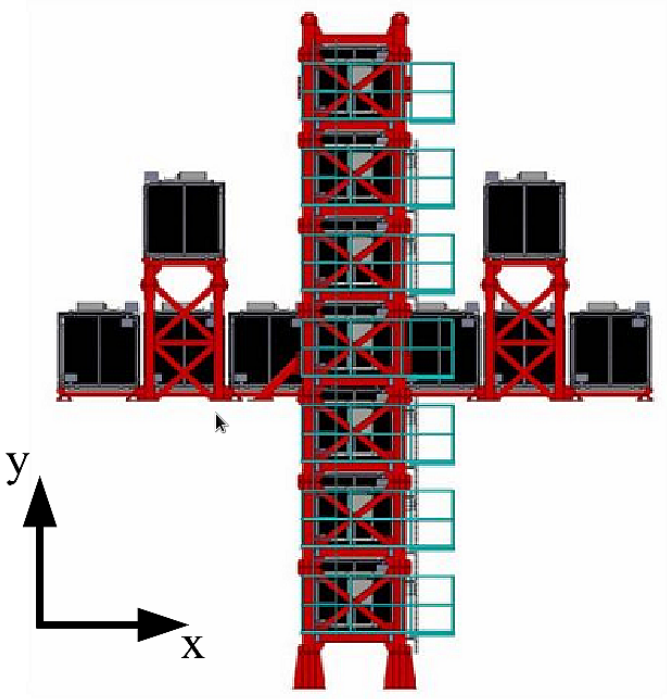
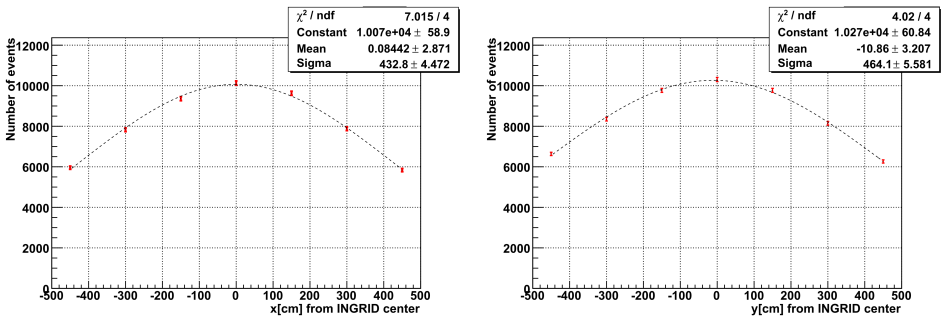


Figure 2.10: INGRID on-axis detector.



(a) Horizontal profile

(b) Vertical profile

Figure 2.11: Neutrino beam intensity on-axis profiles measured by the INGRID detectors in the first T2K run (accumulated exposure of 3.23×10^19 protons on target (POT) [48]. Each data point represents the beam intensity measured by one INGRID module. The fitted beam profiles provide a measurement of the beam centre with precisions of respectively 2.9 cm and 3.2 cm. This level of precision exceeds requirements.

an inner detector (ID) and outer detector (OD) separated by a stainless steel structure.

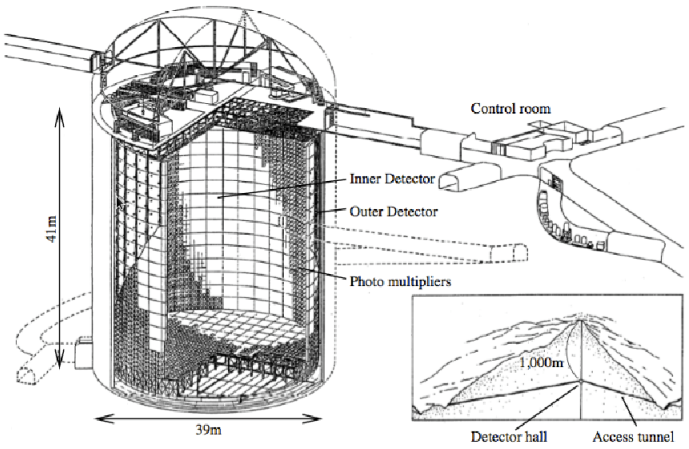


Figure 2.12: Super-Kamiokande detector.

Inner Detector (ID).

The ID is itself a cylindrical volume of diameter 33.8 m and height 36.2 m instrumented on the inner wall with 11,129 50 cm diameter Hamamatsu photomultiplier tubes (PMT) giving a total of 40 % surface coverage.

Outer Detector (OD).

The OD is instrumented on its outer wall to act as an active veto of cosmic ray muons and other backgrounds. In contrast to the ID, the OD is only scarcely populated with 1,885 20 cm PMTs that give a total surface area coverage of approximately 1 %. This is sufficient for a rejection efficiency of cosmic ray backgrounds from the OD of almost 100% given the typical sizes of Cherenkov rings produces when charge particles exceed the speed of light in water. [34].

GPS synchronisation system

The T2K experiment also uses a GPS synchronisation system to select beam events within a time window of $\pm 500\ \mu s$ from a beam spill. The combination of OD veto and GPS synchronisation results in an excellent background rejection for the T2K analysis (see 5.5.2 for more details).

Super-K operation.

Charged particles travelling in the water result in the production of a cone of Cherenkov photons if their energy is above the Cherenkov threshold in water. This is the case when $\beta \geq \frac{1}{n}$ where n is the refractive index of the medium.

The energy threshold for a particle of rest mass m is given by:

$$\begin{aligned} E_{th} &= \gamma m, \\ &= \frac{m}{\sqrt{1-\beta^2}}, \\ &= \frac{m}{\sqrt{1-\frac{1}{n^2}}}, \end{aligned} \tag{2.2}$$

Table 2.3 presents the rest mass and theoretical Cherenkov energy threshold of some relevant particles.

	mass (MeV)	Cherenkov threshold (MeV)
$e^{\pm}$	0.51	0.78
$\mu^{\pm}$	105.66	160.26
$\pi^{\pm}$	139.57	211.70
p	938.272	1423.13
$\tau^{\pm}$	1776.84	2695.04

The Cherenkov light reaches the PMTs on the ID walls in ring-shape patterns from which information can be reconstructed such as the event vertex position or the momentum of the charged particle. The ring topology also allows distinguishing between different lepton flavors:

- Muons have a relatively large rest mass so they experience only small changes in momentum and travel in a relatively straight line. As a result, they produce well-defined cones of Cherenkov light which translate into sharp rings of PMT hits.

- Electrons in contrast have a smaller mass and tend to produce electromagnetic showers at the energies relevant to Super-Kamiokande. This results in "fuzzier" PMT ring patterns on the ID walls.

The difference between an electron and a muon-like event is presented in figure 2.14 for illustration.

The strategy used by Super-K for the measurement of ν_μ oscillation to either ν_e or ν_τ is to use the distinction between lepton rings to count the number of charged current quasi-elastic (CCQE) interactions of each flavor. Statistical intervals of probabilistic values of the oscillation parameters are obtained by comparing the reconstructed energy spectrum of the neutrino parents to the expected observations.

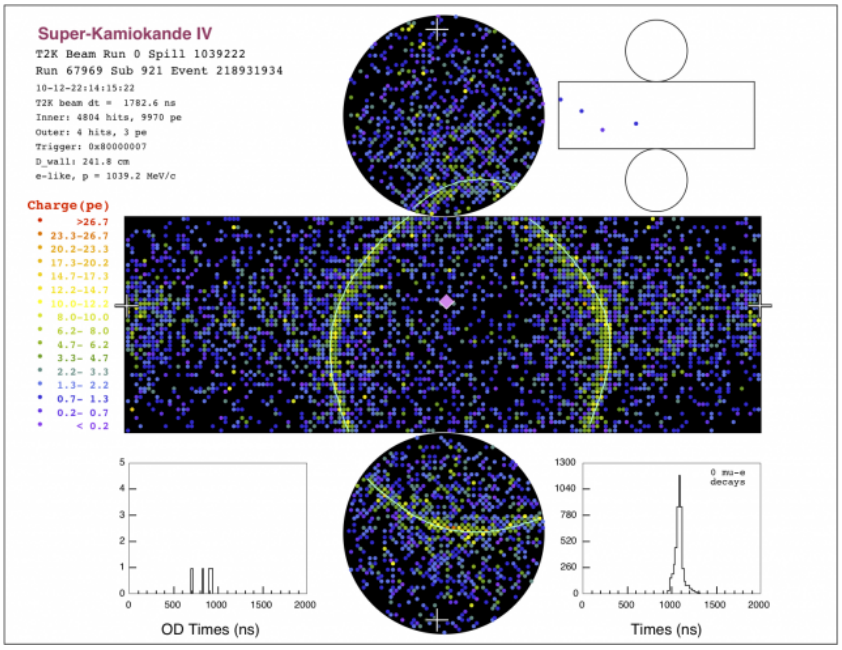


Figure 2.13: Event display of an electron-like ring at Super-Kamiokande [51].

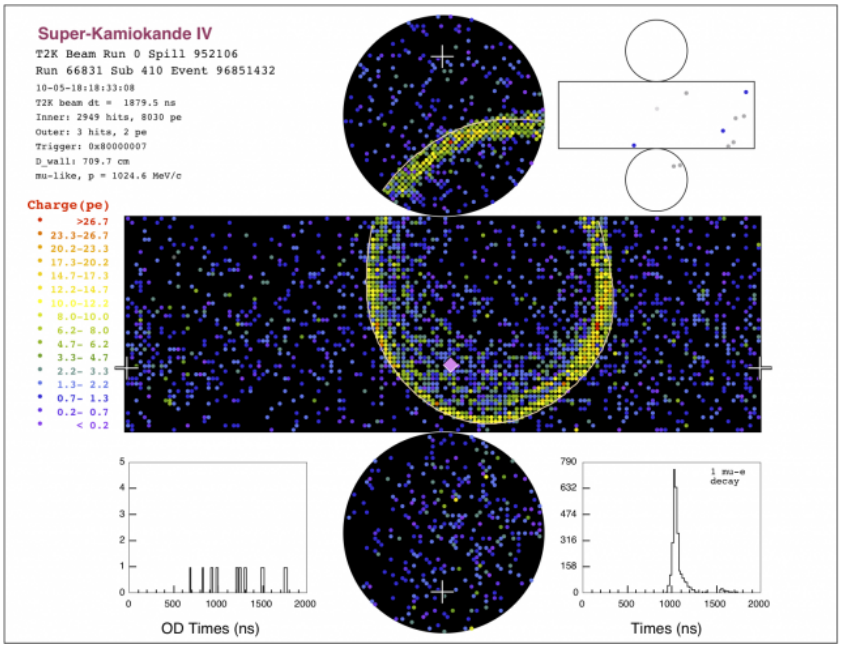


Figure 2.14: Event display of a muon-like ring at Super-Kamiokande [51].

3 Characterisation of the MPPC response with Y-11 fibre and TRIP-T electronics

The P0D, FGD, ECal, SMRD and INGRID sub-detectors all use plastic scintillation material as an active volume to sample the energy of daughter particles in neutrino interactions.

Charged particles that go through these detectors deposit energy in the scintillator bars by ionisation. A fraction of that energy is released as photons when the excited molecules fall back to their ground state. Those photons are then collected by wavelength shifting fibres and channelled to the MPPC, and, finally, the MPPC signal is read out by the electronics (AFTER ASIC for the FGD and TRIP-T front end board for all others).

Energy calibration accurate to a few percent is required for particle identification (PID) purposes with some specificity depending on the sub-detector. Because of the novelty of MPPCs and lack of existing literature, a test bench experiment was designed at Imperial College by A.Vacheret in order to characterise the MPPC's (non linear) response accurately enough to calibrate the measurement of the light signal.

Time calibration is required for the purpose of track reconstruction: timing and hit position are used to match hits into showers or tracks. The test bench experiment was also used to characterise the time-walk effects caused by the fibre and TRIP-t front end boards for calibration purposes.

Energy Calibration

The electronic readout of each sub-detector provides a signal for every channel in an ADC unit. The corresponding energy lost in scintillator material is recovered by applying a succession of calibration and reconstruction steps.

Calibration of the TRIP-t front end board converts the raw ADC signal into a "pixel equivalent unit" (PEU) used to represent the MPPC output charge in a meaningful way. Calibration of the MPPC linearity response provides the conversion from PEU to surface illumination in number of photons. Finally the wavelength shifting (WLS) fibre is calibrated to get the light signal in the scintillator bar. A conversion factor to recover the energy from the fibre signal

is determined using the minimum ionising properties of cosmic muons.

Properties of the MPPC, TFB and WLS fibre are summarised in sections 3.1, 3.3 and 3.4. The test bench procedure and measurements to characterise the MPPC linearity response are described in section 3.5.

3.1 MPPC properties

MPPCs are an array of independent avalanche photo-diodes (APDs) connected in parallel on a single substrate with a common resistive layer and common electrodes, and operated in Geiger mode (with a reverse bias voltage above the breakdown voltage). An MPPC has a finite number of APDs and each APD has a single photon detection capacity: whether one or n photons hit a single pixel at time t, the output signal of that APD will be identical. As a consequence, the MPPC response is intrinsically non linear. Properties are summarised below. T2K used MPPC arrays of 667 APDs. More details about the MPPC properties can be found in [52].

3.1.1 Operating principles

An incident photon striking the MPPC in Geiger mode creates an avalanche discharge with probability given by the photon detection efficiency (PDE). A very large current flow with an amplification gain up to 10^6 can be obtained from a single photon so single photon events can be counted by the device. Each diode has a quenching resistor (R_Q) in series in order to stop the avalanche current and then restore the initial bias condition. In steady state, the pixel capacitances, C_i, are charged at $V_{\text{bias}} > V_{\text{br}}$. When an avalanche discharge is initiated, C_i discharges down to the breakdown voltage with a time constant $\tau = R_Q C$ ns. The output signal of the MPPC is the sum of the charge from all Geiger mode APDs.

3.1.2 Photon Detection Efficiency (PDE)

The photon detection efficiency is defined as the probability for an incident photon to produce an avalanche and is the product of the quantum efficiency, the geometrical efficiency and the Geiger efficiency:

$$PDE = QE * \epsilon_{\text{geom}} * \epsilon_{\text{Geiger}}, \tag{3.1}$$

Quantum efficiency: QE is the quantum efficiency or the probability for a photon to generate an electron-hole pair in the active area of the device. It takes into account the reflection on the surface, which is minimised by implementing an anti-reflective coating, the optical absorption as a function of wavelength and the effective thickness of the depletion layer.

Geometrical efficiency: The geometrical efficiency ϵ_{geom} is the ratio of the active area over the total area of the device.

Geiger efficiency: The Geiger efficiency ϵ_{Geiger} is the probability for a photo-electron to trigger an avalanche.

3.1.3 Gain

The gain, G, for a single pixel is defined by the total charge, Q, collected during the discharge of the pixel capacitor:

$$G = \frac{Q}{|e|} = \frac{C_i}{|e|} \times (V_{\text{bias}} - V_{\text{breakdown}}), \tag{3.2}$$

where $V_{\text{bias}} - V_{\text{breakdown}}$ is the over-voltage.

3.1.4 Dark noise

Dark noise hits are avalanches caused by electrons generated due to thermal noise.

3.1.5 After-pulse

When an avalanche occurs, carriers may be trapped in defects in the silicon. These carriers are then released with a statistical delay and can re-trigger an avalanche. The probability of having an after-pulse increases with the amount of charge during an avalanche and therefore increases with bias voltage. Measurements have characterised that this effect occurs with an exponentially falling time distribution with two decay constants: short and long. At $25°C$, the short decay constant τ_S is 17.6 ± 2.1 ns and the long decay constant τ_L is 71.4 ± 8.3 ns. The intensities of long and short after-pulses are almost equal with total probability of about 0.16 at an over-voltage of $\Delta V = 1.4$ V [52].

3.1.6 Crosstalk

Crosstalk describes the process by which optical photons produced during the avalanche process of the MPPC propagate to neighbouring pixels and trigger further avalanches [53]. As a result, two or more pixels can be fired almost simultaneously with only one initial photo-electron.

3.1.7 Recovery

When a pixel is fired, the voltage across the pixel is reduced from the operating voltage to the breakdown voltage. Recovery mechanisms describe how the pixel voltage is brought back to its initial voltage V_{bias}. For the case of a single pixel fired, charge is drained from other pixels and the recovery happens with

an exponential time constant $R_Q \times C \sim 13.5$ ns with $R_Q = 150\ k\Omega$ and $C = 90$ fF given by Hamamatsu.

3.1.8 Temperature effect

Temperature was measured to affect the breakdown voltage of the MPPC according to equation 3.3:

$$V_{\mathrm{br}}(T^{\circ}C) = V_{\mathrm{br}}(25^{\circ}C) - 0.056\Delta T, \tag{3.3}$$

where $V_{\mathrm{br}}(T^{\circ}C)$ is the breakdown voltage at temperature T, $V_{\mathrm{br}}(25^{\circ}C)$ is the known breakdown voltage at 25° and ΔT is the temperature difference $T - 25^{\circ}C$).

A temperature increase causes a drop in breakdown voltage. For an MPPC held at a constant bias voltage, this results in an increase in Gain and all that this implies (increase in PDE, dark noise, crosstalk, after-pulse...). Control and monitoring of temperature are therefore highly desired features to keep the calibration procedure more manageable.

3.2 MPPC Response

Characterisation of the MPPC response with an analytical function is necessary for the calibration of the light intensity. Due to the complexity of the device, an empirically constructed model was proposed and tested using a test bench measurement. This section describes the construction of this model and justifies approximations made.

3.2.1 The Ideal Response

The device used at the near detector is composed of 667 pixels. In an ideal case using simple statistics, each pixel can be described as a binary device and the MPPC response is obtained by summing up the individual pixel response. Since the number of pixels is finite, the photon detection capacity saturates at high intensity. The MPPC saturation response is described as ideal and following "simple statistics" if characterised by the following assumptions:

- all photons hit the MPPC simultaneously
- the beam cross section is larger than the MPPC
- the distribution of photons is uniform over the array
- the response of the MPPC is Poissonian;

The response is measured as an average number of pixel equivalents (PEU) which is a unit that converts the output signal into a number of pixels triggered at full charge. In the first order approximation, the saturation formula derived in equation 3.4 is a function of the number of effective pixels and the photon detection efficiency

$$\langle N_{\text{pe}} \rangle = N_{\text{pix}} \left(1 - \exp\left(\frac{-N_{\text{phot}} \times PDE}{N_{\text{pix}}} \right) \right), \tag{3.4}$$

where the number of effective pixels N_{pix} is the number of illuminated pixels, N_{phot} is the number of photons hitting the MPPC surface and PDE is the photon detection efficiency.

Figure 3.1 presents a plot of the deviation from linearity in this idealised case as a function of the number of photons times the PDE for MPPCs with respectively 100, 400, 667 and 1600 pixels.

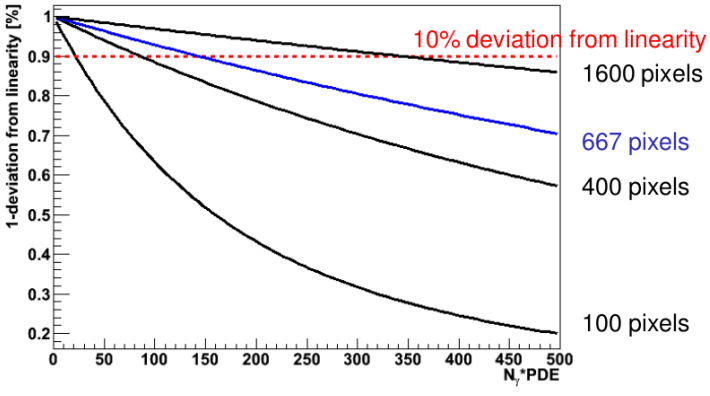


Figure 3.1: Simulation of the 10% deviation from linearity as a function of the number of photons times the PDE for MPPCs with respectively 100, 400, 667 and 1600 pixels.

3.2.2 Correlated noise effects

Correlated noise effects cause the response to deviate from an ideal behaviour.

Crosstalk increases the total number of pixels triggered and skews the charge distributions towards higher values.

After-pulse increases the output charge of affected pixels by creating a new avalanche. The combination of after-pulse exponential decay time and pixel recovery time results in avalanche with varying gain and skews the charge distribution to higher charge values.

Crosstalk and after-pulse are two correlated noise effects that impact the response of the MPPC and break the ideal Poissonian assumption used to derive equation 3.4. Other remaining first order assumptions: timing simultaneity,

active pixel assumption and uniformity of the distribution of photons over the array are broken by the Y-11 fibre and the finite acceptance coupling between the MPPC and fibre.

Although the assumptions used to derive the analytical formula (equation 3.4) clearly don't hold under experimental conditions, we have verified empirically that it can still be used to describe the MPPC saturation accurately with minor changes (section 3.7). The saturation function with correlated noise effects is given by equation 3.5.

$$\langle N_{\text{pe}} \rangle = N_{\text{eff}} \left(1 - \exp\left(\frac{-N_{\text{phot}} \times CN}{N_{\text{eff}}} \right) \right), \tag{3.5}$$

where:

- N_{phot} is the number of photons hitting the MPPC.
- N_{eff} is the number of effective pixels and an unknown non linearity component (due to after pulse effects, fibre to MPPC connection, beam profile on the MPPC effective area and fibre decay time).
- CN is the correlated noise factor, equivalent to the PDE with some excess due to non-linearity effects (crosstalk and after-pulse).

3.2.3 Signal Resolution

The signal resolution $\frac{\sigma_\gamma}{\langle N_\gamma \rangle}$ of the MPPC response is given by:

$$\frac{\sigma_\gamma}{\langle N_\gamma \rangle} = \frac{f^{-1}\left(\langle N_{\text{pe}} \rangle + \sigma_{\text{pe}}\right) - f^{-1}\left(\langle N_{\text{pe}} \rangle\right)}{f^{-1}\left(\langle N_{\text{pe}} \rangle\right)}, \tag{3.6}$$

where σ_{pe} is the resolution on the pe value and $f^{-1}\left(\langle N_{\text{pe}} \rangle\right)$ is the inverted MPPC response function, or linearised signal.

The linearised signal gives the average value $\langle N_\gamma \rangle$ of the number of photons that hit the MPPC. Inverting equation 3.5:

$$\langle N_\gamma \rangle = -\frac{N_{\text{eff}}}{CN} \ln\left(1 - \frac{\langle N_{\text{pe}} \rangle}{N_{\text{eff}}} \right), \tag{3.7}$$

Assuming that σ_{pe} from equation 3.6 is small,

$$\sigma_{\text{pe}} = N(pe + \sigma_{\text{dist}}) - N(pe) = N(pe) - N(pe - \sigma_{\text{dist}}), \tag{3.8}$$

where σ_{dist} is the sigma of the skew Gaussian fit to the ADC signal in the low gain channel, divided by the peak value.

The final signal resolution approximation is given by:

$$\frac{\sigma_\gamma}{\langle N_\gamma \rangle} \simeq \frac{\ln\left(1 - \frac{N(pe+\sigma_{\rm dist})}{N_{\rm eff}}\right) - \ln\left(1 - \frac{N(pe)}{N_{\rm eff}}\right)}{\ln\left(1 - \frac{N(pe)}{N_{\rm eff}}\right)}, \tag{3.9}$$

3.3 TRIP-t front end board (TFB)

3.3.1 TFB Operation

TRIP-t front end boards with four TRIP-t ASICs chips of 16 channels each are used to read out the MPPCs which are connected to the TRIP-t board by miniature coaxial cables. An FPGA (Field Programmable Gate Array) implements time-stamping with a discriminator, sequences the operation, digitises and reads out the TRIP-t chips and formats and transmits the data off-board. The boards have two channels in order to accommodate the whole range of signals while allowing sufficient precision in the discriminator setting: the signal capacity is divided between high gain and low gain channels with a 10 to 1 gain ratio. Figure 3.2 shows schematically how the MPPC interfaces with the TRIP-t electronics. The high to low channel gain ratio is set with the C_{HI} and C_{LO} capacitances and the value of C_G is used to adjust the overall gain. The front end boards integrate the signal in cycles. A cycle has 23 integration periods of 480 ns (540 ns in our setup) with reset periods of 100 ns and two readout modes: cosmic and beam mode. In cosmic mode, a sequence of cycles is continuously running. A trigger stops the cycling and the last 23 integration periods are kept. In beam mode, a cycle of 23 integration periods is triggered at the target and the six bunches of the beam fall in consecutive integration periods at the near detector starting from the 5th period (figure 3.3).

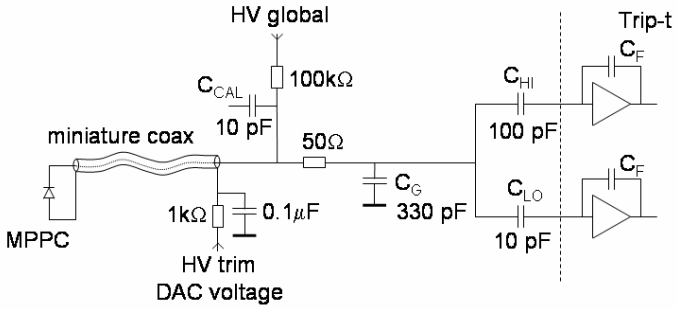


Figure 3.2: Schematic of the MPPC/TRIP-t channel architecture interface showing the charge splitting between low and high gain channels and MPPC bias scheme [43].

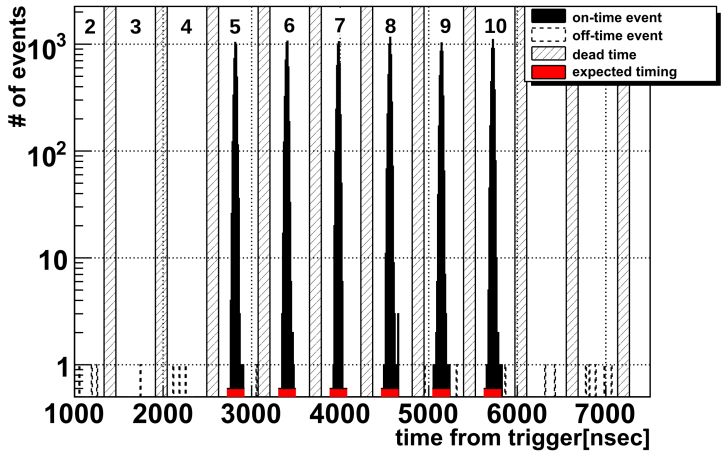


Figure 3.3: Ingrid event timing distribution for T2K runs from 2010 with six bunches. [48].

3.3.2 Electronics time walk

When a signal is received, the TRIP-t discriminator sets the time-stamping after the charging capacitor exceeds the threshold value. The charging of a capacitor is a function of the voltage V across the capacitor and the capacitance C and impedance R of the circuit:

$$Q = C \times V, \tag{3.10}$$

The charge reaches the discriminator threshold Q_{Th} value after a time t_{Th}:

$$Q_{Th} = C \times V_0 \left(1 - \exp\left(-\frac{t_{th}}{\tau}\right)\right), \tag{3.11}$$

$$t_{Th} = -\tau \times \ln\left(1 - \frac{Q_{th}}{Q_0}\right), \tag{3.12}$$

where $\tau = RC$ is the decay constant.

Figure 3.4 shows the charge as a function of time for 3 potential voltage values that correspond to final charges in the capacitor of Q1, Q2 and Q3, where $Q1 < Q2 < Q3$. This figure illustrates that the charge of the capacitor exceeds the threshold value faster for larger signals: $t1 > t2 > t3$. This effect is called the electronic time-walk.

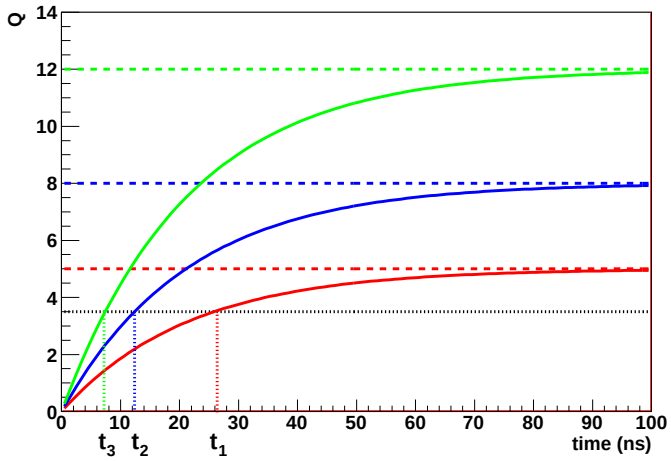


Figure 3.4: Charge as a function of time for 3 potential voltage values that correspond to final charges in the capacitor

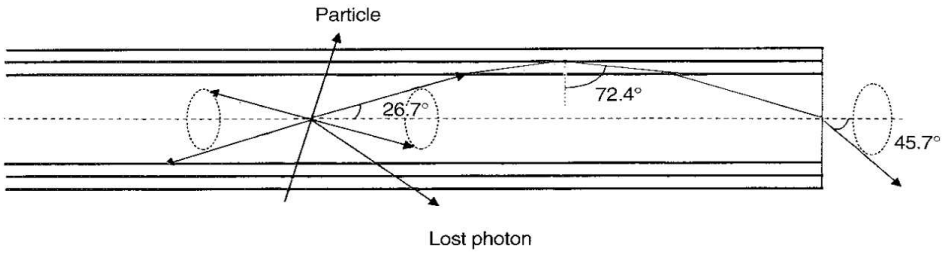


Figure 3.5: Y-11 scintillating fibre diagram. The Y-11 fibre is composed of a Polystyrene (PS) core of refractive index 1.59, a Polymethylmethacrylate (PMMA) inner cladding of refractive index 1.49 and a Fluorinated polymer (FP) outer cladding of refractive index 1.42. The critical angle inside the core is 26.7° and the cone of light angle is 45.7° (taken from Kuraray documentation)

3.4 Y-11 fibre

3.4.1 Transmission mechanism

T2K ND280 sub-detectors use Kuraray Y-11 wavelength shifting fibres with core diameter of 1 mm and multi-cladding. A diagram of the fibre with transmission mechanism is presented in figure 3.5.

Photons absorbed in the fibre are re-emitted isotropically with an emission peak of 476 nm and an exponential time decay of an order of 9 ns [42]. Those that are emitted with an angle smaller than the critical angle in the core (26.7%) then diffuse along the fibre to the MPPC, or, in the case of single ended bars, are reflected at the dead end back towards the MPPC.

The cone of light angle of photons leaving the fibre is 45.7^o. The thickness

of the fibre is 1 mm and the square length of the MPPC is 1.3 mm. The transmission mechanism and coupling distance between the MPPC and fibre therefore affect the distribution of photons, or light profile, over the active area of the MPPC.

3.4.2 Fibre time-walk

Coupled with the TRIP-t discriminator threshold, the fibre transmission mechanism creates a second time-walk effect. The discriminator threshold timestamps hits when enough photons have created a signal. Because of the exponential re-emission decay and diffusion path along the fibre, the time-stamp delay after the first photon has been detected will be affected by the number of photons above threshold of the hit, i.e the charge. The effect the re-emission decay probability has on the electronics has been studied and simulated with the following model:

Assuming that the photon distribution out of the fibre follows an exponential decay with time constant τ, the probability that the n^{th} photon out of N is detected at time t is given by:

$$P(t,\tau,N,n) = \exp\left(-\frac{t(N-n)}{\tau}\right)\left(1-\exp\left(-\frac{t}{\tau}\right)\right)^{n}, \tag{3.13}$$

where n is the threshold value in pe and N is the number of photo-electrons. Time-walk is then calculated from the first moment:

$$m_{\mu}(\tau,N,n) = \frac{\int tP(t,\tau,N,n)dt}{\int P(t,\tau,N,n)dt}. \tag{3.14}$$

This equation has analytical solutions that can be calculated for each threshold value n and are function of N and τ only.

3.5 MPPC linearity response

Two test bench measurements were used for the following purposes:

- Measurement of the photon detection efficiencies;
- Determination of empirical fit parameters to match the MPPC response function with TFB and TFB plus Y-11 fibre;
- Calculation of the signal resolution given by the linearised MPPC response function;
- Determination of fit parameters for the electronics and fibre time-walk models.

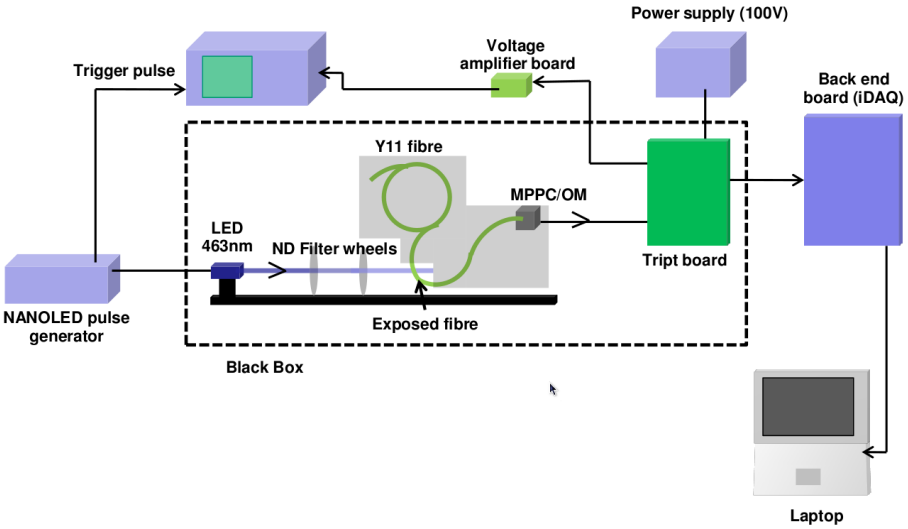


Figure 3.6: Diagram of the 2^{nd} stage of the experiment

The MPPC response, photon detection efficiencies and signal resolutions were characterised as a function of over-voltage with temperature variations in the lab kept below $1°C$.

The first test bench was set up to illuminate the MPPC uniformly and in short pulses in order to measure the response in a near ideal case and decoupled from any fibre effect. The second test bench was set up to illuminate the MPPC surface with light from a Y-11 fibre in order to measure the MPPC response in conditions that approximate the ND280 environment.

3.6 Experimental setup and calibration

Figure 3.6 presents a diagram of the second experimental setup built by Antonin Vacheret at Imperial College in order to characterise the MPPC response with Y-11 fibre and TRIP-t Front end Board (TFB).

An LED of wavelength 463 nm and FWHM of approximately 1 ns coupled with a pulse generator of frequency $10^5 Hz$ controlled by an external trigger were used to make short pulses of light on a small section of a Y-11 fibre. An MPPC of dimensions $1.3\text{mm} \times 1.3\text{mm}$ with 667 pixels was mounted on the end of a Y-11 fibre and connected for read out to a TFB and Read-Out Merger Module (RMM). The light intensity was controlled using two neutral density filter wheels placed on railings with 6 filters on each wheel allowing for a total of 49 configurations. Transmissivity values provided by the constructor are given below:

- wheel 1: 0.7, 0.5, 0.4, 0.25, 0.125, 0.003
- wheel 2: 0.63, 0.32, 0.20, 0.10, 0.03, 0.01

A 918D-UV-OD3 Newport optical metre with 10 mm diameter circular area was used to measure the beam power and verify the transmissivity of the density filters. Finally, the whole apparatus was placed in a black box to prevent contamination from external light.

The first test bench experiment was set up similarly. Instead of a fibre, a diffuser was placed in the path of the beam in front of an MPPC mounted on an X coordinate control stage (figure 3.7).

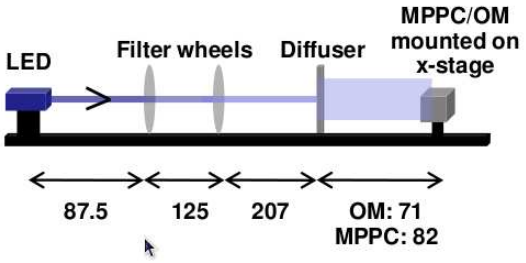


Figure 3.7: Diagram of the optical beam-line for the first experimental setup.

3.6.1 Optical Meter Calibration

The OM measures the power of the LED pulse on a circular surface of diameter 10 mm and is used to calculate the number of photons per pulse using the formula:

$$N_{\gamma}^{OM} = \frac{P}{E \cdot f} = \frac{P}{f} \times \frac{\lambda}{hc}, \tag{3.15}$$

where P is the power measured with the Optical meter in Watts, f is the frequency of the LED in Hertz and λ is the mean wavelength of the LED.

The number of corresponding photons per pulse hitting the MPPC surface before correction for flux non uniformity is:

$$N_{\gamma}^{MPPC} = N_{\gamma}^{OM} \cdot S_R, \tag{3.16}$$

where S_R is the ratio of active surfaces between the MPPC and the OM:

$$S_R = \frac{S_{MPPC}}{S_{OM}} = \frac{l_{MPPC}^2}{\pi R_{OM}^2}, \tag{3.17}$$

where $l_{MPPC} = 1.3$ mm is the square side of the active surface of the MPPC and $R_{OM} = 5$ mm is the radius of the OM.

An additional flux non uniformity correction is applied by measuring the flux variations in the X-direction using the MPPC. The non uniformity over a 40 mm range in the X direction is plotted in figure 3.8 and a Gaussian fit is applied close to the maximum of the profile.

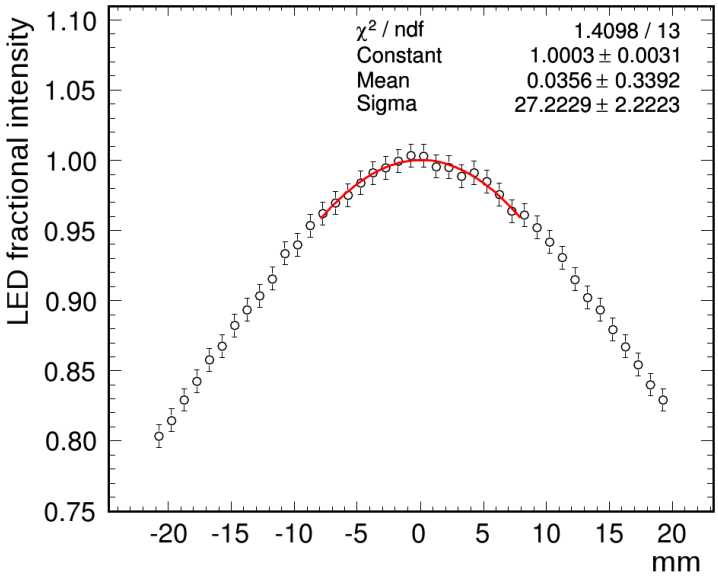


Figure 3.8: LED pulse X-profile over a length of 40 mm measure with the MPPC. A Gaussian fit is applied close to the maximum of the profile

The integrated power on the MPPC and OM are measured by integrating the area under the Gaussian fit in the X direction and assuming circular symmetry of the beam shape. For the MPPC, the area is integrated from -0.65 to 0.65 mm and from -5 to 5 mm for the OM. The ratio of integrated power from the circular optical meter to the square MPPC is given by:

$$\frac{F_{MPPC}}{F_{OM}} = \frac{I^2_{MPPC}}{\pi(I_{OM}/2)^2} = \frac{4}{\pi} \times \frac{I^2_{MPPC}}{I^2_{OM}}, \tag{3.18}$$

where F_{MPPC} is the MPPC square integrated profile and F_{OM} is the OM circular integrated profile. The final flux correction is given by:

$$F_{\text{cor}} = \left(\frac{R_{MPPC}}{R_{OM}}\right) / \left(\frac{F_{MPPC}}{F_{OM}}\right) \approx 0.98 \tag{3.19}$$

3.7 Results

Results are presented for both test bench measurements.

3.7.1 Error Analysis

Statistical errors:

Intrinsic PDE is measured using the fraction of pedestal to total events: $\lambda = -ln\left(\frac{N_0}{N_{\text{total}}}\right)$ (see section 3.7.3). The error in the fraction of pedestal to

total events is given by:

$$\sigma_s = \sqrt{\frac{1}{N_{\text{zero}}} + \frac{1}{N_{\text{total}}}} \tag{3.20}$$

Due to correlations, an error on the saturation and resolution measurements is determined from upper and lower bounds on the function parameters

Systematic errors:

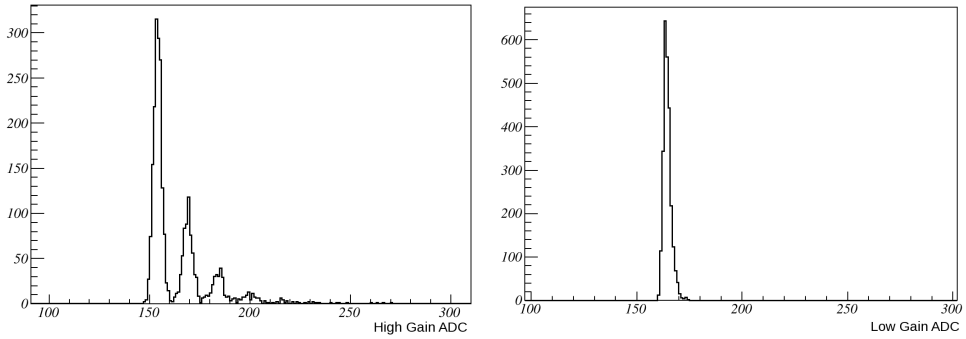
An uncertainty over the range of the OM measurements of 1% was given by the manufacturer detector calibration report. In addition, fluctuations in the power and baseline of 5.10^{-13} W, corresponding to approximately 0.25 photons were determined empirically.

Temperature in the experimental environment was monitored on a regular basis for the determination of temperature induced systematic uncertainties. Fluctuations of $\pm 1^0 C$ were measured, corresponding to a $0.056V$ maximal variation in breakdown voltage. By comparison, daily and seasonal variations in temperature of several degrees are observed in the ND280 pit.

Temperature, which affects the breakdown voltage of the MPPC, was monitored in the laboratory. Variations in a range of $\pm 1^0 C$ were measured corresponding to a $0.056V$ variation in breakdown voltage.

3.7.2 Pedestal data

The pedestal is the average TRIP-t electronic response of the MPPC in the absence of any signal. The response is non-zero in part due to intrinsic systemic noise and in part due to a deliberate bias introduced in order to avoid low ADC channels where the response is less linear. Typical pedestal peaks in the high and low gain channels are presented in figure 3.9. The photo-electron structure is visibly resolved in the high gain channel. In the low gain channel, additional peaks are merged with the pedestal and therefore indistinguishable.



(a) Hi-Gain channel (b) Low-Gain channel

Figure 3.9: Typical pedestal peaks in the high and low gain channels. Several photo-electron peaks are visible in the high gain channel. In the low gain channels, additional peaks are merged with the pedestal and therefore indistinguishable.

3.7.3 PDE extraction

The PDE was extracted at low photon intensity based on the fit of the average number of Pixel Equivalent Unit, $\langle N_{\text{peu}} \rangle$, versus the estimated number of photons hitting the active area of the MPPC N_γ. At low intensity, saturation is negligible so the relation is linear and the gradient of the fit to the data gives the PDE.

$\langle N_{\text{peu}} \rangle$'s were estimated for each intensity N_γ, measured with the OM (section 3.6.1), by first taking repeated MPPC readings in PEU at each intensity and assuming that they follow a Poisson distribution:

$$P(n, \lambda) = \frac{\lambda^n e^{-\lambda}}{n!}$$

where λ is the average number of PEU and n is the PEU value.

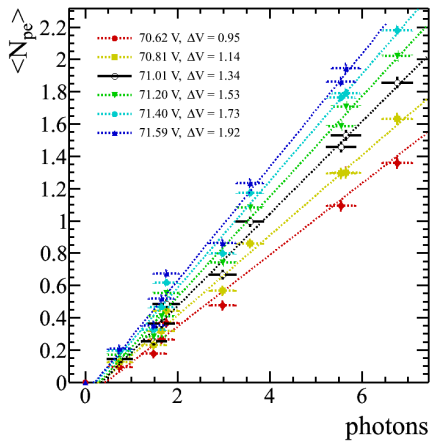
Using the pedestals which are not affected by crosstalk and after-pulse, the average intrinsic PDE value un-affected by correlated noise was measured:

$$\begin{aligned} P(0, \lambda) &= e^{-\lambda} \\ \Rightarrow \lambda &= -\ln\left(\frac{N_0}{N_{\text{total}}}\right) \end{aligned} \tag{3.21}$$

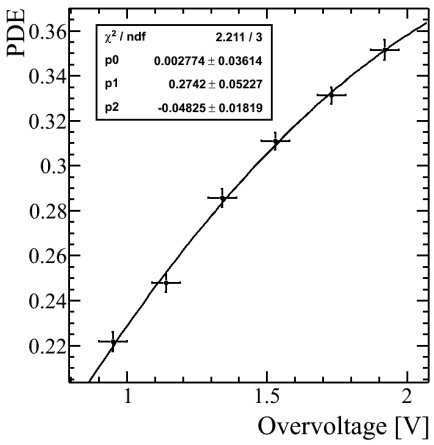
where 0 stands for the pedestal, N_0 is the number of events in the pedestal and N_{total} is the total number of events.

Figures 3.10(a) and 3.10(c) present measurements of $\langle N_{\text{peu}} \rangle$ versus N_γ for over-voltages of 0.95 V, 1.14 V, 1.34 V, 1.53 V, 1.73 V and 1.92 V for both experimental setups. The fibre length to MPPC is 50 cm for the setup with

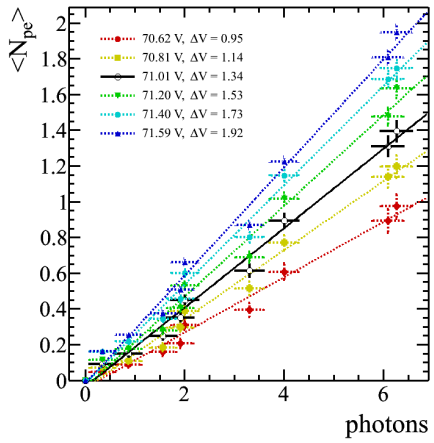
fibre. First order polynomials are used to fit the data. The PDEs are extracted from the fit, plot in figures 3.10(b) and 3.10(d) versus over-voltage and fitted with second order polynomials.



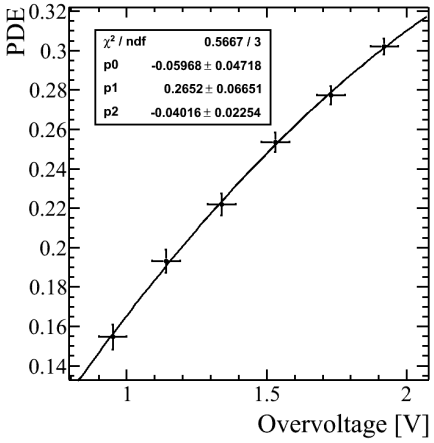
(a) 1st setup (no fibre)



(b) 1st setup (no fibre)



(c) 2nd setup



(d) 2nd setup

Figure 3.10: (a) and (c): Average number of pe as a function of the number of photons hitting the active surface of the MPPC for over-voltages of 0.95 V, 1.14 V, 1.34 V, 1.53 V, 1.73 V and 1.92 V, fitted with first order polynomials. (b) and (d): PDE as a function of over-voltage fitted with second order polynomials.

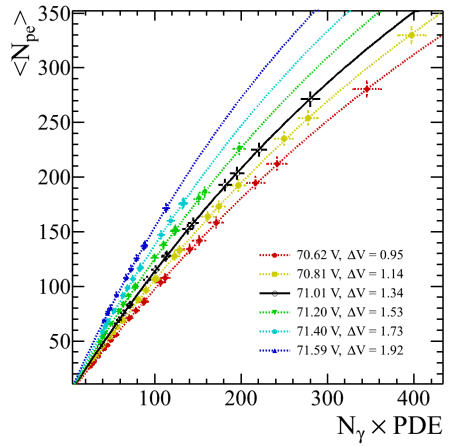
The PDE for the second experimental setup with Y-11 fibre was found to be smaller than for the first by several percentage points. This difference was attributed to a loss of photons caused by the coupling between the MPPC and fibre.

3.7.4 Saturation response

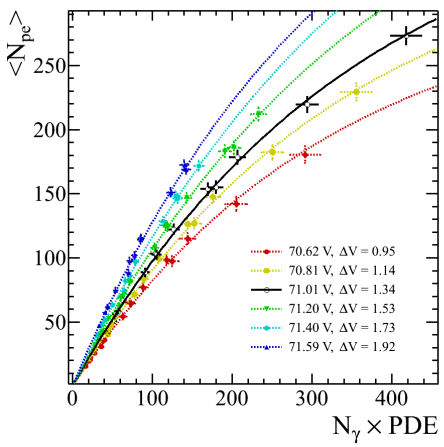
The saturation response was characterised by fitting the saturation functions from equation 3.5 to $\langle N_{\text{peu}} \rangle$ versus N_γ for high photon intensities (over 50 pe) and extracting the fitted model parameters. For the first experimental setup, the number of effective pixels N_{eff} was fixed to 667: the number of pixels on the MPPC surface.

Figures 3.11(a) and 3.11(b) present the fitted saturation curves for overvoltages of 0.95 V, 1.14 V, 1.34 V, 1.53 V, 1.73 V and 1.92 V for both experimental setups. The fibre length to MPPC is 50 cm for the setup with fibre. The fit residuals for the first experimental setup with no fibre and uniform illumination, plotted on figure 3.11(c), are below 1% and uncorrelated. Residuals for the second experimental setup with Y-11 fibre, plotted on figure 3.11(d), are mostly below 5% and also uncorrelated indicating. The test bench measurements results therefore demonstrate that the derived saturation model is a good approximation for the description of the MPPC response.

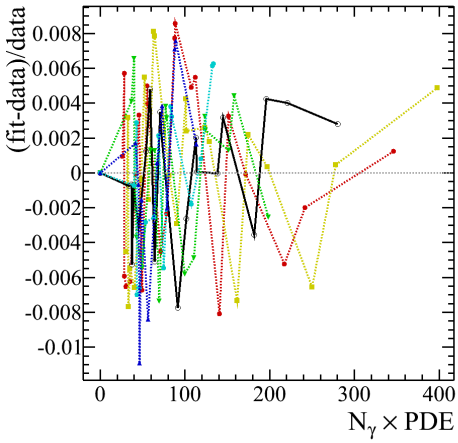
The CN fit parameters for both experimental setups are plotted as a function of voltage and fitted with second order polynomials. The N_{eff} fit parameter for the second experimental setup is plotted on figure 3.12(c) and fitted with a first order polynomial. Both fit parameters are plotted as a function of voltage and fitted with respectively first and second order polynomials (see figures 3.12(b) and 3.12(c)).



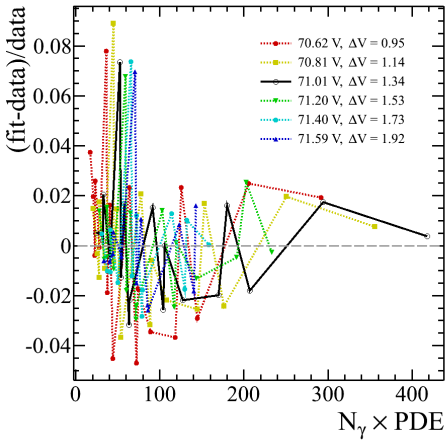
(a) 1st setup (no fibre)



(b) 2nd setup

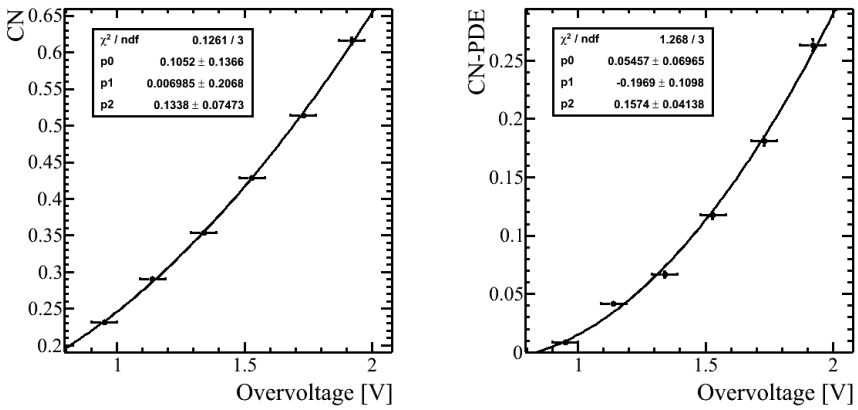


(c) 1st setup (no fibre)

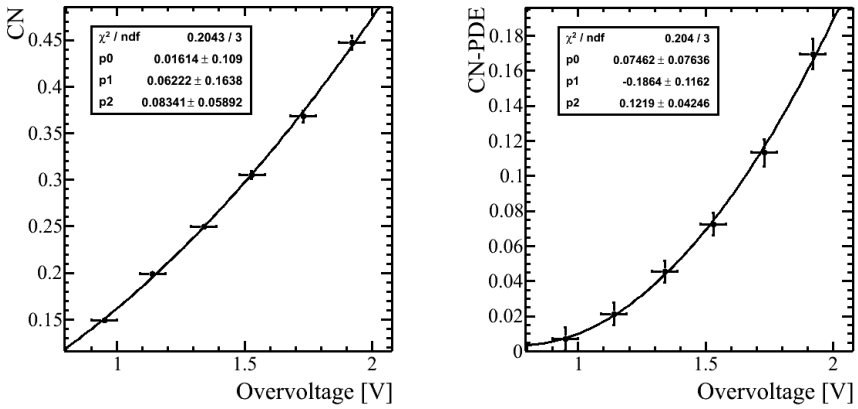


(d) 2nd setup

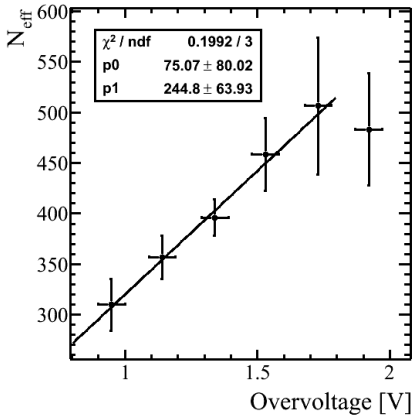
Figure 3.11: (a) and (b): Average number of pe as a function of the number of photons hitting the active surface of the MPPC for over-voltages of 0.95 V, 1.14 V, 1.34 V, 1.53 V, 1.73 V and 1.92 V, fitted with equation 3.5.
(c) and (d): Residuals of the saturation fits. The X axis is the number of photons and the Y axis is the fractional residual (fit-data)/data. Residuals for the first experimental setup (c) are below 1% and uncorrelated. This shows excellent agreement between the derived saturation model and test bench data with no fibre and uniform illumination of the MPPC surface. Residuals for the second setup (d) are mostly inferior to 5% and uncorrelated which also implies good agreement with the model.



(a) 1st setup (no fibre)



(b) 2nd setup



(c) 2nd setup

Figure 3.12: (a) and (b): Correlated Noise (CN) fit parameter of the saturation data as a function of voltage fitted with second order polynomials on the left. CN with intrinsic PDE subtracted on the right.
(c): Fit parameter N_{eff} of the saturation data as a function of voltage fitted with a first order polynomial from 0.8 V to 1.8 V.

3.7.5 Signal resolution

The signal resolution was calculated using the approximation derived in section 3.2.3 and fitted based on the particle data group energy resolution function for photon detectors (eq 3.22):

$$\frac{\sigma(E)}{E} = \sqrt{\frac{f_N}{n_\gamma PDE} + \left(\frac{N_e}{G n_\gamma PDE}\right)^2}, \tag{3.22}$$

where E is the energy, N_e the electronic noise, f_N the excess noise factor and n_γ, G and PDE are respectively the number of photons, the gain and the PDE.

The actual function used for the fit expresses the resolution in terms of the number of photons and 3 variable parameters:

$$\frac{\sigma(N)}{N} = \sqrt{\frac{\mathrm{t_{f_N}}}{n_\gamma} + \left(\frac{\mathrm{t_{N_e}}}{n_\gamma}\right)^2 + \mathrm{C}}, \tag{3.23}$$

where:

- t_{f_N} stands for excess noise term and is equivalent to the excess noise factor f_N over the PDE,
- t_{N_e} stands for electronic noise term and is equivalent to the electronic noise N_e over the gain and PDE,
- C is an extra constant term

The MPPC resolution combines both statistical and detector effects. The irreducible statistical component can be obtained separately using equation 3.24 below:

$$\frac{\sigma_{\mathrm{stat}}}{N_\gamma} = \frac{1}{\sqrt{N_\gamma PDE}}, \tag{3.24}$$

The detector resolution effects can then be inferred by subtracting the statistical component of the resolution from the total.

The signal resolution for the setup with fibre is plotted as a function of the number of photons for over-voltages of 0.95 V, 1.14 V, 1.34 V, 1.53 V, 1.73 V and 1.92 V and fitted with equation 3.23 (figure 3.13). The statistical component for an over-voltage of 1.34 V (dashed black line) and the detector resolution are also added to the plot.

Two competing effects shape the detector resolution: crosstalk and saturation. The fraction of pixels triggered by crosstalk is larger at low luminosity where the saturation effect is smaller. At high luminosity, the MPPC saturates and most pixels are triggered by optical photons. Overall, for an over-voltage value of 1.34 V, the resolution is better than 5%.

The fit parameters t_{f_N}, t_{N_e} and C are plotted as a function of over-voltage and fitted with second order polynomials (figure 3.14).

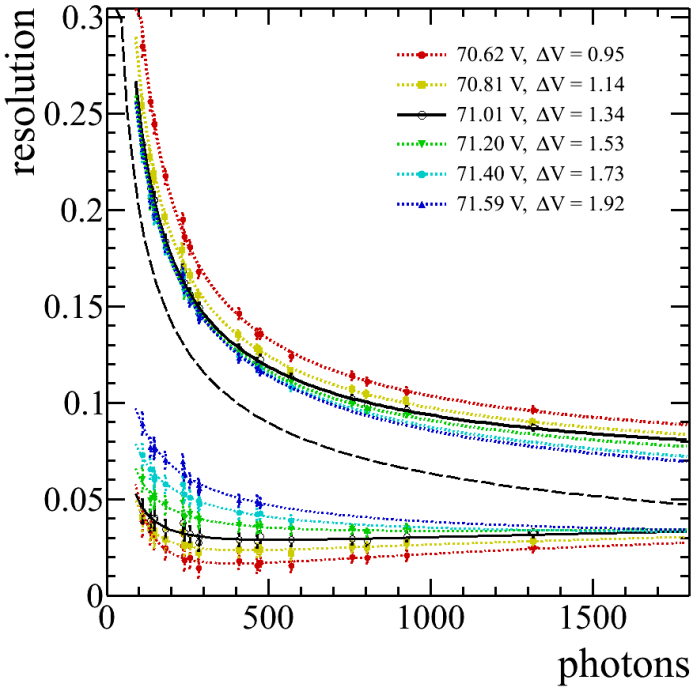


Figure 3.13: Total Signal resolution (top) and statistical (dashed black line) and detector components (bottom) for over-voltages of 0.95 V, 1.14 V, 1.34 V, 1.53 V, 1.73 V and 1.92 V.

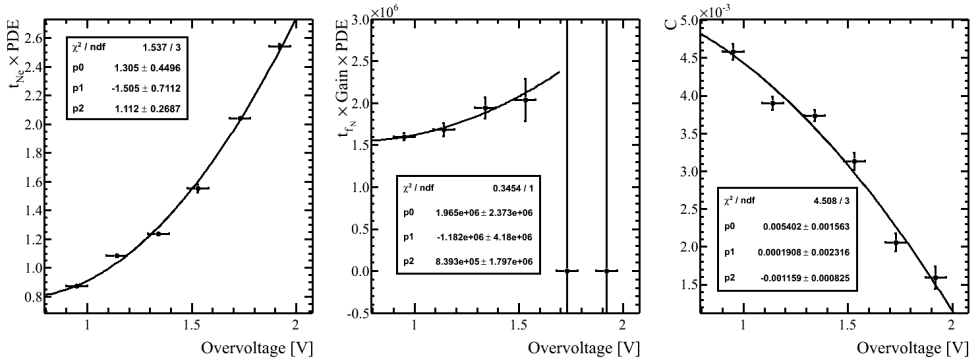


Figure 3.14: Parameters of the resolution fit, fitted with second order polynomials. (a): electronic noise, (b): excess noise factor (c): constant term

3.8 Time Calibration

3.9 TRIP-t time-walk

Figure 3.15(a) presents the results of the electronics time-walk in the hi gain channel, fitted with the three variable parameter fit below:

$$t_{Th} = A - \tau \ln\left(1 - \frac{Q_{Th}}{Q_0}\right), \tag{3.25}$$

where Q_0 is the charge with pedestal subtracted and the three fit parameters are:

- A, the initial time of the event,
- τ, the decay constant from the electronics RC value (expected τ value from the circuit impedance and capacitance: 21.5 ns),
- Q_{Th}, the threshold charge with pedestal subtracted in ADC.

A 3.5 pe threshold was used in the test bench measurements so the threshold charge in ADC is given by

$$Q_{Th} = \text{pedestal} + 3.5 \times G$$

, where G is the gain. The data were fitted from 0.5 pe after the threshold up to 500 ADC in order to avoid threshold effects and electronics non-linearity effects. The fitted parameter value for τ is 21.4 ± 0.2, in good agreement with the expected value. Residuals are presented in figure 3.15(b) and are all below 10%, apart from one outlier. The electronics model is therefore considered to be a good approximation to the data. For longer cable lengths, a parametrisation of MPPC cable reflection effects might be required to improve the validity of the model.

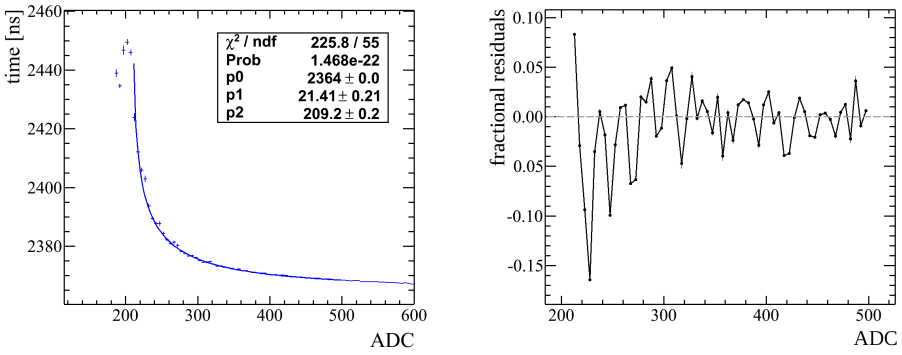
3.10 Fibre time-walk

Figure 3.15(a) presents the results of the electronics and scintillation time-walk in the hi gain channel for fibre lengths of 100, 200 and 300 cm. The scintillation time-walk fit is derived using equation 3.14 for a 3.5 pe threshold:

$$Y11Fit = \frac{\frac{1}{N^2} - \frac{4}{(N-1)^2} + \frac{6}{(N-2)^2} - \frac{4}{(N-3)^2} + \frac{1}{(N-4)^2}}{\frac{1}{N} - \frac{4}{(N-1)} + \frac{6}{(N-2)} - \frac{4}{(N-3)} + \frac{1}{(N-4)}} \times \tau_{\text{fibre}}, \tag{3.26}$$

where N is the number of photons and τ_{fibre} is the fibre decay constant.

The combined fit for both electronics and scintillation time-walk is the sum of equations 3.25 and 3.26 and is a function of four parameters. To reduce

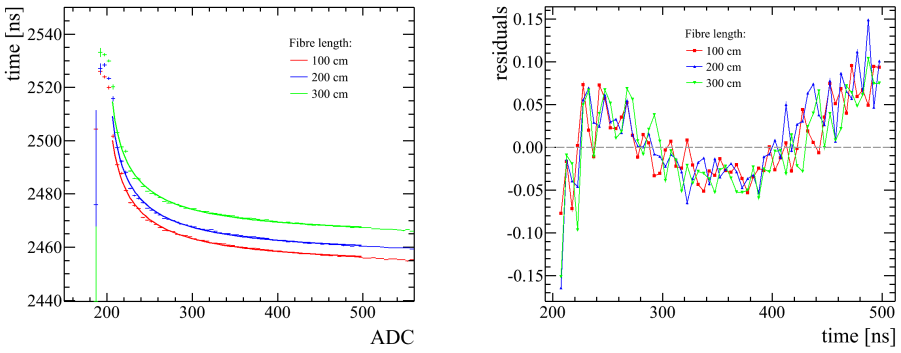


(a) (b)

Figure 3.15: (a) MPPC Electronics time-walk effect in the high gain channel, fitted with a three parameter fit derived from the discriminator threshold effect. (b) Residuals between the electronics time-walk fit and data.

the number of free parameters, τ is set to the experimentally measured value of 21.5 ns and Q_{Th} is calculated for each set of data. The remaining two free parameters are the initial time constant A and the fibre decay constant τ_{fibre}.

The fitted fibre decay constant values for the three fibre lengths are, in order of increasing length, 7.97 ± 0.08, 7.13 ± 0.09 and 6.41 ± 0.08 where the errors on the fit are statistical only. Residuals for the fit are plotted in figure 3.16(b) and show a clear bias indicative of either a flaw in the model or a non-linear systematic error in the measurements. A study of the time properties of different WLS fibres including Kuraray models found results of the same order of magnitude [54].



(a) (b)

Figure 3.16: (a) MPPC electronics plus scintillation time-walk effects in the high gain channel for fibre lengths of 100 cm, 200 cm and 300 cm, fitted with a 4 parameter fit given by the sum of equations 3.25 and 3.26. (b) Residuals between the electronics plus scintillation time-walk fit and data for the 3 fibre lengths used.

4 Simulation of the near detector response

Monte-Carlo studies require a working simulation of the detector response. A software package was therefore developed in order to simulate the electronics response to energy deposited in active parts of the ND280 detectors. More specifically, the package was designed to convert simulated GEANT4 hits into digitised hits, in analogue to digital charge (ADC) unit, ready for calibration with three intermediary key steps identified: simulation of the active medium, sensor response and electronics read-out (figure 4.1).

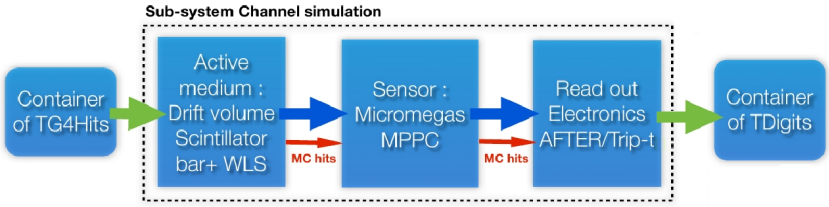


Figure 4.1: Diagram of the detector electronics simulation package.

In order, the response of the active medium to energy deposited in the form of GEANT4 hits was simulated, then the sensor response and finally the electronics read-out. Three different chains account for all variations between sub detectors summarised in table 4.1. Table 4.2 indicates the correspondence between chains and sub-detectors. 4.2.

Chain	active medium	sensor	read-out
1	Scintillator bars	MPPC	TRIP-t
2	Scintillator bars	MPPC	After ASIC
3	Drift volume	Micromegas	After ASIC

Table 4.1: Chains of active media, sensor and electronics read-out.

This section presents data to Monte-Carlo comparisons for validation purposes in the case of chain 1 sub-detectors.

Chain	Detectors
1	ECAL, P0D, SMRD, INGRID
2	FGD
3	TPC

Table 4.2: Correspondence between chains and sub-detectors.

4.1 Fibre model

The fibre model in the electronics simulation package simulates the re-emission time of absorbed photons with an exponential decay. The direction of the re-emitted photons is generated from an isotropic distribution and the critical angle of the fibre then determines whether photons escape or diffuse along to the MPPC. In case of diffusion, a constant drift velocity is assumed.

The scintillator bars of ND280 detectors are instrumented with an MPPC on either one or both ends. In the former case, Y-11 fibres with mirroring on the non-instrumented end are used to increase the light yield. Both fibre models are implemented in the simulation.

The initial GEANT4 hits are characterised by their timing and spatial co-ordinates. The spatial coordinates are used to calculate the channel ID and position of hits in the scintillator bar. The fibre response is then simulated using this information, the timing and a mapping to relate each channel with it's corresponding fibre model and length.

The timing information and channel ID of photons leaving the fibre is passed on to the MPPC model. Details of individual photon paths in the fibre are not simulated and the light profile is left as a property of the MPPC model.

4.2 MPPC model

A fully featured simulation of the MPPC response was implemented in the Monte-Carlo with total control over the customisation of both external and internal parameters provided in a separate configuration file. An extensive study and validation of the simulation of all the important MPPC parameters was done for low light yields [42].

Light profile

The fibre model simulation provides a list of hits with timing and channel ID information. For each hit, a probability density function that describes the illumination profile from the fibre to the MPPC is used to calculate X and Y potential hit coordinates on the MPPC surface. By default, a two dimensional uniform distribution is used.

Dark noise

The coordinate of pixels hit by dark noise and time in the integration window are generated respectively with a two dimensional uniform distribution in position and uniform distribution in time. Two dark noise models were implemented to satisfy different needs.

model 1: first principles

Description: Monte-Carlo model of the dark noise using the known dark noise frequency to generate hits.

Advantages: Correct implementation.

Drawbacks: A time-stamp is given by the TFB after the accumulated charge in an integration window exceeds the TDC threshold. If there are too few dark noise hits in an integration window to exceed the TDC threshold, then these hits are discarded. Since dark noise decreases linearly as a function of PEU on a log scale, depending on the threshold value, orders of magnitude of computations are wasted.

model 2: time-stamp probability method

Description: Pedestal data runs were used to determine the probability of getting a time-stamp for one integration window as a function of the TRIP-t front end board TDC threshold value. Then, given a TDC threshold n_{th} in PEU and a known probability of time-stamp at thresholds of n_{th} and $n_{\text{th}} - 1$:

- A first order approximation for the noise is obtained by generating noise based on the time-stamp probability at n_{th}.
- A second order approximation for the noise is obtained by generating noise based on the time-stamp probability at $n_{\text{th}} - 1$.

Advantages: Orders of magnitude of computations are saved.

Drawbacks: Approximate method needs to be validated.

Operation

When a pixel is triggered, its bias voltage is set to the breakdown voltage. If the same pixel is triggered a second time, the photon detection efficiency is automatically updated to take into account the pixel recovery. As a result, saturation is an inherent property of the simulation.

Crosstalk and after-pulse

Crosstalk and after-pulse are simulated probabilistically from first principles based on the probability functions derived in [52].

4.2.1 MPPC saturation and light illumination profile.

The MPPC simulation and fibre light illumination profile model were tested by comparing results from test bench measurements to Monte-Carlo results obtained by running the MPPC model in standalone mode. In the first test bench measurement, fibre and MPPC effects were decoupled and the MPPC was illuminated with a uniform light profile. The second test bench measurement used a Y-11 fibre connected to an MPPPC. This first measurement was used to test the intrinsic MPPC model while the second was used to test light illumination profile models.

MPPC saturation

Figure 4.2 presents a data to Monte-Carlo comparison of saturation decoupled from fibre effects for over-voltage values of 0.95 V, 1.14 V, 1.3 V, 1.53 V, 1.73 V and 1.92 V. The Monte-Carlo was tuned to match with the data at lower over-voltage values since values of 1.53 V and above are not relevant for the T2K experiment. Very good agreement is observed at low over-voltage with

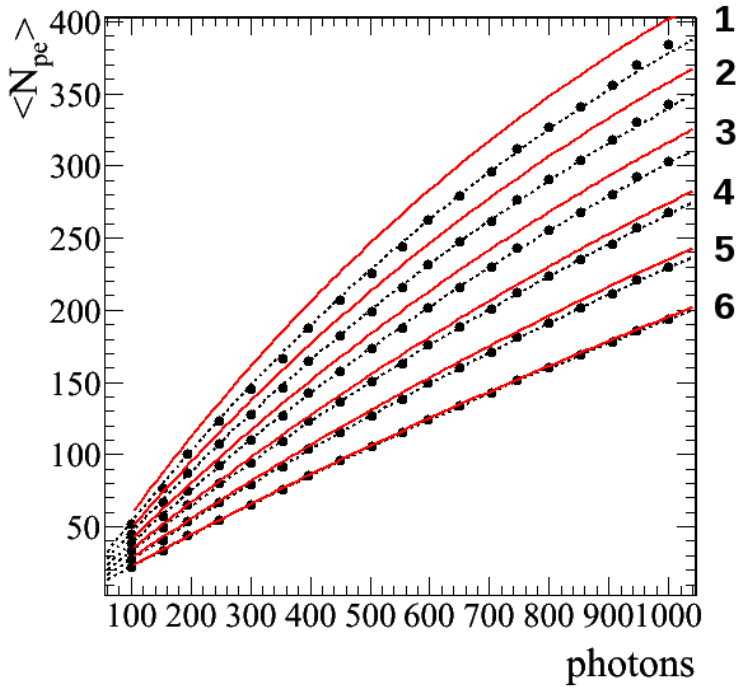


Figure 4.2: Data to Monte-Carlo comparison of saturation decoupled from fibre effects for over-voltage values of 0.95 V (6), 1.14 V (5), 1.3 V (4), 1.53 V (3), 1.73 V (2) and 1.92 V (1) tuned for a best fit at low over-voltage values.Black: Monte Carlo data, red: fit extracted from the test bench measurements (see figure 3.11(a)).

deviations contained within 5%.

Light illumination profile

The light illumination profile from the fibre to the MPPC is not yet well understood. The second test bench measurement with both fibre and MPPC was used to test different models empirically. Using the MPPC model in standalone, a time delay was added with exponential decay 9 ns in order to simulate the fibre time-walk effect.

Two different light illumination models were tested by reproducing the saturation curve and comparing the parameter values of the saturation function.

Two dimensional Gaussian function

Starting with minimal complexity, a two dimensional Gaussian function was investigated.

The width of the Gaussian was optimised manually to obtain the best reproduction of the saturation data with Y-11 fibre. A width of 0.25 mm was found to give very good agreement between data and Monte-Carlo at high over-voltage but a poorer agreement at low over-voltage which is the region of interest (figure 4.3).

The first fitted parameter of the saturation function, CN, represents the PDE and correlated noise. It shows a reasonable agreement with data with a 10% deviation at lower over-voltage values but better agreement at higher over-voltages (see figure 4.4).

The second fitted parameter, N_{eff}, also illustrated in figure 4.4, represents the number of effective pixels of the MPPC. Unlike CN, it shows a poor agreement between data and Monte-Carlo with over 30% deviation at low over-voltages and a large difference in the gradient. For N_{eff}, the Monte-Carlo result gives a flatter response compared to the data that is indicative of a more "uniform" illumination over the MPPC surface. Reducing the width of the Gaussian provides a less "uniform" response but degrades the overall saturation agreement.

In order to account for the disparities observed, a second model implementing a Gaussian function in cylindrical coordinates was investigated.

Gaussian function in cylindrical coordinates

A Gaussian function in cylindrical coordinates with mean value of radius R was tested according to the following algorithm: For each potential hit,

1. a radial coordinate, r, is chosen based on a 1 dimensional Gaussian probability density function of mean R and sigma σ_R,

2. an angular position, θ, is chosen uniformly in the range $[0, 2\pi]$,

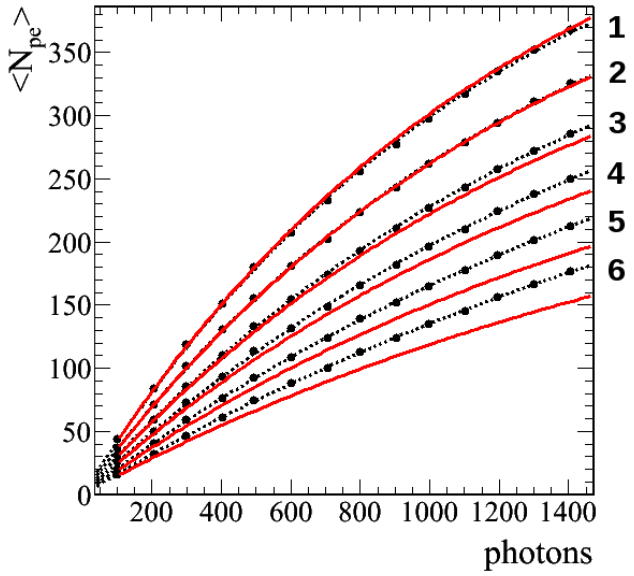


Figure 4.3: Comparison of Data to Monte-Carlo saturation for test bench measurement with Y-11 fibre and Gaussian illumination model with σ of 0.25 mm for over-voltage values of 0.95 V (6), 1.14 V (5), 1.3 V (4), 1.53 V (3), 1.73 V (2) and 1.92 V (1). Black: Monte Carlo data, red: fit extracted from the test bench measurements (see figure 3.11(b)).

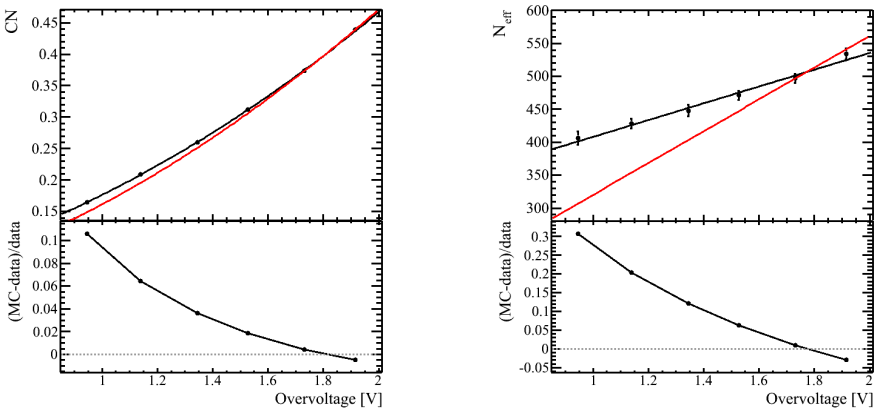


Figure 4.4: Comparison of the parameters of the data and Monte-Carlo saturation fits, CN (left) and N_{eff} (right) versus over-voltage for the Gaussian illumination model with σ of 0.25 mm. Residuals between the Monte Carlo data and test bench fits are also shown. Black: Monte Carlo data, red: fit extracted from the test bench measurements (see figure 3.12(b) and (c)).

3. the corresponding X and Y-coordinates of the hit on the MPPC surface are calculated trivially:

$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta. \end{aligned} \tag{4.1}$$

4. The potential hit is kept if the X and Y-coordinates fall in an active area of the MPPC.

Figure 4.5 presents some illumination profile examples for different values R and σ_R. Such a configuration is symmetric in circular coordinates. This is

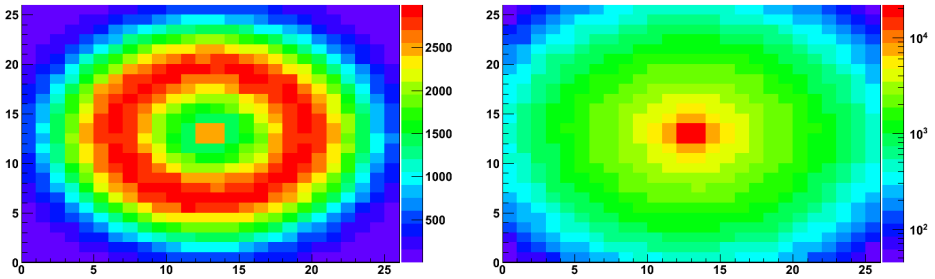


Figure 4.5: Monte-Carlo distributions over the MPPC surface of 10^6 hits for a Gaussian probability density function in cylindrical coordinates with radii R of respectively 0.4 mm and 0.3 mm and widths of respectively 0.15 mm and 0.30 mm. 1 bin represents 1 pixel.

consistent with the fact that the fibre model under consideration has a circular cross-section and presents two advantages compared to the simple Gaussian configuration:

- More hits are concentrated on a smaller surface which results in a smaller value of N_{eff} at lower illuminations,
- outer pixels on the MPPC still, none the less, get more illumination than with a simple Gaussian model resulting in a larger value of N_{eff} at higher illuminations.

In the case of a Gaussian in cylindrical coordinates, both the radius and width of the Gaussian were optimised manually to obtain a good fit between the data and Monte-Carlo saturation results. A radius of 0.3 mm and width of 0.25 mm were found to give the best agreement between data and Monte-Carlo with less than 10% deviation in the worst case (figure 4.6). The first fit parameter, CN, of the saturation function was found to agree with data to the level of 5% for over-voltages below 1.3 V (figure 4.7). A similar slope

between data and Monte-Carlo was found for the second parameter N_{eff} but with a deviation of no less than 25% for the same over-voltage values.

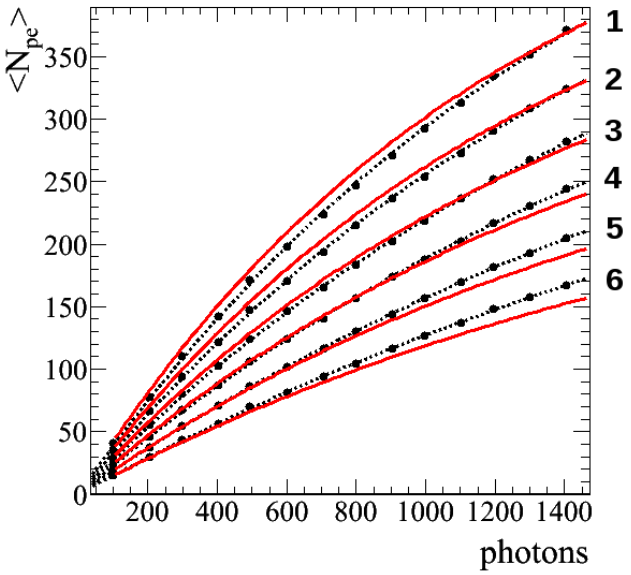


Figure 4.6: Comparison of Data to Monte-Carlo saturation for test bench measurement with Y-11 fibre and Gaussian model in cylindrical coordinates with radius of 0.3 mm and σ of 0.25 mm for over-voltage values of 0.95 V (6), 1.14 V (5), 1.3 V (4), 1.53 V (3), 1.73 V (2) and 1.92 V (1). Black: Monte Carlo data, red: fit extracted from the test bench measurements (see figure 3.11(b) and (c)).

The MC simulation is used for the measurement of systematic uncertainties. Since the first T2K runs are dominated by statistical uncertainties, a 10% deviation in the saturation response between data and Monte-Carlo was considered acceptable. Further studies requiring better agreement might consider different models, such as a Gaussian with kurtosis, and study possible optical and reflection effects caused by the epoxy of the MPPC.

4.3 Full electronic simulation model

The full electronics simulation model is the combined implementation of the fibre, MPPC and TRIP-t front end board. To validate the simulation, Monte-Carlo runs were compared to data taken with the ND280.

Pedestal runs

The pedestal of a TFB board is the average intrinsic electronic response of the TFB in ADC in the absence of any signal, obtained by integrating the charge in one of the 23 integration windows. In so called “pedestal” runs, signal from

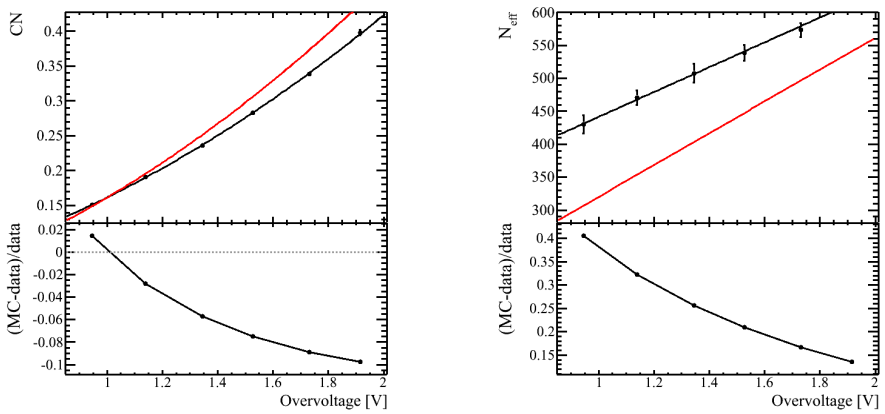


Figure 4.7: Comparison of the parameters of the data and Monte-Carlo saturation fits, CN (left) and N_{eff} (right) versus over-voltage for the Gaussian model in cylindrical coordinates with radius of 0.3 mm and σ of 0.25 mm. Residuals between the Monte Carlo data and test bench fits are also shown. Black: Monte Carlo data, red: fit extracted from the test bench measurements (see figure 3.12(b).

the TFB integration windows is recorded in the absence of beam data. These runs serve various purposes such as:

- Alignment and calibration of pedestals amongst all the TFB boards,
- Verification of the MPPC gains (given by the separation between two peaks in the hi-gain channel),
- Tuning of the dark noise simulation and the PEU structure in the high gain channel (figure 4.8),
- Tuning of the time structure of noise hits in the TFB integration windows (figure 4.9).

Here in particular, we concentrate on the study of the last two items. The PEU structure of the dark noise simulation in the high gain channel, presented in figure 4.8, shows a good agreement between data and Monte Carlo with some divergence in the tail where more events are observed in the data than the Monte-Carlo. In pedestal runs, all photo-electron peaks arise probabilistically from a combination of dark noise followed by crosstalk or after-pulse. Any discrepancy between the probability values of these underlying quantities between the model and reality get compounded for higher photo-electron peaks in the tail. However, because of the logarithmic form of the event distribution, these higher order effects are not consequential.

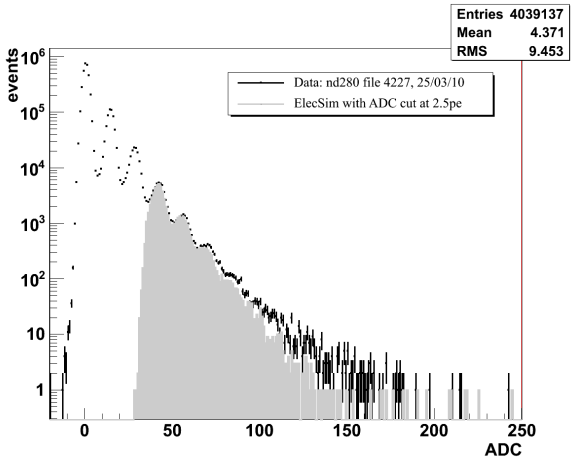


Figure 4.8: Comparison of Dark Noise in the high gain channel between an ND280 pedestal data run in the downstream ECAL and a Monte Carlo electronics simulation noise only run in the downstream ECAL with ADC cut of 2.5 PEU.

The time structure of noise hits in the TFB integration windows is presented in figure 4.9. A slight excess of timestamped events in the Monte Carlo (grey) is visible close to the end of each window.

The general shape of the distribution is a function of the TDC threshold value because a time-stamp is created only if and when the total accumulated charge in the integration window has exceeded the TDC threshold. For a threshold corresponding to 1 PEU, hits are timestamped with a uniform distribution over the integration window. As the threshold increases, probabilistically, hits will be given a time-stamp increasingly towards the end of the integration window.

Given a constant TDC threshold, the time of time-stamp will depend on the frequency of dark noise and the probability of crosstalk and after-pulse. For a lower dark noise frequency or smaller probability of crosstalk or after-pulse, timestamps would be given, on average at a later time.

A reasonable assumption is therefore that the dark noise, crosstalk or after-pulse models are underestimated in the simulation. We also observe that this corroborates with the observations made based on figure 4.8. Dark noise probabilities are well measured but the crosstalk measurement in particular is subject to more uncertainty. The models implemented in the Monte Carlo were taken from results published in [42].

Beam runs

Beam runs are used to verify the matching of calibrated hit charges (figure 4.10) , and, verify the bunch structure of simulated beam events and alignment

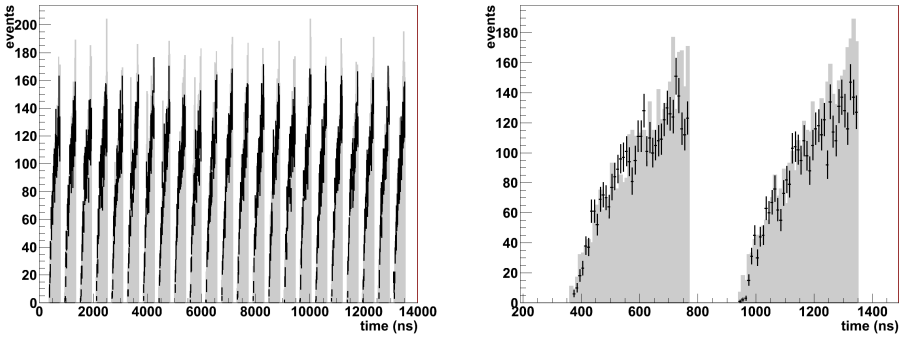


Figure 4.9: Comparison of the time distribution of pedestals in one TFB cycle between a data run and Monte-Carlo.

of sub-detector beam timing (figure 4.11).

Data is collected in beam runs using a trigger that coincides with beam spill times. Out of all the channels in the detector, most will only record pedestals or noise during the spill coincidence window. The vast majority of calibrated hit charges shown in figure 4.10 are therefore below 10 PEU. The region of interest in this case lies above 10 PEU.

Most of the region of interest shows good agreement between data and Monte Carlo apart from the region between 10 and 20 PEU where a large discrepancy, still under investigation by the calibration group, can be clearly observed.

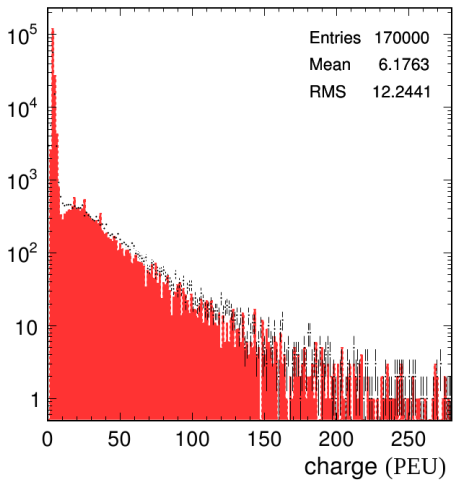


Figure 4.10: Comparison of the calibrated hit charge in ADC for the downstream ECAL between data and Monte-Carlo for beam data runs.

The time structure of simulated beam events on the other hand shows that the bunch times of the three TRIP-t sub-detector are well aligned. The first

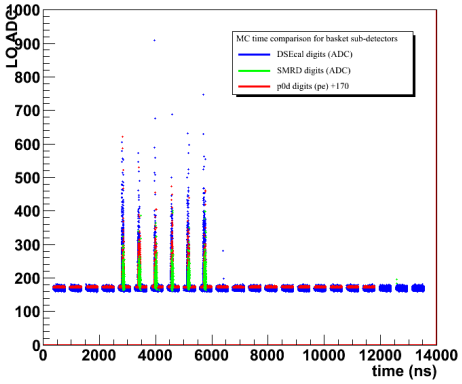


Figure 4.11: Monte-Carlo plot of the Lo gain ADC as a function of time in the TFB cycle for the downstream ECAL, P0D and SMRD for beam data runs.

bunch also falls inside the first half of the 5^{th} integration window as expected.

Cosmic runs

Cosmic runs were used to verify the Monte-Carlo reconstructed MIP spectrum (figure 4.12). The cosmic muon distributions were corrected for channel by channel and attenuation variations and fitted with a convolved Landau-Gaussian fit. The most probable value for the fit was obtained and used as a scaling factor to recover the energy deposited in scintillator bars.

Poor agreement was found between data and Monte-Carlo with variations of up to 100% in the normalised distributions. In particular, the width of the MC spectrum was found to be smaller than the width of the data spectrum. Ongoing work is being carried out in order to account for the divergence observed. For instance, the light illumination model used in the Monte-Carlo simulation results presented in figure 4.12 was the uniform distribution. A Gaussian illumination model has since been set as default in the software.

The calibration effort by the calibration group is still very much a work in progress at the time of writing. The first T2K runs are highly statistically limited however so uncertainties due to calibration have little impact on the final physics results. The first T2K run in particular, presented in the next section, has a stronger focus on the far detector measurement at Super-K which is a very well understood detector.

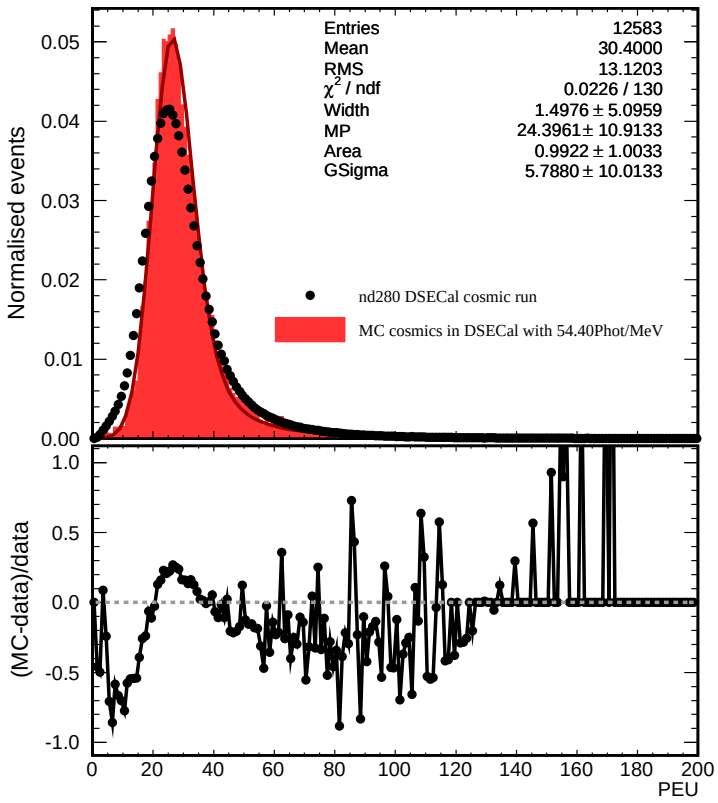


Figure 4.12: Comparison between the data and Monte-Carlo reconstructed MIP spectrum where the MC spectrum is fitted with a convolved Landau-Gaussian fit.

5 ν_μ disappearance in T2K

T2K uses a beam of muon neutrinos to investigate ν_e appearance and ν_μ disappearance. The disappearance measurement compares the observed number of muon neutrinos at the far detector with the expected number at Super-K. If fewer events are observed than expected, as predicted by the mathematical formalism, this is interpreted as ν_μ having oscillated to other flavours of neutrinos. In the two-flavour approximation, all the muon neutrinos are assumed to have oscillated to τ neutrinos according to the oscillation probability given by equation 1.25.

The ultimate aim of the disappearance analysis is to achieve the most sensitive measurement to date of the oscillation parameters $\sin^2 2\theta_{23}$ and Δm^2_{32}.

This chapter provides a detailed description of all the steps of a ν_μ disappearance analysis. In section 5.1, the reconstructed energy spectrum at Super-K is formally derived and potential sources of systematic error are discussed. Section 5.2 introduces the maximum likelihood method used to estimate the oscillation parameters based on the measured and predicted data. Sections 5.3 and 5.4 compare two candidate methods for the construction of the statistical intervals in the parameter space of the oscillation parameters. The first method explores techniques of minimisation, optimisation with constraints and the construction of confidence regions. The second method approaches the problem with a different angle and introduces marginalisation, the sampling of the likelihood function using a Markov Chain Monte Carlo (MCMC) method and the construction of Bayesian credible intervals. The signal and backgrounds to the disappearance analysis at the far and near detectors are introduced respectively in sections 5.5 and 5.7. Section 5.6 describes the PDF construction procedure for the Super-K and ND280 inputs. The effect of the oscillation parameters on the expected number of events and final reconstructed spectrum at Super-K is studied. Section 5.8 introduces the systematic error sources and illustrates how the variation of each systematic affects the expected energy spectrum at Super-K. Sections 5.9 presents the results of an MCMC simulation in the favoured oscillation hypothesis for an accumulated exposure of 3.23×10^{19} POT. The T2K results for the first physics run are finally presented in section 5.10.

5.1 Derivation of expected energy spectra at Super-K

The muon neutrino disappearance analysis in the two neutrino flavour approximation is a function of the two oscillation parameters and the energy spectra of the neutrinos at the far detector. The distance travelled by each neutrino is constant and so does not need to be considered.

The T2K experiment aims to make a precision measurement of the atmospheric oscillation parameters by comparing the Monte Carlo prediction of the reconstructed energy spectrum at Super-K with the collected data. The understanding and simulation of physical processes that affect the expected energy spectrum at Super-K are therefore critical to the success of the T2K experiment. Where our understanding is lacking and uncertainties creep in, it is necessary for the validity of our results to assign appropriate systematic errors.

This section goes through a step by step formal derivation of the expected energy spectrum at Super-K using a top down approach and based on the derivation provided in [55]. Sources of systematic uncertainties are then clarified based on this derivation.

5.1.1 Formal derivation of the expected energy spectrum at Super-K

Step 1: Separation of neutrino flavours and interaction modes.

The predicted reconstructed energy spectrum at Super-K: $\frac{dN^{\rm exp}_{\rm SK(}E_r)}{dE_r}$, consists of all the events that have passed a set of selection cuts designed to select ν_μ's from Charged Current Quasi Elastic (CCQE) interactions. ν_μ CCQE events are therefore the signal sample. However, other types of events also pass the chosen set of selection cuts and are backgrounds to the analysis (see section 5.5 for an exhaustive description of signal and backgrounds). $\frac{dN^{\rm exp}_{\rm SK}(E_r)}{dE_r}$ must therefore be equal to the sum of the reconstructed energy spectrum of all parent neutrino flavours "f" and interaction mode "m" that are present after the selection cuts are applied:

$$\frac{dN^{\rm exp}_{\rm SK}(E_r)}{dE_r} = \sum_f \sum_m \frac{dN^{\rm exp}_{\rm SK;f;m}(E_r)}{dE_r}. \tag{5.1}$$

The selection efficiency determines the fraction of events of each parent neutrino flavor and interaction mode to appear in the final reconstructed energy spectrum. Equation 5.1 is re-written below taking into account the selection

efficiency $\epsilon^{\rm sel}_{\rm SK;f;m}(E_r)$:

$$\frac{dN^{\rm exp}_{\rm SK}(E_r)}{dE_r} = \sum_f \sum_m \frac{dN^{\rm tot}_{\rm SK;f;m}(E_r)}{dE_r} \epsilon^{\rm sel}_{\rm SK;f;m}(E_r), \tag{5.2}$$

where $\frac{dN^{\rm tot}_{\rm SK;f;m}(E_r)}{dE_r}$ is the total reconstructed energy spectrum of parent neutrino flavour "f" and interaction mode "m".

Step 2: Expression of reconstructed energy spectrum in terms of true energy.

We can then express the expected reconstructed energy spectrum in terms of the total true neutrino energy spectrum $\frac{dN^{\rm tot}_{\rm SK;f;m}(E_t)}{dE_t}$:

$$\frac{dN^{\rm exp}_{\rm SK}(E_r)}{dE_r} = \sum_f \sum_m \int dE_t \frac{dN^{\rm tot}_{\rm SK;f;m}(E_t)}{dE_t} R_{\rm SK;f;m}(E_t, E_r) \epsilon^{\rm sel}_{\rm SK;f;m}(E_r) \epsilon^{\rm reco}_{\rm SK;m}(E_t), \tag{5.3}$$

where $\epsilon^{\rm reco}_{\rm SK;m}(E_t)$ and $R_{\rm SK;f;m}(E_t, E_r)$ are defined below:

- $\epsilon^{\rm reco}_{\rm SK;m}(E_t)$: Super-K efficiency for reconstructing a neutrino event of mode m and true energy E_t.
- $R_{\rm SK;f;m}(E_t, E_r)$: Super-K energy transfer matrix for events of neutrino parent flavour f and mode m; gives the probability for an event of true energy E_t to be reconstructed with energy E_r.

$R_{\rm SK;f;m}(E_t, E_r)$ is defined such that:

$$\frac{dN^{\rm exp}_{\rm SK}(E_r)}{dE_r} = \int dE_t \frac{d^2 N_{SK;f;m}}{dE_r dE_t} = \int dE_t \frac{dN_{SK;f;m}}{dE_t} R_{\rm SK;f;m}(E_t, E_r), \tag{5.4}$$

and therefore:

$$R_{\rm SK;f;m}(E_t, E_r) = \frac{d^2 N_{SK;f;m}}{dE_r dE_t} \Big/ \frac{dN_{SK;f;m}}{dE_t} \, . \tag{5.5}$$

For each true energy value E_t, a probability density function normalised to 1 can be computed that describes the probability for an event of true energy E_t to be reconstructed with energy E_r.

Step 3: Separation of the oscillation probability function dependence.

The dependence on the ν_μ oscillation probability as a function of true neu-

trino energy $P^{\rm osc}_{f;m}(E_t)$ is removed below from equation 5.3:

$$\frac{dN^{\rm exp}_{\rm SK}(E_r)}{dE_r} = \sum_f \sum_m \int dE_t \frac{dN^{\rm tot;unosc}_{\rm SK;f;m}(E_t)}{dE_t} \times P^{\rm osc}_{f;m}(E_t) R_{\rm SK;f;m}(E_t, E_r) \epsilon^{\rm sel}_{\rm SK;f;m}(E_r) \epsilon^{\rm reco}_{\rm SK;m}(E_t), \quad (5.6)$$

where $\frac{dN^{\rm tot;unosc}_{\rm SK;f;m}(E_t)}{dE_t}$ is the total non oscillated neutrino energy spectrum for a neutrino of flavour f and interaction mode m.

Step 4: Expression of true neutrino energy spectrum in terms of neutrino flux at Super-K.

The true neutrino energy spectrum is a function of the estimated neutrino flux at SK $\frac{d\Phi_{\rm SK}(E_t)}{dE_t}$:

$$\frac{dN^{\rm tot;unosc}_{\rm SK;f;m}(E_t)}{dE_t} = N^{\rm POT}_{\rm SK} M^{\rm fid}_{\rm SK} \frac{N_A}{A_{H_2O}} \frac{d\Phi_{\rm SK}(E_t)}{dE_t} \sigma^{\rm tot}_{H_2O;f;m}(E_t) f_m(E_t), \quad (5.7)$$

where the terms $N^{\rm POT}_{\rm SK}$, $M^{\rm fid}_{\rm SK}$, $\frac{N_A}{A_{H_2O}}$, $\sigma^{\rm tot}_{H_2O}(E_t)$ and $f_m(E_t)$ are defined below:

- $N^{\rm POT}_{\rm SK}$: Number of protons on target corresponding to the Super-K data sample.
- $M^{\rm fid}_{\rm SK}$: Mass of the Super-K fiducial volume.
- $\frac{N_A}{A_{H_2O}}$: Avogadro's number divided by the water mass number.
- $\sigma^{\rm tot}_{H_2O;f;m}(E_t)$: Total neutrino interaction cross section in water at energy E_t for a neutrino parent f and interaction mode m.
- $f_m(E_t)$: Fraction of events of interaction type m.

Substituting equation 5.7 into equation 5.6 gives:

$$\frac{dN^{\rm exp}_{\rm SK}(E_r)}{dE_r} = N^{\rm POT}_{\rm SK} M^{\rm fid}_{\rm SK} \frac{N_A}{A_{H_2O}} \sum_f \sum_m \int dE_t \left(\frac{d\Phi_{\rm SK}(E_t)}{dE_t} \sigma^{\rm tot}_{H_2O;f;m}(E_t) f_m(E_t) \times P^{\rm osc}_{f;m}(E_t) R_{\rm SK;f;m}(E_t, E_r) \epsilon^{\rm sel}_{\rm SK;f;m}(E_r) \epsilon^{\rm reco}_{\rm SK;m}(E_t) \right). \quad (5.8)$$

The estimation of the true neutrino flux at SK $\frac{d\Phi_{\rm SK}(E_t)}{dE_t}$ is a cause of large systematic uncertainties. One of the key factors to a successful analysis in the T2K experiment is the appropriate use of ND280 measurements in order to constrain these systematic errors and demonstrate that we understand the neutrino beam. This leads to the next step in the formal derivation of the final reconstructed energy spectra at Super-K: the use of the far to near ratio $R^{\rm F/N}(E_t)$ to constrain the Super-K neutrino flux.

Step 5: Use of far to near ratio to constrain the Super-K neutrino flux.

The far to near ratio is defined as the ratio of neutrino fluxes at Super-K and ND280:

$$R^{\mathrm{F/N}}(E_t) = \left(\frac{d\Phi_{\mathrm{SK}}(E_t)}{dE_t}\right) \Big/ \left(\frac{d\Phi_{\mathrm{ND280}}(E_t)}{dE_t}\right). \tag{5.9}$$

In order to constrain the Super-K neutrino flux, $R^{\mathrm{F/N}}(E_t)$ is approximated with Monte Carlo and then used in combination with ND280 data in order to constrain the Super-K neutrino flux:

$$\frac{d\Phi_{\mathrm{SK}}(E_t)}{dE_t} = R^{\mathrm{F/N}}(E_t)\frac{d\Phi_{\mathrm{ND280}}(E_t)}{dE_t}. \tag{5.10}$$

The use of $R^{\mathrm{F/N}}(E_t)$ is justified because the systematic uncertainty associated with its calculation is considerably smaller than the uncertainty in the calculation of the far detector neutrino flux due to cancellations in the near to far ratio. The near detector flux is of course determined to better accuracy since the neutrino flux is orders of magnitude greater than at the far detector. Therefore the far to near ratio has constraining power over the far detector neutrino flux. The next steps detail how ND280 data are used to measure the neutrino flux at ND280.

Step 6: Expression of the ND280 neutrino flux in terms of the ND280 neutrino energy spectra.

The neutrino flux at ND280 is a function of the true neutrino energy spectra at ND280 $\frac{dN_{\mathrm{ND280}}(E_t)}{dE_t}$:

$$\frac{d\Phi_{\mathrm{ND280}}(E_t)}{dE_t} = \frac{A_{\mathrm{ND280}}}{N^{\mathrm{POT}}_{\mathrm{ND280}} M^{\mathrm{fid}}_{\mathrm{ND280}} N_A} \frac{1}{\sigma_{\mathrm{ND280}}(E_t)} \frac{dN_{\mathrm{ND280}}(E_t)}{dE_t}, \tag{5.11}$$

where the terms $N^{\mathrm{POT}}_{\mathrm{ND280}}$, $M^{\mathrm{fid}}_{\mathrm{ND280}}$ and $\sigma_{\mathrm{ND280}}(E_t)$ are defined below:

- $N^{\mathrm{POT}}_{\mathrm{ND280}}$: Number of protons on target corresponding to the ND280 data sample.
- $M^{\mathrm{fid}}_{\mathrm{ND280}}$: Mass of the ND280 fiducial volume.
- $\sigma_{\mathrm{ND280}}(E_t)$: Total neutrino interaction cross section on scintillator at energy E_t.

In turn, the true neutrino energy spectrum at ND280 is expressed as a function of the reconstructed energy spectrum of sample "s" $\frac{dN_{\mathrm{ND280;s}}(E_r)}{dE_r}$:

$$\frac{dN_{\mathrm{ND280}}(E_t)}{dE_t} = \sum_s \int dE_r \left(\frac{dN^{\mathrm{Data}}_{\mathrm{ND280;s-like}}(E_r)}{dE_r} - \frac{dN^{\mathrm{MC}}_{\mathrm{ND280;s-bkg}}(E_r)}{dE_r} \right) \times \frac{R_{\mathrm{ND280;s}}(E_r, E_t)}{\epsilon^{\mathrm{sel}}_{\mathrm{ND280;s}}(E_r)\epsilon^{\mathrm{reco}}_{\mathrm{ND280;s}}(E_t)}, \tag{5.12}$$

where the Monte Carlo prediction on the "s"-like background $\frac{dN^{\mathrm{MC}}_{\mathrm{ND280;s-bkg}}(E_r)}{dE_r}$ is subtracted from the "s"-like data sample $\frac{dN^{\mathrm{Data}}_{\mathrm{ND280;s-like}}(E_r)}{dE_r}$ to get the reconstructed energy spectrum of sample "s" $\frac{dN_{\mathrm{ND280;s}}(E_r)}{dE_r}$.
The terms $\epsilon^{\mathrm{sel}}_{\mathrm{ND280;s}}(E_r)$, $\epsilon^{\mathrm{reco}}_{\mathrm{ND280;s}}(E_t)$ and $R_{\mathrm{ND280}}(E_r, E_t)$ are defined below:

- $\epsilon^{\mathrm{sel}}_{\mathrm{ND280;s}}(E_r)$: ND280 selection efficiency determines the fraction of events that have passed selection cuts associated with sample "s" to appear in the final reconstructed energy spectrum.

- $\epsilon^{\mathrm{reco}}_{\mathrm{ND280;s}}(E_t)$: ND280 efficiency for reconstructing a neutrino event of true energy E_t.

- $R_{\mathrm{ND280;s}}(E_r, E_t)$: Gives the probability for an event of reconstructed energy E_r to have true energy E_t. Is equivalent to the transpose of the energy transfer matrix: $R^T_{\mathrm{ND280;s}}(E_t, E_r)$.

Step 7: Bringing it all together.

Combining equations 5.10, 5.11, 5.12 shows how the Super-K neutrino flux is constrained by ND280 data and Monte Carlo prediction:

$$\frac{d\Phi_{\mathrm{SK}}(E_t)}{dE_t} = \frac{R^{\mathrm{F/N}}(E_t) A_{\mathrm{ND280}}}{N^{\mathrm{POT}}_{\mathrm{ND280}} M^{\mathrm{fid}}_{\mathrm{ND280}} N_A \sigma_{\mathrm{ND280}}(E_t)} \frac{dN_{\mathrm{ND280}}(E_t)}{dE_t} \times \sum_s \int dE_r \left(\frac{dN^{\mathrm{Data}}_{\mathrm{ND280;s-like}}(E_r)}{dE_r} - \frac{dN^{\mathrm{MC}}_{\mathrm{ND280;s-bkg}}(E_r)}{dE_r} \right). \tag{5.13}$$

The final expression for the Super-K reconstructed energy spectrum is obtained by combining equation 5.13 with equation 5.8:

$$\begin{aligned}\frac{dN_{\mathrm{SK}}^{\mathrm{exp}}(E_r)}{dE_r} &= \frac{N_{\mathrm{SK}}^{\mathrm{POT}} M_{\mathrm{SK}}^{\mathrm{fid}} A_{\mathrm{ND280}}}{N_{\mathrm{ND280}}^{\mathrm{POT}} M_{\mathrm{ND280}}^{\mathrm{fid}} A_{H_2O}} \sum_f \sum_m \sum_s \int dE_t \int dE_{r'} \Big(R^{\mathrm{F/N}}(E_t) P_{f;m}^{\mathrm{osc}}(E_t) \times \\ &\quad \frac{\sigma_{H_2O;f;m}^{\mathrm{tot}}(E_t) f_m(E_t)}{\sigma_{\mathrm{ND280;s}}(E_t)} \left(\frac{dN_{\mathrm{ND280;s-like}}^{\mathrm{Data}}(E_{r'})}{dE_{r'}} - \frac{dN_{\mathrm{ND280;s-bkg}}^{\mathrm{MC}}(E_{r'})}{dE_{r'}} \right) \times \\ &\quad R_{\mathrm{SK;f;m}}(E_t, E_r) R_{\mathrm{ND280;s}}^{T}(E_t, E_{r'}) \frac{\epsilon_{\mathrm{SK;f;m}}^{\mathrm{sel}}(E_r) \epsilon_{\mathrm{SK;m}}^{\mathrm{reco}}(E_t)}{\epsilon_{\mathrm{ND280;s}}^{\mathrm{sel}}(E_{r'}) \epsilon_{\mathrm{ND280;s}}^{\mathrm{reco}}(E_t)} \Big). \end{aligned} \quad (5.14)$$

5.1.2 Discussion of systematic uncertainties based on the derived Super-K reconstructed energy spectrum

With each new term introduced during the derivation of the Super-K reconstructed energy spectrum, associated systematic uncertainties must be considered. Overall, sources of systematic uncertainties affect both the far and near detectors and can be subdivided into the following categories:

Selection efficiencies:

The selection efficiencies are a function of reconstruction energy for each interaction mode and neutrino parent. The systematic uncertainty in the selection efficiencies therefore describe the uncertainty in the shape and normalisation of the spectrum of each interaction mode after selection cuts.

Energy reconstruction efficiencies:

Energy reconstruction efficiency uncertainties are introduced when converting from true to reconstructed energy or vice versa.

Neutrino flux shape and normalisation:

Neutrino flux uncertainties represent the uncertainty in our knowledge of the shape and normalisation of the flux of neutrinos at the near and far detectors.

Nuclear interactions:

Nuclear interaction uncertainties refer to uncertainties in the cross sections of neutrino interactions with detector target material and uncertainties in the final state of an interaction.

An in-depth description of each systematic error included in the analysis is provided in section 5.8

5.1.3 Practical implementation

The formal derivation described in section 5.1.1 is useful, especially for the visualisation of systematic effects that enter the analysis. In practice the analysis is much simplified by the use of sets of templates that describe the inputs from Super-K and ND280.

A first set of templates describes the true energy spectrum of events of each interaction mode and neutrino parent that pass the Super-K selection cuts. A second set gives the conversion from true to reconstructed energy for each interaction mode and neutrino parent that pass the Super-K selection cuts. A third set describes the ND280 reconstructed energy spectrum for each interaction mode and neutrino parent that passes ND280 selection modes.

In order to produce the first two sets of templates, neutrino generators NEUT [56] and Genie [57] are used in the Super-K software in combination with flux information. The Super-K software creates a ROOT [58] NTuple that describes Super-K variables for each neutrino interaction simulated in the Monte Carlo process. Information provided in the NTuple includes, for each neutrino interaction, the number of rings produced at Super-K, the visible energy, inner and outer detector number of PMT hits, particle type, fitted direction, three momentum and many more parameters. The first two sets of templates are generated by applying selection cuts on Monte Carlo neutrino interaction events based on the Super-K NTuples. Similarly, the third set of templates is constructed by processing the NEUT or Genie Monte Carlo through the ND280 software.

Using these templates and the oscillation probability formula is sufficient to build the final reconstructed energy spectrum at Super-K. Systematics are then included by determining how they affect these templates.

The construction of these templates is detailed in sections 5.5 and 5.7.

5.2 Fitting method

The "observable" quantity in the disappearance measurement is the reconstructed energy spectra of the ν_μ CCQE-like events at the Super-K detector. This "observable" is parametrised as a function of the oscillation parameters to be determined. Once data are collected, either experimentally or by Monte Carlo simulation, the parameters can be estimated and statistical intervals drawn.

Section 5.2.1, based on [31], describes the statistical method used, a maximum likelihood method, to find the oscillation parameters that give the best fit between Super-K data and the Super-K reconstructed oscillated energy spectrum probability density function from Monte Carlo.

5.2.1 Maximum Likelihood method

Suppose we observe a set of N measured neutrino events of reconstructed energies $\mathbf{E_R} = (E_{R1}, ..., E_{RN})$ described by the same probability density function (PDF) $f(\mathbf{E_R}|\sin^2\theta_{23}, \Delta m^2_{32})$, where $\sin^2\theta_{23}$ and Δm^2_{32} are the oscillation parameters to be determined. The likelihood function is given by the PDF evaluated with the data $\mathbf{E_R}$ as a function of the oscillation parameters. Since the measurements for each neutrino event are statistically independent and "drawn" from the same PDF, the likelihood function is:

$$L(\sin^2\theta_{23}, \Delta m^2_{32}) = \prod_{i=1}^{N} f(E_{Ri}|\sin^2\theta_{23}, \Delta m^2_{32}), \tag{5.15}$$

The estimators $\widehat{\sin^2\theta_{23}}$ and $\widehat{\Delta m^2_{32}}$ give us the best guess for the values of the oscillation parameters. They are found by finding the values of $\sin^2\theta_{23}$ and Δm^2_{32} that maximise the likelihood function $L(\sin^2\theta_{23}, \Delta m^2_{32})$.

In fact we use the log of the likelihood for convenience. This is equivalent as both L and $\ln L$ are maximised for the same values of the oscillation parameters. The maximum likelihood estimators can be found by solving the likelihood equations:

$$\frac{\partial \ln L}{\partial(\sin^2\theta_{23})} = 0, \tag{5.16}$$

$$\frac{\partial \ln L}{\partial(\Delta m^2_{32})} = 0. \tag{5.17}$$

An unbinned maximum likelihood method is advantageous when the number of events in the data sample is small because binning results in a loss of information associated with larger statistical errors for the parameter estimates.

An extended maximum likelihood method is used if the sample size is treated as a Poisson distributed variable. For the extended maximum likelihood method, equation 5.15 becomes

$$\begin{aligned} L(\sin^2\theta_{23}, \Delta m^2_{32}) &= \frac{\left(N^{\mathrm{SK}}_{\mathrm{exp}}(\sin^2\theta_{23}, \Delta m^2_{32})\right)^{N^{\mathrm{SK}}_{\mathrm{obs}}}}{N^{\mathrm{SK}}_{\mathrm{obs}}!} \exp\left(-N^{\mathrm{SK}}_{\mathrm{exp}}(\sin^2\theta_{23}, \Delta m^2_{32})\right) \\ &\quad \times \prod_{i=1}^{N^{\mathrm{SK}}_{\mathrm{obs}}} f(E_{Ri}|\sin^2\theta_{23}, \Delta m^2_{32}). \end{aligned} \tag{5.18}$$

which we re-write as

$$L(\sin^2\theta_{23}, \Delta m^2_{32}) = L_{\mathrm{Norm}}\left(\sin^2\theta_{23}, \Delta m^2_{32}\right) \times L_{\mathrm{Shape}}\left(\sin^2\theta_{23}, \Delta m^2_{32}\right), \tag{5.19}$$

where the first term L_{Norm} corresponds to the normalisation and the second term L_{Shape} is the multinomial shape of the reconstructed neutrino energy spectrum.
In case of a large number of expected events at Super-K, the exponential term in the Poisson distribution becomes a source of numerical instability. Under the same condition, however, a Gaussian distribution provides a good approximation for the Poisson distribution. Consequently, the Poisson normalisation term is replaced with a Gaussian term for large numbers of expected events at Super-K:

$$L_{\text{Norm}}\left(\sin^2\theta_{23}, \Delta m^2_{32}\right) = \exp\left(-\frac{1}{2}\frac{(N^{\text{SK}}_{\text{exp}} - N^{\text{SK}}_{\text{obs}})^2}{N^{\text{SK}}_{\text{exp}}}\right) \Big/ \sqrt{2\pi N^{\text{SK}}_{\text{exp}}}\,. \quad (5.20)$$

In contrast, a binned maximum likelihood can be used for large data samples in order to reduce computation time. In this case, we have a set of n bins, $\mathbf{n} = (n_1, ..., n_N)$, with expectation values $\boldsymbol{\nu} = E[\mathbf{n}]$ and probabilities $f(\mathbf{n}; \boldsymbol{\nu})$. Maximising the likelihood function based on the contents of the bins gives the estimators for the oscillation parameters. This is equivalent to maximising the likelihood ratio

$$\lambda(\sin^2\theta_{23}, \Delta m^2_{32}) = f\left(\mathbf{E_B}; \boldsymbol{\nu}(\sin^2\theta_{23}, \Delta m^2_{32})\right) \big/ f(\mathbf{E_B}; \mathbf{E_B}) \quad (5.21)$$

or to minimising $-2\ln\lambda(\sin^2\theta_{23}, \Delta m^2_{32})$ where:

$$-2\ln\lambda(\sin^2\theta_{23}, \Delta m^2_{32}) = 2\sum_{i=1}^{N}\left[\nu_i(\sin^2\theta_{23}, \Delta m^2_{32}) - E_{Bi} + E_{Bi}\ln\frac{E_{Bi}}{\nu_i(\sin^2\theta_{23}, \Delta m^2_{32})}\right]. \quad (5.22)$$

In the limit of zero bin width, it can be verified that maximising equation 5.22 is equivalent to maximising the unbinned maximum likelihood function.

The method used in this particular analysis is the extended unbinned maximum likelihood described in equation 5.19. Systematic errors will be added to equation 5.19 in the next section and the complete Likelihood will be derived.

5.3 Minimisation procedures and construction of confidence regions

Confidence regions can be constructed by minimising the likelihood in the multi-dimensional parameter space of the oscillation parameters and nuisance terms. Two techniques in particular are described for the construction of confidence regions: a Gaussian approximation, as recommended by the Particle Data Group (PDG) [31], and the Feldman-Cousins method [59].

One particularity of the disappearance analysis is that the favoured oscil-

lation angle $\sin^2 2\theta_{23}$ lies on the boundary between a physical and unphysical region. A few of the most popular minimisation algorithms were tested for bias and execution speed. Problems arising from boundary conditions are described and an approach of optimisation with inequality constraint is presented as a work-around.

5.3.1 Construction of confidence intervals

The Neyman construct [60] is used to find a sensitivity region for the oscillation parameters. There are some subtle differences between a frequentist approach such as the Neyman construct and a Bayesian formalism. It is important to understand the differences between both approaches in order to interpret appropriately the results of the frequentist method used.

In the case of a Bayesian construct, the Bayesian region gives the probability based on some prior assumptions that the true parameters lie in said region. Quoting a 90% Bayesian region therefore indicates a belief given prior assumption that the true parameters lie in the quoted region with 90% probability.

The frequentist method calculates the probability of observing a data-set given values of the parameters. The estimators are the values of the parameters for which the probability density function evaluated with the data is maximised. The Neyman construct specifies a procedure to build frequentist confidence intervals. For all possible true values, the interval in the estimators with a specified probability of occurring is calculated and a confidence belt of these intervals is built. Following a measurement, the estimators are calculated and the corresponding confidence interval of true values with probability as specified during the construct of the confidence belt is read out from the confidence belt. As opposed to the Bayesian construct, a 90% confidence interval does not mean that the confidence interval contains the true values with a 90% probability. For any one set of data, the probability could be less or even zero if the confidence interval lies in the unphysical region. However, on average, 90% of generated confidence intervals would contain the true value.

Constructing proper frequentist confidence intervals is often difficult so approximations are generally used. The Gaussian approximation recommended by The Particle Data Group (PDG) is described below.

Gaussian approximation

An approximate confidence interval for the oscillation parameters can be built based on the probability content of a χ^2 with d degrees of freedom by finding

the contour of values $(\mathbf{\sin^2\theta_{23}})'$ and $(\mathbf{\Delta m^2_{32}})'$ such that

$$\ln L((\mathbf{\sin^2\theta_{23}})', (\mathbf{\Delta m^2_{32}})') = \ln L_{\max} - \frac{\Delta\chi^2(d)}{2}, \tag{5.23}$$

where $\ln L_{\max}$ is the value of the maximum of the log likelihood. Table 5.1 presents some relevant χ^2 values used in the Gaussian approximation.

	$\Delta\chi^2$		
Confidence Level	DOF = 2	DOF = 3	DOF = 4
68 %	2.28	3.51	4.70
90 %	4.61	6.25	7.78
95 %	5.99	7.82	9.49
99 %	9.21	11.35	13.28

Table 5.1: $\Delta\chi^2$ for common Confidence Levels for 2, 3 and 4 Degrees Of Freedom (DOF)

The Gaussian approximation is valid in the large sample limit or in the case of a linear model with Gaussian errors. This method does not however give proper coverage if these conditions are not met or near a physical boundary Coverage is defined as the proportion of the time that the interval contains the true value of interest. In other words, an interval constructed with 90% confidence level under an invalid Gaussian approximation will not contain the true values of the parameters 90% of the time.

Past experiments such as MINOS have measured the θ_{23} mixing angle and found maximal mixing. At maximal mixing, that oscillation parameter $\sin^2 2\theta_{23}$ is 1, right at the physical boundary. Since the Gaussian approximation recommended by the PDG doesn't have the right coverage near a physical boundary, the Feldman-Cousins method is considered instead.

Feldman-Cousins method

The Neyman confidence interval construction allows freedom in the procedure used to define the region that contains the desired percentage (say 90%) of the probability content for a fixed value of the parameters. Popular choices include the following where in all cases the chosen region should contain 90% of the probability:

- Drawing a central region with equal probability of the measurement falling either above or below.
- Using an upper or lower limit: all parameters with value greater than or less than a given value are added to the region.

- For binned distributions with no particular shape, points can be added in decreasing order of maximum probability until the region contains 90% of the probability.

The Feldman-Cousins method proposes a different ordering procedure that is meant to fix some issues with classical confidence intervals. Namely, the possibility for confidence intervals to return the empty set, issues with unphysical regions and discontinuities caused by "flip-flopping" (changing the ordering procedure from an upper or lower limit to a central region).

The ordering principle proposed by Feldman and Cousins uses the likelihood ratio:

$$R = \frac{P(\mathbf{x}|\boldsymbol{\mu})}{P(\mathbf{x}|\boldsymbol{\mu}_{\mathbf{best}})}, \tag{5.24}$$

where $\mathbf{x}$ is the measured value of the parameters, $\boldsymbol{\mu}$ is the true value and $\boldsymbol{\mu}_{\mathbf{best}}$ is the best fit to the data. Since we don't have probability density functions as required, a Monte Carlo procedure is used instead to re-normalise the $\Delta\chi^2$ value from the Gaussian method at each point in the parameter space. The implementation of this Monte Carlo procedure is the following:

- Many fake experiments are done for one set of true parameters.
- For each fake experiment, the $\Delta \log L$ between the best fit and true value is calculated.
- The required number of experiments (say 90%) are ordered in terms of increasing $\Delta \log L$. The threshold is then set by $\Delta \log L$ which contains 90% of the fake experiments.
- This procedure is repeated over all possible sets of true parameters.

As mentioned previously, the main advantage of this method compared to the Gaussian approximation is that it get rids of physical boundary issues. On the other hand the Monte Carlo procedure used to re-normalise the $\Delta\chi^2$ value at each point in the parameter space is costly in terms of computing time.

5.3.2 Minimisation of the likelihood

One method to find the estimators of the oscillation parameters and construct confidence regions is to minimise the negative log likelihood function in the multi-dimensional parameter space. Constraints for the viability of a particular method are determined by usual factors: accuracy, speed and cost.

The bottleneck of minimisation algorithms in terms of computational speed is generally the evaluation of the likelihood function for a set of parameter values. In the case at hand however, the likelihood function first needed to be re-calculated and then evaluated for every parameter change. In particular,

the likelihood function was found to take an order of seconds to compute. It was therefore critical that any algorithm used minimised the likelihood with as few evaluations as possible.

Two general categories of algorithms were identified: those that require the computation of the gradient of the likelihood function at each step and those that don't. In general, algorithms that use gradients such as "conjugate gradient" or "quasi-Newton" methods are significantly more efficient than methods that don't such as a "simplex" method [61]. However, in the first case, the gradient of the likelihood function at each point must be computed using finite differences since the functions analytical form is not known. This implies that the likelihood function be calculated at least an order of n times more at each evaluation for gradient methods and seriously deteriorates the efficiency.

Rather than re-inventing the wheel, minimisation algorithms from the GNU Scientific Library (GSL) [62] and MINUIT [63] package were tested.

Overall, both categories of algorithm were found to share the same flaw for the problem considered: poor scaling with increasing dimensions.

5.3.3 Optimisation with inequality constraint

The most likely value for the oscillation parameter $\sin^2 2\theta_{23}$ sits at 1.0 on the boundary between the physical and unphysical region. When minimising the likelihood function with regards to the oscillation parameter $\sin^2 2\theta_{23}$, one of two scenarios must be chosen: either

1. constrain the oscillation parameter to the physical region when minimising the likelihood: $\sin^2 2\theta_{23} <= 1.0$,

2. or, do not constrain the parameter and let the minimiser find a minimum in the non physical region.

However, both choices have underlying issues. Because fast multi-dimensional minimisers calculate the gradient of the likelihood function, they do not work well close to physical boundaries if a discontinuity is introduced by constraining the likelihood. On the other hand, minimisers were found to frequently diverge in the non physical region. We may also wish to find a local minimum in the physical region.

Resolution of minimisation problems with such constraints are commonly referred to as optimisation problems with inequality constraint. These problems are often resolved by transforming the bounded problem into an unbounded one.

Penalty function method

With the penalty function method, a continuous penalty function is added to the likelihood with the aim of ensuring that the minimiser does not reach a stable state in the non physical region. The likelihood is then minimised with no constraint. A quadratic penalty function is often used to ensure that the derivative of the function is also continuous. This is particularly relevant for minimisation algorithms that calculate the Hessian matrix at each point, such as quasi-Newton methods.

The updated likelihood function with quadratic penalty is given by:

$$L^{'}(\sin^2\theta_{23}, \Delta m^2_{32}, \rho) = L(\sin^2\theta_{23}, \Delta m^2_{32}) + \rho(\sin^2\theta_{23} - 1)^2, \tag{5.25}$$

where:

$$\rho = \begin{cases} 0 & \text{if } \sin^2\theta_{23} \leq 0, \\ \rho \gg 0 & \text{if } \sin^2\theta_{23} > 0. \end{cases} \tag{5.26}$$

With very large values of ρ, the penalty function effectively forces the minimiser into the physical region whilst keeping the likelihood function well behaved.

The sequential unconstrained minimisation technique (SUMT) [64] was used to minimise the likelihood in order to avoid instabilities and inefficiencies associated with very large values of ρ. Using this technique, a small value of ρ is chosen initially and then incrementally increased as the solution is found.

Augmented Lagrange multiplier method

In the penalty method approach, the constrained minimum approximately coincides with the minimum of the penalised function but high accuracy is not guaranteed due to numerical instabilities associated with large values of ρ.

The augmented Lagrange multiplier method combines the classical Lagrangian method with the penalty function approach and does not suffer from the same ill-conditioning.

The classical Lagrangian method defines a Lagrangian function, $L(x, y, \lambda)$, based on the constrained function to minimise: $f(x, y)$ with constraint $g\,(x, y) \leq b$, such that the minimum of the constrained problem coincides with a stationary point of the Lagrangian function. The classical Lagrangian in the example outlined is given by:

$$L(x, y, \lambda) = f(x, y) - \lambda g\,(x, y)\,, \tag{5.27}$$

and the minimum found by solving the set of equations

$$\frac{\partial L}{\partial x} = 0,\ \frac{\partial L}{\partial y} = 0,\ \lambda(g\,(x,y) - b) = 0. \tag{5.28}$$

The augmented Lagrange multiplier method, on the other hand, is based on the use of approximations of the Lagrange multipliers. If a good approximation to the classical Lagrange multipliers is found, then it is possible to approach the optimum through the unconstrained minimisation of the augmented Lagrange function without using large values of ρ for the penalty term. The value of ρ must only be sufficiently large to ensure that the augmented Lagrangian has a local minimum point with respect to the parameters rather than simply a stationary point at the optimum [64]. A comparison of the partial derivatives of the classical and augmented Lagrange function suggests the following iterative approximation to determine an approximation for the Lagrange multiplier:

$$\lambda_{k+1} = \langle \lambda_k + 2\rho_k g(x_k) \rangle \tag{5.29}$$

The following augmented Lagrange likelihood function was used for the disappearance analysis:

$$L^{'}(\sin^2\theta_{23}, \Delta m^2_{32}, \lambda, \rho) = L(\sin^2\theta_{23}, \Delta m^2_{32}) + \left\langle \lambda/2 + \rho(\sin^2\theta_{23} - 1) \right\rangle^2, \tag{5.30}$$

where the Lagrange multiplier, λ, and penalty factor ρ were approximated iteratively using respectively equation 5.29 and

$$\rho_{k+1} = 2\rho_k, \tag{5.31}$$

if $||\lambda_k - \lambda_{k-1}|| < 0.5$.

5.4 Sampling of the likelihood function with MCMC

The main issue common to the minimisation, optimisation and Feldman-Cousins methods introduced in chapter 5.3 is the heavy reliance on computing power and poor scalability for problems of high dimensionality. This section explores alternative solutions that don't suffer from these drawbacks.

The simple two flavour disappearance model is a function of two oscillation parameters and various systematic uncertainty or "nuisance" terms. Marginalisation techniques exist to effectively integrate over the nuisance terms. Section 5.4.1 introduces the concept and justifies why the problem is best solved by using a Monte Carlo method. Section 5.4.3 introduces the MCMC method used to sample the likelihood function in the whole parameter space and

marginalise over the nuisance terms. Numerous MCMC methods exist but the methodology of a random-walk MCMC in particular is described. Finally, the construction of credible regions from the MCMC method is detailed in section 5.4.4.

5.4.1 Marginalisation of nuisance terms

The concepts of nuisance term and marginalisation can be illustrated using a fully specified probabilistic model in discrete space. A fully specified probabilistic model gives the joint probability for every combination of values that its variables can take. Table 5.4.1 presents a simplistic joint distribution for a model of the day's weather (example reproduced from [65]).

x1	x2	x3	P(x1,x2,x3)
clear	rising	dry	0.40
clear	rising	wet	0.07
clear	falling	dry	0.08
clear	falling	wet	0.10
cloudy	rising	dry	0.09
cloudy	rising	wet	0.11
cloudy	falling	dry	0.03
cloudy	falling	wet	0.12

Table 5.2: Joint distribution for a model of the day's weather. x1 = Morning sky (clear or cloudy), x2 = Morning barometer (rising or falling), x3 = Weather in the afternoon (dry or wet)

Say, for the purpose of illustration, that only the morning sky parameter, x1, is of concern. Then the morning barometer, x2, and afternoon weather, x3, are called "nuisance" terms. To find the probability that the sky is, for example, "clear" in the morning, we marginalise over all the other parameters by summing over them:

$$\begin{aligned} P(\text{clear}) &= P(\text{clear}|\text{rising},\text{dry}) + P(\text{clear}|\text{rising},\text{wet}) + \\ &\quad P(\text{clear}|\text{falling},\text{dry}) + P(\text{clear}|\text{falling},\text{wet}) \end{aligned} \tag{5.32}$$

In general, marginal probabilities for subsets of the variables are found by summing over all possible combinations of values for the other variables in a discrete case, or, integrating over them if they are continuous.

In the case of the disappearance likelihood, the oscillation parameters and nuisance terms (the systematics uncertainties) are all continuous. Numerical integration method could be used to integrate over the nuisance parameters but this option is too costly computationally and impractical because the likelihood function needs to be re-calculated for each change of parameter.

Monte Carlo integration methods can be significantly more efficient for the calculation of multi-dimensional integrals. Section 5.4.2 presents the techniques of simple Monte Carlo integration and importance sampling. Section 5.4.3 presents the superior class of MCMC and the particular algorithm selected: a random walk MCMC.

5.4.2 Monte Carlo integration

Monte Carlo integration

The simple Monte Carlo integration method can be used to integrate a function, f, over a complicated domain D. To perform the integration, a simple domain, $D^{'}$, which contains D is chosen and sample points are generated randomly over $D^{'}$, as illustrated in figure 5.1. The volume V of D is estimated

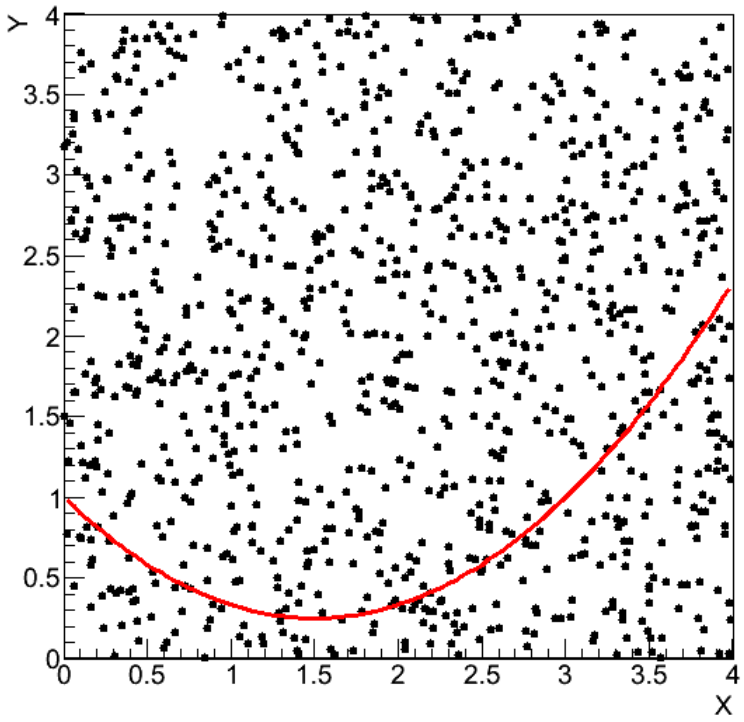


Figure 5.1: Illustration of Monte Carlo integration in 1 dimensions (extensible to multi-dimension integration). Some complicated (undefined here) function f is to be integrated over the domain D delimited by the x axis and the red curve. A simple domain $D^{'}$, represented here, which contains D is chosen and sample points are generated randomly over $D^{'}$.

as the volume $V^{'}$ of $D^{'}$ times the fraction of points falling in D. The integral

of f over D is then given by:

$$\int f dV \simeq v\langle f\rangle \pm \sqrt{\frac{\langle f^2\rangle - \langle f\rangle^2}{N}}, \tag{5.33}$$

where:

$$\langle f\rangle \equiv \frac{1}{N}\sum_{i=1}^{N} f(x_i) \text{ and } \langle f^2\rangle \equiv \frac{1}{N}\sum_{i=1}^{N} f^2(x_i) \tag{5.34}$$

The main issue with the simple Monte Carlo method is one of efficiency: a large fraction of generated sample points are wasted and regions of high and low density are sampled with equal probability.

These issues are partially resolved by the method of importance sampling.

Importance sampling

In the importance sampling method, also called weighted sampling, a distribution h(x) that approximates the region of integration is sampled instead. Samples are then weighted by f(x)/h(x), where f(x) is the function to be integrated. This results in fewer wasted points evaluating the function where it is known to be small. On the other hand, it may be difficult to find a suitable approximate distribution h(x), Monte Carlo average may be dominated by a few samples (high variance), or, none of the high weight samples may be found.

Both methods and in particular importance sampling could potentially deliver good results for the disappearance study. None the less, an MCMC algorithm, traditionally used to sample from high dimensional distributions, was selected instead.

5.4.3 Markov Chain Monte Carlo (MCMC) method

Using MCMC, sequences of random numbers that accurately reflect complicated and multi-dimensional probability density functions can be generated.

Markov chain definition: A Markov chain refers to a sequence of random variables $(X_0, X_1, \cdots, X_n)$ generated by a Markov process.

Markov process definition: Let x_t denote the value of a random variable at time t, and let the state space refer to the range of possible X values. The random variable is a Markov process if the transition probabilities between different values in the state space depend only on the random variables current state, i.e:

$$P(X_{t+1}|X_0, \ldots, X_t) = P(X_{t+1}|X_t) \tag{5.35}$$

A random walk is a good example of a Markov process: the probability for the next step in the random walk to end at a particular location depends entirely on the current position. Card games such as blackjack or poker on the other hand are not: the whole history of the game, i.e which cards were already drawn, is needed to calculate to probability for the card to come next.

Some notation

A Markov process is entirely defined by its transitional probability. For a discrete process, the transition matrix $\mathbf{P}$ gives the probability $P(j,i)$ that a process at state space s_i moves to state s_j in a single step:

$$P(j,i) = P(X_{t+1} = s_j | X_t = s_i). \tag{5.36}$$

For a continuous process, the transition matrix becomes a probability kernel $P(x,y)$ that satisfies:

$$\int P(x,y)dy = 1. \tag{5.37}$$

The probability that the chain is in state j at time t for the discrete case is given by:

$$\pi_j(t) = P(X_t = s_j), \tag{5.38}$$

while $\boldsymbol{\pi}(t)$ denotes the row vector of the state space probabilities at step t. The probability that the chain has state value s_i at step $t+1$ is given by the Chapman-Kolomogrov equation:

$$\pi(t+1) = \pi(t)\mathbf{P}, \tag{5.39}$$

and by extension,

$$\pi(t) = \pi(0)\mathbf{P}_t. \tag{5.40}$$

The Chapman-Kolomogrov equation for a continuous process is given by:

$$\pi_t(y) = \int \pi_{t-1}(x)P(x,y)dy = 1. \tag{5.41}$$

More details can be found in [66]

The Metropolis-Hastings algorithm

The general Metropolis algorithm was first proposed in 1953 for the Monte Carlo simulations of atomic and molecular systems [67].

The underlying principle behind this algorithm and its derivatives, such as Metropolis Hastings, is the generation of a Markov chain whose equilibrium density is the density of the function being sampled.

Given a distribution $f(\mathbf{x})$ where $\mathbf{x}$ represents a state vector of n parameters: $\mathbf{x} = [x_1, x_2, \cdots, x_n]$, the Metropolis algorithm generates a sequence of draws from $f(\mathbf{x})$ according to the procedure outlined:

1. A set of initial values for the state vector parameters, $\mathbf{x_0} = [x_1^0, x_2^0, \cdots, x_n^0]$, is first chosen such that $\mathbf{x_0}$ satisfies the condition $f(\mathbf{x_0}) > 0$. From there on, repetitively,

2. a sample candidate $\mathbf{x}_{t+1}$ is chosen according to a proposal distribution $q(\mathbf{x}_{t+1}, \mathbf{x}_t)$ that gives the probability of returning a value, $\mathbf{x}_{t+1}$, given a previous value, $\mathbf{x}_t$. In the case of the Metropolis algorithm, this proposal distribution is required to be symmetric: $q(\mathbf{x}_{t+1}, \mathbf{x}_t) = q(\mathbf{x}_t, \mathbf{x}_{t+1})$.

3. The density ratio of the distribution $f(\mathbf{x})$ between the sampled candidate point $\mathbf{x}_{t+1}$ and current point $\mathbf{x}_t$ is calculated:

$$\alpha = \frac{f(\mathbf{x}_{t+1})}{f(\mathbf{x}_t)}. \tag{5.42}$$

4. If $\alpha > 1$, the proposed point increases the density, is automatically accepted, and the current point is replaced by the sample candidate point. If $\alpha < 1$, the proposed point decreases the density and is only accepted with probability α. In both cases, the chain is continued by returning to the second step.

The Metropolis algorithm generates a Markov chain $(\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n})$ since the transition probabilities only depend on the current element of the chain.

The algorithm was generalised by Hastings such that the proposal distribution from step 2 need not be symmetric [68]. In the general Metropolis-Hastings algorithm, the density ratio from step 3 is replaced by:

$$\alpha = \frac{f(\mathbf{x}_{t+1})q(\mathbf{x}_{t+1}, \mathbf{x}_t)}{f(\mathbf{x}_t)q(\mathbf{x}_t, \mathbf{x}_{t+1})}. \tag{5.43}$$

In the case that the proposal distribution is symmetric, the Metropolis algorithm is automatically recovered.

Metropolis-Hastings for the disappearance analysis

The distribution we wish to sample in the disappearance analysis is the likelihood, $L(\sin^2\theta_{23}, \Delta m^2_{32}, \mathbf{f})$ which is a function of the oscillation parameters and the systematic uncertainties $\mathbf{f}$. If we call $\boldsymbol{\theta}_t$ the vector of parameters, the density ratio for the Metropolis algorithm is given by:

$$\alpha = \frac{L(\boldsymbol{\theta}_{t+1})}{L(\boldsymbol{\theta}_t)}. \tag{5.44}$$

By taking the natural logarithm of equation 5.44, the likelihood distribution can be sampled by using the negative of the log likelihood instead:

$$\ln \alpha = \ln L(\boldsymbol{\theta}_{t+1}) - \ln L(\boldsymbol{\theta}_t). \tag{5.45}$$

The so called proposal distribution used is a random perturbation in each parameter:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \boldsymbol{\epsilon}, \tag{5.46}$$

where $\boldsymbol{\epsilon}$ is a vector of random numbers generated from a multi-dimensional uniform distribution with boundaries defined by a vector $\mathbf{t}$: $[-\mathbf{t}, +\mathbf{t}]$. The resulting Metropolis algorithm takes a random walk along the likelihood function density.

Tuning of the MCMC

Given sufficient time, Markov chains generated with the Metropolis algorithm approach a stationary distribution (the one that is being sampled from). However, poor choice of starting values or proposal distribution can greatly increase the "burn-in" period required to reach stationarity. Samples generated during the "burn-in" period are discarded resulting in longer chains being required.

Unlike other Monte Carlo techniques, such as importance sampling, successive draws in a random walk MCMC are correlated. However, correlated draws still provide an unbiased picture of a stationary distribution provided that the sample size is sufficiently large. The variance of a correlated sample is greater than for an uncorrelated sample by the "sample size inflation factor" [69]:

$$\sqrt{\frac{1+\rho}{1-\rho}}, \tag{5.47}$$

where ρ gives the correlation.

The acceptance probability α of the Metropolis-Hastings algorithm is linked with the mixing (defined below) and correlation of a Markov chain. The time series of Markov chains that are characterised by a low acceptance probability have long flat periods corresponding to strings of rejection of proposed points. On the other hand, chains with very high acceptance often move in very small steps and take time to explore the whole space. In both scenarios, the chains are said to be poorly mixing and display high levels of correlation. As a consequence, much longer chains and "burn-in" periods are required.

Tuning of the random walk Metropolis-Hastings MCMC is therefore sensitive and requires care in order improve the efficiency and reduce the computational cost.

The range of the proposal distribution for a uniform random walk is used as a handle to tune the acceptance, and by extension, the mixing and correlation

of the Markov chains generated. Figure 5.2 illustrates the results of a change of boundaries for a random walk Metropolis algorithm applied to a simple 1 dimensional Gaussian probability function with mean of 0 and standard deviation of 1. Three different boundaries $[-t, +t]$ are tested for the uniform proposal distribution. Respectively $[-1, +1]$, $[-0.5, +0.5]$ and $[-0.1, +0.1]$. A Markov chain of 5000 elements is generated in each case, and, in each case, the distribution of values is fitted with a Gaussian function. The convergence of the mean of the chain and the mixing are also presented.

The Gaussian fit to the Markov chain generated with $t = 1.0$ gives a reasonable approximation for the mean and sigma of the underlying distribution and a χ^2 of 1.80. The mean also appears to converge after a few thousand chain elements and the mixing plot shows that the whole parameter space is being sampled uniformly.

The fit to the Markov chain with $t = 0.5$ has an even better χ^2 of 1.18 but the approximation for the mean and sigma are poor with respectively 0.07 ± 0.01 and 0.94 ± 0.01. In this case, the convergence of the mean of the distribution is clearly slower than for the first case and the mixing shows that the chain takes longer to sample the parameter space.

The last Markov chain with $t = 0.5$ presents an example of poor mixing with high acceptance (close to 100%) and no convergence.

To illustrate that Markov chains will eventually converge even in cases where bad decisions lead to poor mixing and high correlation, the same example is run for chains 500,000 elements and the results presented on figure 5.3.

5.4.4 Construction of Bayesian credible regions

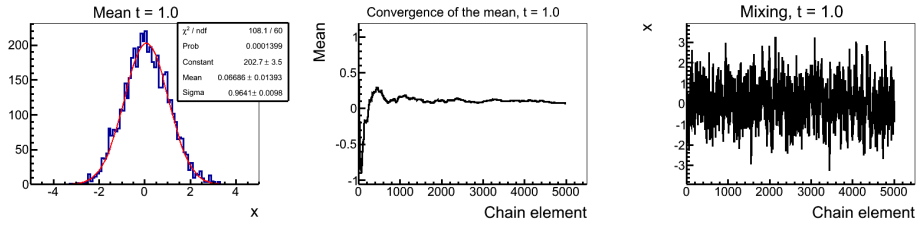
Bayesian credible intervals are built using the following procedure:

1. Some binning width is chosen and the unbinned Markov chain data is transformed into a binned histogram. Prior information can be used to restrict the range of the histogram.

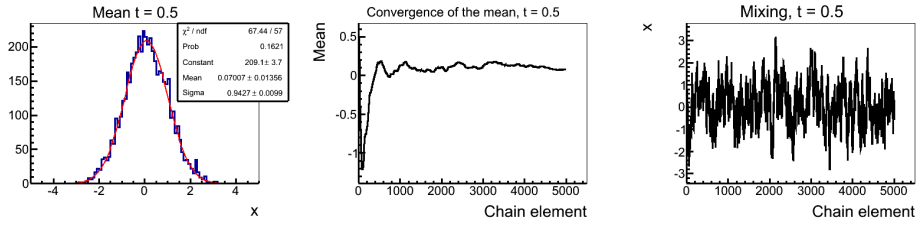
2. Confidence intervals are then constructed by picking bins in order of decreasing number of events until the fraction of events corresponding to the required sensitivity is reached (68%, 90% or 99% for example).

5.5 Signal event selection and Backgrounds at Super-K

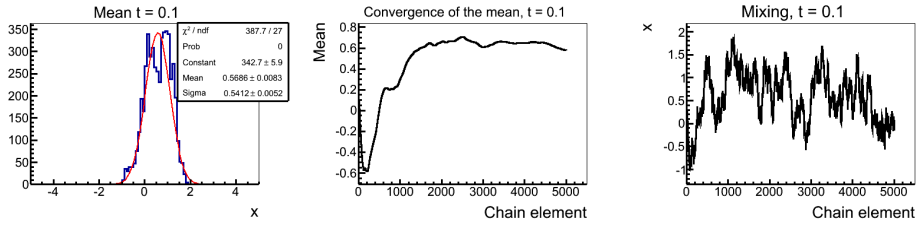
Data collected at the Super-K detector are sorted using a set of selection cuts. The same set of selection cuts is applied to the Super-K NTuples in order to build the Super-K templates. This section looks at the cuts used and how



(a)



(b)



(c)

Figure 5.2: The random walk Metropolis algorithm is applied to a simple 1 dimensional Gaussian probability function with mean of 0 and standard deviation of 1. The proposal distribution is a uniform distribution in the range $[-t, +t]$. A Markov chain of 5000 elements is generated for 3 cases. In each case, the distribution of values is fitted with a Gaussian function (left), and the convergence of the mean (middle) and mixing (right) are presented. The three cases have different boundary conditions for the uniform proposal distribution with: (a): $t = 1.0$, (b): $t = 0.5$, (c): $t = 0.1$.

they improve the signal efficiency and purity. Sources of background are also discussed and categorised.

5.5.1 Signal

The signal event at Super-K is a muon-like ring-shaped hit pattern, produced when Cherenkov photons hit the PMT's on the surface of the Super Kamiokande detector, initiated by a ν_μ charged current quasi elastic event (CCQE). A CCQE event is an exchange of a W boson between a neutrino and nucleus leading to the emission of a charged lepton.

The selection cuts applied to select muon rings are the so called "Fully Contained Fiducial Volume (FCFV) 1 ring μ-like tight cuts" [70] summarised

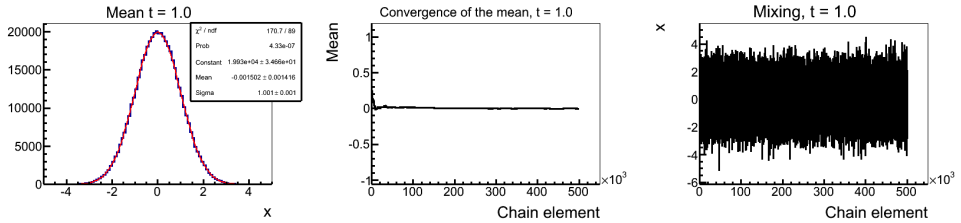


Figure 5.3: Same as figure 5.2 with Markov chains of 500,000 elements instead of 5,000.

below:

1. Number of PMT hits in highest charge Outer Detector cluster <= 15;
2. Reconstructed vertex from Inner Detector wall > 2m;
3. Visible energy > 30MeV;
4. Exactly 1 ring as reconstructed by the Super-K software;
5. Reconstructed particle ID of a muon;
6. Reconstructed momentum of the muon like ring > 200 MeV;
7. Reconstructed signal for 0 or 1 time delayed decay electrons.

The first cut on the number of PMT hits in the highest charge OD cluster is used in order to remove events that originated outside of the fiducial volume.

The second cut defines the fiducial volume as starting from two metres inwards from the walls of the inner detector. The resulting loss in efficiency

can be calculated using the diameter d_{ID} and height h_{ID} of the inner detector cylinder:

$$\frac{d_{ID}^2 h_{ID}}{(d_{ID}-4)^2(h_{ID}-4)} = 0.69.$$

This cut reduces the fiducial volume by more than 30% but events that occur close to the wall are poorly reconstructed and are associated with large systematic errors [71].

The visible energy cut is used to reduce contamination from the photomultiplier tubes and corner events (events that enter and exit the tank in a "corner" and are not removed by the outer detector cut) and to cut events with poor energy reconstruction.

The 4^{th} and 5^{th} cuts are used to select events with only one ring where the PID of the ring is that of a muon. These two cuts are self explanatory given that the signal is the ν_μ CCQE interaction mode.

The momentum cut of the muon like ring is used to remove events with reconstructed momentum below 200 MeV where the PID probability between data and Monte Carlo is below requirements.

The final cut on the decay electron increases the signal purity at no cost for the efficiency by requiring that there are less than 2 decay electrons. Decay electrons cause time delayed secondary rings and come from a muon decay. The CCQE signal is one muon so at most one decay electron is expected for the situation when the muon decays before it is absorbed. Some backgrounds such as charged current single pion production can appear like signal if only one of the particles in the final state is above Cherenkov threshold. However, pions quickly decay to muons and there is a chance that both muons will decay to a Michel electron. The Cherenkov threshold of the electron is very low such that both electrons are guaranteed to give a signal. Basically, observing more than 1 decay electron is a strong indication that the reaction was not CCQE.

The purity and efficiency of the selection cuts on the signal from the Monte Carlo based on Super-K NTuples are presented in figure 5.4 below:

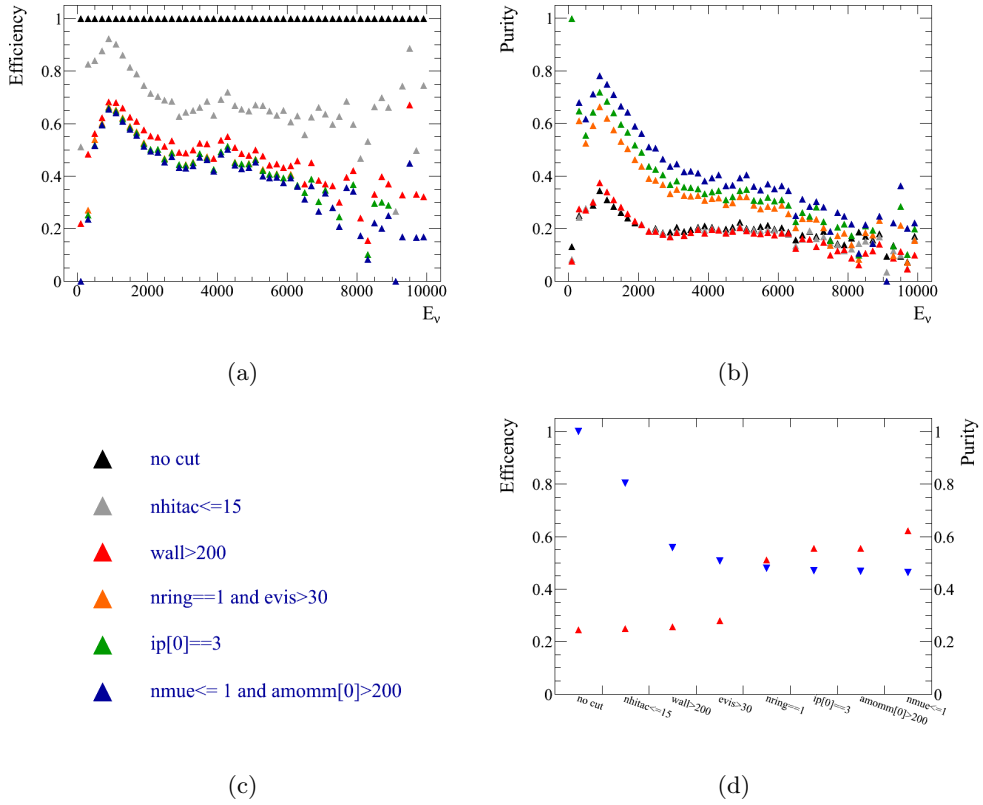


Figure 5.4: (a) Efficiency of the CCQE signal selection as a function of the oscillated reconstructed energy after each cut (Monte Carlo). (b) Purity of the CCQE signal over background from other ν_μ interactions as a function of the oscillated reconstructed energy after each cut (MC). (c) Legend for 5.4(a) and 5.4(b). (d) Total efficiency and purity after each successive cut.

The energy and momentum of the "assumed" muon and the angle of the particle with respect to the beam direction are reconstructed based on the charge and timing of the hits from the muon like ring. The energy of the parent neutrino particle is then reconstructed using two body kinematics with the neutron and neutrino particles as the initial state and the muon and proton as the final state:

$$E_{\nu_\mu} = \frac{m_p^2 - (m_n - V)^2 + 2(m_n - V)E_\mu - m_\mu^2}{2(m_n - V - E_\mu + pcos\theta_\mu)}, \tag{5.48}$$

where the parameters are listed below:

- m_p: mass of proton;
- m_n: mass of neutron;
- m_μ: mass of muon;
- E_μ: Energy of muon;
- p_μ: momentum of muon;
- θ_μ: angle of muon with respect to the beam direction;
- V: Nuclear potential energy (27 MeV for ^{16}O).

Backgrounds to the CCQE signal distort both the count and the energy spectrum of events passing the 1-ring μ-like cuts at Super-K. Understanding these backgrounds and the associated systematic errors is therefore crucial to the disappearance analysis.

5.5.2 Backgrounds

Backgrounds are any events that pass the FCFV cuts but don't come from a ν_μ CCQE reaction. We identify 3 different sources of potential background:

1. The beam
2. Cosmic events
3. Radioactivity from rocks surrounding the detector.

Each source is investigated independently below.

The beam

The beam is created from decaying pions and kaons as described in section 2.1 and contains 95% ν_μ's with a contamination of 4% from $\overline{\nu}_\mu$'s and 1% from ν_e's

Backgrounds from the beam can therefore be subdivided as a function of neutrino flavour.

ν_μ backgrounds: The signal reaction in the disappearance measurement is a ν_μ CCQE reaction but muon neutrinos interact with matter via many other modes.

CC1π: The main source of background to the disappearance analysis is charged current resonant single pion production (CC1π). The final state of a CC1π reaction is identical to a CCQE if either the π or the muon is not observed but the neutrino energy is mis-reconstructed to a lower energy value. The additional statistical information gained from using a CC inclusive analysis rather CCQE only is more than counter balanced by the loss of constraining power due to the spread of the peak and through of the distribution. The advantage of CCQE over CC inclusive is similar in effect to the advantage gained by using an off axis rather than on axis beam.

Figure 5.5 shows the energy of the pion versus the energy of the muon after the FCFV 1 ring μ like tight cuts have been applied. As expected, most events are below threshold for either the pion or the muon. For events where both particles are above Cherenkov threshold, the angular difference $\cos\theta_{\pi\mu}$ between the direction of the π and μ is plotted in figure 5.6. We see that in most cases, both particles travel in the same direction which increases the likelihood that the light signal is reconstructed as only one ring.

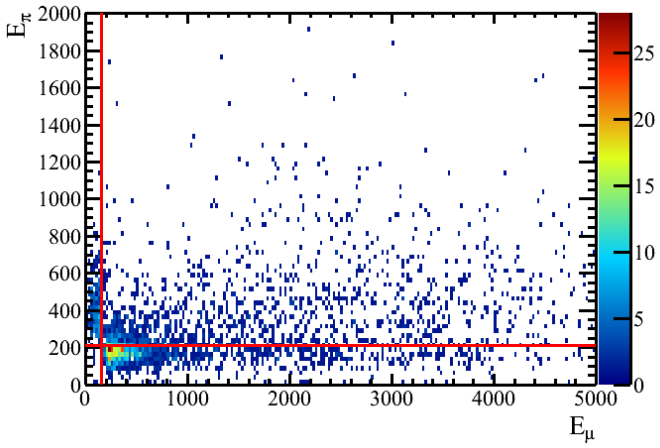


Figure 5.5: Pion energy versus muon energy of CC1π events after FCFV 1 ring μ-like tight cuts. The red lines indicate the pion and muon theoretical Cherenkov thresholds in water.

CC Other: This category includes other types of charged current reactions such as multi-π in the final state, charged current resonant single eta or kaon production or deep inelastic scattering.

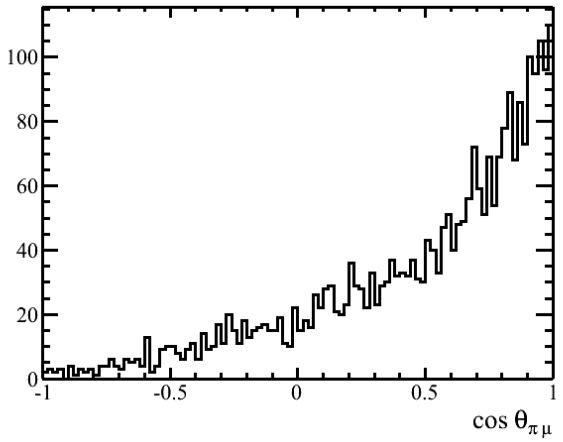


Figure 5.6: Angular difference between the directions of the pion and muon when both are above Cherenkov threshold

NC events: The largest contribution from neutral current events comes from neutral current resonant single charged pion production. If the pion is above Cherenkov threshold, its ring can be mis-identified as a muon ring. Other contributing reactions include charge current multi-π, charged current resonant single π^0 or deep inelastic scattering. In the case of the CCπ^0 reaction, the final state of the reaction is identical to the CCQE signal if the γ from the pion decay is misidentified as a muon, or if the π^0 is absorbed in the nucleus.

$\overline{\nu}_\mu$ backgrounds: The $\overline{\nu}_\mu$ backgrounds are identical to the ν_μ backgrounds with the addition of the CCQE reaction which gives a μ^+ in the final state. Super-K has no sensitivity to the sign of charged particles so the final state if the $\overline{\nu}_\mu$ CCQE reaction looks identical to the signal.

ν_e backgrounds: Again the ν_e backgrounds are identical to the ν_μ backgrounds with the addition of the CCQE reaction which gives an e^- in the final state. The ν_e CCQE reaction looks identical to the signal of the electron is mis-identified for a muon.

Cosmic events

The average number of daily cosmic events that enter the fully contained fiducial volume atmospheric data sample is 8.11 ± 0.16 events/day for Super-K IV [72]). Most are neutrinos since the OD cut gets rid of almost 100% of the cosmic ray muon background [34]. However the T2K experiment uses a GPS synchronisation system to cut out events that do not fall within a time window of ± 500 μs from a spill. The duration of a spill that contains 8 bunches is approximately 5 μs and the time between 2 spills is approximately 3 s. An

additional cut on beam events removes events that miss the beam spill by approximately 2 μs. The effective "active" period is therefore approximately 10 μs for every 3 seconds and a quick calculation reveals that the average daily equivalent cosmic contamination in the FCFV sample is negligible:

$$\text{Cosmics} = \frac{10\,(\mu s)}{3\,(s)} \times 8.11 = 2.7 \times 10^{-5}\ \text{Events/Day}. \tag{5.49}$$

Radioactive background

The radioactive backgrounds are mainly low energy events cut by the selection cuts. With the GPS timing window and OD cuts included, these events are also negligible.

5.6 Super-K PDF Construction and event rate prediction

Section 5.2 introduced the Super-K reconstructed neutrino energy spectra as the observable quantity measured and used in the fitting and minimisation procedure. In this section, a step by step construction of the final reconstructed energy spectra of expected events observed at Super-K parametrised as a function of the oscillation parameters to be determined is presented. An initial overview of the steps illustrated in section 5.6.1 is given below.

1. True energy spectra are constructed using the SK NTuples for each interaction mode and neutrino parent flavour of single ring μ-like events.
2. Flux correction re-weighting is applied to the true energy spectra (if necessary).
3. True oscillated spectra which are functions of the oscillation parameters are obtained by multiplying the true Charged-Current spectra with the neutrino survival probability PDF.
4. Conversion matrices of the true to reconstructed energy are built for each interaction mode and neutrino parent flavour.
5. Reconstructed oscillated spectra are built from the true oscillated energy spectra and the conversion matrices
6. The final oscillated reconstructed energy spectrum, function of the oscillation parameters, is obtained by summing up the signal and background spectra

Predictions of the expected number of events and energy spectra at Super-K are presented for different oscillation hypotheses and amounts of collected data in section 5.6.2.

5.6.1 True and reconstructed energy spectra

1. True energy spectra:

The true energy spectra from the SK NTuples after application of the FCFV single ring μ-like tight selection cuts are presented in figure 5.7. These plots show the prediction for the shape of the signal and background spectrum in the null oscillation hypothesis for the ν_μ signal and backgrounds and the $\overline{\nu}_\mu$ and ν_e backgrounds for normalisations of respectively 269 years, 632 years and 1367 years. A year is defined as 750 kW for a 30 GeV beam for 1.07×10^7 s.

2. Flux re-weighting:

The Monte Carlo energy spectra prediction for truth information are re-weighted to correct for changes in the Flux expectations due to changes in experimental conditions at J-PARC and improvements in the flux MC. The alternative to re-weighting for flux correction would be to re-run the whole Monte Carlo. However, because of the associated costs, Monte Carlo productions are done sparingly.

The Beam group provided three histograms to re-weight interactions from the three different true neutrino parent modes of muon like events (ν_μ,$\overline{\nu}_\mu$ and ν_e parents). These histograms are reproduced in figure 5.8.

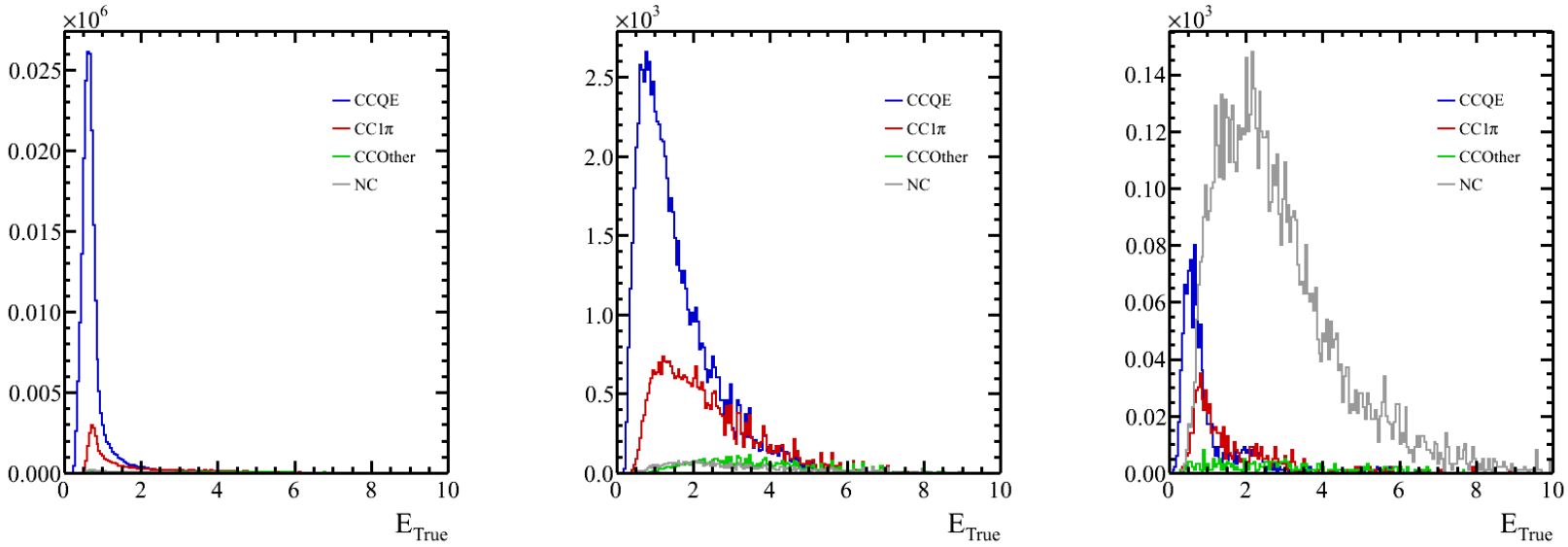
3. True oscillated spectra:

The neutrino survival probability is applied to the true re-weighted charged current energy spectra. Figure 5.9 illustrates the effect of varying the oscillation parameters $\sin^2 2\theta_{23}$ and Δm^2_{32} over the relevant energy range of 0 GeV to 10 GeV.

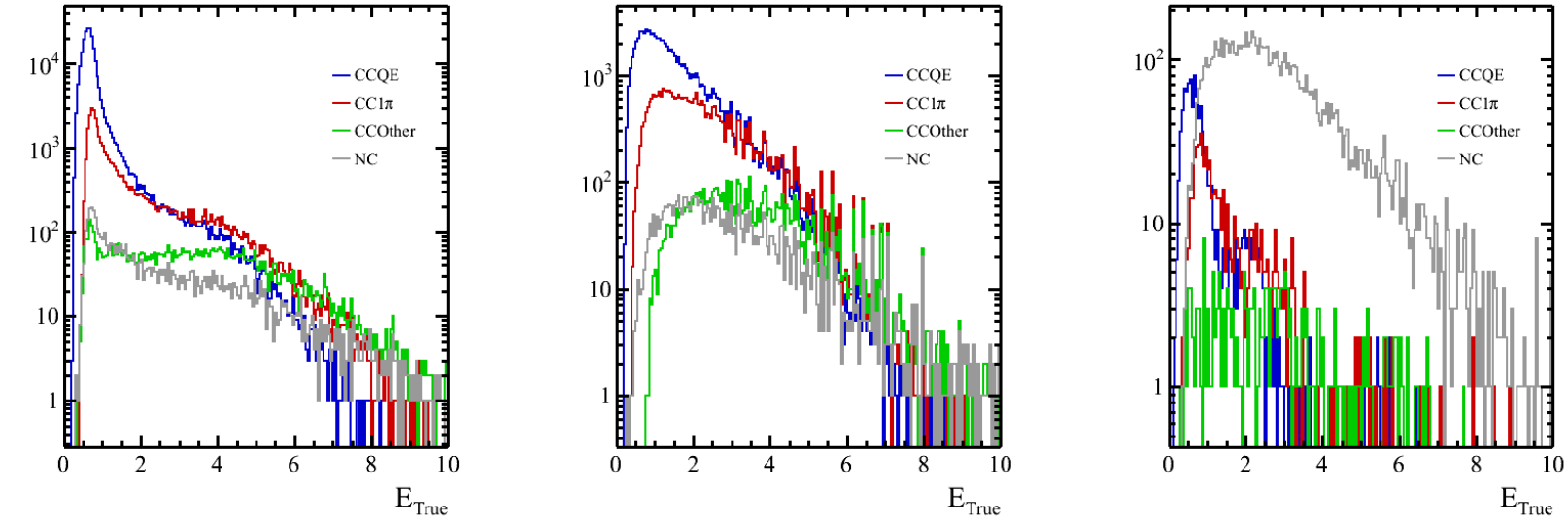
Qualitatively, variations of $\sin^2 2\theta_{23}$ affect the amplitude of the oscillation dips. For any constant value of Δm^2_{32}, increasing the $\sin^2 2\theta_{23}$ parameter results in lower neutrino survival probability. In the case where $\sin^2 2\theta_{23}$ is 0, there is no oscillation and the survival probability is 1.

The positions of the oscillation dips over the energy range are determined by the Δm^2_{32} parameter as illustrated in figure 5.9. The angle of the off-axis beam was chosen such that the dip in the neutrino survival probability for the favoured oscillation hypothesis ($\Delta m^2_{32} = 2.4 \times 10^{-3}$) coincides with the peak of the neutrino spectra in the null oscillation hypothesis.

The true re-weighted oscillated spectra for the ν_μ signal and backgrounds and the $\overline{\nu}_\mu$ and ν_e backgrounds in the case of the favoured oscillation hypothesis are presented in figure 5.10.

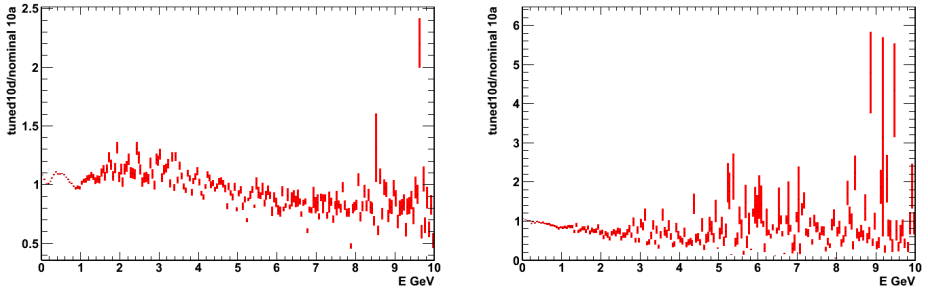


(a) True energy spectra at SK in null hypothesis for ν_μ signal and backgrounds (left), $\overline{\nu}_\mu$ backgrounds (centre) and ν_e backgrounds (right) for normalisations of 269 years, 632 years and 1367 years where a year is defined as 750 kW for a 30 GeV beam for 1.07×10^7 s.

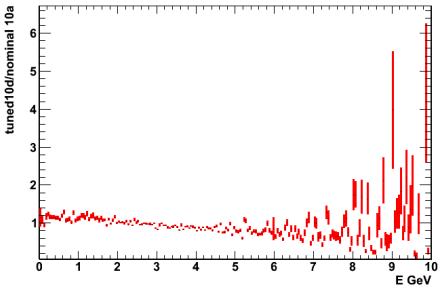


(b) Same as above in logarithmic scale.

Figure 5.7: Histograms of the Monte Carlo truth (non reconstructed) information from SK NTuples in null oscillation hypothesis



(a) Neutrino parent: ν_μ (b) Neutrino parent: $\overline{\nu}_\mu$



(c) Neutrino parent: ν_e

Figure 5.8: Super-K Flux correction for the three neutrino parent modes

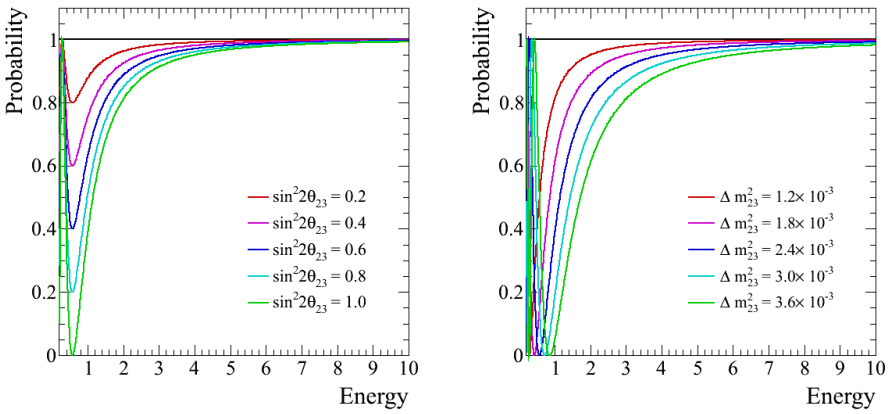
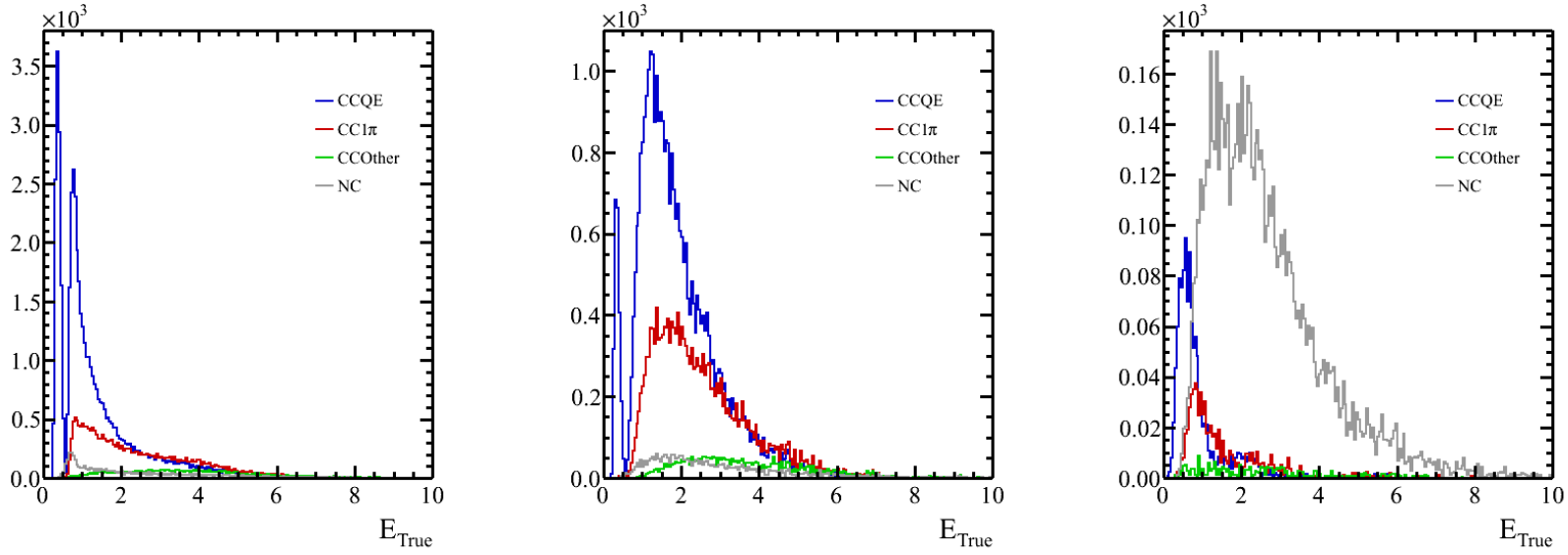


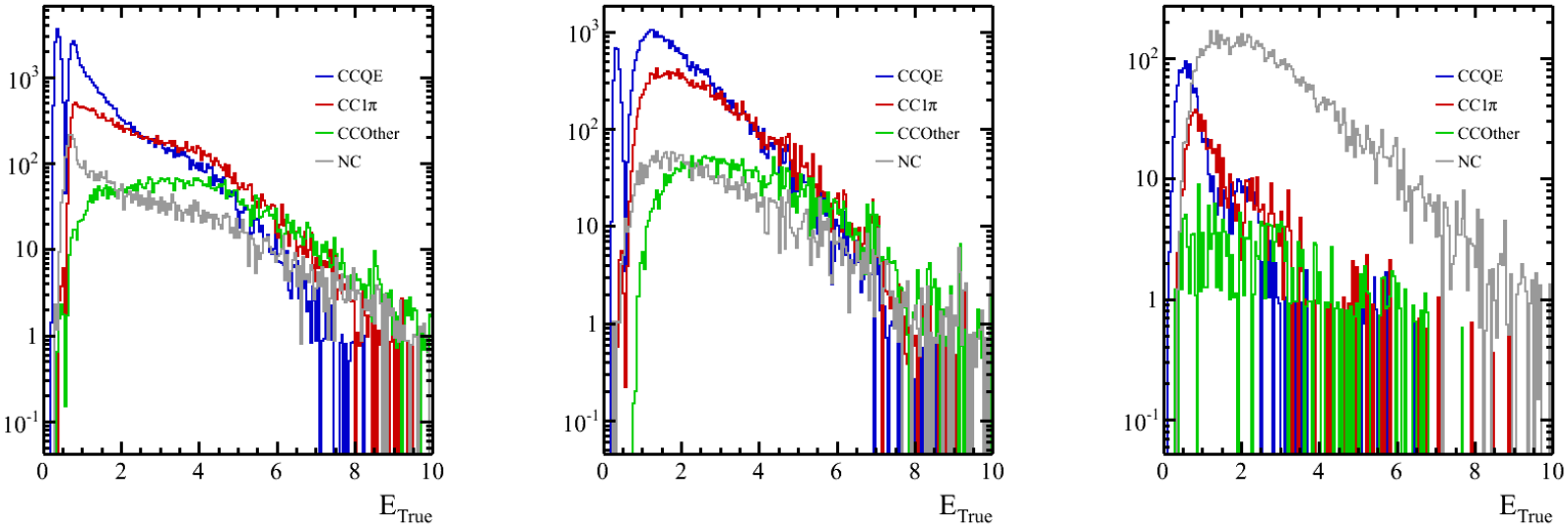
Figure 5.9: Left: Survival probability function as a function of neutrino energy for different values of $\sin^2 2\theta_{23}$ and constant Δm^2_{32} of 2.4×10^{-3}. Right: Survival probability function as a function of neutrino energy for different values of Δm^2_{32} and constant $\sin^2 2\theta_{23}$ of 1.0

4. Conversion matrices:

The energy of the parent neutrino of each μ-like ring at Super-K is reconstructed assuming two body kinematics (see equation 5.48)



(a) ν_μ signal and backgrounds (left), $\overline{\nu}_\mu$ backgrounds (centre) and ν_e backgrounds (right)



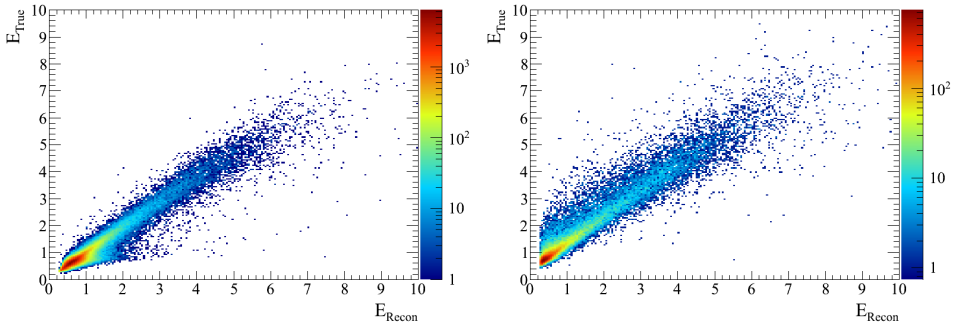
(b) Same as above in logarithmic scale.

Figure 5.10: True re-weighted oscillated histograms in favoured oscillation hypothesis ($\sin^2 2\theta_{23} = 1.0$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$) for normalisations of 269 years, 632 years and 1367 years where a year is defined as 750 kW for a 30 GeV beam for 1.07×10^7 s.

The "measured" true and reconstructed energy of simulated μ-like Monte Carlo interactions is plotted in a 2 dimensional histogram in order to build the conversion matrices which give the true to reconstructed energy conversion. A separate conversion matrix is built for each parent neutrino flavour and neutrino interaction mode of the μ-like events. The conversion matrices for each parent neutrino flavour and interaction mode are shown in figures 5.11 and 5.12.

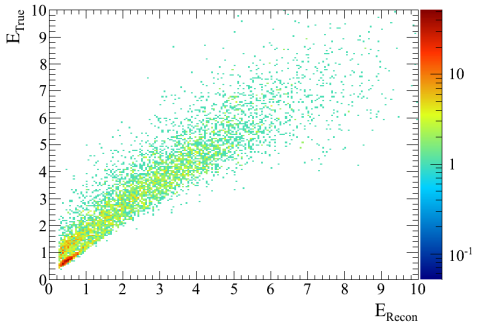
The expectation for the 2-body interaction ν_μ CCQE events is to observe a straight line along the diagonal. However, the reconstructed neutrino energy formula assumes that the hit nucleon was at rest, which is not the case. The other interactions or backgrounds are generally not 2-body collisions so the reconstructed energy is lower than the true energy. In particular, in the case of neutral current events, most of the energy is taken by the invisible neutrino in the final state and the mis-reconstructed ring only carries a small fraction of the parent neutrino. As a result, the reconstructed energy is highly biased towards low energies.

Once the conversion matrices are built, the reconstructed neutrino energy from the Monte Carlo is no longer used. Instead, the oscillation probability is applied to the true spectrum and the conversion matrices are then used to calculate the reconstructed oscillated neutrino parent energies.

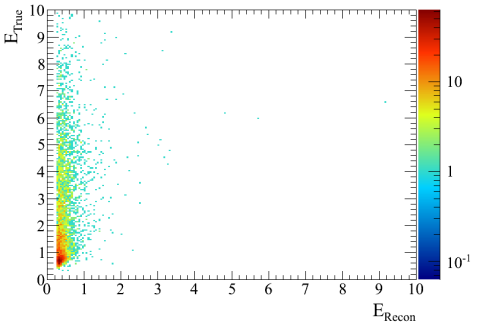


(a) SK ν_μ CCQE conversion matrix

(b) SK ν_μ CC1π conversion matrix

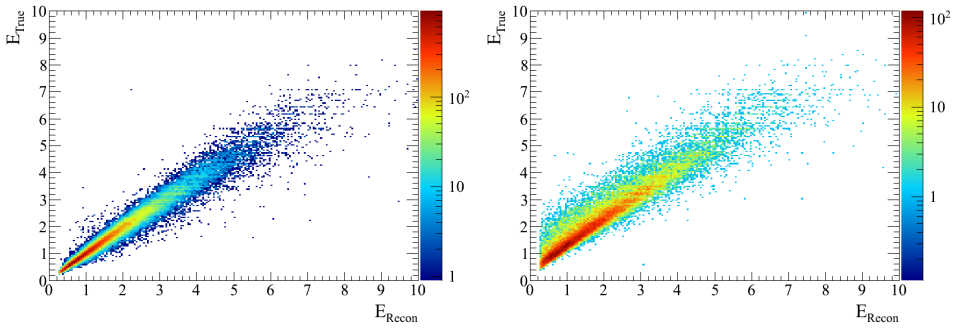


(c) SK ν_μ CCOTH conversion matrix



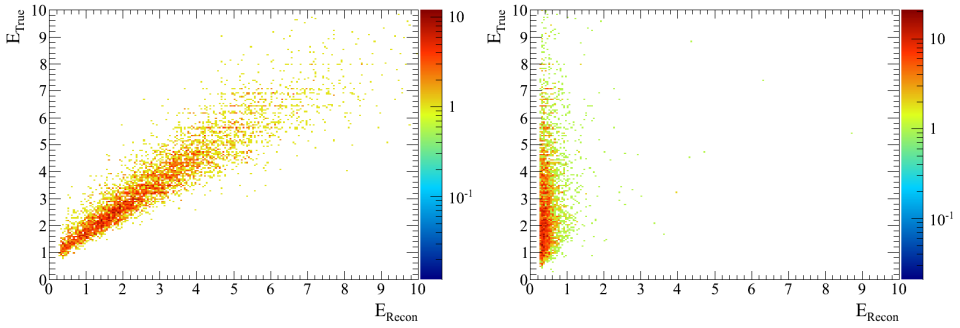
(d) SK ν_μ NC conversion matrix

Figure 5.11: Super-K ν_μ conversion matrices



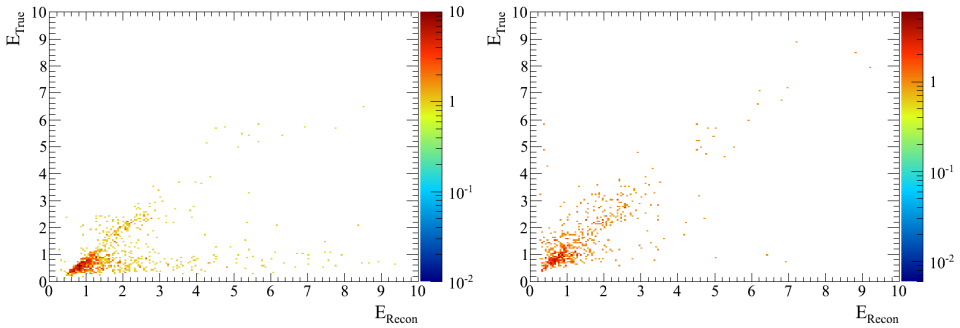
(a) SK $\overline{\nu}_\mu$ CCQE conversion matrix

(b) SK $\overline{\nu}_\mu$ CC1π conversion matrix



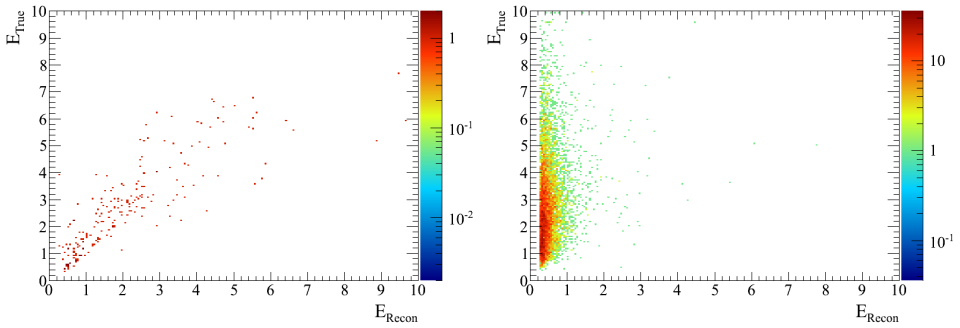
(c) SK $\overline{\nu}_\mu$ CCOTH conversion matrix

(d) SK $\overline{\nu}_\mu$ NC conversion matrix



(e) SK ν_e CCQE conversion matrix

(f) SK ν_e CC1π conversion matrix



(g) SK ν_e CCOTH conversion matrix

(h) SK ν_e NC conversion matrix

Figure 5.12: Super-K $\overline{\nu}_\mu$ and ν_e conversion matrices

5. Reconstructed oscillated spectra:

The reconstructed re-weighted oscillated spectra for the ν_μ signal and backgrounds and the $\overline{\nu}_\mu$ and ν_e backgrounds, in the case of the null and favoured oscillation hypothesis, are presented in figures 5.13 and 5.14.

6. Final reconstructed oscillated spectrum:

The final reconstructed oscillated spectrum is obtained by taking the sum of all other spectra and presented for various oscillation hypothesis in figure 5.15. The spectrum for each parent neutrino flavour is normalised to the same exposure. In figure 5.15 the exposure chosen corresponds to the T2K physics run 1 described in section 5.6.2 below.

The oscillation parameters are shown to have a strong influence on both the shape and normalisation of the final neutrino spectra. This perfectly illustrates the strong constraining power that the off-axis method has on the oscillation parameters

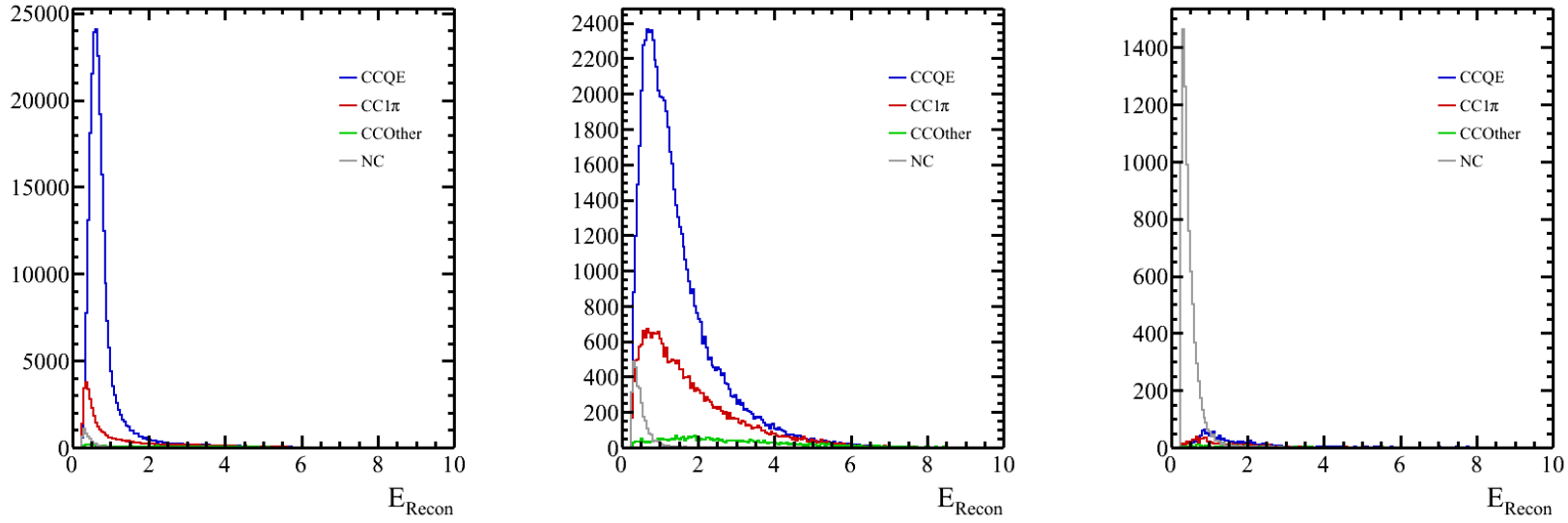
5.6.2 Super-K event rate prediction

T2K physics run 1:

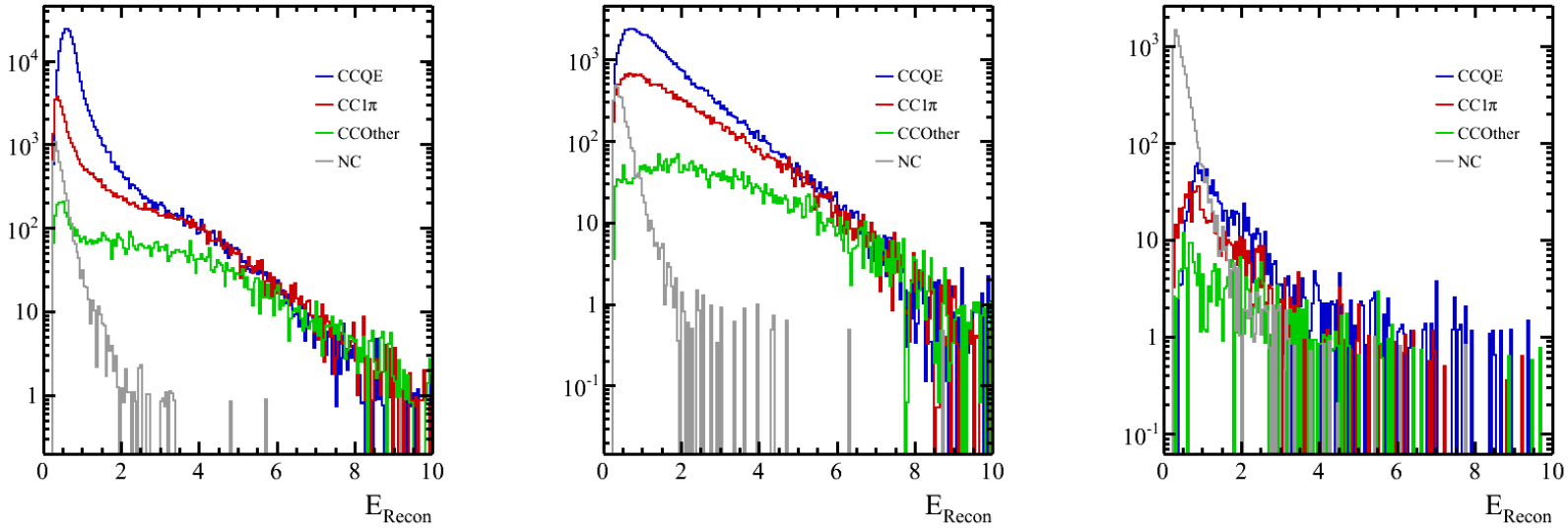
The first T2K physics data-set was collected from January to June 2010 and corresponds to an integrated J-PARC neutrino beam exposure of 3.23×10^{19} POT.

The Super-K event rate prediction in the null hypothesis and the favoured oscillation hypothesis broken down as a function of parent neutrino flavour and interaction mode are presented in table 5.3 for the T2K physics run 1. The fraction of signal and background events change for each sample because they all have a different energy response.

The total event rate prediction for the T2K run 1 as a function of the oscillation parameters $\sin^2 2\theta_{23}$ and Δm^2_{32} is presented in figure 5.16.

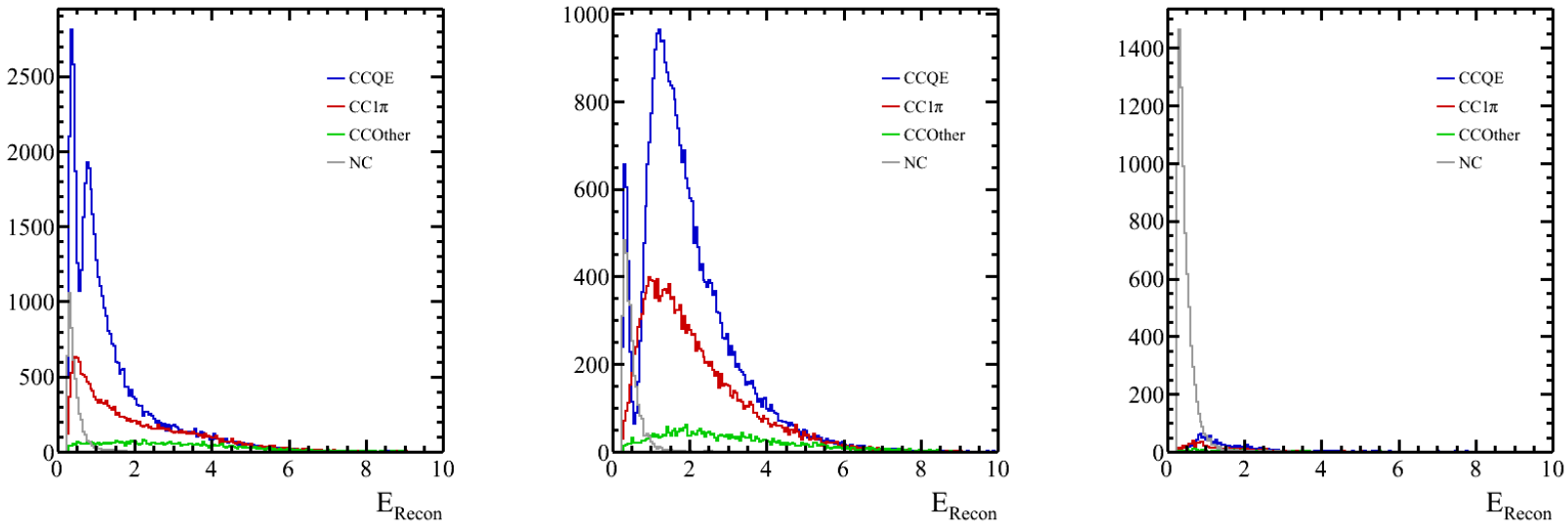


(a) ν_μ signal and backgrounds (left), $\overline{\nu}_\mu$ backgrounds (centre) and ν_e backgrounds (right)

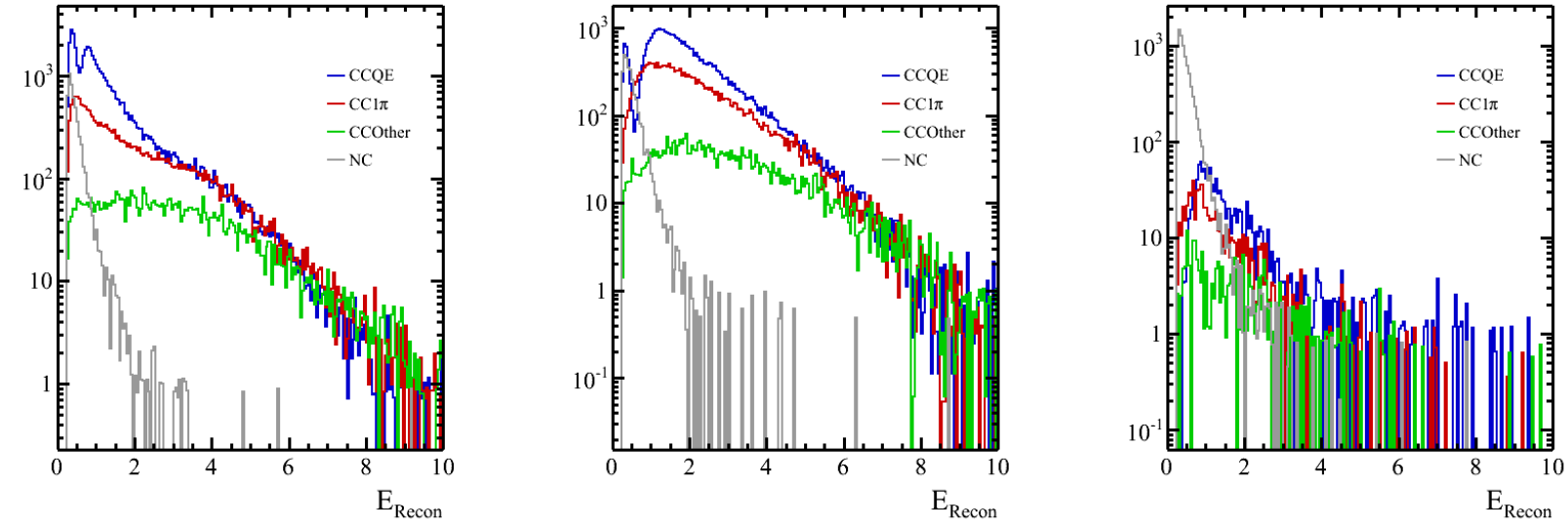


(b) Same as above in logarithmic scale.

Figure 5.13: Reconstructed re-weighted oscillated histograms in null oscillation hypothesis for normalisations of 269 years, 632 years and 1367 years where a year is defined as 750 kW for a 30 GeV beam for 1.07×10^7 s.



(a) ν_μ signal and backgrounds (left), $\overline{\nu}_\mu$ backgrounds (centre) and ν_e backgrounds (right)



(b) Same as above in logarithmic scale.

Figure 5.14: Reconstructed re-weighted oscillated histograms in favoured oscillation hypothesis ($\sin^2 2\theta_{23} = 1.0$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$) for normalisations of 269 years, 632 years and 1367 years where a year is defined as 750 kW for a 30 GeV beam for 1.07×10^7 s.

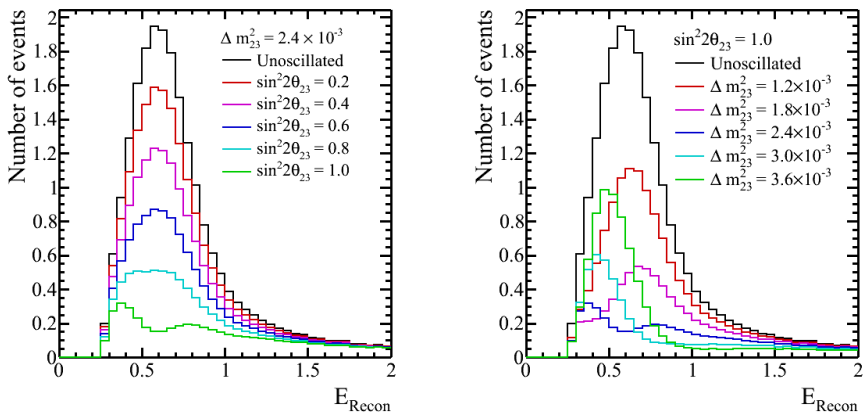


Figure 5.15: Final reconstructed oscillated spectrum as a function of the oscillation parameters $\sin^2 2\theta_{23}$ and Δm^2_{32}. Left: Effect of the variation of $\sin^2 2\theta_{23}$ at a constant value of $\Delta m^2_{32} = 2.4 \times 10^{-3}$. Right: Effect of the variation of Δm^2_{32} at a constant value of $\sin^2 2\theta_{23} = 1$.

	Un-Oscillated	Oscillated
Total	22.8	6.3
	ν_μ	
$CCQE$	18.1 (79.2%)	3.6 (56.7%)
$CC1\pi$	3.1 (13.5%)	1.5 (23.9%)
$CCOther$	0.5 (2.3%)	0.4 (6.4%)
NC	0.4 (1.7%)	0.4 (6.0%)
	$\bar{\nu}_\mu$	
$CCQE$	0.5 (2.2%)	0.3 (4.0%)
$CC1\pi$	0.2 (0.8%)	0.1 (2.1%)
$CCOther$	0.0 (0.1%)	0.0 (0.4%)
NC	0.0 (0.1%)	0.0 (0.3%)
	ν_e	
$CCQE$	0.0 (0.0%)	0.0 (0.0%)
$CC1\pi$	0.0 (0.0%)	0.0 (0.0%)
$CCOther$	0.0 (0.0%)	0.0 (0.0%)
NC	0.0 (0.1%)	0.0 (0.2%)

Table 5.3: Super-K event rate prediction in the null hypothesis and the favoured oscillation hypothesis ($\sin^2 2\theta_{23} = 1$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$) broken down as a function of the parent neutrino flavour and the interaction mode.

5.7 ND280 PDF Construction and event rate prediction

The near detector measurement, as mentioned previously, plays a critical role in the constraint of systematic uncertainties that affect the Super-K energy

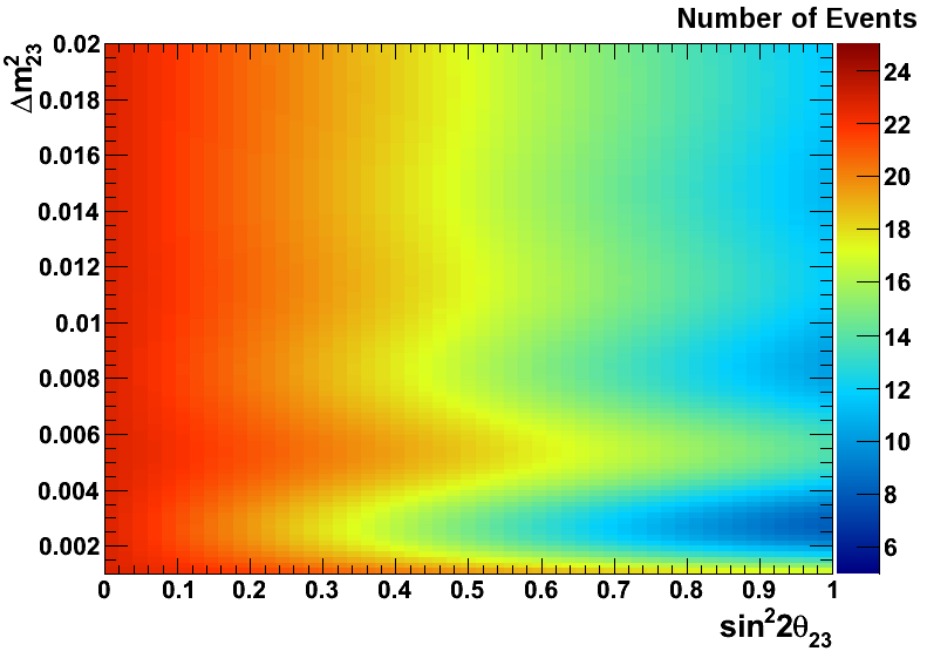


Figure 5.16: Final reconstructed oscillated number of events as a function of the oscillation parameters $\sin^2 2\theta_{23}$ and Δm^2_{32}.

spectra. In particular, normalisation and shape of the neutrino spectra and cross section uncertainties. This section details how the PDF for the ND280 measurement is constructed.

The expected reconstructed neutrino energy spectra at the near detector are simulated step by step based on the following procedure:

1. The neutrino beam flux is predicted in T2K based on GEANT3: neutrino vectors are provided as inputs to the near detector simulation.

2. The NEUT or GENIE neutrino interaction simulation packages are used to simulate the interaction of neutrinos at the near and far detectors.

3. GEANT4 [73] is used for the simulation of the final state particles from neutrino interactions at the near detector and the simulation of cosmic rays and "sand" neutrinos.

4. The ND280 detector response is simulated including the active media, sensor and electronic technologies, calibration, shower/track determination, vertex finding and particle identification.

5. The reconstructed information from the ND280 software is analysed in order to determine signals and backgrounds relevant to the construction of the ν_μ neutrino reconstructed energy spectra.

The expected reconstructed neutrino energy spectra at the near detector are the result of a combined effort from the ND280 group. Figure 5.17 presents the signal and background events for a generated Monte Carlo sample of events normalised to the T2K run 1 data that pass a set of cuts designed to select charge current events in the ND280 tracker [74] For the T2K physics run 1, the operational time of the ND280 corresponds to 2.88×10^{19} POT. The amount of Monte Carlo generated corresponds to 49.8×10^{19} POT which is equivalent to approximately 17.12 times the amount of data collected. The signals for the charged current like (CC-like) sample are all charged current interactions. The backgrounds are all neutral current interactions and "out-of-FGD events".

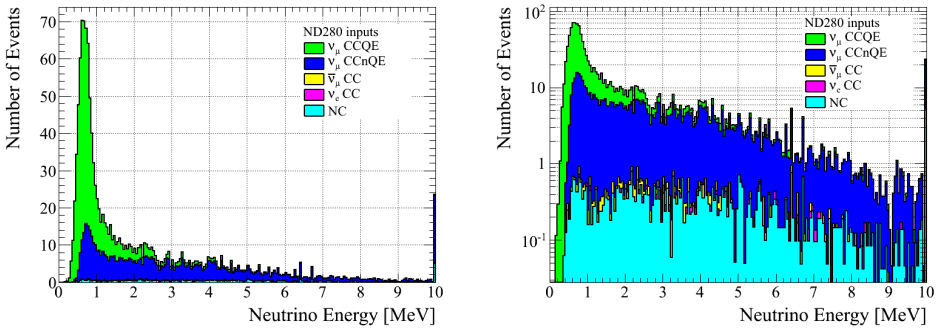


Figure 5.17: Stacked histogram of expected reconstructed neutrino energy spectra from Monte Carlo for events that pass charged current like cuts in the ND280 tracker. The normalisation corresponds to the T2K physics run 1 data.

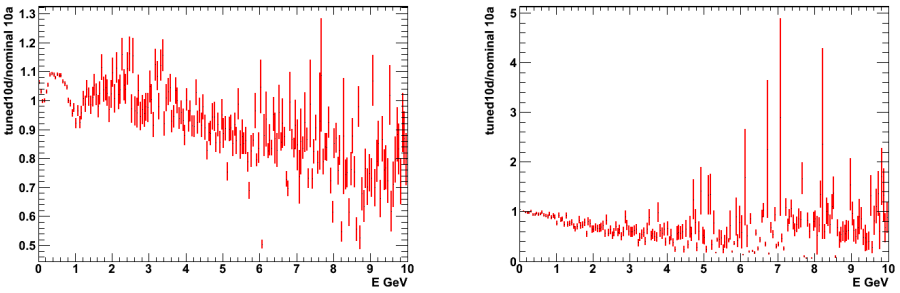
Similarly to the Super-K spectra, the ND280 Monte Carlo has been re-weighted to take into account flux corrections. The re-weighting histograms for the near detector neutrino flux shape and normalisation are presented for each parent neutrino flavour in figure 5.18. The spectra presented in figure 5.17 already include this flux re-weighting.

For the T2K physics run 1, the near detector data are used to re-normalise the far detector neutrino flux using a single correction factor independent of energy. The correction factor is the ratio of observed over expected number of charged current like events in the ND280 tracker.

The expected number of events that pass Monte Carlo CC-like cuts with NEUT is 1372. The corresponding number of selected events in the data 1456. The ND280 data driven correction factor by which the Super-K spectra are re-normalised is therefore:

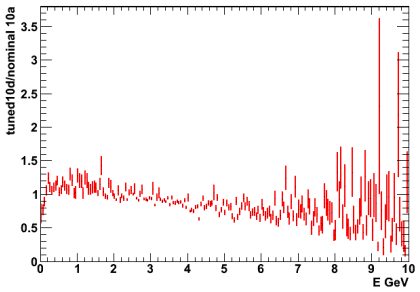
$$C_{\mathrm{ND280}} = \frac{N^{\mathrm{obs}}_{\mathrm{ND280;CC-like}}}{N^{\mathrm{MC}}_{\mathrm{ND280;CC-like}}} = 1.06, \tag{5.50}$$

where C_{ND280} is the ND280 data driven correction factor, $N^{\mathrm{obs}}_{\mathrm{ND280;CC-like}}$ is the



(a) Neutrino parent: ν_μ

(b) Neutrino parent: $\overline{\nu}_\mu$



(c) Neutrino parent: ν_e

Figure 5.18: ND280 Flux correction for the three neutrino parent modes

observed number of selected CC-like events and $N^{\mathrm{MC}}_{\mathrm{ND280;CC-like}}$ is the expected number of selected CC-like events.

The far to near ratio $R^{F/N}(E_t)$ is a function of energy so does not appear in the normalisation only correction provided by the ND280 at this stage.

5.8 Systematic errors

The Super-K and ND280 expected reconstructed neutrino energy spectra do not include any kind of systematic uncertainty at this stage. However, it is of course necessary to take them into account in the analysis in order to build correct confidence regions for the oscillation parameters. What these systematic errors are, how they are included and how they affect the final reconstructed energy spectra at Super-K is presented in detail in this section.

Sources of uncertainties are separated into the following four different categories based on the derived expected energy spectra (equation 5.14):

1. Super-K detector uncertainties
2. ND280 detector uncertainties
3. Neutrino flux uncertainties

4. Neutrino generator model and cross section uncertainties

The uncertainties described below are applied to the T2K physics run 1 analysis and are generally conservative. This is deemed acceptable since the first T2K run was highly dominated by statistical errors. As more data are collected, continuous efforts are made by the various groups working on systematic errors to improve our knowledge of these uncertainties and the way that they can be best incorporated into the oscillation analysis.

5.8.1 Super-K detector uncertainties

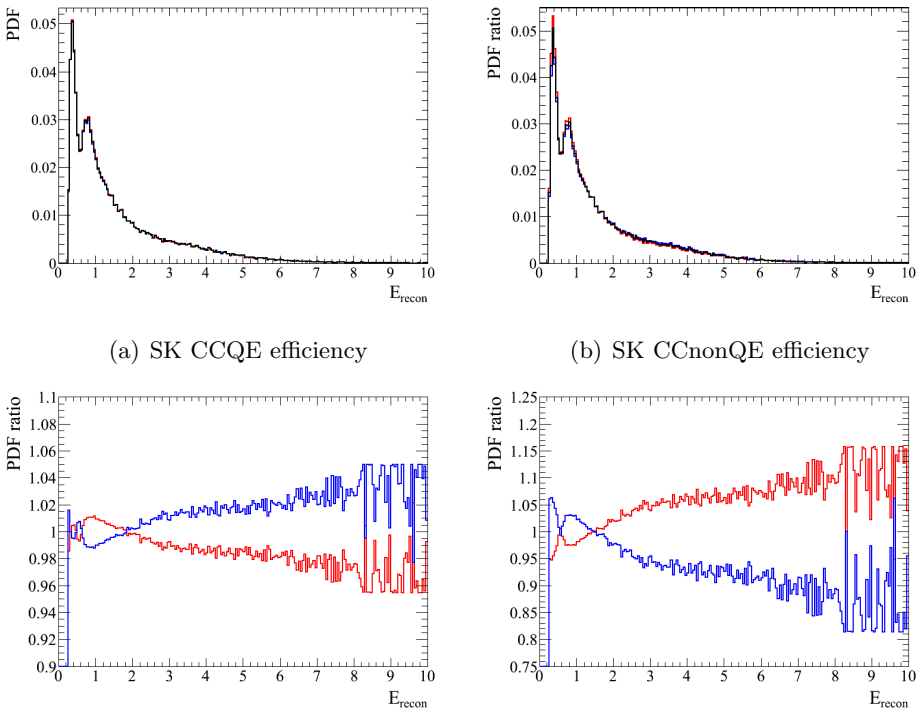
The Super-K detector uncertainties are implemented via the error on the efficiencies of the selection cuts as a function of interaction mode and the uncertainty on the reconstructed energy scale.

For each interaction mode, a 1-σ energy independent error on the efficiency is provided by the Super-K group and applied uniformly to the whole energy range. The energy scale error on the other hand is a function of reconstructed energy and affects only the shape of the distribution. As an illustration, if an uncertainty on the reconstructed energy scale of +1% is applied, then the reconstructed energy of each event is scaled by a factor of 1.01. The normalisation is constant since no events are added or removed from the distribution. Table 5.4 below summarises the 1-σ error of each Super-K detector uncertainty.

Systematic uncertainty	1-σ error
$f^{\mathrm{SK}}_{\epsilon_{\mathrm{CCQE}}}$	7.8%
$f^{\mathrm{SK}}_{\epsilon_{\mathrm{CCnnQE}}}$	25.5%
$f^{\mathrm{SK}}_{\epsilon_{\mathrm{NC}}}$	115.1%
$f^{\mathrm{SK}}_{\epsilon_{\nu_e\mathrm{CC}}}$	100%
$f^{\mathrm{SK}}_{E_{\mathrm{Scale}}}$	1%

Table 5.4: 1-σ error for each Super-K detector uncertainty for the T2K physics run 1 analysis.

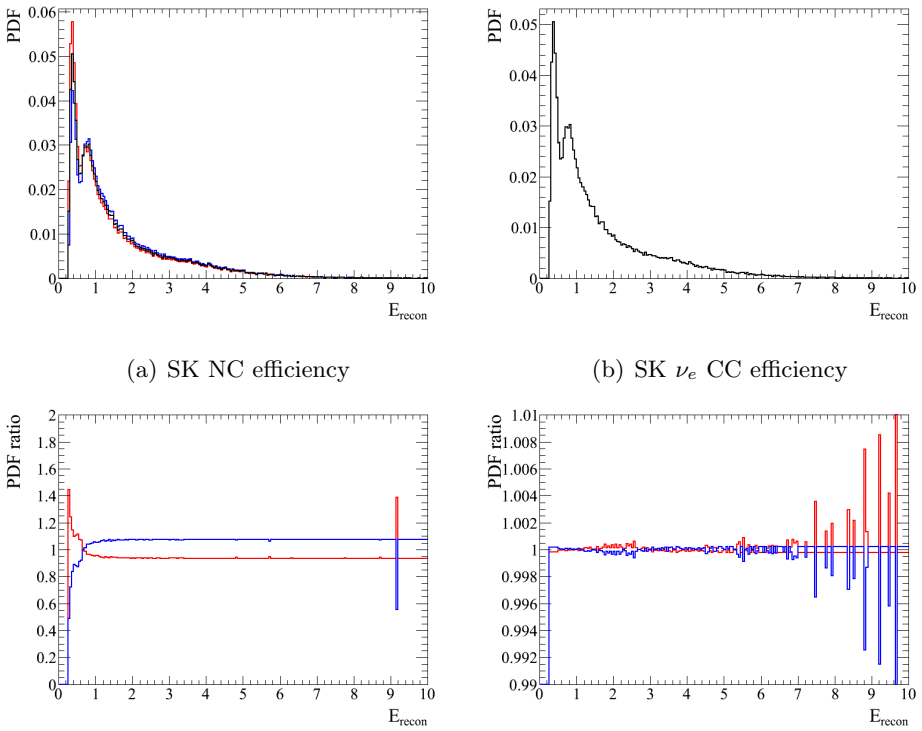
Figures 5.19, 5.20 and 5.21 illustrate how the total reconstructed neutrino energy spectrum in the favoured oscillation hypothesis is affected by a 1-σ standard deviation shift where each Super-K efficiency uncertainty is varied independently. Since the fraction that each neutrino mode contributes to total energy spectrum is a function of energy, a 1-σ shift of any Super-K efficiency uncertainty actually affects the total energy spectrum in an energy dependent manner. The favoured oscillation hypothesis is defined by $\sin^2 2\theta_{23} = 1.0$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$. The dominant error in the favoured oscillation hypothesis is the neutral current efficiency with 1-σ deviations for bins below 1 GeV of over 10 %.



(a) SK CCQE efficiency

(b) SK CCnonQE efficiency

Figure 5.19: Effect of ± 1-σ deviation of CCQE and CCnonQE SK systematic efficiencies on the expected reconstructed energy spectrum at SK in the favoured oscillation hypothesis. Blue: +1 σ, Red: -1 σ



(a) SK NC efficiency

(b) SK ν_e CC efficiency

Figure 5.20: Effect of $\pm$ 1-σ deviation of NC and ν_e CC SK systematic efficiencies on the expected reconstructed energy spectrum at SK in the favoured oscillation hypothesis. Blue: +1 σ, Red: -1 σ

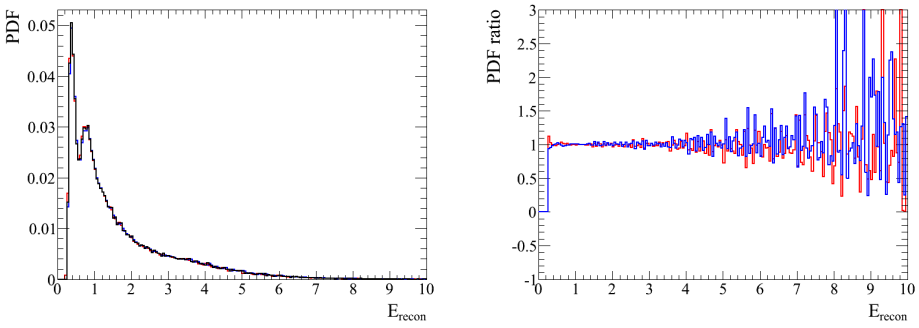


Figure 5.21: Effect of ± 1-σ deviation of the energy scale uncertainty on the expected reconstructed energy spectrum at SK in the favoured oscillation hypothesis. Blue: +1 σ, Red: -1 σ

5.8.2 ND280 uncertainties

The near detector measurement in the T2K run 1 analysis uses the ratio of the number of events between data and Monte Carlo in a charged current inclusive measurement in the tracker (TPC+FGD) [74] in order to provide a near detector uncertainty. The analysis has no spectral dependence so the resulting uncertainty provides only a normalisation error for the number of events at Super-K. Uncertainties include detector systematics and efficiency differences between data and Monte Carlo such as charge mis-identification and time determination errors. Other systematics such as selection cut uncertainties, reconstruction efficiencies and backgrounds from cosmics or neutral current events form a non-exhaustive list of all the included uncertainties. The combined error on the near detector normalisation measurement, summarised in table 5.5, was estimated to be 5%.

Systematic uncertainty	1-σ error
f^{ND280}	5%

Table 5.5: 1-σ combined error for the ND280 normalisation uncertainty for the T2K physics run 1 analysis.

Since no spectral information was provided by the near detector measurement for the run 1 conditions, f^{ND280} only affects the normalisation of the Super-K spectrum and the probability density function is therefore unchanged.

5.8.3 Flux uncertainties

Flux uncertainties refer to the uncertainties in the normalisation and shape of the neutrino energy spectra at Super-K. The beam production facilities and monitors at the J-PARC accelerator complex were mentioned in section

2.1. Knowing precisely the composition of the beam in terms of the neutrino flavours content and energy spectra is certainly a challenge. For the T2K physics run 1, hadron interactions from the 30 GeV protons hitting the target were simulated with FLUKA2008 [36] and tuned using results from the NA61/SHINE experiment based at CERN [75, 76]. The results from FLUKA are then used as inputs to the T2K neutrino beam simulation software, based on GEANT3 [37]. The simulation tracks the particles that leave the target and their interactions with the three horn magnets, decay volume, beam dump and muon monitors until they decay to neutrinos. The neutrino tracks are then simply extrapolated to the near and far detectors.

This section summarises the systematic errors on the flux uncertainties calculated by the T2K beam group.

Flux normalisation uncertainty ($f^{\mathrm{FLUX}}_{\mathrm{SK/ND}}$):

The flux normalisation uncertainty represents the uncertainty in the normalisation of the far and near neutrino energy spectra. Contributing sources are listed below.

- hadron production cross section (p, $\pi^{\pm}$,$K^{\pm}$)
- off-axis angle
- proton beam position/angle
- absolute horn current
- horn alignment
- target alignment

Since the uncertainty on the flux normalisation is a function of the oscillation parameters, it must be re-calculated by the fitter each time a new set of parameters is tried. A function was provided by the beam group to calculate this uncertainty in a transparent way. Using the near and far detector energy spectra as inputs, this function returns a normalisation factor that is used to reweight the far detector energy spectra. In the favoured oscillation hypothesis, the value of this uncertainty is approximately 6.2 % and independent of energy.

Flux shape uncertainty ($f^{\mathrm{FLUX}}_{\mathrm{Shape}}$):

The flux shape uncertainty describes the uncertainty on the shape of the ν_μ energy spectrum at Super-K. Contributing sources listed below are the same as for the flux normalisation ratio uncertainty with the addition of the hadron production multiplicities.

- π production multiplicity
- K production multiplicity
- Production cross section
- Off-axis angle
- Proton beam position/angle
- Horn absolute current
- Horn alignment
- Target alignment

The total uncertainty on the shape of the ν_μ energy spectrum at Super-K is provided in the format of a fractional covariance matrix. The element at the i^{th} and j^{th} position in the matrix give the covariance between the i^{th} and j^{th} bins of the ν_μ energy spectrum. This matrix is itself the sum of covariance matrices calculated independently for each contributing source. Figure 5.22 shows the binned covariance matrix for the flux shape uncertainty provided by the beam group.

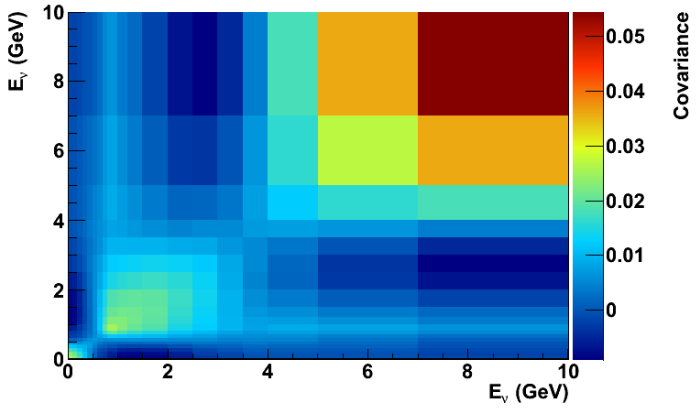


Figure 5.22: Beam shape covariance matrix

For an element at position i, j, the covariance matrix gives the covariance between the i^{th} and j^{th} elements of a vector of random variables:

$$\mathrm{cov}(\mathrm{x_i}, \mathrm{x_j}) = \mathrm{E}[(\mathrm{x_i}, \mu_\mathrm{i})(\mathrm{x_j}, \mu_\mathrm{j})] \tag{5.51}$$

where $\mu_i = E(x_i)$ and $E(x_i)$ is the expected value of x_i

Given a covariance matrix of parameters (in this case bins of the energy spectrum at Super-K), a vector containing the fluctuation of each parameter

can be obtained by sampling the matrix using linear algebra techniques as described below.

step 1:

The Cholesky decomposition method [77] is used to obtain a triangular decomposition of the covariance matrix in order to solve sets of linear equations more efficiently. Triangular matrices are square matrices were all the entries either below or above the diagonal are zero. The Cholesky method is a special case of LU decomposition that can be used when the square matrix is symmetric and positive definite.

LU decomposition [78] is a well known linear algebra factorisation method that expresses a square matrix $\mathbf{A}$ as the product of a lower and upper triangulation matrix $\mathbf{L}$ and $\mathbf{U}$:

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{U}.$$

$\mathbf{L}$ and $\mathbf{U}$ are then used instead of $\mathbf{A}$ to solve the set of linear equations $\mathbf{A} \cdot \mathbf{x}$:

$$\mathbf{A} \cdot \mathbf{x} = (\mathbf{L} \cdot \mathbf{U}) \cdot \mathbf{x} = \mathbf{L} \cdot (\mathbf{U} \cdot \mathbf{x}) = \mathbf{b}.$$

A matrix $\mathbf{A}$ is symmetric if, for all elements, $a_{i,j} = a_{j,i}$ and is positive definite if $\mathbf{v} \cdot \mathbf{A} \cdot \mathbf{v} > 0$ for all vectors $\mathbf{v}$.

The Cholesky method is very similar to the LU method. A symmetric and positive definite square matrix is decomposed into triangular matrices where the upper triangular matrix is the transpose of the lower triangular matrix:

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{L}^T.$$

A factor of 2 is gained in terms of efficiency in comparison to the LU decomposition method. Figure 5.23 shows the lower triangular matrix of the covariance matrix presented in figure 5.22.

step 2:

A vector $\mathbf{v}$ is drawn from a multivariate Gaussian distribution with mean 0 and variance 1 in all dimensions.

step 3:

A vector containing a set of fluctuations for the parameters consistent with the covariance matrix is obtained by taking the product of the lower triangular

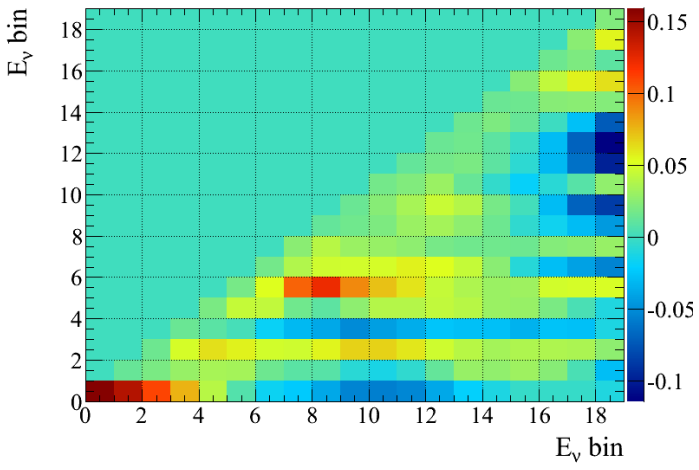


Figure 5.23: lower triangular matrix of the covariance matrix (5.22) using Cholesky decomposition.

matrix $\mathbf{L}$ with the random vector $\mathbf{v}$:

$$\mathbf{x} = \mathbf{L} \cdot \mathbf{v}$$

Vectors containing different sets of fluctuations also consistent with the covariance matrix can be obtained by repeating step 2 and 3.

Figure 5.24 illustrates how the total reconstructed neutrino energy spectrum in the favoured oscillation hypothesis is affected when the flux shape covariance matrix is sampled 1000 times according to the procedure described above. The relative error is also plotted by calculating the standard deviation of the 1000 samples.

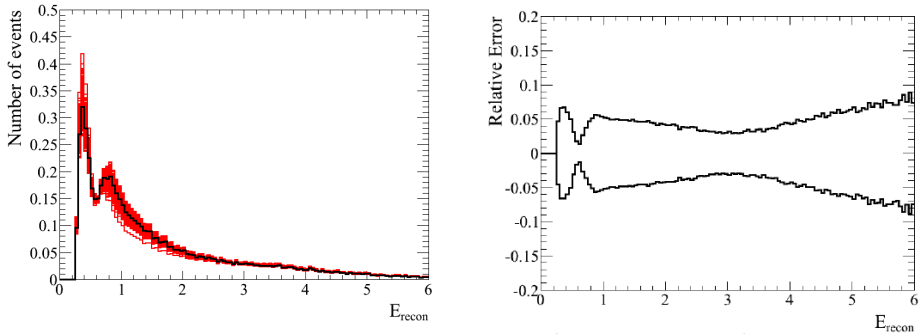


Figure 5.24: Left: Effect of 1000 samplings of the flux shape covariance matrix on the expected reconstructed energy spectrum at SK in the favoured oscillation hypothesis. Right: Relative error calculated by taking the standard deviation of the samples in each energy bin

5.8.4 Neutrino generator model uncertainties

The neutrino interaction generator simulates neutrino interactions with matter at the near and far detector sites.

The neutrino generators predict the rates and kinematics of the interaction modes that contribute to the neutrino events observed at the near and far detectors by using a combination of fundamental physics models. These models, which sometimes have different ranges of validity, are combined or extrapolated in order to generate events over the whole experimental phase space. To make matters worse, current simulations are based on limited data samples with limited coverage in the range of neutrino energy, and often, have been tuned to bubble chamber data taken over 30 years ago [79]. It therefore goes without saying that there are substantial uncertainties associated with neutrino interaction generators.

In order to acknowledge and reduce these uncertainties, two of the most recognised competing neutrino generators are used in T2K: GENIE and NEUT. The analysis presented uses the NEUT generator and therefore the uncertainties quoted in this section relate to NEUT only.

The T2K Neutrino Interaction Working Group (NIWG) calculated systematic errors on the cross sections of the relevant interaction modes in NEUT by comparison with SciBooNE, MiniBooNE, and K2K experiments, comparisons with GENIE and NuWro generators and recent theoretical work [74]. The parameters studies with their central value and 1-σ deviation are reproduced from the NIWG in table 5.8.4.

Parameter	Value	1-σ $(GeV/c)^2$
M_A^{QE}	1.21 $(GeV/c)^2$	0.20 $(GeV/c)^2$
M_V^{QE}	0.84 $(GeV/c)^2$	0.20 $(GeV/c)^2$
p_F	225 MeV/c	10 MeV/c
E_B	27 MeV	3 MeV
M_A^{RES}	1.21 $(GeV/c)^2$	0.20 $(GeV/c)^2$
M_A^{COH}	1.0 $(GeV/c)^2$	0.5 $(GeV/c)^2$
κ	1.00	0.007

Table 5.6: Central value and 1-σ systematic uncertainty on cross section parameters.

The calculated cross section systematic errors based on the variation of the parameters in table 5.8.4 are presented below.

CCQE cross section uncertainty ($f_{\sigma_{\mathrm{CCQE}}}^{\mathrm{Xsec}}$): The normalisation error of the CCQE cross section is cancelled out in the far-to-near ratio. A residual efficiency that includes nuclear model uncertainties and differences in target nuclei between the near and far detector is none the less estimated. This

systematic uncertainty on the CCQE cross section presented in figure 5.25(a) is only applied to the Super-K sample and affects the true neutrino energy.

CC1π / CCQE cross section ratio ($f^{\mathrm{Xsec}}_{\sigma_{\mathrm{CC1}\pi}/\sigma_{\mathrm{CCQE}}}$): The systematic uncertainty on the $CC1\pi/CCQE$ cross section ratio was estimated by comparing data from MiniBooNE and the NEUT MC generator. Figure 5.25(b) presents the 1-σ uncertainty on the cross section ratio as a function of energy.

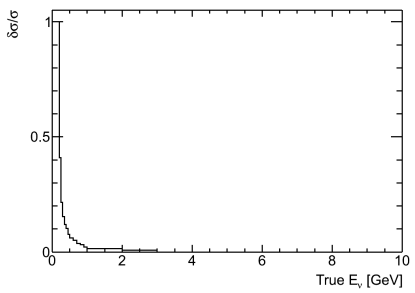
CCOTH / CCQE cross section ratio ($f^{\mathrm{Xsec}}_{\sigma_{\mathrm{CCOTH}}/\sigma_{\mathrm{CCQE}}}$): The 1-$\sigma$ systematic uncertainty on the $CCOTH/CCQE$ cross section ratio as a function of true neutrino energy is presented figure 5.25(c).

NC / CCQE cross section ratio ($f^{\mathrm{Xsec}}_{\sigma_{\mathrm{NC}}/\sigma_{\mathrm{CCQE}}}$): The systematic uncertainty on the $NC/CCQE$ cross section ratio was estimated by comparing SciBooNE data with NEUT MC. The 1-σ uncertainty on this cross section ratio as a function of true neutrino energy is presented figure 5.25(d).

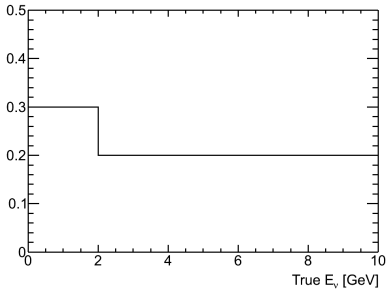
ν_e / ν_μ cross section ratio ($f^{\mathrm{Xsec}}_{\sigma_{\nu_e}/\sigma_{\nu_\mu}}$): The 1-$\sigma$ systematic uncertainty on the ν_e/ν_μ cross section ratio as a function of true neutrino energy is presented figure 5.25(c).

Final State Interaction (FSI) uncertainty (f^{FSI}): FSI are hadronic interactions of reaction products in the nucleus. The 1-σ systematic uncertainty on the FSI cross section ratio as a function of reconstructed neutrino energy is presented in figure 5.25(c).

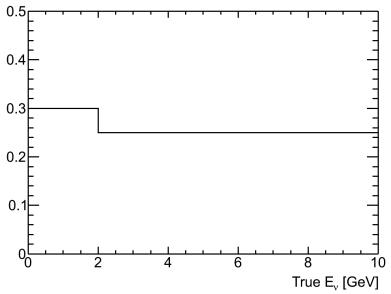
The effect of a 1-σ deviation from the expected value for each cross section error on the reconstructed neutrino energy spectrum at Super-K is presented in figure 5.26.



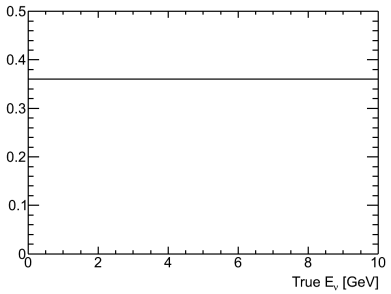
(a) Systematic uncertainty on the CCQE cross section.



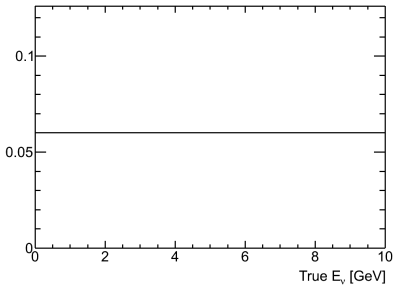
(b) Systematic uncertainty on the CC1π / CCQE cross section ratio.



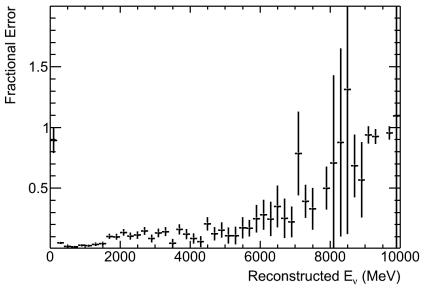
(c) Systematic uncertainty on the CCOTH / / CCQE cross section ratio.



(d) Systematic uncertainty on the NC / CCQE cross section ratio.

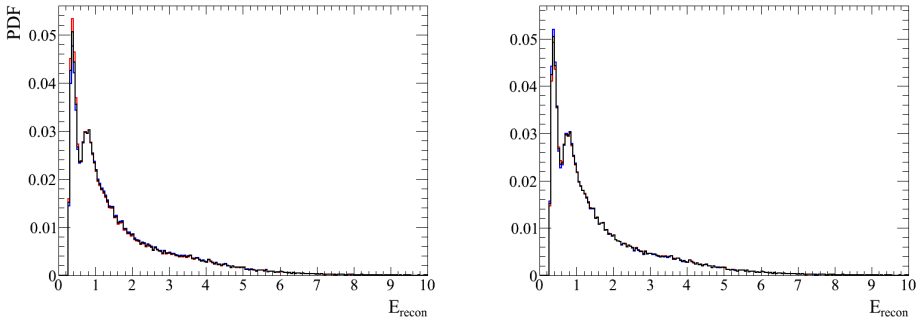


(e) Systematic uncertainty on the ν_e/ν_μ cross section ratio.



(f) Systematic uncertainty on FSI.

Figure 5.25: The 1-σ CCQE cross section, cross section ratios and FSI systematic uncertainties as a function of energy (histograms provided by the T2K beam group).



(a) CCQE cross section uncertainty.

(b) CC1π / CCQE cross section ratio uncertainty.

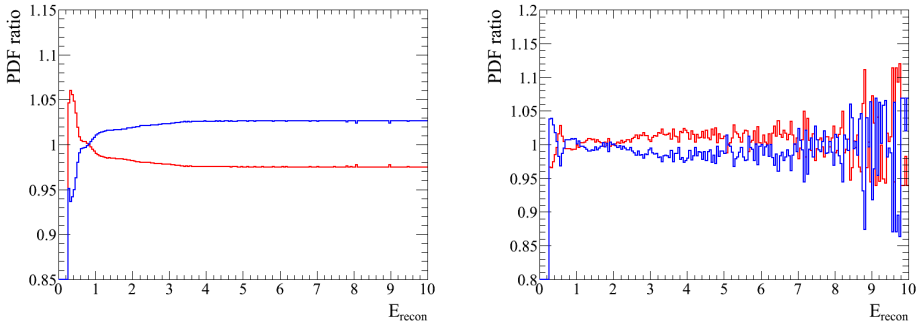
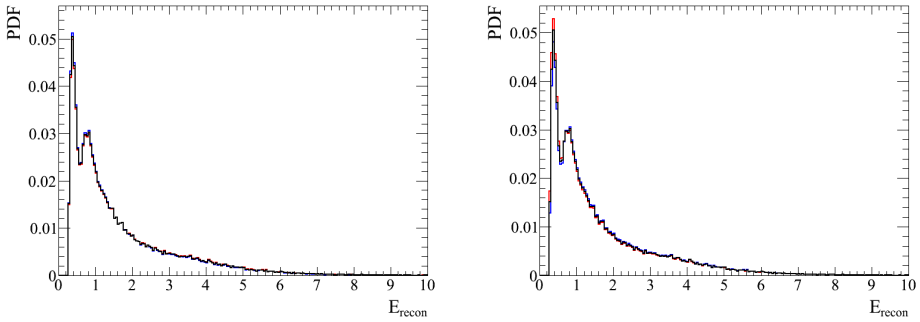


Figure 5.26: Effect of $\pm$ 1-σ deviation of CCQE and CC1π / CCQE cross section ratio systematic efficiencies on the expected reconstructed energy spectrum at SK in the favoured oscillation hypothesis. Blue: +1 σ, Red: -1 σ



(a) CCOTH / CCQE cross section ratio uncertainty.

(b) NC / CCQE cross section ratio uncertainty.

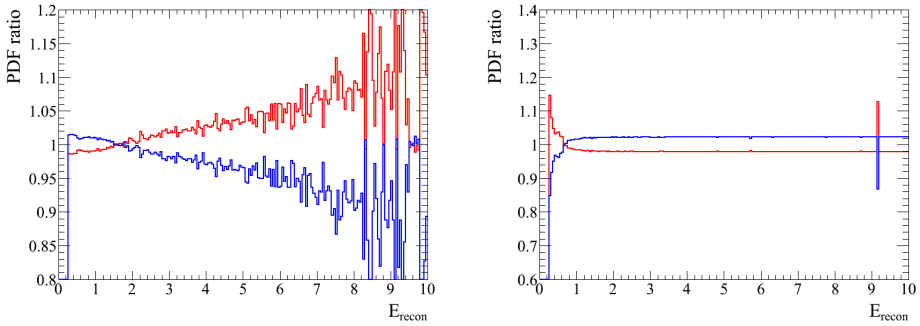


Figure 5.27: Effect of $\pm$ 1-σ deviation of CCOTH / CCQE and NC / CCQE cross section ratio systematic efficiencies on the expected reconstructed energy spectrum at SK in the favoured oscillation hypothesis. Blue: +1 σ, Red: -1 σ

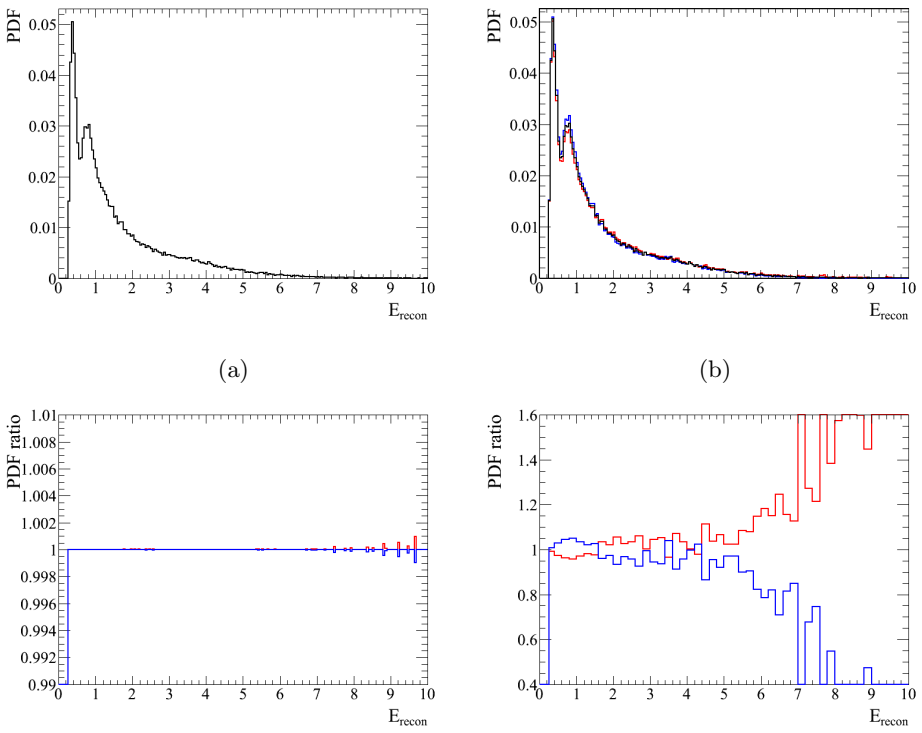


Figure 5.28: Effect of ± 1-σ deviation of ν_e / ν_μ cross section ratio systematic efficiency and FSI uncertainty on the expected reconstructed energy spectrum at SK in the favoured oscillation hypothesis. Blue: +1 σ, Red: -1 σ

5.8.5 Effect of systematic uncertainty on expected number of events at Super-K

The percentage change of the number of expected events in the Super-K reconstructed energy distribution for a 1-σ independent variation of each systematic uncertainty is summarised in table 5.8.5. The total error is obtained by adding each contributing source in quadrature.

Systematic Error	% change of N^{SK}_{exp}
f^{SK}_{CCQE}	+4.74% -4.74%
f^{SK}_{CCnQE}	+8.34% -8.34%
f^{SK}_{NC}	+7.46% -7.46%
$f^{SK}_{CC\nu_e}$	+0.02% -0.02%
$f^{SK}_{E_scale}$	+0.00% 0.00%
f^{ND}	-4.75% 5.25%
$f^{X_{sec}}_{CCQE}$	+2.54% -2.54%
$f^{X_{sec}}_{CC1\pi}$	+0.47% -0.53%
$f^{X_{sec}}_{CCOTH}$	-3.74% +4.20%
$f^{X_{sec}}_{CCNC}$	+0.87% -0.90%
$f^{X_{sec}}_{\nu_e/\nu_\mu}$	+0.00% -0.00%
f^{FSI}	+6.45% -6.32%
$f^{Flux}_{SK/ND}$	+6.28% -6.28%
Total	16.11% 16.30%

Table 5.7: percentage change of the number of expected events in the Super-K reconstructed energy distribution for a 1-σ independent variation of each systematic uncertainty

5.9 Monte Carlo simulation

Monte Carlo simulation was used to statistically investigate the range of possible outcomes, study the effect of variations in the oscillation parameters and systematic errors and determine the sensitivity of the experiment for the T2K run 1.

5.9.1 MCMC simulation

The MCMC Metropolis-Hastings method was used to sample the likelihood function for the T2K run 1 normalisation given the favoured oscillation parameters $\sin^2 2\theta_{23} = 1.0$ and $\Delta m^2_{32} = 2.4 \times 10^3$. Simulations are done with and without systematic uncertainties and compared by computing 68%, 90% and 99% Bayesian credible intervals in both cases. When systematic uncertainties are included, the credible intervals over the oscillation parameters are obtained via the marginalisation of the uncertainties. This is done trivially:

the output of the MCMC simulation is an NTuple containing the sampled values of all the likelihood function parameters $p_1, p_2, \cdots, p_n$; plotting NTUple entry p_1 against entry p_2 results in the effective marginalisation of all other parameters.

5.9.2 Processing MCMC jobs

The MCMC jobs were parallelised and run on the Imperial College High Energy Physics (HEP) computer cluster.

The generation of multiple Markov chains in parallel and subsequent combination of all the generated chains offers a clear computational advantage compared to using a single longer chain. On the downside, the multiple chains need to be monitored more closely in case long "burn-in" periods are required or if the chains are highly correlated.

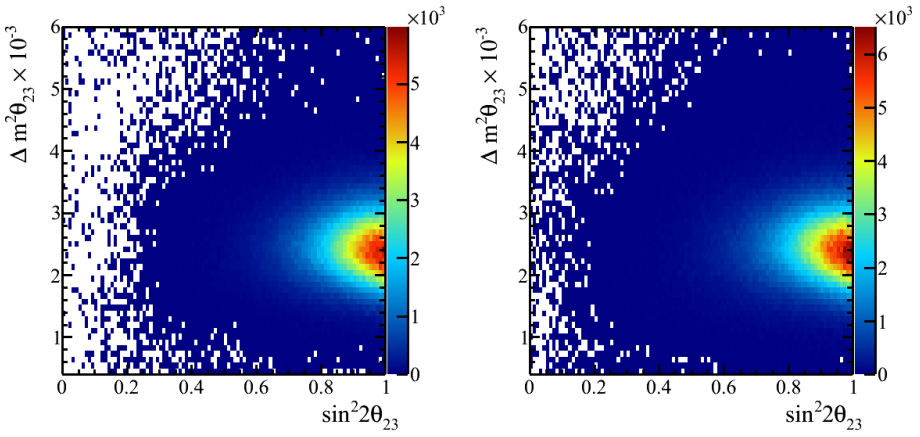
For the MCMC with no systematic uncertainties, 15 chains of 60,000 entries were generated resulting in a total of 900,000 entries. On average, 23 hours and 20 minutes were needed to generate one chain, corresponding to 1.4 seconds per entry computed.

For the case with systematic uncertainties, 40 chains of 20,000 entries were generated resulting in a total of 800,000 entries. On average, 23 hours and 10 minutes were needed for the generation of one chain, corresponding to 4.16 seconds per entry computed.

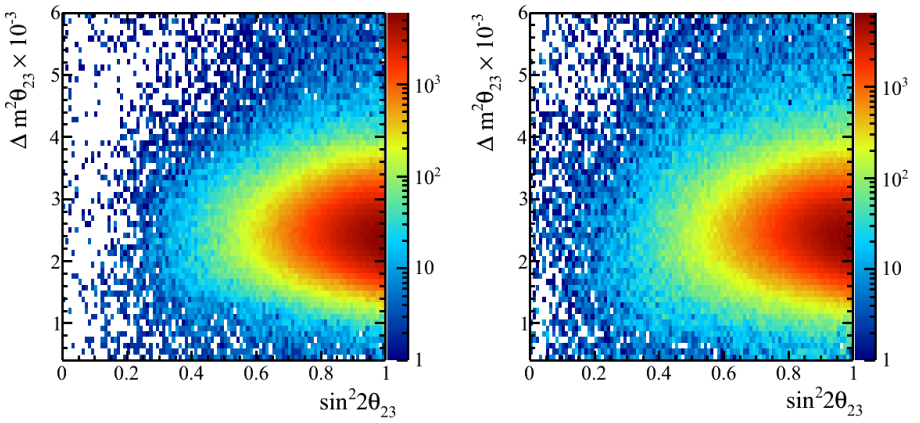
5.9.3 MCMC simulation results

The parameter space of the $\sin^2 2\theta_{23}$ oscillation parameter was limited to the physical region: $0 < \sin^2 2\theta_{23} < 1$ and the Δm^2_{32} parameter was bounded on both sides by: $0 < \Delta m^2_{32} < 6 \times 10^3$ based on current estimation of its value.

Figure 5.29 plots the Markov chain parameters of the sampled disappearance likelihood formula Δm^2_{32} versus $\sin^2 2\theta_{23}$ with and without systematic uncertainties. In the former case, the systematic uncertainties are marginalised. The maximum of the sampled likelihood was found, according to expectations, for oscillation parameter values of $\sin^2 2\theta_{23} = 1.0$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}\ eV^2/c^4$.



(a) left: no systematics, right: with systematics.



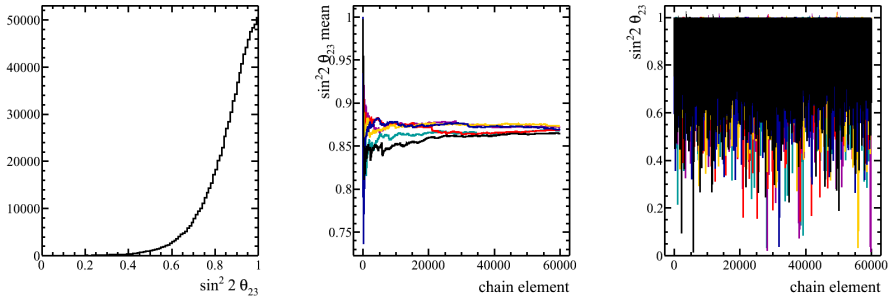
(b) same with log-scale on the Z axis.

Figure 5.29: Sampled disappearance likelihood in the parameter space of the oscillation parameters Δm^2_{32} and $\sin^2 2\theta_{23}$ for input parameters $\Delta m^2_{32} = 2.4 \times 10^3$ and $\sin^2 2\theta_{23} = 1.0$ obtained using the MCMC Metropolis-Hastings method.

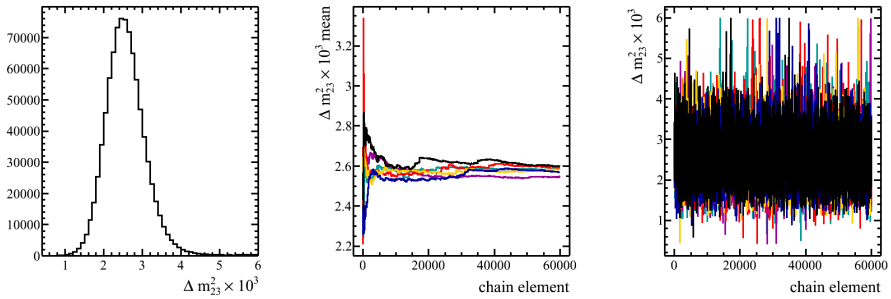
Figures 5.30 and 5.31 show the $\sin^2 2\theta_{23}$ and Δm^2_{32} profiles as well as "time series" of the convergence of the means and sampled parameter values of some Markov chains. Out of the 40 chains for the MCMC with systematics and the 15 chains for the MCMC without, a set of 6 Markov chains were selected randomly in each case for illustration purposes.

The $\sin^2 2\theta_{23}$ parameter is observed to converge in the range between 0.86 and 0.88 for the case with no systematics and to 0.84 to 0.86 in the case where they are included for all Markov chains. The Δm^2_{32} parameter is observed to converge to 2.4×10^3 in both cases.

The "time series" of the sampled parameters show good mixing properties. This indicates that correlations are not too large and that the results obtained are of good quality.

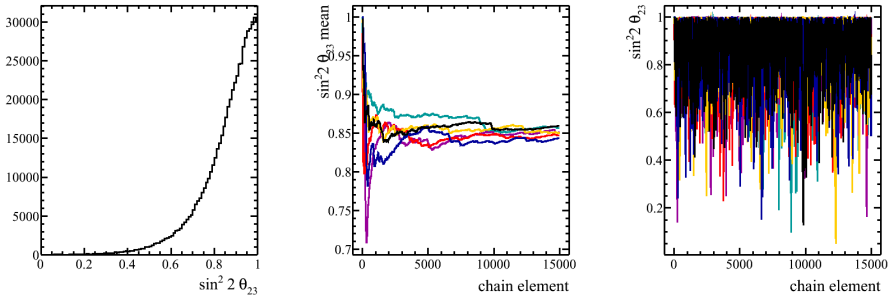


(a) $\sin^2 2\theta_{23}$ profile, aggregate mean and sampled parameter; no systematics.

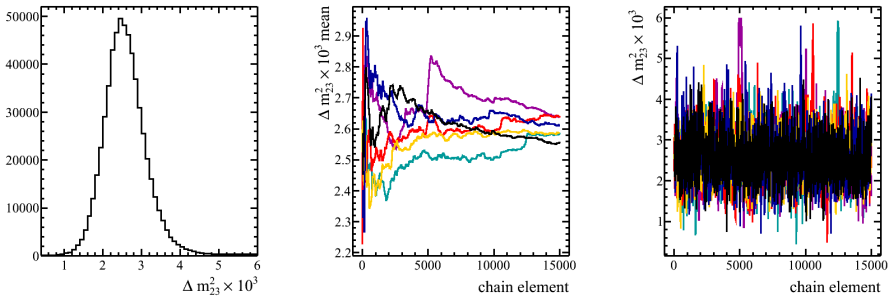


(b) Δm^2_{32} profile, aggregate mean and sampled parameter; no systematics.

Figure 5.30: Left: profile of the $\sin^2 2\theta_{23}$ and Δm^2_{32} Markov chain parameters without systematic uncertainties. Middle: "Time series" of the convergence of the means of $\sin^2 2\theta_{23}$ and Δm^2_{32} for 6 Markov chains chosen randomly amongst the 40. Right: "Time series" of the sampled parameters for the 6 same chains; both exhibit good mixing properties.



(a) $\sin^2 2\theta_{23}$ profile, aggregate mean and sampled parameter; with systematics.



(b) Δm^2_{32} profile, aggregate mean and sampled parameter; with systematics.

Figure 5.31: From Monte Carlo simulation. Same as figure 5.30 above but with systematic uncertainties included.

The posterior profiles of each systematic uncertainty for the relevant MCMC simulations are plotted in figures 5.32 and 5.33 and fitted with Gaussian functions.

In the disappearance likelihood function, the L_{syst} term controls the values of the systematic uncertainties using a multivariate Gaussian prior function. The standard deviations of the multivariate Gaussian in each dimension are given by the 1-σ variation of the underlying systematic uncertainty that were summarised in section 5.8. In the Markov chain data, a value of ± 1, for instance, corresponds to a ± 1-σ change in the systematic uncertainty in the likelihood.

We observe that the posterior systematic profiles of figures 5.32 and 5.33 generally match the prior functions with some deviations in the means and standard deviations observed. In particular, mean values of 0 are expected for the simulated data. However, a mean value of -0.30 sigma units was found for the Super-K CCQE efficiency systematic error, $f^{\text{SK}}_{\text{CCQE}}$, for example. It is possible that the presence of the unphysical boundary leads to some "pull" effects: in the physical region, a lower value of $f^{\text{SK}}_{\text{CCQE}}$ can compensate for a smaller $\sin^2 2\theta_{23}$ value where as a greater value of $f^{\text{SK}}_{\text{CCQE}}$ would be needed to compensate for a value of $\sin^2 2\theta_{23}$ in the unphysical region.

The close agreement regarding the overall Gaussian shape and sigma is only an indication that the T2K run 1 is heavily dominated by statistical errors and that the fit has little constraining power on the systematic uncertainties at this level.

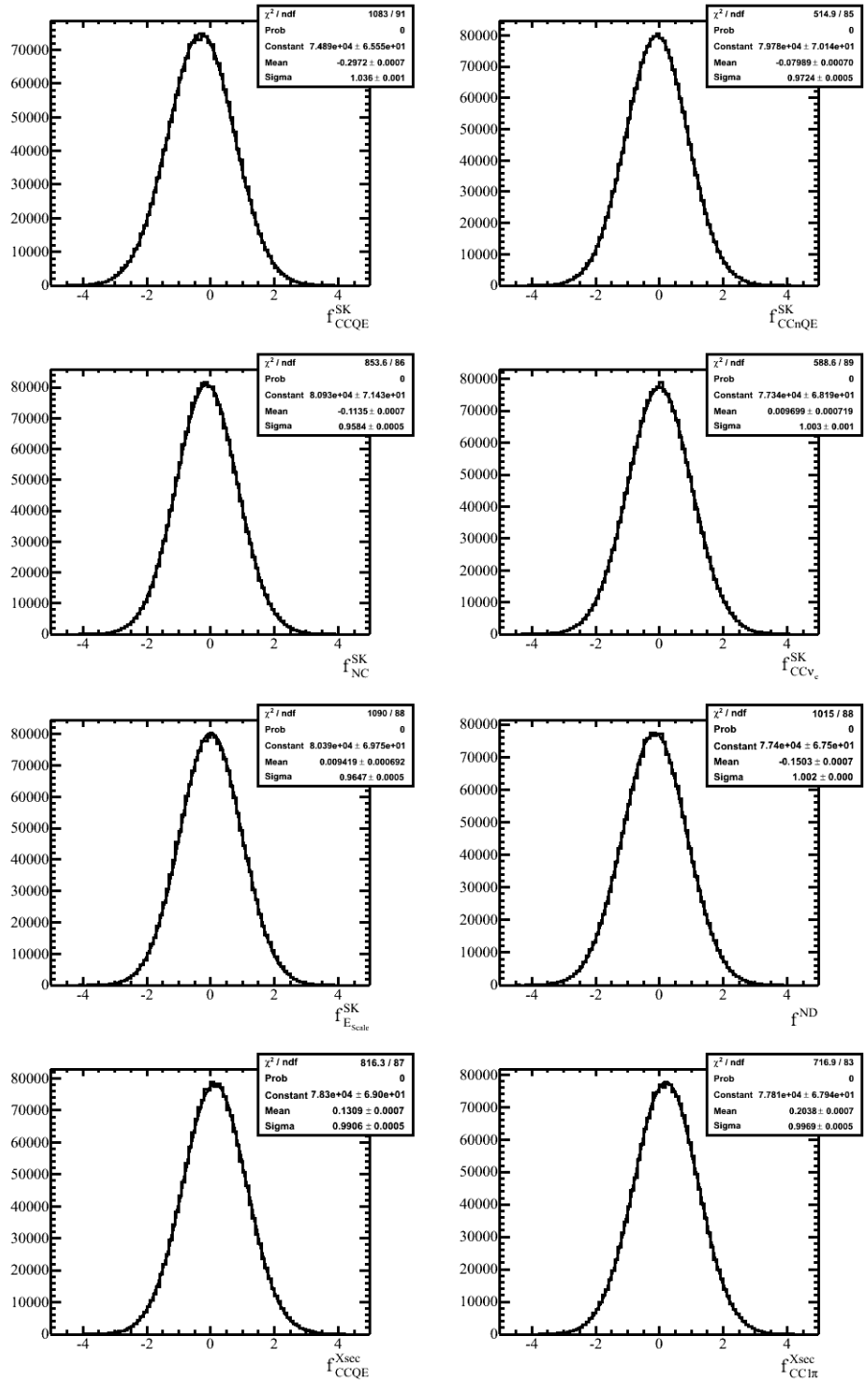


Figure 5.32: Profile of the systematic uncertainty parameters of the simulated Monte Carlo Markov chains fitted with Gaussian functions. Here, from top left to bottom right: Super-K CCQE, CCnQE, NC and CCν_e efficiency uncertainties, Super-K energy scale uncertainty, ND280 normalisation uncertainty, CCQE cross section uncertainty, CC1π cross section ratio uncertainty.

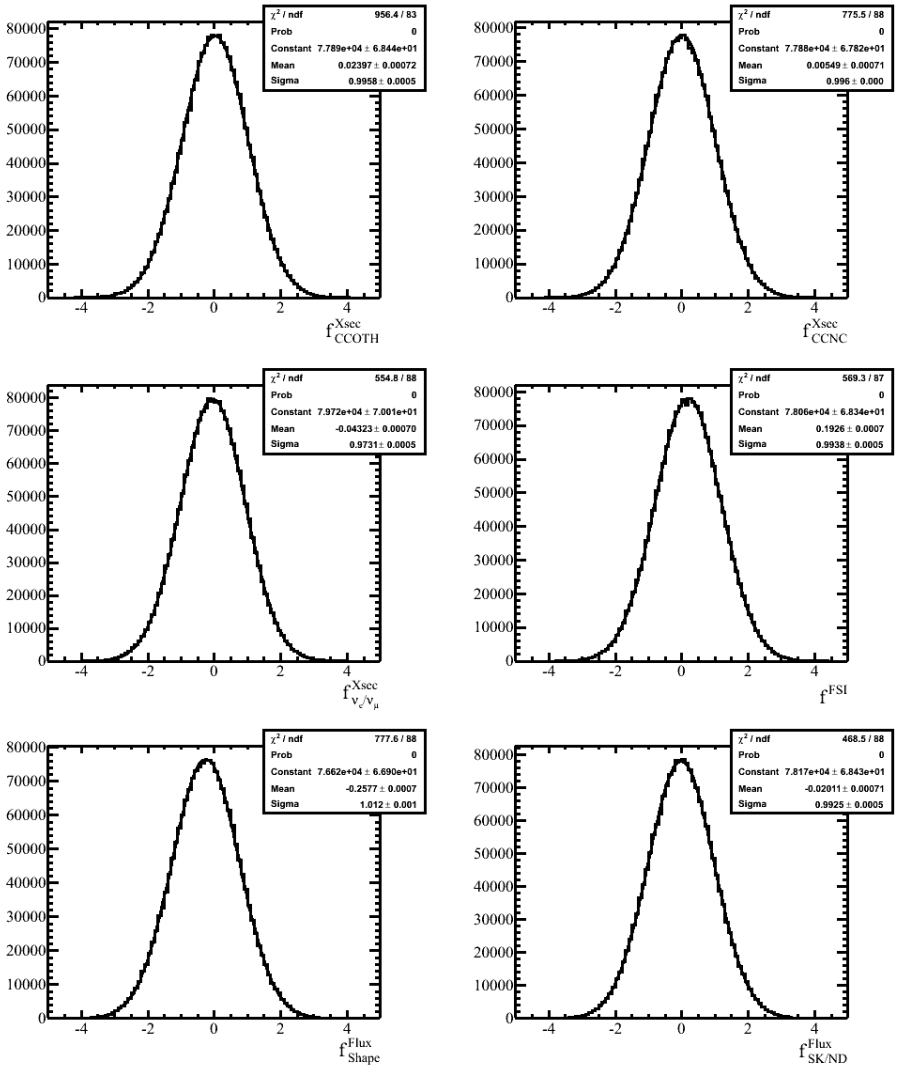


Figure 5.33: Continued systematic uncertainty parameters from simulated Monte Carlo fitted with Gaussian functions. From top left to bottom right: CCOTH and CCNC cross section ratios uncertainties, ν_e/ν_μ cross section ratio uncertainty, FSI uncertainty and flux shape and normalisation uncertainties.

Figure 5.34 presents 68%, 90% and 99% Bayesian credible intervals for T2K run 1 exposure computed using the data from figure 5.29 and the method outlined in section 5.4.4.

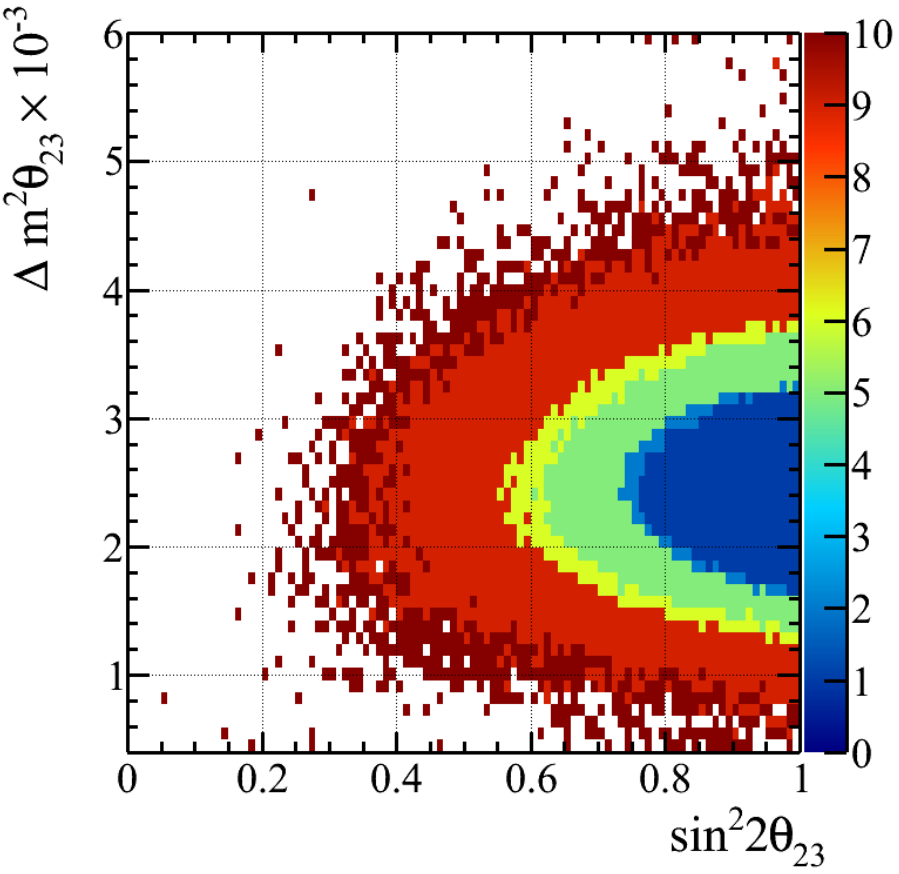


Figure 5.34: 68% (blue), 90% (green/yellow) and 99% (red) Bayesian credible intervals from MC simulation for T2K run 1 exposure and input oscillation parameters $\sin^2 2\theta_{23} = 1.0$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$. With and without systematic uncertainties included.

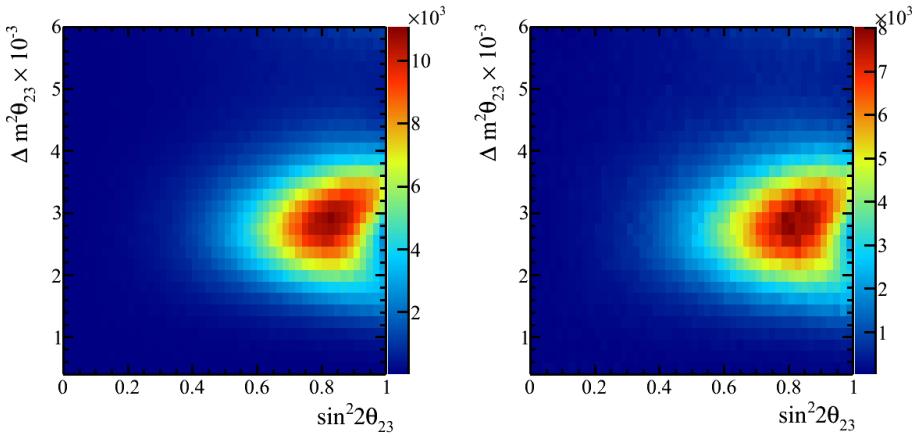
5.10 T2K run 1 results.

A ν_μ disappearance oscillation result was obtained for the T2K physics run 1 data, from an exposure of 3.23×10^{19} POT, using the MCMC method outlined in the previous sections. Boundary conditions identical to those used in the simulation run were applied to the oscillation parameters.

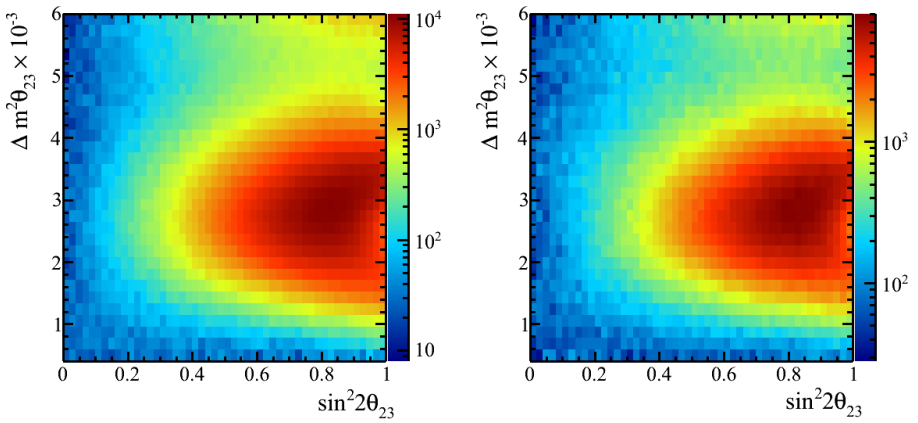
5.10.1 Oscillation parameters

The oscillation parameter values, sampled from the disappearance likelihood for the 8 1-ring μ-like events that passed the selection cuts at the far detec-

tor, are plotted in figure 5.35. Results are presented for both the statistics only and the stats+systs likelihood function. The systematic uncertainties are marginalised in the case of the full likelihood functions. This analysis found best-fit values of $\sin^2 2\theta_{23} = 0.83$ and $\Delta m^2_{32} = 2.9 \times 10^{-3}\ eV^2/c^4$.



(a) left: no systematics, right: with systematics.



(b) same with log-scale on the Z axis.

Figure 5.35: T2K run1 data sampled disappearance likelihood in the parameter space of the oscillation parameters Δm^2_{32} and $\sin^2 2\theta_{23}$ for the 8 1-ring μ-like events that passed the selection cuts at the far detector obtained using the MCMC Metropolis-Hastings method.

The reconstructed neutrino energy distribution for the 8 1-ring μ-like events along with the best fit and null oscillation predictions are shown in figure 5.36.

The marginal $\sin^2 2\theta_{23}$ and Δm^2_{32} profiles are shown in figures 5.37 and 5.38 along with the convergence of the means of 6 randomly selected Markov chains and the sampled parameter values of these chains over "time".

The $\sin^2 2\theta_{23}$ parameters are observed to converge for all Markov Chains in the range between 0.75 and 0.77 for the case with no systematics and between

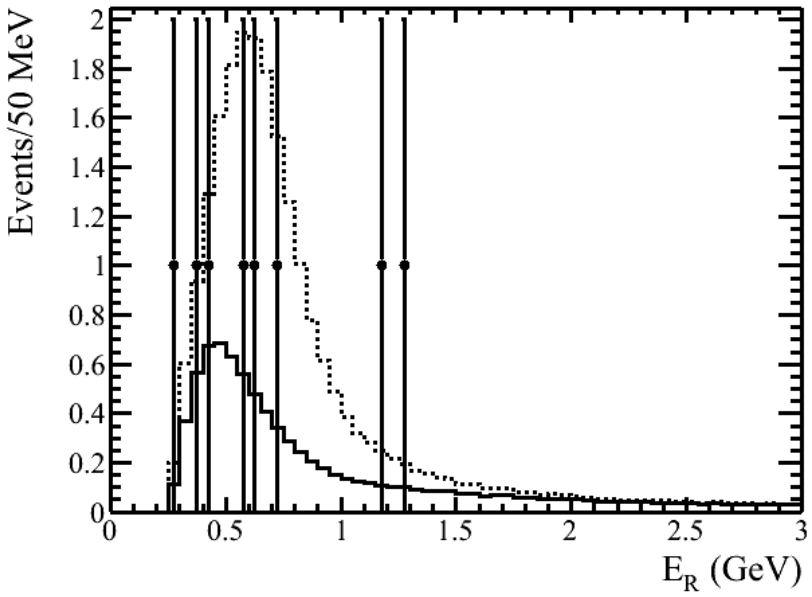
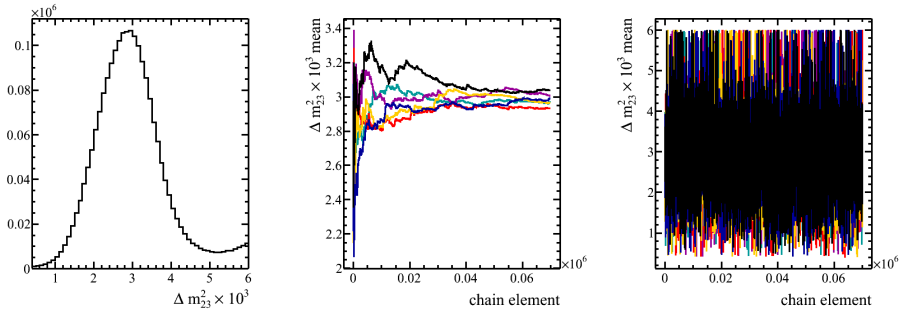


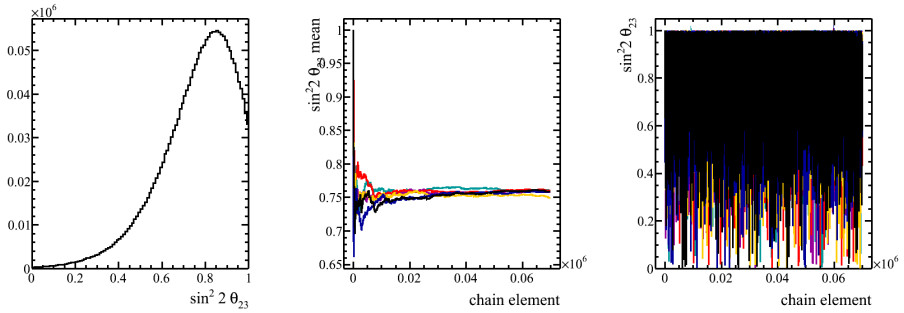
Figure 5.36: Reconstructed neutrino energy distribution for the 8 1-ring μ-like events collected during the first T2K physics run (3.23×10^{19} POT exposure). The reconstructed energy distributions in the null hypothesis (dashed line) and best fit hypothesis (solid line) are overlayed. The oscillation parameters for the best fit hypothesis are $\sin^2 2\theta_{23} = 0.83$ and $\Delta m^2_{32} = 2.9 \times 10^{-3}\ eV^2/c^4$.

0.70 and 0.77 for the case where they are included. Conversions in the range $2.9 \times 10^{-3}\ eV^2/c^4$ to $3.1 \times 10^{-3}\ eV^2/c^4$ are observed for Δm^2_{32} in both cases. Different chain lengths were used in each case with lengths of 70,000 events in the first and lengths of 20,000 for the second. In each case, the sampled parameter value plots exhibit good mixing properties.

The data profiles with systematic uncertainties included, presented in figure 5.38, are used to calculate a 1-σ error on the best fit values. For both $\sin^2 2\theta_{23}$ and Δm^2_{32}, upper and lower limits are found by selecting bins in order of decreasing number of entries until 68.3 % of the data has been included. Using this method, the best fit values with statistical and systematical uncertainties are $\sin^2 2\theta_{23} = 0.83^{+0.16}_{-0.18}$ and $\Delta m^2_{32} = 2.9^{+0.8}_{-1.0} \times 10^{-3}\ eV^2/c^4$.

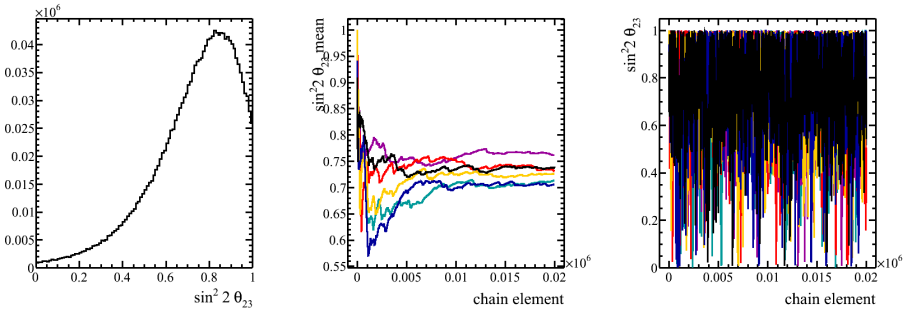


(a) $\sin^2 2\theta_{23}$ profile, aggregate mean and sampled parameter; no systematics.

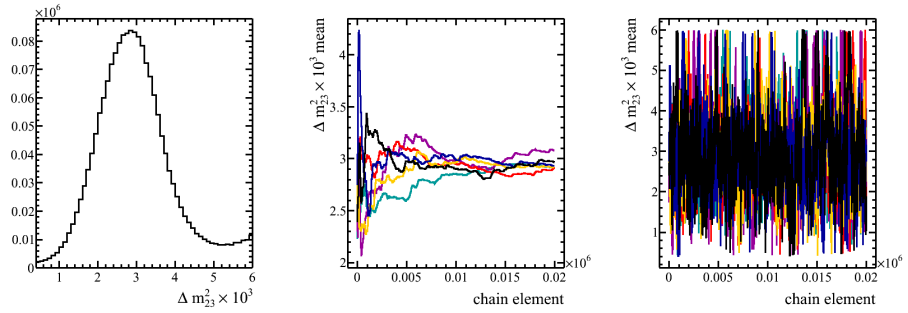


(b) Δm^2_{32} profile, aggregate mean and sampled parameter; no systematics.

Figure 5.37: Left: T2K run1 data profile of the $\sin^2 2\theta_{23}$ and Δm^2_{32} Monte Carlo Markov chain parameters without systematic uncertainties. Middle: "Time series" of the convergence of the means of $\sin^2 2\theta_{23}$ and Δm^2_{32} for 6 Markov chains chosen randomly. Right: "Time series" of the sampled parameters for the 6 same chains; both exhibit good mixing properties as the parameter space is sampled very fast.



(a) $\sin^2 2\theta_{23}$ profile, aggregate mean and sampled parameter; with systematics.



(b) Δm^2_{32} profile, aggregate mean and sampled parameter; with systematics.

Figure 5.38: T2K run1 data. Same as figure 5.37 above but with systematic uncertainties included.

5.10.2 Systematic uncertainties

The marginalised posterior profiles of each systematic uncertainty, for the MCMC with systematics, are plotted in figures 5.39 and 5.40 and fitted with Gaussian functions For each uncertainty, the nominal value from the prior Gaussian penalty term of the likelihood is centred on zero and the x-axis scale is in units of σ.

Most profiles retain a Gaussian shape with some small shifts in the mean and standard deviation. This behaviour is expected: because the T2K run 1 is heavily dominated by statistical errors, the likelihood fit has little constraining power over the uncertainties.

5.10.3 Bayesian credible intervals

The 68%, 90% and 95% Bayesian credible intervals for the T2K run 1 data are constructed using the results in figure 5.41 and the method outlined in section 5.4.4. For comparison, results from MINOS, Super-K and K2K were presented in figure 1.3.

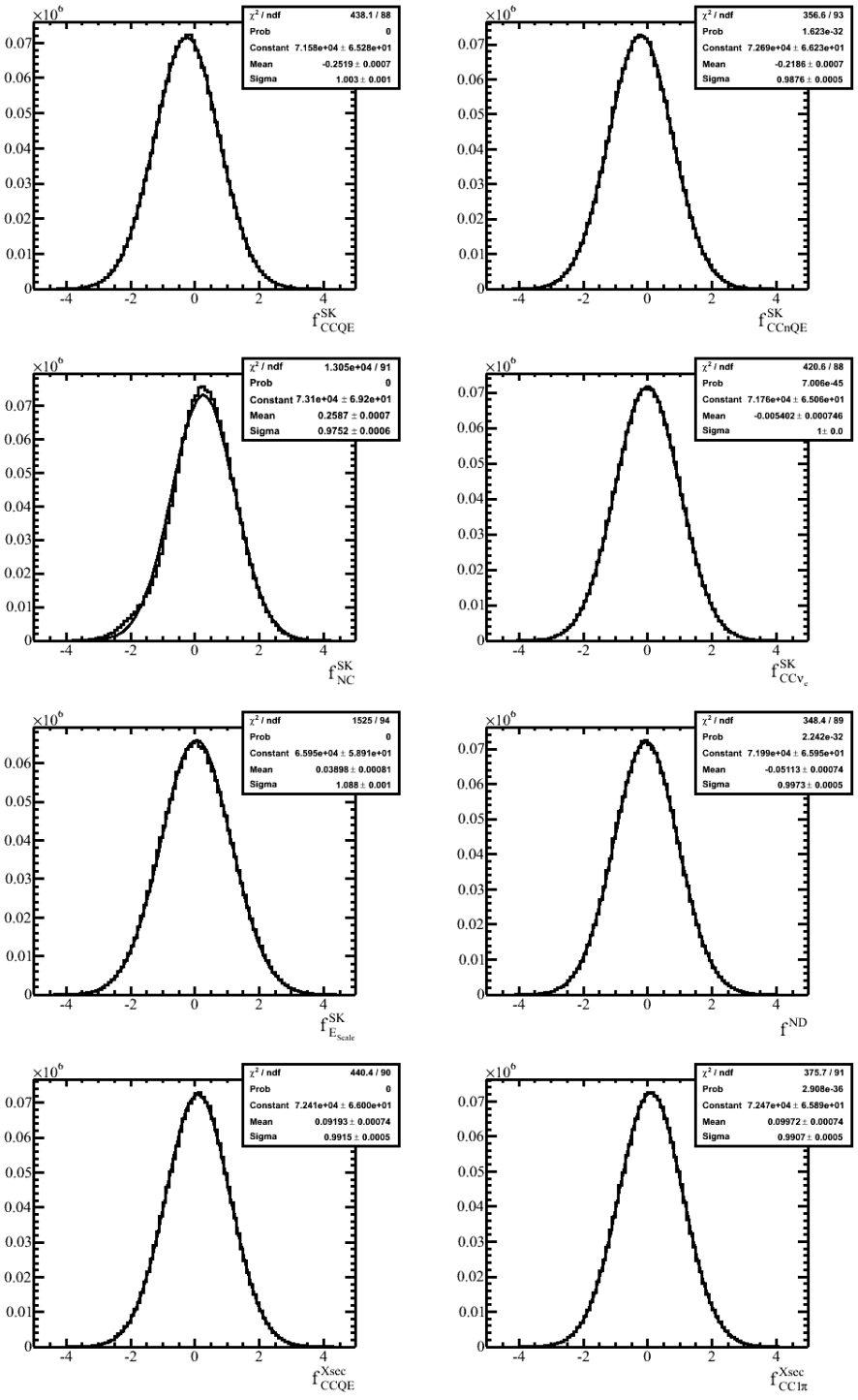


Figure 5.39: Profile of posterior marginalised systematic uncertainty parameters of the Markov chain fitted with Gaussian functions for the T2K run1 data. Here, from top left to bottom right: Super-K CCQE, CCnQE, NC and CCν_e efficiency uncertainties, Super-K energy scale uncertainty, ND280 normalisation uncertainty, CCQE cross section uncertainty, CC1π cross section ratio uncertainty.

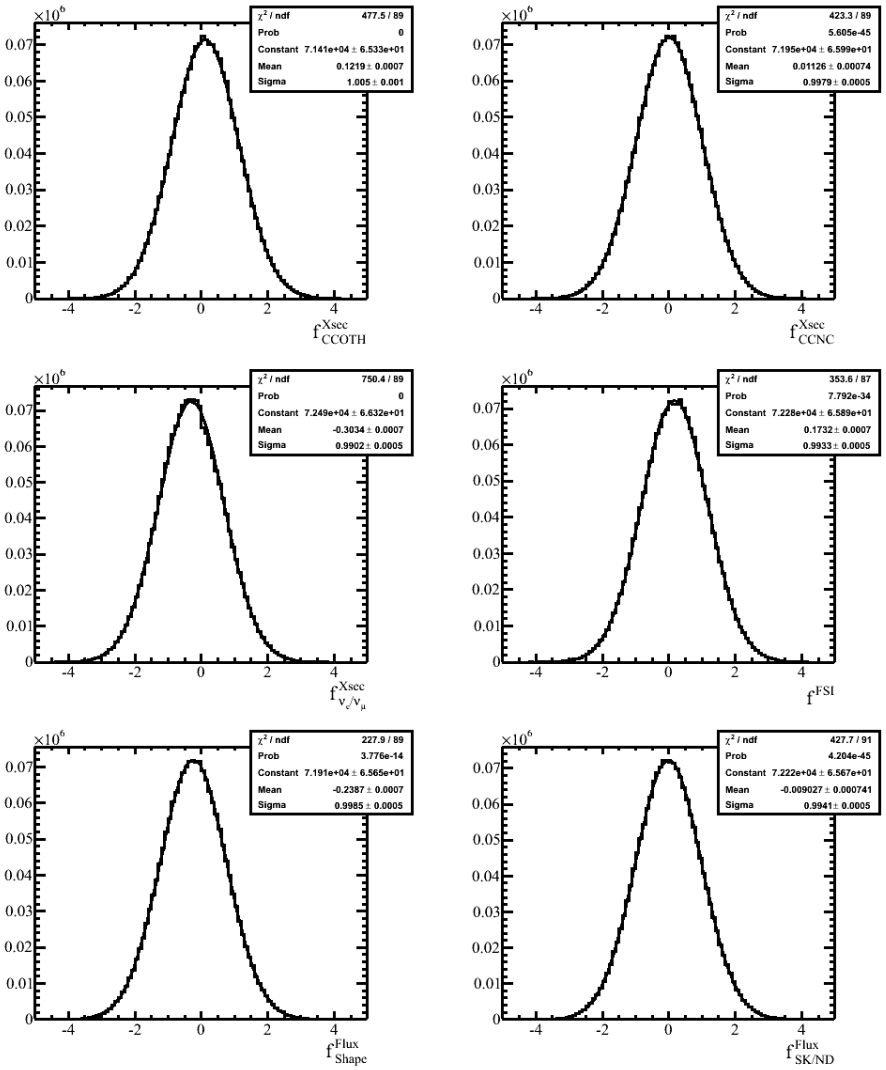


Figure 5.40: Continued systematic uncertainty parameters fitted with Gaussian functions for T2K run 1 data. From top left to bottom right: CCOTH and CCNC cross section ratios uncertainties, ν_e/ν_μ cross section ratio uncertainty, FSI uncertainty and flux shape and normalisation uncertainties.

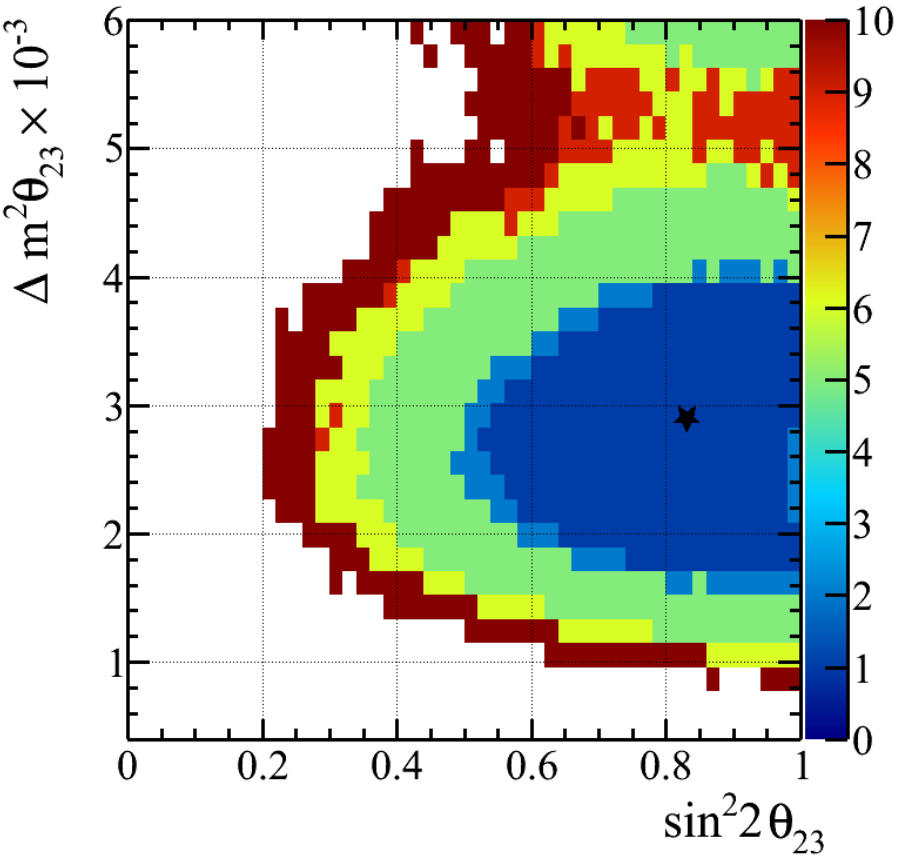


Figure 5.41: 68% (blue), 90% (green/yellow) and 99% (red) Bayesian credible intervals for the 8 1-ring μ-like events collected during the first T2K physics run (3.23×10^{19} POT exposure) with and without systematic uncertainties included. The best fit point, at $\sin^2 2\theta_{23} = 0.83$ and $\Delta m^2_{32} = 2.9 \times 10^{-3}\ eV^2/c^4$, is marked.

6 Conclusion

The atmospheric mixing angle, $\sin^2 2\theta_{23}$, and mass splitting, Δm^2_{32}, were previously measured by the MINOS, Super-K and K2K experiments. Current limits set on these parameters give a mixing angle value greater than 0.92 at 90% confidence and a mass splitting value of $2.43 \pm 0.13 \times 10^{-3}\ eV^2/c^4$. The T2K experiment is set to give the most sensitive measurements of both parameters in the next few years.

This thesis presented the results of a ν_μ disappearance analysis for the first T2K physics data set. This first data set was collected from January to June 2010 and comprised an integrated J-PARC neutrino beam exposure of 3.23×10^{19} POT. In total, eight 1-ring μ-like events had passed the Super-K selection cuts compared to an expectation of $22.8 \pm 4.8(\text{stat})^{+3.7}_{-3.7}(\text{syst})$ in the null oscillation hypothesis.

The Markov chain Monte Carlo Metropolis-Hastings algorithm was used in combination with a maximum likelihood method in order to construct 68%, 90% and 95% Bayesian credible intervals for the oscillation parameters. Oscillation parameter values of $\sin^2 2\theta_{23} = 0.83^{+0.16}_{-0.18}$ and $\Delta m^2_{32} = 2.9^{+0.8}_{-1.0} \times 10^{-3}$ eV^2/c^4 were found to give the best fit between the likelihood function and the 8 events in the data sample. The results presented in this thesis are in agreement with the main T2K analysis.

This analysis included a complete set of systematic uncertainties with contributions from Super-K, ND280, neutrino interactions and the beam provided by various T2K working groups. These uncertainties were marginalised using the MCMC method. Although most of the uncertainties used in the T2K run1 analysis were preliminary estimations, this did not matter because the first T2K run was heavily dominated by statistical uncertainties. As the amount of data collected increases, however, the analysis becomes increasingly limited by the systematical uncertainties.

A second data run was collected by the T2K experiment, before the March 2011 earthquake interuption, giving a combined run1 and run2 exposure of 1.431×10^{20} POT. Figure 6.1 presents simulated Bayesian contours for the T2K run1 and run2 combined exposure for best fit parameters of $\sin^2 2\theta_{23} = 1.0$ and $\Delta m^2_{32} = 2.4 \times 10^{-3}$ using the MCMC method. The T2K run1 systematic uncertainties were used.

A comparison of figures 5.34 and 6.1 shows that the systematic uncertainty

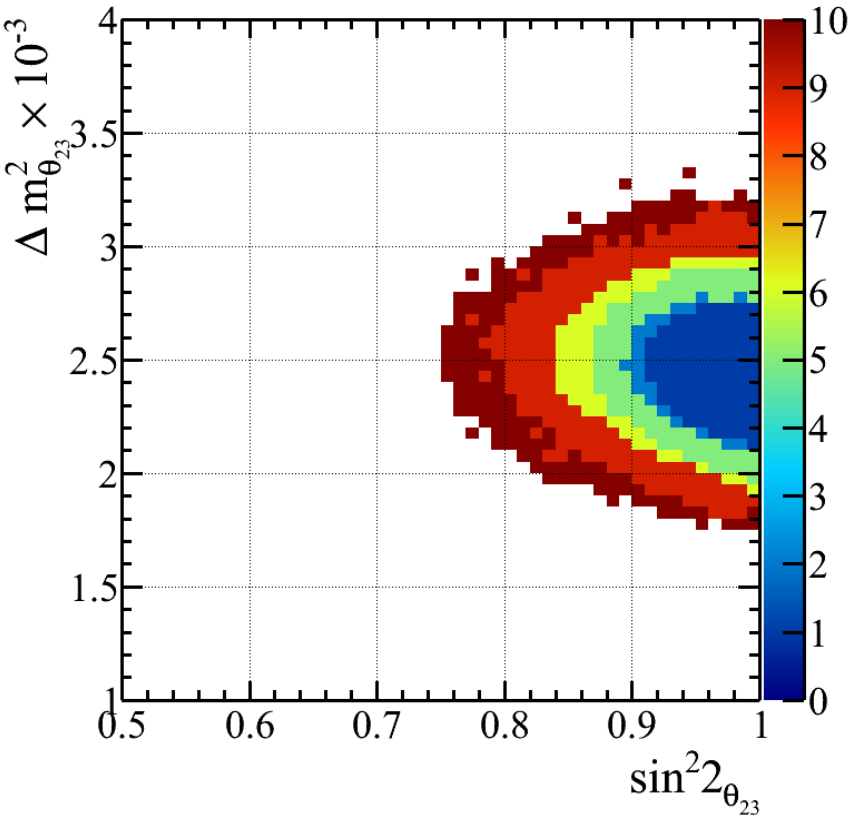


Figure 6.1: 68% (blue), 90% (green/yellow) and 99% (red) Monte Carlo Bayesian credible intervals for the first and second T2K physics runs (1.431×10^{20} POT exposure) with and without systematic uncertainties included.

accounts for an increasing fraction of the credible intervals in the second case.

The MCMC method will prove very useful for future high statistics T2K dataset by providing a posterior distribution on the shape of the systematic uncertainties that are presently assumed to be Gaussian.

Bibliography

[1] J. Chadwick, 1914. Distribution in intensity in the magnetic spectrum of the β-rays of radium, Ver. Dtsch. Physik. Ges. **16**, 383-391.

[2] W. Pauli. Letter to the physical society of Tubingen (1930). The letter is reproduced in *No time to be brief-A scientic biography of Wolfgang Pauli*, Charles Enz, Oxford university press, (2002).

[3] F. Reines and C. L. Cowan. Detection of the free neutrino. *Phys. Rev.*, 92:830–831, 1953.

[4] F. Reines and C. L. Cowan. Detection of the free neutrino: a confirmation. *Science*, 124:103–04, 1956.

[5] J. Schwinger. *Ann. Phys*, 2:407, 1957.

[6] G. Danby et al. Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos. *Phys. Rev. Lett.*, 9:36–44, 1962.

[7] Nobel prize press release (1988), available on the official website, *http* : *//www.nobelprize.org/nobel_prizes/physics/laureates/*1988/*press.html*.

[8] Martin L. Perl et al. Evidence for anomalous lepton production in e+ e- annihilation. *Phys. Rev. Lett.*, 35:1489–1492, 1975.

[9] J. Burmester et al. Anomalous muon production in e+ e- annihilation as evidence for heavy leptons. *Phys. Lett.*, 68B:297, 1977.

[10] Angela Barbaro-Galtieri et al. Electron-Muon and electron-Hadron Production in e+ e- Collisions. *Phys. Rev. Lett.*, 39:1058, 1977.

[11] K. Kodama et al. Observation of tau-neutrino interactions. *Phys. Lett.*, B504:218–224, 2001.

[12] Precision electroweak measurements on the Z resonance. *Phys.Rept.*, 427:257–454, 2006.

[13] M. Goldhaber, L. Grodzins, and A. W. Sunyar. HELICITY OF NEUTRINOS. *Phys. Rev.*, 109:1015–1017, 1958.

[14] Boris Kayser. Neutrino Mass, Mixing, and Flavor Change. 2008.

[15] Boris Kayser. Neutrino physics. *eConf*, C040802:L004, 2004.

[16] Harry J. Lipkin. What is coherent in neutrino oscillations. *Phys.Lett.*, B579:355–360, 2004.

[17] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 28:870–880, 1962.

[18] B. Pontecorvo. Neutrino experiments and the question of leptonic-charge conservation. *Sov. Phys. JETP*, 26:984–988, 1968.

[19] Jr. Davis, Raymond, Don S. Harmer, and Kenneth C. Hoffman. Search for neutrinos from the sun. *Phys. Rev. Lett.*, 20:1205–1209, 1968.

[20] J. N. Abdurashitov et al. Measurement of the solar neutrino capture rate by the Russian-American gallium solar neutrino experiment during one half of the 22-year cycle of solar activity. *J. Exp. Theor. Phys.*, 95:181–193, 2002.

[21] W. Hampel et al. GALLEX solar neutrino observations: Results for GALLEX IV. *Phys. Lett.*, B447:127–133, 1999.

[22] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002.

[23] T. Araki et al. Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion. *Phys. Rev. Lett.*, 94:081801, 2005.

[24] S. Abe et al. Precision measurement of neutrino oscillation parameters with kamland. *Phys. Rev. Lett.*, 100:221803, Jun 2008.

[25] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998.

[26] M. H. Ahn et al. Measurement of Neutrino Oscillation by the K2K Experiment. *Phys. Rev.*, D74:072003, 2006.

[27] D. G. Michael et al. Observation of muon neutrino disappearance with the MINOS detectors and the NuMI neutrino beam. *Phys. Rev. Lett.*, 97:191801, 2006.

[28] Y. Ashie et al. Evidence for an oscillatory signature in atmospheric neutrino oscillation. *Phys. Rev. Lett.*, 93:101801, 2004.

[29] P. Adamson et al. A Study of Muon Neutrino Disappearance Using the Fermilab Main Injector Neutrino Beam. *Phys. Rev.*, D77:072002, 2008.

[30] Patricia L. Vahle. Neutrino Oscillation Studies with MINOS. 2008.

[31] K. Nakamura et al. Review of particle physics. *J. Phys.*, G37:075021, 2010.

[32] M. Apollonio et al. Search for neutrino oscillations on a long baseline at the CHOOZ nuclear power station. *Eur.Phys.J.*, C27:331–374, 2003.

[33] Y. Fukuda et al. The Super-Kamiokande detector. *Nucl. Instrum. Meth.*, A501:418–462, 2003.

[34] K. Abe et al. The T2K experiment. *Nuclear Instruments and Methods in Physics Research A*, 659:106–135, December 2011.

[35] D. Beavis et al. Long Baseline Neutrino Oscillation Experiment at the AGS (Proposal E889). *Physics Design Report*, BNL 52459, 1995.

[36] FLUKA. version 2008.3c, 2008. http://www.fluka.org/fluka.php.

[37] GEANT3. A detector description and simulation tool, 1993. Application Software Group, Computing and Networks Division, CERN, Geneva.

[38] T2K internal note T2K-TN-038, 039, and 040. Accumulated flux at SK and ND280 in run 1 (2011).

[39] C. Amsler et al. Review of particle physics. *Phys. Lett.*, B667:1, 2008.

[40] Anna Pla-Dalmau. Extruded plastic scintillator for the MINOS calorimeters. *Frascati Phys. Ser.*, 21:513–522, 2001.

[41] M. Yokoyama et al. Performance of Multi-Pixel Photon Counters for the T2K near detectors. *Nucl. Instrum. Meth.*, A622:567–573, 2010.

[42] A. Vacheret et al. Characterization and simulation of the response of Multi-Pixel Photon Counters to low light levels. *Nuclear Instruments and Methods in Physics Research A*, 656:69–83, November 2011.

[43] A. Vacheret et al. The front end readout system for the T2K-ND280 detectors. *Nuclear Science Symposium Conference Record. NSS 07. IEEE*, 3:1984–1991, 2007.

[44] I. Giomataris et al. Micromegas in a bulk. *Nuclear Instruments and Methods in Physics Research A*, 560:405–408, May 2006.

[45] P. Baron et al. AFTER, an ASIC for the Readout of the Large T2K Time Projection Chambers. *IEEE transactions on nuclear science*, Vol 55, No 3:1744–1752, 2008.

[46] N. Abgrall et al. Time Projection Chambers for the T2K Near Detectors. *Nucl. Instrum. Meth.*, A637:25–46, 2011.

[47] M. Yokoyama et al. Application of Hamamatsu MPPCs to T2K neutrino detectors. *Nuclear Instruments and Methods in Physics Research A*, 610:128–130, October 2009.

[48] T2K INGRID. Official beam analysis plots.

[49] Y. Fukuda et al. Measurement of the solar neutrino energy spectrum using neutrino electron scattering. *Phys. Rev. Lett.*, 82:2430–2434, 1999.

[50] Y. Ashie et al. A Measurement of Atmospheric Neutrino Oscillation Parameters by Super-Kamiokande I. *Phys. Rev.*, D71:112005, 2005.

[51] T2K SK. Public plots available at http://t2k-experiment.org/gallery/super-kamiokande/.

[52] F. Retiere et al. Characterization of Multi Pixel Photon Counters for T2K Near Detector. *Nucl. Instrum. Meth.*, A610:378–380, 2009.

[53] Alessandro Spinelli and Andrea L. Lacaita. Physics and numerical simulation of single photon avalanche diodes. *IEEE*, 44(11):1931–1943, 1997.

[54] R. Djeliadine. LHCb TDR calorimiter note 14, LHCb 2000-039.

[55] C. Andreopoulos et al. T2K internal note T2K-TN-036.

[56] Yoshinari Hayato. A neutrino interaction simulation program library NEUT. *Acta Phys. Polon.*, B40:2477–2489, 2009.

[57] C. Andreopoulos, A. Bell, D. Bhattacharya, F. Cavanna, J. Dobson, et al. The GENIE Neutrino Monte Carlo Generator. *Nucl.Instrum.Meth.*, A614:87–104, 2010.

[58] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth.*, A389:81–86, 1997.

[59] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, Apr 1998.

[60] J. Neyman. Philos. Trans. R. Soc. London Sect. A 236, 333 (1937), reprinted in A selection of Early Statistical Papers on J. Neyman, University of California, Berkeley, 1967, p. 250.

[61] *Numerical Recipes. The Art of Scientific Computing*, pages 487–490. Cambridge Press, third edition, 2007.

[62] M. Galassi et al. *Gnu Scientic Library Reference Manual - Second Edition* (Network Theory Ltd, 2003).

[63] F. James. *Minuit*, CERN Program Library Long Writeup D506.

[64] *Practical Mathematical Optimization. An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*, pages 57–96. Springer, 2005.

[65] R.M. Neal. *Probabilistic Inference Using Markov Chain Monte Carlo Methods* , Technical Report CRG-TR-93-1, Dept. of Computer Science, University of Toronto.

[66] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2nd edition, 2004.

[67] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys*, 21:1087–1092, june 1953.

[68] W.K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[69] B. Walsh. *Markov Chain Monte Carlo and Gibbs Sampling* , Lecture Notes for EEB 581, version 26 (2004), University of California.

[70] Y. Ashie et al. Measurement of atmospheric neutrino oscillation parameters by super-kamiokande i. *Phys. Rev. D*, 71:112005, Jun 2005.

[71] M. Malek. Private communication.

[72] J. Hignight. presentation, T2K collaboration meeting, September 2001.

[73] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[74] K. Abe et al. First muon-neutrino disappearance study with an off-axis beam. *Phys. Rev. D*, 85:031103, Feb 2012.

[75] NA61 Collaboration. CERN-SPSC-2011-035. SPSC-SR-091, CERN, Geneva, 2011.

[76] N Abgrall et al. Measurements of Cross Sections and Charged Pion Spectra in Proton-Carbon Interactions at 31 GeV/c. *Phys.Rev.*, C84:034604, 2011.

[77] *Numerical Recipes. The Art of Scientific Computing*, pages 100–102. Cambridge Press, third edition, 2007.

[78] *Numerical Recipes. The Art of Scientific Computing*, pages 48–55. Cambridge Press, third edition, 2007.

[79] Andreopoulos, C. and others. The GENIE Neutrino Monte Carlo Generator Physics and User Manual.