Queen Mary University of London

Developing the T2K Neutrino Oscillation Analysis by Using Pion Samples at the Near Detector

Nauman Akhlaq

Submitted in partial fulfilment of the requirements of the Degree of Doctor of Philosophy

in the

Particle Physics Research Centre, School of Physics and Astronomy
Queen Mary, University of London
United Kingdom

September 2023

Declaration

I, Nauman Akhlaq, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date:

Abstract

The T2K (Tokai to Kamioka) experiment is a long baseline neutrino experiment in Japan that consists of a neutrino source beamline (from J-PARC), a near detector (ND280) and a far water Cherenkov detector (Super-Kamiokande). The neutrino rates between the two detectors are measured and this is used to determine the Pontecorvo–Maki-Nakagawa-Sakata Matrix (PMNS) parameters that describe neutrino oscillations. The current oscillation analysis aims to constrain the δ_{CP} parameter using the Markov Chain for 3 flavour fitting (MaCh3) framework. In this thesis, I added pion information at the near detector (ND280) to better restrict the cross section parameters and potentially improve the neutrino oscillation sensitivity. I implemented a new feature that introduced pion tagged samples to the oscillation framework after studying the kinematics of pions at ND280. The impact of this new feature was studied using likelihood scans and fake data studies at the near detector, followed by spectra predictions at the far detector. Finally, I performed a joint data fit and extracted neutrino oscillation parameter sensitivities. During my time on T2K I served as a Data Quality (DQ) expert for the ND280, the work that I did is also described in this thesis.

Acknowledgements

بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ

The journey in starting a PhD to submitting a final thesis is long one, one that is filled with adventure and unexpected turns. The journey can be exciting, addictive, and, at times, rewarding. But it can also seem to be long, arduous, and, at times, lonely. Although this work conveys (I hope) important information, it contains just some of the fruit of my thought over four years. A period of over four years in which I started out as a PhD student, then a year later became a husband to a superb wife, and three years after, a father to a beautiful daughter. I have been blessed with so much in such a short period of time. I cannot help but feel that time has moved by so quickly. As the Arabic poet said:

ما مَضَتِ الدهورُ فقد ساءَتْ وبَدَتْ ولكنَّ العمرَ يُبقي ويُمحَى الإنسانُ

The ages have passed, they appeared and departed,

But life persists, while man fades away.

Indeed. Like everyone else, I will move forward. Like 2018, 2023 will become a distant memory, time will march on, with our growing older as every day passes. Until only our memory remains, as we can never abide in this world eternally. In making this journey, I was blessed with great company and amazing individuals. People who were very willing to help, and went out of their way to make sure I was doing well. I will try to mention them here, but I may not remember some names, I ask you to forgive me if you are not mentioned and deserve to be.

All praise is for the Most High. I owe a debt of gratitude to my parents: Dr Akhlaq and Dr Naseem, who supported me throughout this journey. I also owe this to my wife, Iman, who was with me every step of the way from when we got married in 2019. She put up with my odd working hours and is a great mother to our daughter. I would also like to thank my sister and my brother for their important support during this time. I would like to thank my in-laws for being ever so supportive, kind, and welcoming whether I was in London, UK or Amman, Jordan. Thank you all so very much.

I would like to thank my supervisors Dr Linda, Dr Ulla, and Prof Francesca for their support during my PhD, the thesis writing process, and beyond. Thank you all for supervising this thesis from start to finish. Thank you Linda for being there for discussions and issue debugging at a moment's notice. I would like to thank Dr Sophie King for all her help and support during my PhD especially early on when I was trying to find my feet.

Thank you to the MaCh3 group: Dr Clarence, Dr Patrick, and Dr Kamil. The help and support you provided was much appreciated and helpful. A sincere thank-you to my fellow Near Detector Analysers, in particular: Dr Joe Walsh (who was always ready to help), Dr Tristan Doyle (who possessed much patience in answering numerous questions), and Dr Matt Lowe (who helped me in Japan and with ND detector work). Thanks to Dr Tom Holvey, from the Far Detector analysers, who helped with the joint-fit and was ready to answer questions at any time.

Finally, I would like to thank my examiners Dr Seth Zenz and Dr Andy Blake for reading through this thesis, assessing my work, and suggesting valuable improvements.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis focuses on the T2K (Tokai to Kamioka) experiment, specifically looking at developing the oscillation analysis by using pion information at the near detector. The T2K experiment is a long-baseline neutrino oscillation experiment established to make cutting edge measurements of the PMNS oscillation parameters: $\sin^2\theta_{23}$, Δm^2_{32} and $\sin^2\theta_{13}$. Designed to run with a muon neutrino beam, T2K has been looking at the two channels of muon neutrino disappearance and electron neutrino appearance. Between 2010 and 2013, T2K collected data and provided the most precise measurement of $\sin^2\theta_{23}$ in the world at that time [1]. The T2K experiment is still taking data with an additional goal: the determination of the CP violating phase in neutrinos, δ_{CP}. This is the parameter that captures the flavour oscillation difference between antineutrinos and neutrinos.

My thesis will be structured as follows: in Chapter 2, the Neutrino Physics chapter, the history of neutrinos and neutrino oscillations will be discussed. I will recap the formalism for describing neutrino oscillations, alongside looking at the most recent values of the PMNS parameters (as of November 2022). I will also look at the different neutrino interactions at T2K, and how understanding these allow us to reconstruct the neutrino energy more accurately and constrain systematic parameters in order to improve our sensitivity to neutrino oscillations. This will be followed by a description of nuclear effects and open questions in neutrino physics.

In Chapter 3, The T2K Neutrino Experiment, I will describe the different detectors which are used within the experiment, alongside the neutrino source at T2K. A section describing the neutrino simulation at the near and far detectors is included at the end. In the following Chapter 4, Trip-T Calibration and Electromagnetic Calorimeter Data Quality, I will speak about calibration and the data quality tasks I performed as my service task.

In Chapter 5, The Oscillation Analysis at T2K, I will describe the methods of fitting used in T2K, including the mechanics of fitting with oscillation and non-oscillation parameters. A description of methods used to extract results from the multidimensional MaCh3 fit used in this thesis will also be presented. In the following chapter, Chapter 6, The Near Detector Selections and Systematic Sources, the selections used at ND280 will be described along with the novel work done in this thesis, namely to split the CC1Pi No Photon sample by pion tag in lepton kinematics, generate a set of systematic uncertainties for this sample, and fit this new feature to data.

In final chapter, Chapter 7, Results, I will look at the results of the feature implementation. A comparison between the split feature and the no-split feature will be carried out to assess the implementation and constraints seen in cross-section parameters. Additionally, neutrino oscillation parameters extracted from a joint fit with the near and far detector will be shown. Finally, a conclusion and summary of my work will be presented. My work develops the oscillation analysis, such that pion kinematics can be incorporated in the future to help constrain systematic sources, with the ultimate aim to constrain the CP violation parameter, δ_{CP}, and learn more about the nature of our universe.

Chapter 2

Neutrino Physics

Neutrinos are fundamental particles with a spin of $\frac{1}{2}$ (fermions) which have a small non-zero rest mass and interact via the weak interaction and gravity. They are neutral in charge and do not interact via the strong force. Excluding massless particles, neutrinos have the lightest rest mass of all the particles in the Standard Model. The gravitational interaction is therefore negligible, with interactions via the weak force the most prominent interactions considered.

Neutrinos are found in three flavours, namely the electron, muon, and tau neutrinos, which are accompanied by their associated charged lepton. Although a separate distinct mass could be expected for each neutrino flavour, there are three distinct mass states which are not in one-to-one correspondence with the three distinct flavour states. In actuality, each flavour state is a quantum superposition of three mass states. This chapter looks at neutrino physics, in particular the history from Wolfgang Pauli's postulation of the neutrino to the 2015 Nobel Prize winning results from SNO [2] and Super-K [3]. The discovery of neutrino oscillations is discussed along with developments until the present day. Alongside this, the current theory framework underpinning neutrino oscillations will be discussed with current open questions and a view towards future neutrino research.

2.1 Neutrino History

In the early 1900s, the model of the atom transitioned from a 'Plum Pudding' model proposed in 1907 by J. J. Thomson after the discovery of the electron a decade prior [4], to an atomic nucleus based model pioneered by Ernst Rutherford. Rutherford's experiments looking at the deflection angle of alpha particles demonstrated that positive charge was not spread throughout the atom, but was rather centred in a small central point within the atom surrounded by negatively charged electrons [5].

It was understood at the time that protons and electrons were all the components of the atom, and the structure of the atom consisted of a positively charged nucleus holding together the outer shell of negatively charged electrons. However this model was not able to account for a surprising phenomenon at the time. Some atoms were observed to 'break down' through the emission of particles; this was unexpected as atoms were assumed to be stable. A key observation illustrating this was the existence of beta decay [6], where a negatively charged particle (an electron) would be emitted from the atom, in a sense removing some energy from it.

In the following years, β decay became an intense area of study. It was found that an emitted β particle could have a range of energies on a spectrum. This was puzzling as the current model of β decay as a two-body decay involving a proton and an electron was not sufficient to explain the continuous energy spectrum of the emitted electron. In 1930, Wolfgang Pauli postulated the existence of, as he called it, a new particle coined the 'neutron' to explain what carried away additional energy as a *'verzweifelten Ausweg'* - a desperate way out [7]. Pauli's postulation of a new particle maintained the principle of the conservation of energy at a time when the possibility of it being violated at small scales was discussed. The word neutron was taken in 1932 by the discovery of the neutron proper by James Chadwick in 1932 [8]. Hence, Pauli's particle was renamed by Enrico Fermi as 'neutrino', meaning 'little neutral one'.

Enrico Fermi was a well renowned theoretical and experimental physicist who took the idea of the neutrino and developed the theory surrounding it in 1934 [9]. His theory described the neutron in the nucleus decaying to a proton and emitting an electron and a neutrino:

$$n \rightarrow p + e^{-} + \nu \,. \tag{2.1}$$

Atomic theory and β decay were now better understood due to the detection of the neutron, however the neutrino was thought too elusive to ever be detected. With the advent of nuclear technology and fission reactors, the experimental outlook changed substantially such that $\sim$20 years later the first neutrinos were detected.

Neutrinos are extremely light and hard to detect due to rarely interacting with matter. In order to increase the chances of observing a neutrino, a large neutrino flux is required to yield a significant probability of detection. In the 1950s, Frederick Reines and Clyde Cowan working at Los Alamos in New Mexico looked at detecting neutrinos initially by setting up sensors in a shaft next to a nuclear bomb explosion [10], their hope was that the large amount of β decays producing a huge flux of neutrinos would allow for these elusive particles to be detected. However, they were convinced to use a nuclear reactor to ensure easier repeatability of results. For this they relied on the inverse β decay process:

$$p + \bar{\nu} \rightarrow n + e^{+} . \tag{2.2}$$

In 1953, Reines and Cowan tried their scintillator setup near the Hanford plutonium reactor in Hanford, Washington. Relying on neutrinos from the reactor interacting via inverse β decay in their detector filled with 300 litres of liquid scintillator, they would observe a characteristic trace of two flashes of light, one from neutron capture and the second from the positron annihilation. However, due to cosmic ray backgrounds, it was hard to get more than a speculative result. Following this, Reines and Cowan improved their setup in their consequent Savannah River Experiment in 1956. They provided their detector, which was now 12 m underground, with better shielding from cosmic rays. Reines and Cowan observed characteristic light traces of inverse β decay from neutrinos, confirming their results from the Hanford experiment. The elusive neutrino had finally been discovered. Wolfgang Pauli was informed at the time on June 1956 via telegram of the results; what he had theorised had now been shown to exist by Reines and Cowans' ingenious experiments [11].

2.1.1 Neutrino Flavours and Oscillation

Independent verification of Reines and Cowans' detection of the free neutrino took eight years to occur, perhaps in part due to the secretive nature of their reactor neutrino source [12]. However, in the 1960s different neutrino sources were being explored with other experiments. In the Homestake Gold Mine in Lead, South Dakota, Ray Davis and his collaborator John Bahcall were attempting to measure the number of solar neutrinos produced by the sun. Deep underground at 1490m, Davis filled a 100,000 gallon tank with C_2Cl_4, perchloroethylene (dry-cleaning fluid), in order to detect solar neutrinos [13]. This detection made use of another well-known type of inverse β decay: neutrino capture. The specific process they used was elucidated theoretically by Pontecorvo back in 1946 [14]:

$$\nu_e + {}^{37}Cl \rightarrow e^{-} + {}^{37}Ar. \tag{2.3}$$

Neutrino capture occurs when a neutrino is captured by a neutron, which changes into a proton with an electron being emitted. Experiments which make use of neutrino capture, known as radiochemical experiments, have a lower energy threshold and are therefore able to detect low energy neutrinos (less than 10 MeV). Davis and Bahcall waited 2-3 months for interactions to occur in their tank [15], after which they bubbled helium gas through the tank to isolate radioactive isotopes of Argon-37 formed by neutrino capture. Through this, they were able to determine the solar neutrino flux rate [16].

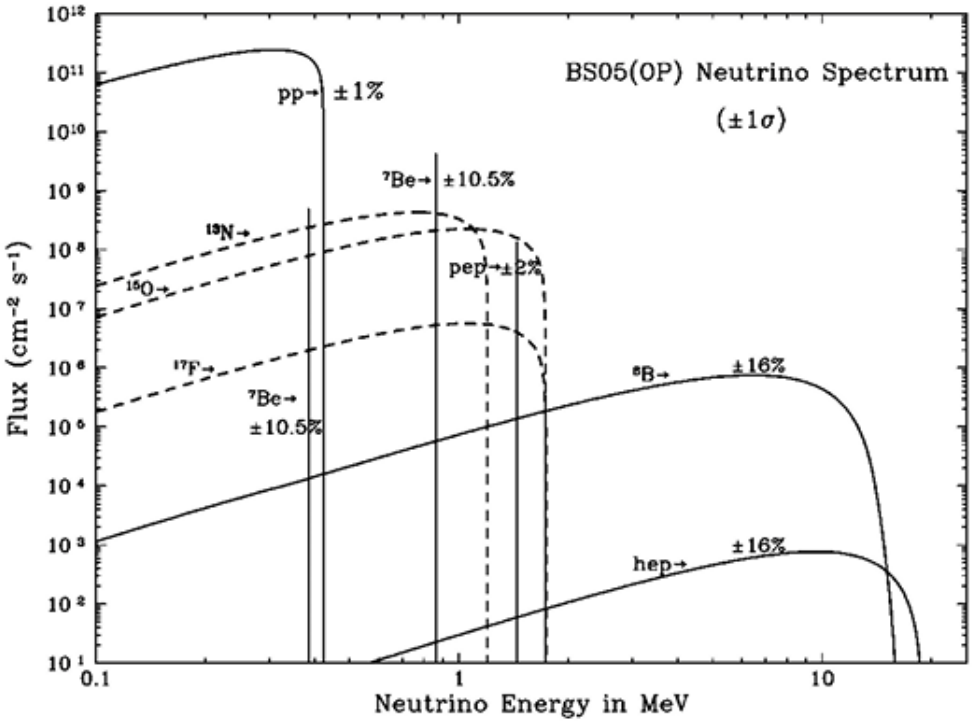


FIGURE 2.1: The neutrino flux rate and energy from earth expected by the Standard Solar Model (SSM). The neutrinos observed within the Homestake detector are from the ^8B SSM process. Around 99% of the neutrinos from the sun come from the pp chain, with the other processes constituting $\sim 1\%$. Diagram from [17].

However, when Davis and Bahcall compared their results to the predictions of the Standard Solar Model (SSM) at the time (see Figure 2.1), they found that there was a deficit in the number of neutrinos detected. They looked at neutrinos from the ^8B SSM process, where the energy of the neutrinos were compatible with the experiment threshold for neutrino capture. Their observations were roughly $\frac{1}{3}$ of what was expected from the Solar Neutrino Model. In 1968, Davis and Bahcall published their result and the reported deficit resulted in the famous Solar Neutrino Problem [18]. This was the first experiment to show a deviation from the Standard Solar Model.

In order to solve the solar neutrino problem, in 1969 (after the muon neutrino flavour discovery in 1962) Pontecorvo and V. Gribov proposed that electron neutrinos had oscillated into muon neutrinos on their way from the sun and were thus being missed by the Homestake experiment [19]. This became a possible explanation, alongside the possibility that the solar model or the experiment itself was incorrect [20]. The Solar Neutrino Problem became an important fixture of the 1970s - 1990s in neutrino experimental and theoretical research, with many experiments attempting to understand the observed deficit. Similar radiochemical experiments, such as GALLEX/GNO [21] and SAGE [22], used Gallium instead of C_2Cl_4 for neutrino capture. This allowed them to probe neutrinos of lower energy from the sun, specifically from the *pp*

process as Gallium has a lower threshold for neutrino capture (0.233 MeV). These experiments also reported a deficit. With more and more data confirming the deficit, neutrino oscillations became the leading hypothesis to explain the discrepancy between theory and experiment.

Ten years before the Homestake experiment, the fact that a neutrino could change its type or oscillate (albeit to an antineutrino) was first introduced by Bruno Pontecorvo in 1958 based on an analogy with Neutral Kaon mixing [23]. The theory of neutrino oscillations was later elaborated and worked on by Maki, Nakagawa and Sakata in 1962, leading to the development of the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) formalism showing how neutrino flavours oscillate based on mass eigenstates [24]. Maki, Nakagawa, and Sakata proposed that neutrinos interact through their flavour eigenstate via the weak interaction but travel in their mass eigenstate (see Section 2.2). This could only occur if at least 2 of the neutrino mass states were non-zero, meaning that neutrino oscillations were due to mixing between neutrino mass and flavour eigenstates which necessitated the novel idea that neutrinos have mass.

The model of three neutrino flavours is widely accepted, with a fourth active flavour being ruled out through precise measurements at CERN from the ALEPH, DELPHI, L3 and OPAL collaborations [25]. The muon neutrino was discovered early by Schwartz, Steinberger and Lederman in 1962 [26] using their novel method of accelerating particles to produce a beam of neutrinos, six years after the detection of the (electron) neutrino by Reines and Cowan. The tau neutrino was discovered in 2000, from the results of the DONUT experiment [27]. Each neutrino flavour corresponds to an electron, muon, or tau particle which is the signature lepton observed when the corresponding neutrino interacts with a nucleon in a charged current interaction. At the time, this set of discoveries completed the full set of fundamental particles that the Standard Model of Particle Physics predicted (excluding the Higgs Boson which was discovered in 2012 [28]). However, understanding why neutrinos have mass, and the fact that these experiments do not rule out other types of neutrinos interacting in a different way (such as sterile neutrinos), are open questions which are discussed below.

2.1.2 Neutrino Oscillation Confirmation

To settle the solar neutrino problem once and for all, a new generation of large experiments in the 1980s and 1990s paved the way for the confirmation of neutrino oscillations. The Kamiokande experiment, located 1000 m underground in Kamioka, Japan, and the SNO (Sudbury Neutrino Observatory) experiment, located 2100 m underground in Ontario, Canada were two such experiments that contributed significantly to understanding neutrino oscillations. Kamiokande (Kamioka Neutrino Experiment) was originally constructed in 1983 in Kamioka, Japan to study proton decay [29]. Kamiokande was an innovative water Cherenkov detector which was later

adapted to be able to detect solar neutrinos in 1985 after proton decay was not observed [30]. In 1984, SNO was established to study neutrinos from the sun, and to make the first model independent measurement of neutrinos [31]. The SNO heavy water Cherenkov detector started data taking in 1999. Both experiments focused on the ^8B SSM process at neutrino energies greater than 1 MeV.

The SNO experiment is an acrylic vessel that serves as a water Cherenkov detector filled with 1,000 tonnes of ultra-pure deuterium (D_2O). Located underground in Sudbury, Canada, SNO detects the Cherenkov light cone from high momentum interacting particles [32]. SNO detects neutrinos through 3 interaction types: charged current, neutral current, and elastic scattering. The charged current interaction is sensitive to electron neutrinos, with the elastic scattering reaction sensitive to all neutrinos but less sensitive to tau and muon neutrinos. The neutral current interaction is sensitive to all neutrinos. The neutral current interaction (with a energy threshold of 2.2 MeV) results in a neutron which is captured by the deuterium nucleus in heavy water. This produces a 6.25 MeV gamma ray that Compton scatters producing electrons, these high momentum electrons go on to produce Cherenkov light [33]. Ultimately, by observing the flux of each neutrino flavour within its interaction type, and comparing this to the SSM; the first direct evidence of neutrino oscillation in the early 2000s was found. The significant graph is shown in Figure 2.2 below:

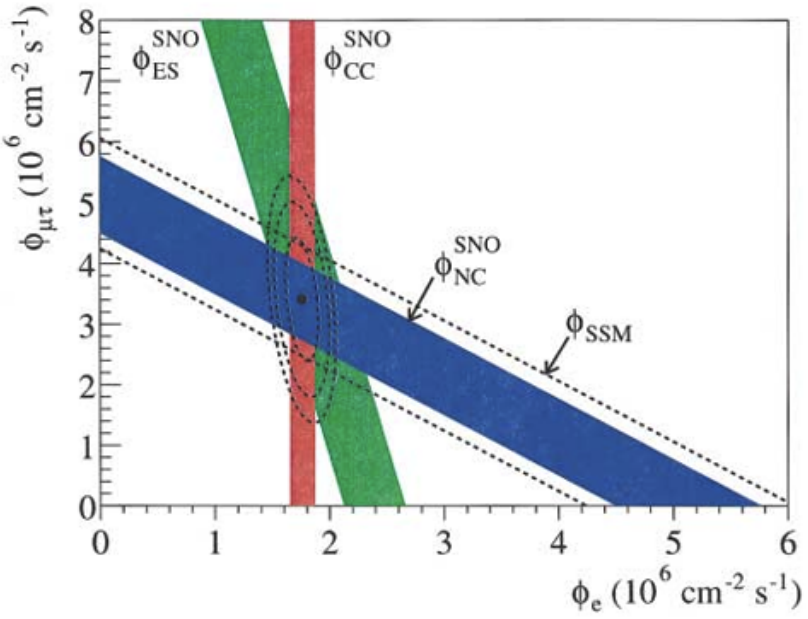


FIGURE 2.2: The seminal graph from the paper SNO published in 2002, showing the flux of muon or tau neutrino flavours given the flux of electron neutrinos. The 3 solid coloured strips show the 3 separate interactions. The thickness of these strips represent $\pm 1\,\sigma$ uncertainties. The dotted black line around the blue band represents the flux prediction of the ^8B process from the SSM. The bands intersect with each other at a best fit point, showing that neutrino flavour oscillation occurs. Diagram from [33].

The Kamiokande experiment had its final run in 1995. Alongside looking at solar and atmospheric neutrinos, it also detected neutrinos from the first well studied supernova SN 1987A [34]. In 1996, the successor and the new next generation Super-Kamiokande (Super-K) experiment was switched on [35]. Super-K was unique in the fact that it could distinguish the direction of incoming neutrinos and their flavour using the direction of the Cherenkov light cone and its signature. Super-K was also able to look at atmospheric neutrinos, which became another important source for observing oscillations [36]. Super-K was built with a bigger tank filled with 50,000 tonnes of ultra pure water (47,000 more tonnes than Kamiokande) and had $\sim 10,000$ more light detectors (PMTs) lining the inner detector compared to Kamiokande's 1,000 PMTs.

Super-K had particular success with understanding atmospheric neutrinos. These neutrinos are formed by cosmic rays interacting with Earth's atmosphere. In particular, understanding the atmospheric neutrino anomaly observed by Kamiokande prior was important [37]. This anomaly arose due to observing a deficit of certain neutrino flavours reaching the detector when compared to the simulation. An expected number of ν_e events were observed, however there was a deficit in ν_μ events [38]. Super-K tackled this problem by using their innovative direction information from events; by observing neutrinos from the near side of the earth (leaving Cherenkov radiation on the bottom half of the detector) and comparing this to neutrinos from the far side of the earth which had to travel through earth (leaving Cherenkov radiation on the top half of the detector), deficits could be seen as a function of the zenith angle. Measuring the difference in events, direct evidence of neutrino oscillations in atmospheric neutrinos for three flavours was established. [39]

Neutrino oscillations became widely accepted and were experimentally confirmed by the Super-Kamiokande collaboration, for atmospheric neutrino oscillations, in 1998 and by SNO, for solar neutrino oscillations, in 2001, earning Takaaki Kajita (Super Kamiokande) and Arthur MacDonald (SNO) the Nobel Prize in Physics in 2015 [40]. The theory of neutrino oscillations is an addition to the original Standard Model of Particle Physics; therefore, its study has become a fast-moving and exciting field to be part of. Ever since the 2000s, neutrino physicists have adopted the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) framework as the standard theoretical description of three-flavour neutrino oscillations. As technology and experiments have progressed (T2K, NOvA [41], and, in the future: DUNE [42] and Hyper-K [43]), so has the ability to constrain the parameters of the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) Matrix.

2.2 Neutrino Oscillation Physics and Theory

The parameters of interest for the neutrino oscillation community are contained within the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) Matrix formalism.

Theoretically, flavour states for neutrinos are described as a sum of the different mass states $|\nu_i\rangle$, these can be inverted to give the mass state as a sum of the different neutrino flavour states $|\nu_\alpha\rangle$:

$$|\nu_i\rangle = \sum_\alpha U_{\alpha i} \, |\nu_\alpha\rangle \, , \tag{2.4}$$

where $U_{\alpha i}$ are elements of the unitary matrix U, where $\alpha = e, \mu, \tau$ and $i = 1, 2, 3$, so that the PMNS matrix is generally expressed:

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} . \tag{2.5}$$

One of the possible parameterisations is detailed as such [44]:

$$U = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{CP}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{CP}} & c_{23}c_{13} \end{bmatrix} , \tag{2.6}$$

here $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$. The δ_{CP} parameter is known as the CP violating Dirac phase. The unitary matrix U is usually separated into 3 separate matrices which describe the different sources of neutrino oscillations:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{2.7}$$

Atmospheric neutrino experiments such as Super-Kamiokande are mainly sensitive to θ_{23} parameter in the first matrix. Reactor neutrino experiments such as Daya Bay [45] are mainly sensitive to the θ_{13} and δ_{CP} parameters in the second matrix. Solar neutrino experiments such as SNO are sensitive to the θ_{12} parameter in the final matrix. Long-baseline experiments such as T2K and NOvA are sensitive to parameters in the first two matrices (θ_{23}, θ_{13} and δ_{CP}).

In studying the solar sector, SNO has found out much about solar neutrinos and oscillations, with its successor SNO+ looking for neutrino-less double beta decay and the neutrino mass hierarchy (see Section 2.5). For atmospheric neutrino studies, the Super-Kamiokande detector is looking at θ_{23}. Reactor neutrinos are being studied by CHOOZ and Daya Bay. Daya Bay conclusively showed that the θ_{13} mixing angle is non zero, finding its value consistent with other reactor experiments. Long-baseline experiments include T2K and NOvA. NOvA is looking at precisely measuring the mixing angle θ_{23} to see whether it is larger or smaller than, or equal to 45°. Along with T2K, it is looking to determine the mass splitting for neutrinos and anti-neutrinos. Furthermore, a shared aim is to put strong constraints on the CP-violating phase δ_{CP} and the neutrino mass hierarchy. It is hoped that a joint fit between the T2K and NOvA collaborations in the future will allow for better constraints on δ_{CP}.

Looking at Equation 2.4 above, the superposition of one state with another can then be calculated by taking the scalar product between them [46]. This leads to a probability of oscillation from one flavour to another (Equation 2.8), given as a function of the oscillation distance L, the neutrino energy E and the difference between the squares of the neutrino masses $\Delta m^2_{ij} = m^2_i - m^2_j$, where the index i runs from 1 to 3 and the index j runs from 1 to 3:

$$P\left(v_\alpha \to v_\beta\right) = \left|\langle v_\beta | v_\alpha(t)\rangle\right|^2 = \sum_{k,j} U^*_{ak} U_{\beta k} U^*_{\alpha j} U_{\beta j} \exp\left(-i\frac{\Delta m^2_{ij} L}{2E}\right) . \tag{2.8}$$

By using the squared unitarity relation, one can arrive at:

$$\begin{aligned} P\left(v_\alpha \to v_\beta\right) = \delta_{\alpha\beta} - 4\sum_{i>j} \mathrm{Re}\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin^2\left(\Delta m^2_{ij}\frac{L}{4E}\right) \\ + (-)2\sum_{i>j} \mathrm{Im}\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin\left(\Delta m^2_{ij}\frac{L}{2E}\right) , \end{aligned} \tag{2.9}$$

where the negative sign (-) in front of the imaginary component is for the case of antineutrino oscillations, $P\left(\bar{v}_\alpha \to \bar{v}_\beta\right)$, and $\delta_{\alpha\beta}$ is the Kronecker delta.

Multiple features of neutrino oscillations can be studied in the two flavour approximation which has a simpler analytical form. By reducing the unitary matrix U to two-dimensions, assuming that one mass state is degenerate with the other (essentially a two flavour oscillation), U can be written as follows:

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} . \tag{2.10}$$

The oscillation probability then becomes:

$$P\left(\nu_\alpha \rightarrow \nu_\beta\right) = \delta_{\alpha\beta} - (+)\sin^2(2\theta)\sin^2\left(\frac{1.267\Delta m^2\left[\mathrm{eV}^2\right] L[\mathrm{km}]}{E[\mathrm{GeV}]}\right). \tag{2.11}$$

A disappearance oscillation experiment is described when $\alpha = \beta$ with the probability picking up a negative sign. When $\alpha \neq \beta$, an appearance oscillation experiment is described by the positive sign in Equation 2.11. The sinusoidal nature of the oscillation probability can be seen in Figure 2.3, where the amplitude is controlled by the mixing angle θ and the period is controlled by parameter Δm^2. For a mixing angle that is fixed, the maximum probability occurs at $L/E \sim 1.25/\Delta m^2$ and for $\Delta m^2 \sim 2.5 \times 10^{-3}\mathrm{eV}^2$ (roughly the mass difference between the lowest and highest neutrino mass state). Therefore, the T2K experiment, with an E = 0.6 GeV and an L = 295 km, allows for neutrinos to be studied near the maximum oscillation point [47].

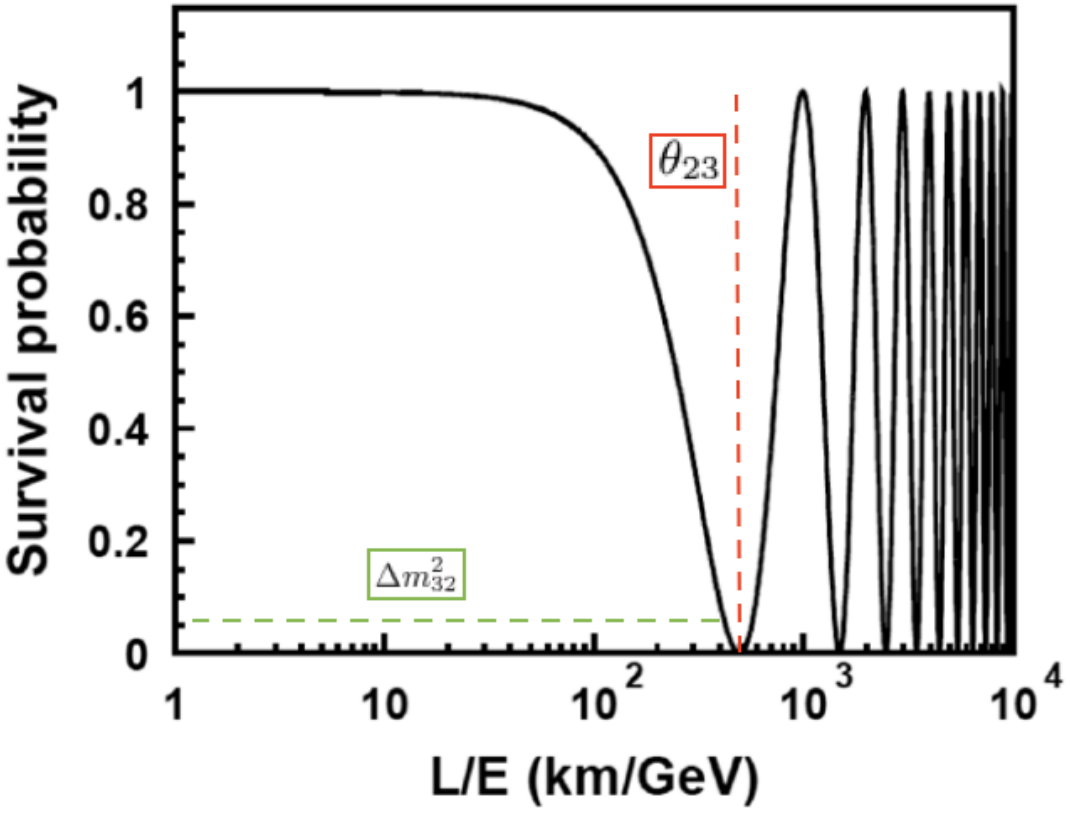


FIGURE 2.3: Survival probability of muon neutrinos as a function of L/E where E is energy measured in GeV, and distance is L measured in km. The neutrino mass state (heavier) is taken as 10^{-7} of the mass of the electron. θ_{23} is given by the magnitude of the amplitude (shown in red above). Δm^2_{32} is given by the distance in phase space in L/E at the survival probability of 1 until the first minimum (shown in green above). Figure edited from [48].

2.3 The PMNS Matrix and Current Parameter Values

The most recent values (November 2022 [49]) for the PMNS Matrix and PMNS parameters are shown below with and without Super-K atmospheric data:

$$|U|_{3\sigma}^{\text{w/o SK-atm}} = \begin{pmatrix} 0.803 \to 0.845 & 0.514 \to 0.578 & 0.142 \to 0.155 \\ 0.233 \to 0.505 & 0.460 \to 0.693 & 0.630 \to 0.779 \\ 0.262 \to 0.525 & 0.473 \to 0.702 & 0.610 \to 0.762 \end{pmatrix} \tag{2.12}$$

$$|U|_{3\sigma}^{\text{with SK-atm}} = \begin{pmatrix} 0.803 \to 0.845 & 0.514 \to 0.578 & 0.143 \to 0.155 \\ 0.244 \to 0.498 & 0.502 \to 0.693 & 0.632 \to 0.768 \\ 0.272 \to 0.517 & 0.473 \to 0.672 & 0.623 \to 0.761 \end{pmatrix} \tag{2.13}$$

W/o SK-atm	Normal Ordering (best fit)		Inverted Ordering $(\Delta\chi^2 = 2.3)$	
	bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
$\sin^2\theta_{12}$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$
$\theta_{12}/^\circ$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$
$\sin^2\theta_{23}$	$0.572^{+0.018}_{-0.023}$	$0.406 \to 0.620$	$0.578^{+0.016}_{-0.021}$	$0.412 \to 0.623$
$\theta_{23}/^\circ$	$49.1^{+1.0}_{-1.3}$	$39.6 \to 51.9$	$49.5^{+0.9}_{-1.2}$	$39.9 \to 52.1$
$\sin^2\theta_{13}$	$0.02203^{+0.00056}_{-0.00059}$	$0.02029 \to 0.02391$	$0.02219^{+0.0000}_{-0.00057}$	$0.02047 \to 0.02396$
$\theta_{13}/^\circ$	$8.54^{+0.11}_{-0.12}$	$8.19 \to 8.89$	$8.57^{+0.12}_{-0.11}$	$8.23 \to 8.90$
$\delta_{\text{CP}}/^\circ$	197^{+42}_{-25}	$108 \to 404$	286^{+27}_{-32}	$192 \to 360$
$\frac{\Delta m^2_{21}}{10^{-5}\text{eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$
$\frac{\Delta m^2_{3\ell}}{10^{-3}\text{eV}^2}$	$+2.511^{+0.028}_{-0.027}$	$+2.428 \to +2.597$	$-2.498^{+0.032}_{-0.025}$	$-2.581 \to -2.408$

With SK-atm	Normal Ordering (best fit)		Inverted Ordering $(\Delta\chi^2 = 6.4)$	
	bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
$\sin^2\theta_{12}$	$0.303^{+0.012}_{-0.012}$	$0.270 \to 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$
$\theta_{12}/^\circ$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$
$\sin^2\theta_{23}$	$0.451^{+0.019}_{-0.016}$	$0.408 \to 0.603$	$0.569^{+0.016}_{-0.021}$	$0.412 \to 0.613$
$\theta_{23}/^\circ$	$42.2^{+1.1}_{-0.9}$	$39.7 \to 51.0$	$49.0^{+1.0}_{-1.2}$	$39.9 \to 51.5$
$\sin^2\theta_{13}$	$0.02225^{+0.00056}_{-0.00059}$	$0.02052 \to 0.02398$	$0.02223^{+0.0058}_{-0.0058}$	$0.02048 \to 0.02416$
$\theta_{13}/^\circ$	$8.58^{+0.11}_{-0.11}$	$8.23 \to 8.91$	$8.57^{+0.11}_{-0.11}$	$8.23 \to 8.94$
$\delta_{\text{CP}}/^\circ$	232^{+36}_{-26}	$144 \to 350$	276^{+22}_{-29}	$194 \to 344$
$\frac{\Delta m^2_{21}}{10^{-5}\text{eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$
$\frac{\Delta m^2_{3\ell}}{10^{-3}\text{eV}^2}$	$+2.507^{+0.026}_{-0.027}$	$+2.427 \to +2.590$	$-2.486^{+0.025}_{-0.028}$	$-2.570 \to -2.406$

Two different sets of parameter results, with and without Super-K atmospheric data are given. This is because the Super-K collaboration has published a χ^2 map of their atmospheric neutrino analysis. Including this atmospheric data significantly impacts the best fit point (bfp) and width of some parameters (such as δ_{CP}). Because this map includes extra data and is hard to reproduce outside of the collaboration, it is included for one set of parameter results (Equation 2.13). This set is not published standalone, due to the difficulty of reproducing the map independently to verify it, and thus another set of results are given without the χ^2 map (Equation 2.12) [50].

2.4 Neutrino Interactions at T2K

There are two main categories of neutrino interaction within the Standard Model: CC (Charged Current) and NC (Neutral Current) interactions. Charged current interactions are unique as they yield a charged lepton that can be directly measured, with its flavour corresponding to the flavour of the interacting neutrino (Figure 2.4, left and centre). Neutral current interactions leave neutrinos in their same flavour (Figure 2.4, right); as the final state is a neutrino, the flavour of the initial interacting neutrinos cannot be deduced.

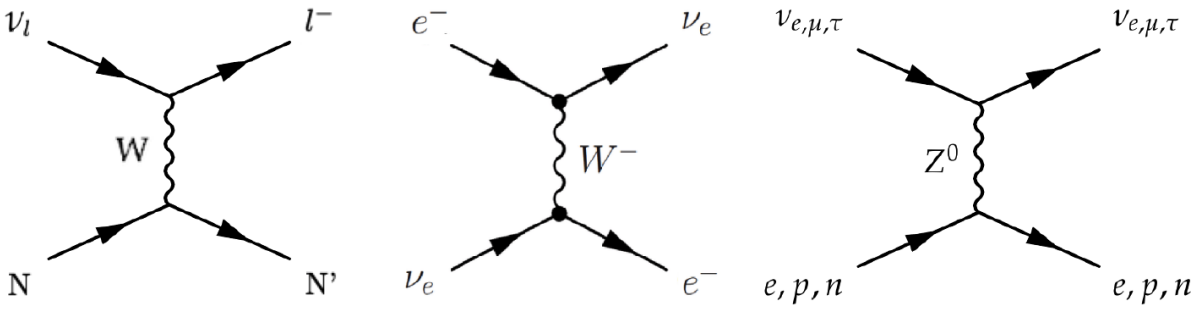


FIGURE 2.4: Feynman diagrams of Charged Current (CC) interactions, on the left a neutrino-nucleus scatter (N, N') is shown, in the centre a neutrino-electron scatter is shown. The Neutral Current Interaction diagram is displayed on the right [51].

Neutrino-nucleon scattering is studied at T2K due to its higher interaction probability and ease of detection. There are three main charged current (CC) interaction modes for neutrino-nucleus scattering at energies relevant to T2K. These are: Quasi-elastic Scattering (CCQE), Resonance Production (CCRES), and Deep Inelastic Scattering (CCDIS). At low energies (<1 GeV) the CCQE interaction dominates, at intermediate energies (~ 1 GeV - 6 GeV) the CCRES interaction dominates, and at high energies the CCDIS interactions are most dominant. The intermediate and high energy regions are not accurately modelled to date, however these interactions are minimised at T2K due to the peak in the neutrino energy spectrum aligning with the CCQE cross section peak. These are relevant at different energies as shown below in Figure 2.5.

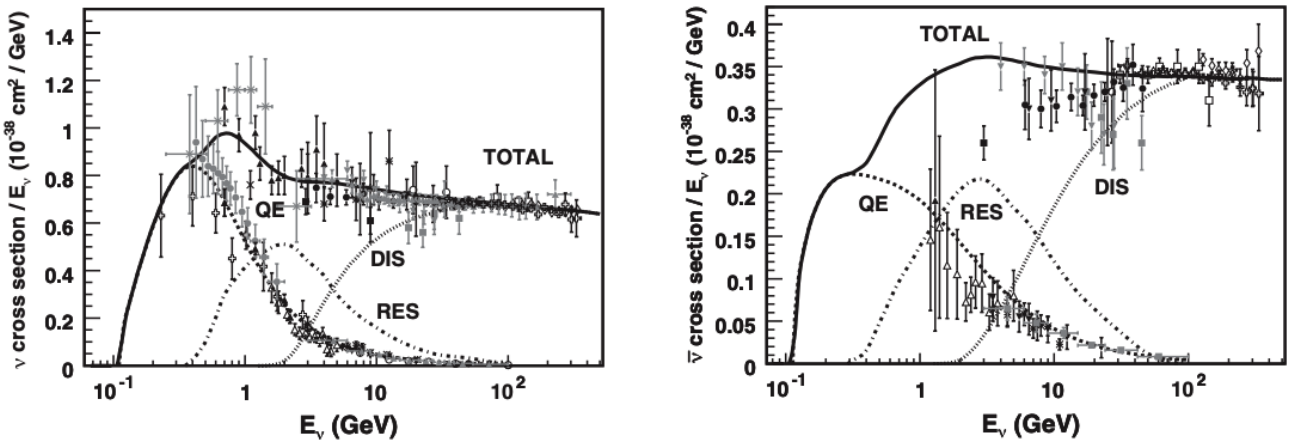


FIGURE 2.5: The graphs show the cross section as a function of neutrino energy for the three main charged current interactions at T2K. These are labelled: Quasi-elastic Scattering (QE), Resonance Production (RES), and Deep Inelastic Scattering (DIS).
On the left: the muon neutrino cross section. On the right: the muon antineutrino cross section. At peak T2K energy (∼0.6 GeV) the CCQE interactions are the most dominant [52].

2.4.1 CCQE

Quasi-elastic scattering is the most dominant interaction at the neutrino energies used at T2K (this is due to the neutrino energy spectrum peaking at ~ 0.6 GeV, as seen in Figure 2.5). The name quasi-elastic is given due to the elastic nature of the scattering between the two incoming particles and the outgoing particles. This means that it can be modelled relativistically quite easily. The neutrino energy can be reconstructed completely with just the outgoing lepton kinematics with the following equation:

$$E_{rec} = \frac{m_p^2 - (m_n - E_b)^2 + m_\mu^2 + 2\,(m_n - E_b)\,E_\mu}{2\,(m_n - E_b - E_\mu + p_\mu \cos\theta_\mu)}\,. \tag{2.14}$$

In Equation 2.14 above, m_p is the mass of the proton and m_n is the mass of the neutron. The mass of the muon is m_μ and the neutron binding energy is E_b. The energy of the muon is E_μ and the angle between the muon and the incoming neutrino is θ_μ with p_μ being the muon momentum.

T2K uses the Benhar Spectral Function (SF) model for modelling the energy and momentum distribution of nucleons within the nucleus. This serves as an improvement over the square distribution of the Fermi-gas (FG) models used previously, giving a more realistic model of the nucleus for CCQE interactions. The spectral function contains a better description of removal energy and Fermi motion, leading to a better modelling of the nuclear ground state. Furthermore, electron scattering data has been used to derive a new shell model with outgoing nucleon kinematics better predicted. The neutrino nucleus cross section can be written as [53]:

$$\frac{d^5\sigma}{dE_\ell d\Omega_\ell d^3p'} = K\sigma_{\ell n} S(E, \mathbf{k}). \tag{2.15}$$

In Equation 2.15 above, $\sigma_{\ell n}$ is the scattering cross section for a single nucleon (which contains the M_A^{QE} term), K is a kinematic factor and the final term, $S(E, \mathbf{k})$, is the spectral function as a function of removal energy (E) and nucleon momentum ($\mathbf{k}$). E_l is the final state lepton energy, Ω_l is the solid angle of the lepton emission, and p' is the momentum of the final state nucleon. The spectral function used in T2K is made up of a mean field term for single nucleons with an additional term describing the correlation of the nucleons within the nuclear medium. The spectral function comes from an Impulse Approximation which looks at the basic interaction of an incoming lepton interacting with a single nucleon that exits the nuclear medium with no final state interactions. The nucleon momentum and removal energy is integrated over, with the spectral function giving the appropriate weight. Therefore, the spectral function serves as a 2D distribution which gives the probability of a nucleon being found with a certain momentum alongside the energy required to remove this nucleon from the potential. There are two spectral functions, one for oxygen and another for carbon, which are implemented according to the target nuclei in the T2K detectors.

2.4.2 CCRES

Alongside the CCQE interaction, the CCRES interaction plays a significant role in the 1 - 10 GeV region (known as the intermediate energy region) where resonant interactions transition to deep inelastic scattering. CCRES dominates from $\sim$2 - 5 GeV. As a charged current interaction, a charged lepton ends up in the final state, however the neutrino-nucleon interaction also yields a baryon; a delta resonance that quickly decays into a pion and another nucleon. This is known as single pion production (SPP) as shown in Figure 2.6 below.

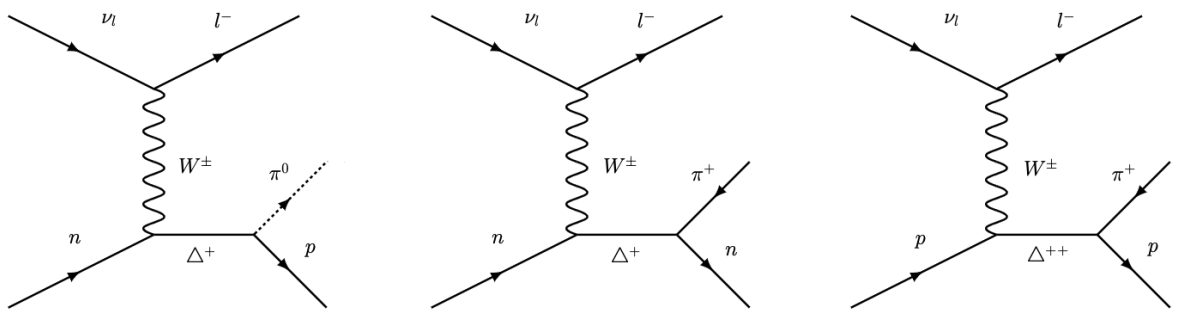


FIGURE 2.6: Feynman diagrams for Single Pion Production. The interactions displayed are:
CC1π^0 (left) - The delta resonance (Δ^+) decaying to a π^0 and a proton,
CC1π^+n (centre) - The delta resonance (Δ^+) decaying to a π^+ and a neutron,
CC1π^+1p (right) - The delta resonance (Δ^{++}) decaying to a π^+ and a proton.

Resonant production interactions are modelled at T2K using the Rein-Sehgal model that describes resonant pion production in charged and neutral current interactions. There are 18 different baryon resonances which are accounted for in the model. The most dominant resonance is the $\Delta(1232)$ baryon resonance, an excited state at 1232 MeV that decays to give a pion (shown in Figure 2.6). A further description of the model is given in Chapter 6, Section 6.7.4. At T2K energies, the $\Delta(1232)$ is the most dominant of the resonances, with the higher resonances at 1440 MeV (the following resonance) playing a smaller role. This regime is generally understood but poorly modelled in comparison to the CCQE interaction. The exact contributions and interference between higher resonances at higher energies are also not well understood [54].

Currently the lepton momentum and angle are measured in order to reconstruct the neutrino energy (Equation 2.14). However, the reconstruction depends on how well the interaction is modelled. If a pion is produced in the interaction then it will be vital to have information on its kinematics as the formula of E_{rec} would also change. Furthermore, some CCRES interactions can be mis-identified for CCQE interactions within the detector due to Final State Interactions (mentioned below) forming an indistinguishable background to CCQE interactions. This can lead to neutrino energy mis-reconstruction. My thesis work is focused on developing the near detector fit, such that T2K can incorporate pion kinematics in the future to help constrain systematics, with the ultimate goal to decrease the error on the parameter δ_{CP}.

2.4.3 CCDIS

In the high energy region (> 10 GeV), CCDIS interactions become the most dominant interaction. High energy neutrinos interact with the quarks of the nucleons directly leading to inelastic scattering, making it one of the hardest interactions to model accurately. The Feynman diagram for this interaction is shown below in Figure 2.7:

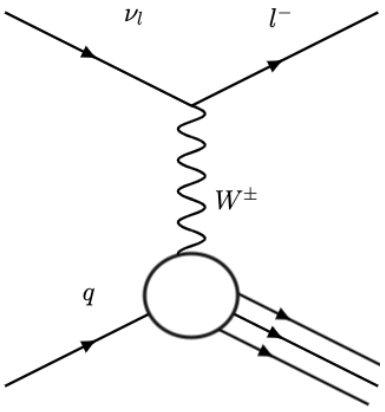


FIGURE 2.7: Feynman diagram for Deep Inelastic Scattering (CCDIS). The diagram shows the high energy interaction of the neutrino with the quarks of the nucleon, leading to multiple nucleons in the final state.

CCDIS interactions dominate neutrino interaction cross sections at high energy, producing a multiplicity of pions. Attempts to model this interaction accurately must resort to methods in hadronic physics, such as PDFs (Parton Distribution Functions) used by proton colliders. CCDIS cross sections are calculated using GRV98 PDFs [55] which describe the probability of finding a quark with a certain fraction of momentum from the nucleon. CCDIS interactions are modelled in T2K using a custom Multi-Pion model with corrections from Bodek and Yang [56] which allow the GRV98 PDFs to be used in the low Q^2 region [57].

2.4.4 Nuclear Effects

When considering the context of neutrino-nucleus interactions, it is understood that a nucleon is bounded within a nucleus. In the T2K experiment, these interactions mainly take place in carbon within plastic scintillator or in oxygen within water from Cherenkov detectors. As interactions occur in nuclear media, the particle produced from a neutrino-nucleon interaction can undergo several interactions in the nucleus before it is detected. At T2K, the most important nuclear effects are 2p2h and Final State Interactions (FSI) which are mentioned below.

2.4.4.1 2p2h

The 2p2h (two-particle-two-hole) interaction consists of two nucleons instead of one being emitted in the final state of a neutrino-nucleon interaction. As seen in Figure 2.8, the weak current boson exchanges to two nucleons and a charged lepton [58]. The signature of 2p2h interactions are similar to the signature of CCQE interactions (which are a subset of 1p1h, one-particle-one-hole, interactions) as no pions are produced in the final state.

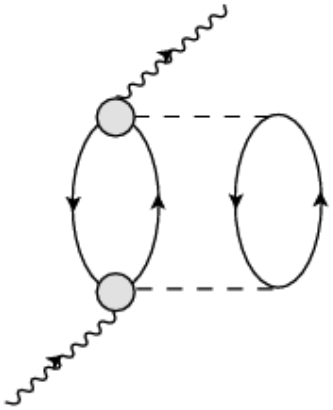


FIGURE 2.8: 2p2h interactions are also known as MEC (Meson Exchange Current) interactions due to the fact that the pair of nucleons (solid lines) connected by a meson propagator (dashed lines) are excited by a boson (squiggly line) leaving a pair of holes in the Fermi sea of nucleons within the nucleus [59].

In T2K, the Nieves et al. model and the Martini et al. model is used to model 2p2h interactions. A further description of the models and their parameters is given in Chapter 6, Section 6.7.4.

2.4.4.2 Final State Interactions (FSI)

Nucleons and pions produced in the nucleus after a neutrino-nucleon interaction can undergo additional interactions prior to emerging from the nucleus, or they can be absorbed and not emerge from the nucleus. Pions and nucleons can be absorbed, elastically scattered, inelastically scattered if energetic enough (producing new pions), or they can exchange charge with a nucleon (Figure 2.9). Pions within the nuclear medium are easily affected by FSI and so the structure within the nuclear medium can easily impact what is detected outside the nucleus.

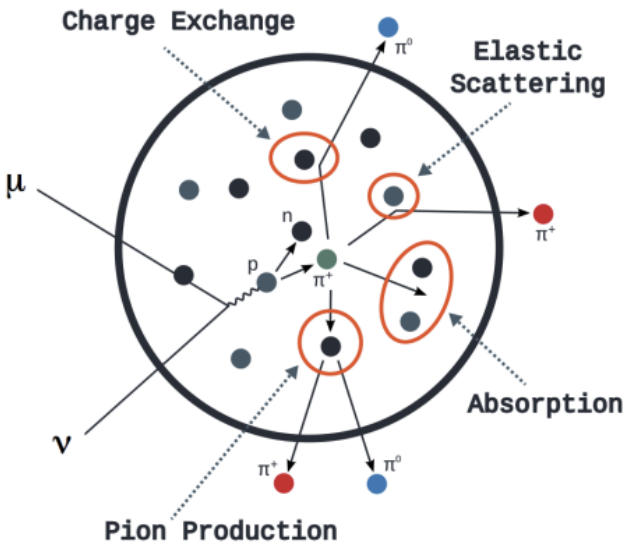


FIGURE 2.9: Diagram illustrating possible final state interactions for a pion (these interactions are similar for a nucleon) [60].

FSI makes it harder to separate interactions when classifying events based on pion multiplicity in sample topologies at T2K. For example, an event may actually be a CCRES interaction event, however due to the final state pion being absorbed by the nucleus due to FSI, it is misidentified as a CCQE interaction event. Therefore it is vital to model this effect reliably and accurately.

Pion FSI is modelled in T2K using a semi-classical cascade model tuned to pion-nuclei scattering data [61]. Nucleon FSI is modelled in an analogous way to pion FSI. A further description of the FSI model and its parameters is given in Chapter 6, Section 6.7.4.

2.5 Open Questions

A number of open questions in the field of neutrino physics still currently remain. The T2K experiment is at the forefront of answering some of these important questions.

2.5.1 Direct Neutrino Mass Measurements

The neutrino mass squared difference Δm^2_{21} is known to a 2.7–2.8% uncertainty (Section 2.3). However, the exact values of the three masses (m_1, m_2, m_3) are not known. Knowing the values of these masses will allow us to understand if there is a hierarchy of neutrino masses, or if these masses are quasi-degenerate. The KATRIN experiment aims to measure neutrino mass in a model independent way. It does this by looking at the spectrum of electrons from tritium decay: $^3\mathrm{H} \rightarrow {}^3\mathrm{He} + \mathrm{e}^- + \bar{\nu}_\mathrm{e}$, in order to put values on the mass of the electron antineutrino directly [62].

2.5.2 Dirac or Majorana

The nature of neutrino masses is also unknown. It is not known whether neutrinos have Majorana masses or Dirac masses (like other fermions). A Dirac particle is a particle which has a distinct antiparticle, a Majorana particle is one which is itself the particle and the antiparticle. The open question is: are neutrinos Dirac or Majorana particles?

If the neutrino is a Majorana particle, the small mass of the neutrino in relation to the larger masses of leptons and quarks can be understood naturally through a mechanism pointing towards a Grand Unified Theory [63]. Furthermore, this mechanism holding true for neutrinos would necessitate the violation of lepton number conservation. Experiments trying to answer this question are looking for neutrinoless double beta decay. This is a forbidden decay in the Standard model, however if observed it would lead to new frontiers in neutrino physics. The SNO+ experiment, alongside SuperNEMO, LEGEND, and others are looking at detecting this hypothesised forbidden interaction in order to answer this question [64].

2.5.3 Neutrino Mass Hierarchy

It is understood that there are three flavour states and three mass states. However, the hierarchy, which is the ordering of the three mass states, is not yet known. The sign of Δm^2_{21} is known through solar neutrino oscillations which are correctly modelled by matter effects in the solar core [65].

It is also known that m_2^2 is greater than m_1^2, but the value of m_3^2 is unknown. It is not known, compared to m_1^2 and m_2^2, if m_3^2 lighter than both (known as the inverted hierarchy) or heavier (known as the normal hierarchy). Both orderings are shown in Figure 2.10 below – claims can be made for both normal and inverted [66].

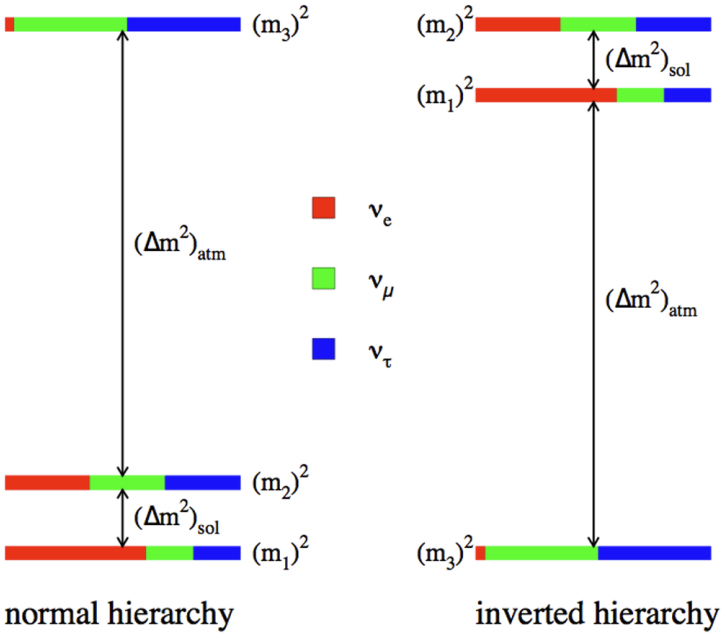


FIGURE 2.10: The neutrino mass hierarchy shown as inverted or normal with the corresponding flavour state contributions [67].

2.5.4 Sterile Neutrinos

An existence of a fourth type of neutrino is a possibility. A sterile neutrino is a neutrino which interacts via gravity and not the weak force. It was suggested to help understand the results of short-baseline experiments such as LSND and MicroBooNE since the 1990s [68]. Neutrinos seemed to be oscillating with a much shorter wavelength than expected; it was theorised that the sterile neutrino could account for this. It is even possible that sterile neutrinos could account for the Dark Matter inferred in the universe. The short-baseline neutrino program (SNB) at Fermilab will help to put constraints on sterile neutrinos [69].

2.5.5 Determination of the θ_{23} Octant

The parameter θ_{23} as mentioned in Section 2.2 determines the extent of neutrino mixing. The mixing angle θ_{23} is said to be in the upper octant if its value is greater than 45 degrees, and in the lower octant if less than 45 degrees. T2K and NOvA are trying to resolve the θ_{23} octant

through ν_μ disappearance experiments (sensitive to $\sin^2(2\theta_{23})$) and ν_e appearance experiments (sensitive to $\sin(\theta_{23})$). The value of θ_{23} being exactly equal to 45 degrees is known as maximal mixing, this results in the third neutrino eigenstate containing equal parts of the tau and muon neutrino flavour states. If true, maximal mixing implies a symmetry which indicates new physics in the form of a conserved current that has yet to be discovered [70].

2.5.6 CP Violation

One key motivation behind neutrino studies is that neutrinos may contain information related to one of the biggest puzzles in physics: why is there more matter than antimatter in the universe? In the Standard Model, the relationship between a particle and its corresponding antiparticle is referred to as CP symmetry. If this symmetry is violated then it could explain why the universe is matter dominated.

CP symmetry violation has been confirmed for the mixing of quarks, however the effect is too small and not substantial enough to describe the matter-antimatter asymmetry that is observed in the universe. The neutrino sector offers a solution to this mysterious question, specifically in the PMNS theory of neutrino oscillations which allows for CP violation through one of its parameters δ_{CP}. The δ_{CP} parameter is studied by looking at the oscillation rate of neutrinos and comparing this to the oscillation rate of antineutrinos. T2K and NOvA are currently collaborating to answer this question through their joint analysis currently in progress.

Chapter 3

The T2K Neutrino Experiment

This chapter describes the neutrino source and detectors used in the T2K (Tokai to Kamioka) experiment in Japan. The T2K experiment is a long-baseline neutrino oscillation experiment that consists of a neutrino source beamline (from J-PARC), Near Detectors (INGRID and ND280) and a far water Cherenkov detector (Super-Kamiokande) [71]. Figure 3.1 shows a schematic of the experiment.

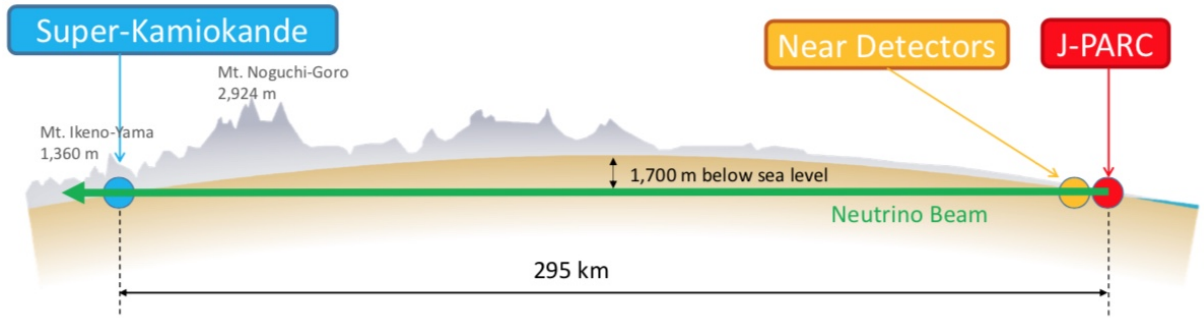


FIGURE 3.1: Schematic of the T2K experiment (T2K Internal).

The neutrino extraction facilities at J-PARC in Tokai were built especially for the T2K Experiment [72]. An intense beam of muon neutrinos is produced at J-PARC. These un-oscillated neutrinos continue towards the near detector (ND280) 280 meters away, where neutrino interaction rates are measured. The neutrinos then travel 295 km roughly west from Tokai to Kamioka to the far detector, Super-K; a water Cherenkov detector containing 50,000 tons of pure water. This detector measures the neutrino interaction rates in the oscillated beam. The neutrino rates between the two detectors are measured and this is used to determine the PMNS parameters that describe neutrino oscillations.

The T2K experiment is sensitive to the PMNS matrix parameters: θ_{13}, θ_{23} and δ_{CP}. θ_{13} is determined through ν_e appearance in the ν_μ beam, alongside $\bar{\nu}_e$ appearance in $\bar{\nu}_\mu$ beam. θ_{23} is determined via muon (anti)neutrino disappearance. The parameter containing information about CP violation, δ_{CP}, is determined by comparing neutrino and antineutrino oscillation rates in the ν_μ to ν_e and $\bar{\nu}_\mu$ to $\bar{\nu}_e$ channels.

3.1 The Japan Proton Accelerator Research Complex J-PARC (Beamline)

J-PARC is an accelerator complex located in Tokai, Japan. The J-PARC beamline consists of a linear accelerator (LINAC) and two synchrotron accelerators; a rapid cycling synchrotron (RCS) and a main ring synchrotron (MR). A H^- beam is accelerated by the linear accelerator up to energies of 400 MeV. It is then fed into the RCS where it changes to a H^+ beam as a result of charge stripping foils at injection. The RCS accelerates the beam up to 3 GeV and cycles at 25 Hz with 2 bunches per cycle, the MR receives 5% of these bunches that are accelerated up to 30 GeV [71].

To produce a proton beam, protons are extracted from the MR using 5 kicker magnets within a single turn. Each spill, or single extraction, contains 8 proton bunches of less than 100 ns each. The bunches are separated by $\sim$ 560 ns and have a total duration of $\sim 5\mu$s. This timing is extremely important as it allows T2K to distinguish non-beam backgrounds such as high energy cosmic-ray muons from the beam signal. A single spill with 8 proton bunches is directed towards the T2K neutrino beamline which is comprised of two parts; a primary and a secondary beamline (Figure 3.2).

The proton beam travels in the primary beamline which orients it towards Kamioka and directs it to collide with a fixed graphite target. The target is a graphite rod 2.6 cm in diameter, 91.4 cm long with a density of 1.8 g/cm^3 [73]. The secondary beamline contains the target station housing the graphite target, the three magnetic horns, and a $\sim$100 m decay volume ending in a beam dump consisting of a 75 ton graphite core.

The beam impinging on the graphite target results in a shower of pions and kaons (Figure 3.3). These charged hadrons continue on through a decay volume and decay to neutrinos from parent mesons. T2K uses a set of 3 magnetic horns in order to focus the beam. These horns focus positively charged pions and kaons in neutrino mode, or negatively charged pions and kaons in antineutrino mode. Positively charged kaons and pions decay to muon neutrinos while negatively charged kaons and pions decay to muon anti-neutrinos.

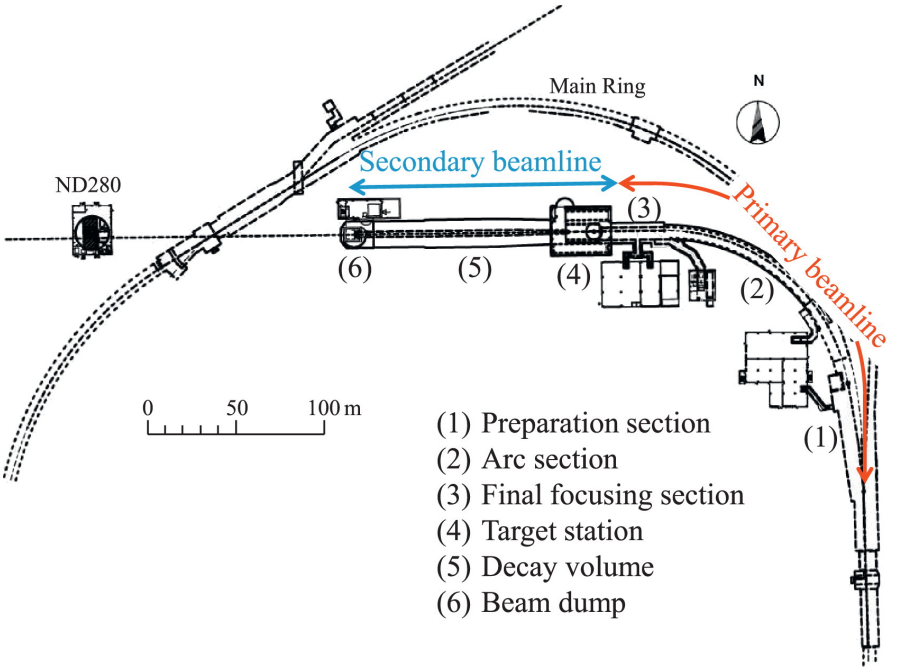


FIGURE 3.2: Picture of the J-PARC neutrino beamline [71].

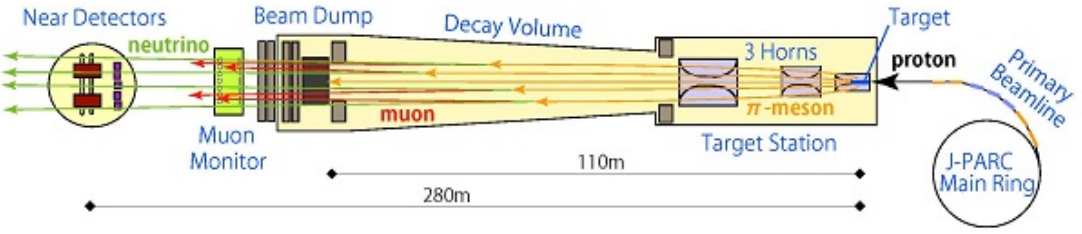


FIGURE 3.3: In-depth look at the secondary beamline [74].

The shower of pions and kaons travels along the decay volume where positive pions and kaons decay into antimuons and muon neutrinos. Some antimuons may then decay into positrons and electron neutrinos. Magnetic horns are used to focus or deflect different-sign mesons. This is in order to enhance neutrinos and reduce anti-neutrinos or vice-versa. Taking data in the neutrino dominated mode is referred to as operating in FHC (Forward Horn Current) mode, and taking data in the anti-neutrino dominated mode is called operating in RHC (Reverse Horn Current) mode. These neutrinos then continue, with the remaining charged particles hitting the beam dump. Many of the mesons are stopped, with high momentum leptons passing through the beam dump which sits at the end of the decay volume. These resultant leptons are monitored by the Muon Monitor so that the neutrino beam direction and intensity can be used to inform the beam simulation group [75].

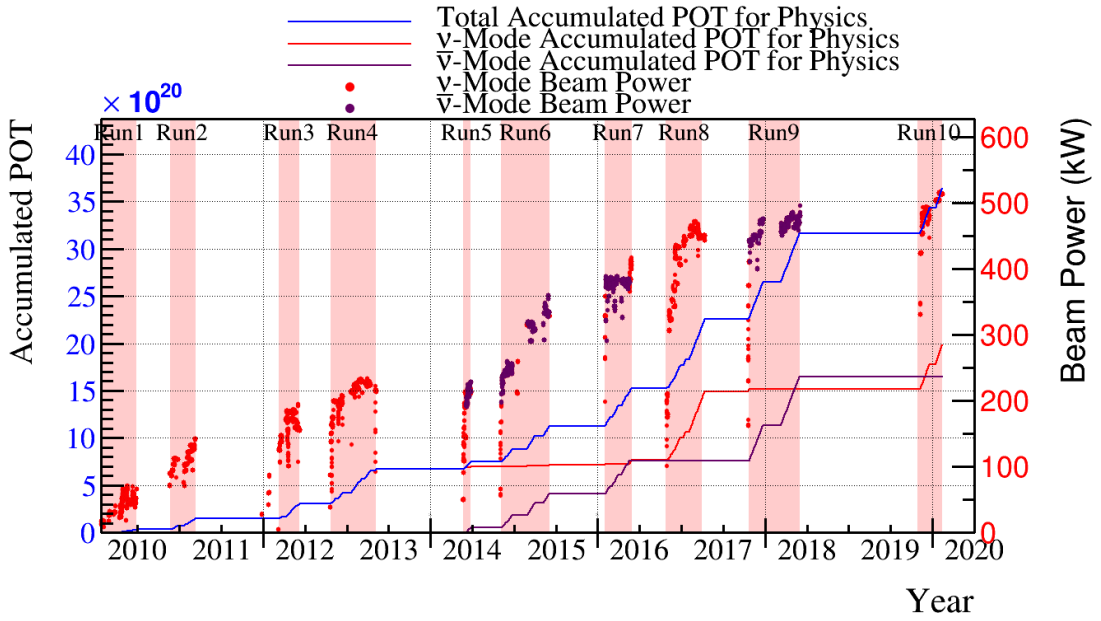


FIGURE 3.4: Total T2K Integrated POT over Run 1–10 (T2K Internal).

It is not possible to precisely measure the number of neutrinos produced from this interaction, therefore the number of protons that have impinged upon the target, called the POT (protons on target), is used as a measure of the provided data statistics. T2K has delivered 3.64×10^{21} POT as of February 2020 (Figure 3.4) [76].

The main meson decays in each operating mode are shown below together with their relevant proportions [47]:

$$\begin{aligned} &\pi^+ \to \mu^+ + \upsilon_\mu \ (99.99\%), \quad K^+ \to \mu^+ + \upsilon_\mu \ (63.6\%) \quad \text{(Foward Horn Current)}, \\ &\pi^- \to \mu^- + \bar{\upsilon}_\mu, \qquad\qquad K^- \to \mu^- + \bar{\upsilon}_\mu \qquad\qquad \text{(Reverse Horn Current)}. \end{aligned}$$

The beam is largely composed of muon (anti)neutrinos, however a small proportion of mesons decay to electron (anti)neutrinos from muon and kaon decay which leaves the beam slightly mixed:

$$\mu^+ \to e^+ + \upsilon_e + \bar{\upsilon}_\mu, \qquad K^+ \to \pi^0 + e^+ + \upsilon_e \,.$$

The resulting neutrinos then travel to the ND280 and Super-K detectors in an off-axis direction; where the centre of the neutrino beam is offset 2.5 degrees from the far detector (Super-K) to near detector axis. This lowers the neutrino flux that reaches the Super-K detector but provides a more desirable neutrino energy spectrum with a narrower spread of energies which peaks at 0.6 GeV (Figure 3.5). The neutrino oscillation probability is maximised at this energy, improving the sensitivity of T2K in measuring the oscillation parameters.

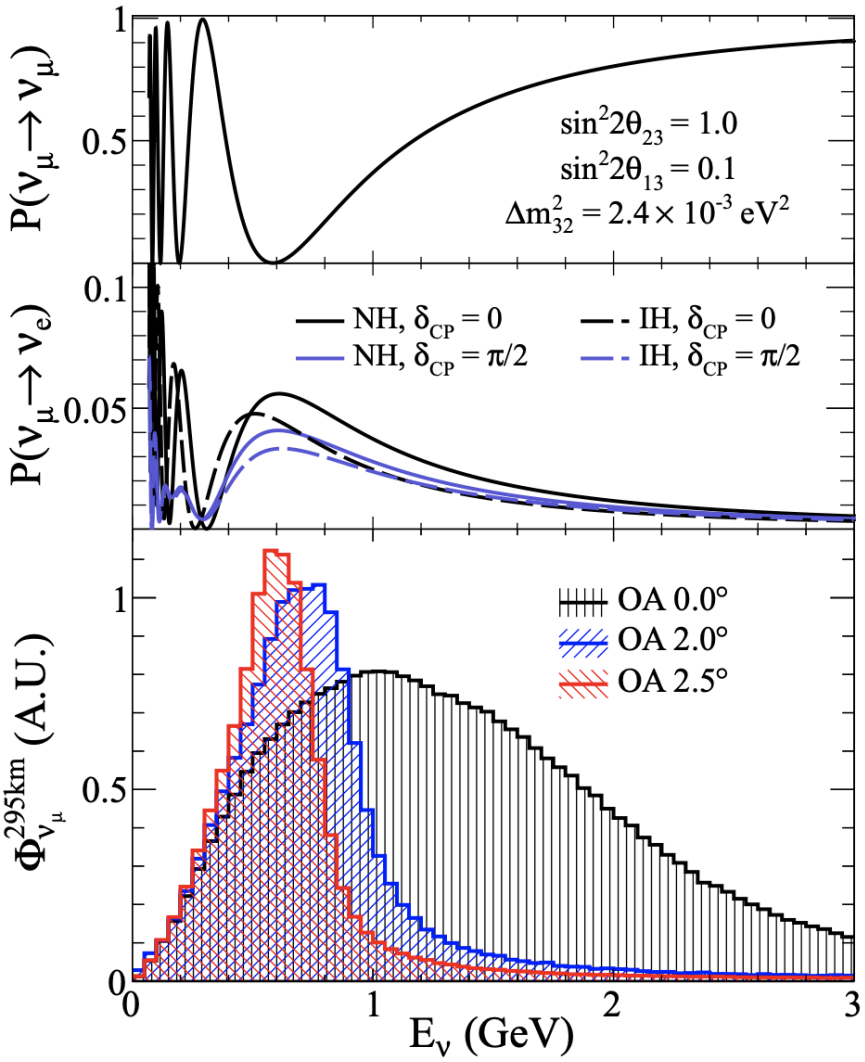


FIGURE 3.5: Un-oscillated ν_μ flux at Super-K at the different off-axis angles (at 0, 2.0 and 2.5°) showing appearance and disappearance probabilities above (T2K Internal).

3.2 The Interactive Neutrino Grid (INGRID)

Alongside the muon monitor, the Interactive Neutrino Grid (INGRID) is the second beam monitor in the T2K experiment which also serves as a target for cross-section measurements given the high flux of neutrinos it receives [71]. As a beam monitor, it measures the neutrino beam intensity and direction. Given the off-axis nature of the experiment, it is paramount that the off-axis angle can be determined precisely alongside the intensity. A beam direction with a 1 mrad uncertainty leads to sizeable 2–3 % uncertainty in the neutrino energy scale [77] which impacts oscillation parameter measurements. To remove these additional uncertainties, INGRID precisely determines the direction of the neutrino beam from J-PARC. The detector sits 280 m downstream from the neutrino beamline in the same pit as ND280, it sits on-axis which is required in order to measure the beam centre [78].

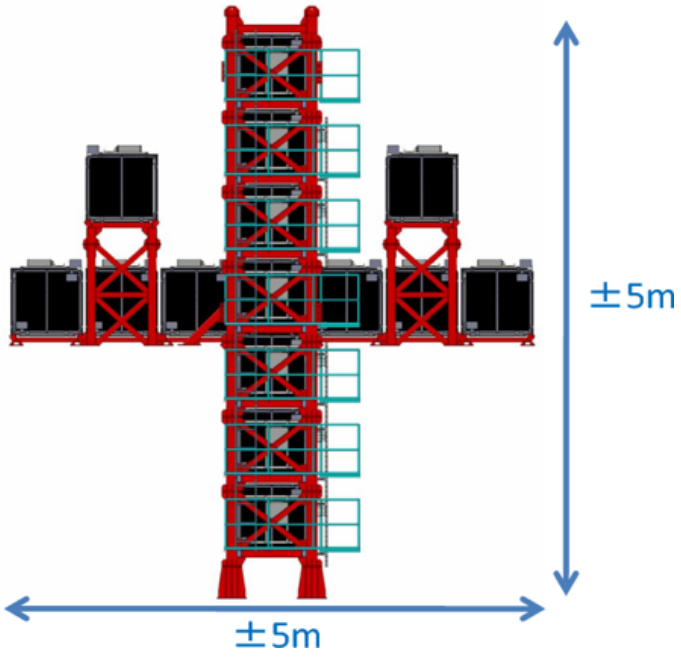


FIGURE 3.6: Schematic of the INGRID detector [71].

INGRID (Figure 3.6) consists of 14 independent modules arranged in a cross-like fashion with the center along the beam axis. Each module is composed of nine iron plates and eleven scintillator trackers; the neutrino beam can be reconstructed by looking at each module and observing lepton tracks from neutrino interactions. INGRID has measured neutrino event rates to within 2 percent of the expected value, with a 0.2 mrad precision on directionality that gives the neutrino beam center within 5 cm. As the neutrino beamline has a spatial width of ∼5 m near INGRID, the detector was constructed to span 10 m by 10 m transverse to the beam [47].

3.3 The Near Detector at 280 m (ND280)

The ND280 detector sits 280 m downstream from the main beamline produced at J-PARC. ND280 is located a short distance from the source in order to measure neutrinos before they have been able to oscillate. It is the off-axis detector for the T2K experiment, its purpose is to accurately reconstruct the neutrino-nucleus interaction that takes place within its center. An off-axis near detector is needed to measure the neutrino flux, energy spectrum and electron neutrino contamination in the direction towards Super-K. This information can then be used to constrain model parameters in the oscillation analysis.

A cross-sectional view of the detector is shown in Figure 3.7 with the UA1 Magnet and Solenoid Coil of the detector displayed. The ND280 detector comprises of various inner detectors. The main central unit, called the 'tracker', consists of two Fine Grain Detectors (FGDs)

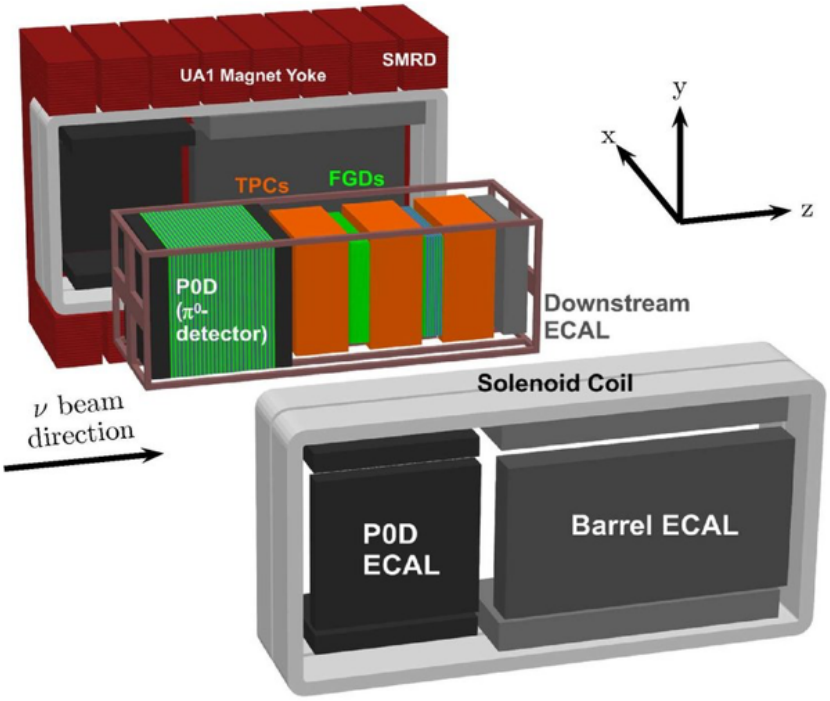


FIGURE 3.7: A cross-sectional view of the ND280 detector - the PØD can be seen, alongside three time-projection chambers (TPCs) which sandwich two fine grained detectors (FGDs) [71].

which are sandwiched between three Time Projection Chambers (TPCs). On either side of the tracker; the PØD (Pi-Zero Detector) is located furthest upstream with the Electromagnetic Calorimeter (ECal) located furthest downstream and a barrel ECal around the tracker made of lead scintillator. The FGDs form a target for the neutrino nucleus interaction, the first FGD is a hydrocarbon (CH) target with organic scintillator. The second FGD contains organic scintillator and water; making it a hydrocarbon water target and introducing an oxygen cross-section model parameter in the oscillation analysis.

The SMRD (Side Muon Range Detector) is also indicated in Figure 3.7. It is composed of scintillation counters with embedded wavelength shifting fibres [79]. The magnetic field strength within the detector is 0.2 T. The entirety of the tracker, ECal and PØD are placed inside this magnetic field, allowing particles to be identified through curvature in their tracks.

3.3.1 ND280 Magnet and SMRD (Side Muon Range Detector)

The coils and magnet yoke in the ND280 which generates the magnetic field was donated by the UA1/NOMAD experiment at CERN [78]. The magnet was repurposed for use in the ND280 with SMRD (Side Muon Range Detector) panel scintillator sandwiched between the steel plates of the magnet flux return yoke. Sixteen 48 mm thick steel plates that are separated by 17 mm gaps comprise each yoke segment, and are held together by long bolts [80]. The magnet generates

a dipole magnetic field of 0.2 T. The combined mass of the magnet and yoke is 850 tons, which makes it an overwhelming contributor to the total ND280 detector mass.

The SMRD is made up of 440 scintillation panels (Figure 3.8) which are embedded within some of the gaps of the magnet yoke. Given that the mass of the magnet is significant, the majority of neutrino interactions take place in the magnetic yoke. The SMRD has three main functions within the detector (1) it serves as a veto for neutrino interactions in the magnet yoke and the rock surrounding it in the pit, (2) it measures the momentum of muons which leave at high angles from ND280, and (3) it acts as a trigger for cosmic rays.

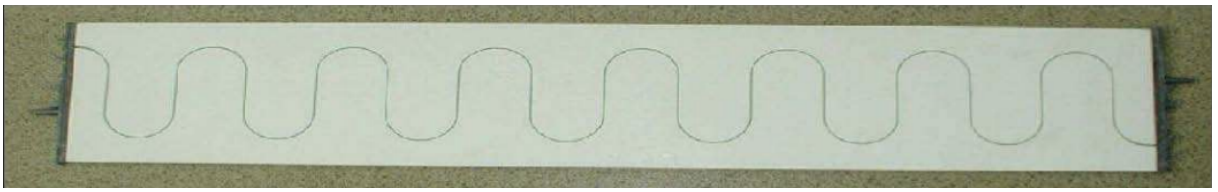


FIGURE 3.8: SMRD Paddle with a snaking WLS fibre shown [80].

3.3.2 Fine-Grained Detectors (FGDs)

The ND280 detector contains two vital Fine-Grained Detectors (FGDs). The measurements taken by these FGDs contribute to understanding the energy spectrum, flux, and the level of electron neutrino contamination in the beam at 2.5 degrees off-axis. Each FGD consists of 1.1 tonnes of target material with dimensions of 186.4 x 186.4 x 2.02 cm per scintillator plane.

The Fine-Grained Detectors (FGDs) are made of dense scintillator bars which, when stacked together, form scintillator planes in an X or a Y configuration as shown in Figure 4.1. These scintillator planes are arranged one after the other in an XY alternating configuration within the detector perpendicular to the beam. When a charged-current interaction occurs within the FGD, scintillated light produced by a charged particle in a dense scintillator bar is collected. This is transmitted to the end of the bar through a WLS (Wave Length Shifting fibre) threaded inside the bar. The WLS fibre is linked to an MPPC (Multi-Pixel Photon Counter) attached at the end for readout [81].

FGD1 is located more up-stream of the two FGDs, closer to the location of beam entry. It is composed of 15 plastic alternating scintillator XY planes, with 2 x 192 bars on each plane; the organic hydrocarbon (CH) scintillator bar provides the target for neutrino-nucleus interactions. FGD2 contains seven plastic scintillator planes sandwiched between six 2.54 cm layers of water, allowing it to serve as the hybrid water-scintillator target. The FGDs are thin enough to allow

muons to pass into the TPCs and for their momenta and charges to be measured in the ∼0.2 T magnetic field provided by the UA1/NOMAD dipole magnet. The architecture of the FGD is shown below in Figure 3.9.

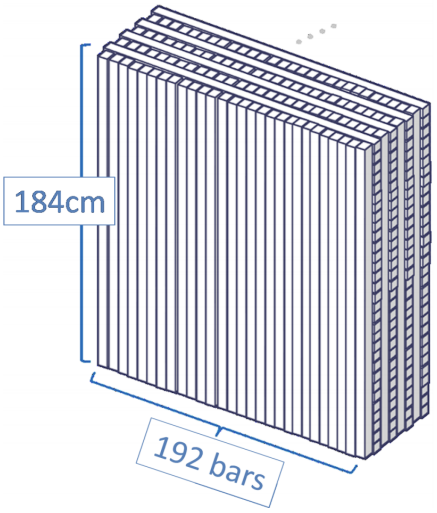


FIGURE 3.9: Architecture of the Fine-Grained Detector [82]. There are 8 scintillator planes displayed here for FGD1. FGD2 is similar but contains water in between planes.

3.3.3 Time Projection Chambers (TPCs)

TPCs are gas based sub-detector modules which allow important particle information such as tracking, momentum, energy loss, and particle identification to be obtained. Momentum measurements are obtained from the curvature of particle tracks in the magnetic field generated by the UA1/NOMAD magnet. A double gas-tight box design is used for the TPC, with an outer box and inner box sitting inside the other. The inner box contains a gas mixture of Ar:CF_4:C_4H_{10} at 95:3:2, which is readily ionised by charged particles. The inner gas mixture is chosen for its high speed, low diffusion, and good performance [83]. The outer box (Figure 3.10) has dimensions of 2302 x 2400 x 974 mm, and the inner box has dimensions of 1808 x 2230 x 854 mm.

Each FGD is surrounded by TPCs (one on each upstream-facing and downstream-facing side) providing good tracking and measurement of forward-going and backward-going tracks. A charged particle entering a TPC ionises the gas and passes through, leaving ionisation electrons along its path. These electrons drift away from the cathode towards to the Micromegas (MM) detector readout panes. Tracks can thus be reconstructed by combining the signal pattern in the readout panes with the arrival timing. This results in a 3D image of the particle trajectory.

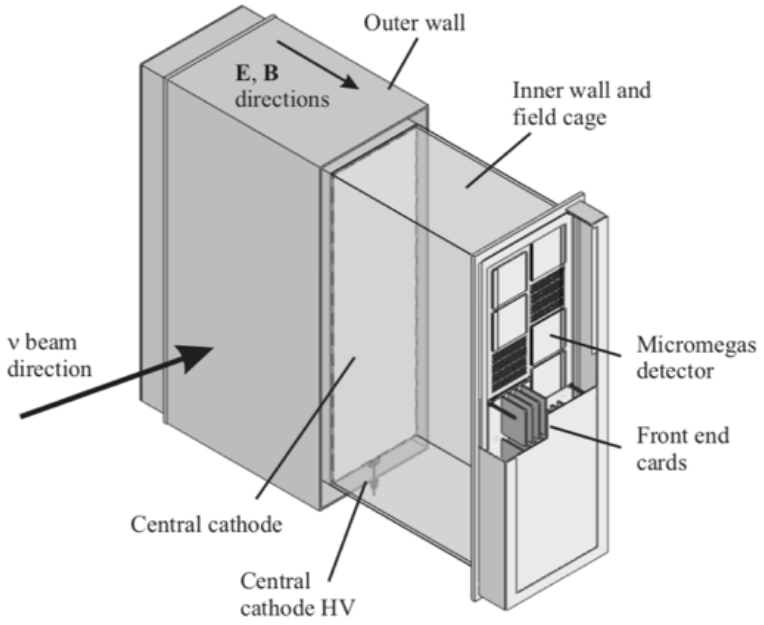


FIGURE 3.10: Diagram showing the main components of a TPC [71].

Looking at Figure 3.11 below, an event with one negative muon candidate is seen in the ND280 event display. In the FGD and the TPC, no charged or neutral pions are seen.

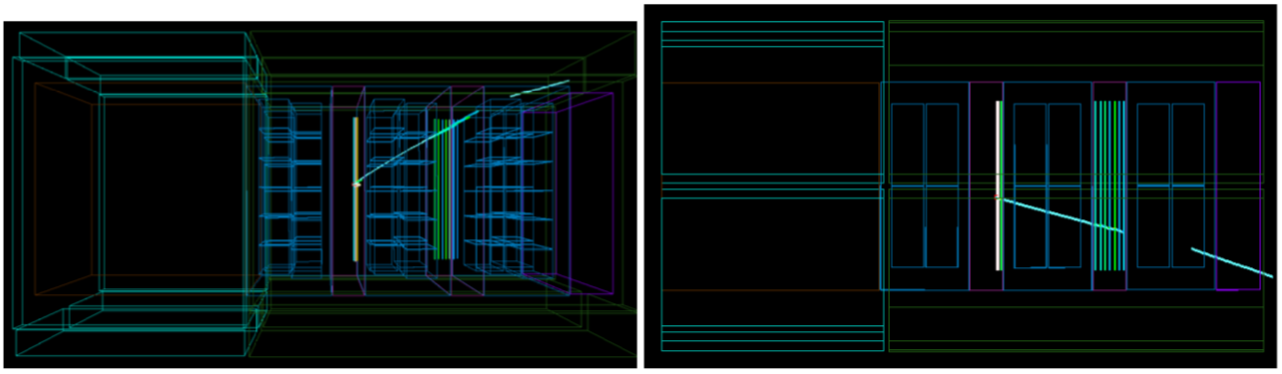


FIGURE 3.11: The ND280 event display showing a charged-current event [83].
Left: Side-view. Right: Top-view.

3.4 Super-Kamiokande

Super-Kamiokande (Super-K) is a large water Cherenkov detector used by T2K as the far detector. It is located in the Mozumi Mine, Kamioka, 295 km west from Tokai where the ND280 is based. The detector cavity is situated 1 km deep under the peak of Mt. Ikeno, with the 2700 m-water-equivalent mean overburden of the mountain reducing cosmic ray muon backgrounds [84].

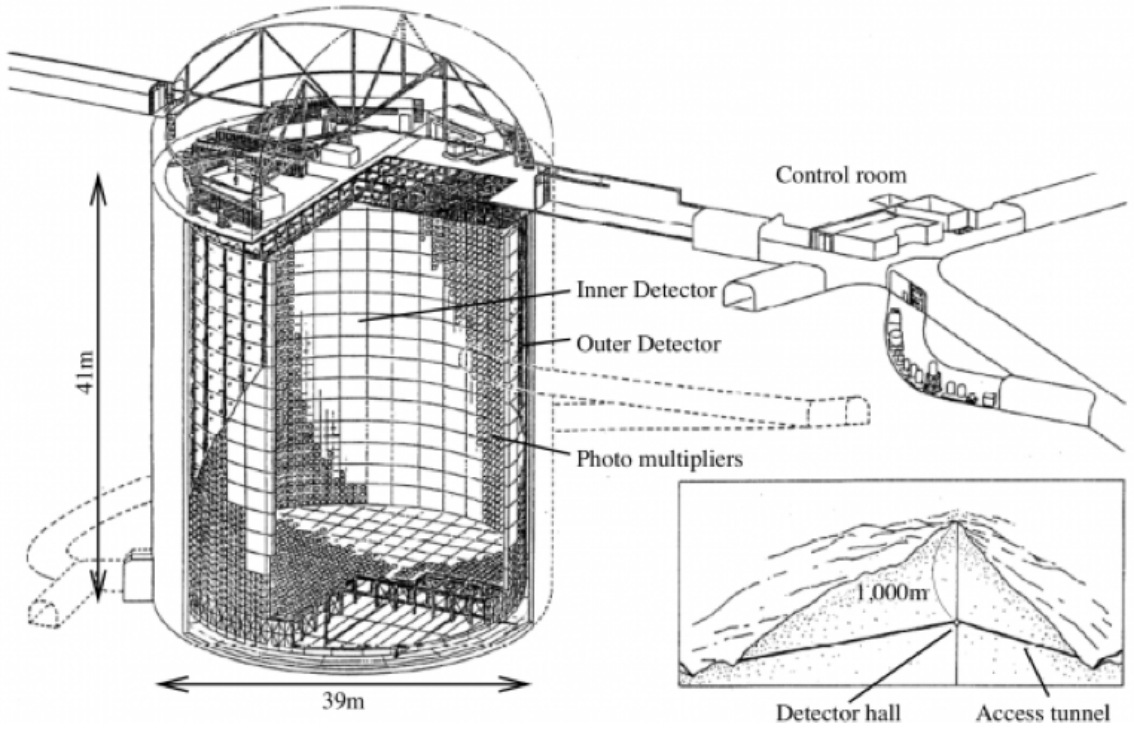


FIGURE 3.12: Super-Kamiokande detector in Kamioka [78].

Super-K is filled with 50,000 tons of ultra-pure water in a cylindrical tank 39 m in diameter and 41 m tall (Figure 3.12). Within this tank, a stainless-steel structure 55 cm in thickness supports separate arrays of inward facing and outward facing Photomultiplier Tubes (PMTs) [84]. This framework is spaced approximately 2–2.5 m from the tank wall on all sides, holding approximately 13,000 photo-multiplier tubes (PMTs) in place. This structure, shown in Figure 3.13, specifies the boundary between the Outer Detector (OD) and the Inner Detector (ID). The ID PMTs (11,129 in total) are 50 cm in diameter and surround the Inner Detector giving around 40% coverage with a timing resolution of ∼2 ns. The OD PMTs (1885 in total) are 20 cm in diameter, being less densely distributed in comparison to the ID PMTs, around the Outer Detector.

The 50 ktons of ulta-pure water in Super-K is reprocessed at a rate of ∼ 30 tons/hour by the water purification system. Water purity is maintained through re-circulation in order to reduce backgrounds and maximise water transparency [84]. As Super-K detects cosmic-ray muons and atmospheric muons, beam timing information from J-PARC must be shared in order to capture data at the correct time corresponding to beamline neutrinos. Each beam spill from J-PARC has a GPS timestamp that allows Super-K to record all interaction information within a 1 ms window [84]. This allows the far detector to focus on the T2K neutrino beam, specifically on registered PMT hits due to leptons produced by neutrino-nucleus interactions.

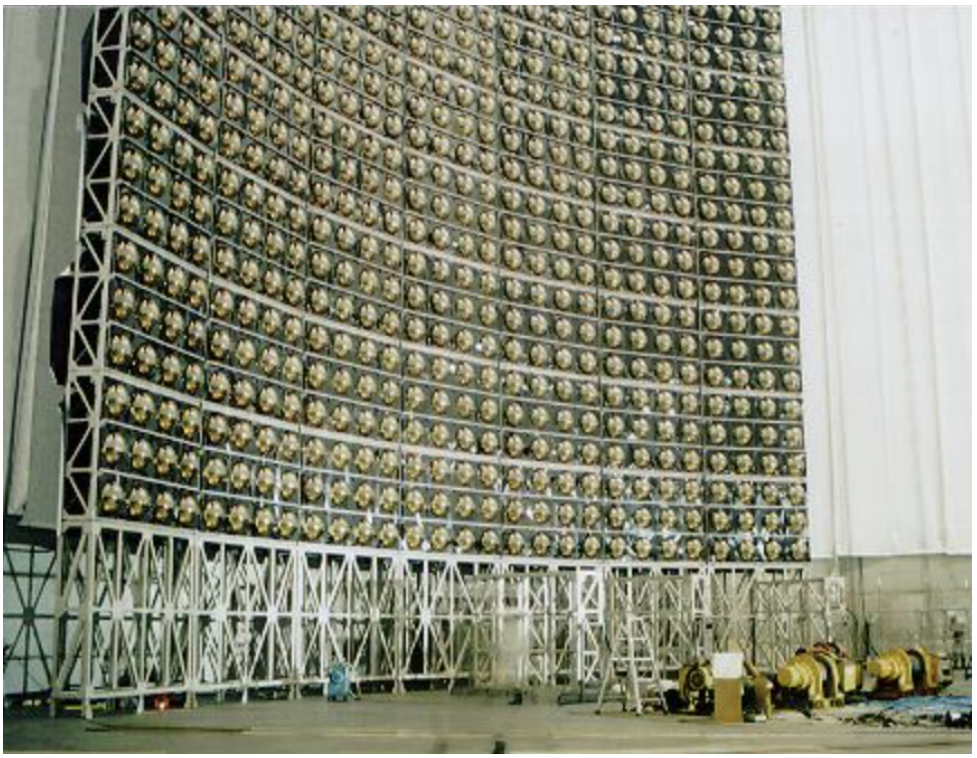


FIGURE 3.13: A PMT wall array during construction showing the Inner Detector (ID) PMTs around the tank wall. The steel framework structure holding ID PMTs can be seen. The Outer Detector (OD) PMTs are on the backside of the steel framework facing the tank wall. [85]

Neutrino interactions produce particles with high energies. Particles travelling at a high velocity in the detector (at a velocity greater than the phase-velocity of light in the medium) will emit Cherenkov radiation. In Super K, Cherenkov radiation is the main signature left by a high-energy particle, such as a muon or an electron, depending on the type of neutrino interaction. This is important because the detection of Cherenkov light in the PMTs allows for a reconstruction of the event that took place. In the case of high energy leptons, this signature is captured by the ID PMTs. The OD PMTs serve as a veto for selecting events. By measuring the radiation signature of selected events, the details of the particle can be determined (Figure 3.14). Muons produce a sharp ring of PMT hits with a distinct cone of Cherenkov radiation. This is due to their large mass and resilience to scattering. On the contrary, electrons produce a "fuzzy" ring as they scatter off particles more frequently, inducing electromagnetic showers. This results in a less distinct ring of PMT hits as multiple Cherenkov light cones are registered, these features are used to classify each particle and therefore distinguish between different types of neutrino interactions.

As mentioned prior, the condition for Cherenkov radiation to be emitted is that the velocity of the particle (v) in the medium (with refractive index n) is greater than the velocity of light in that same medium (given by $\frac{c}{n}$). Cherenkov light is emitted in a cone with an opening angle of $\theta = \arccos(\frac{c}{n.v})$ along the direction of the particle's trajectory. This opening angle is

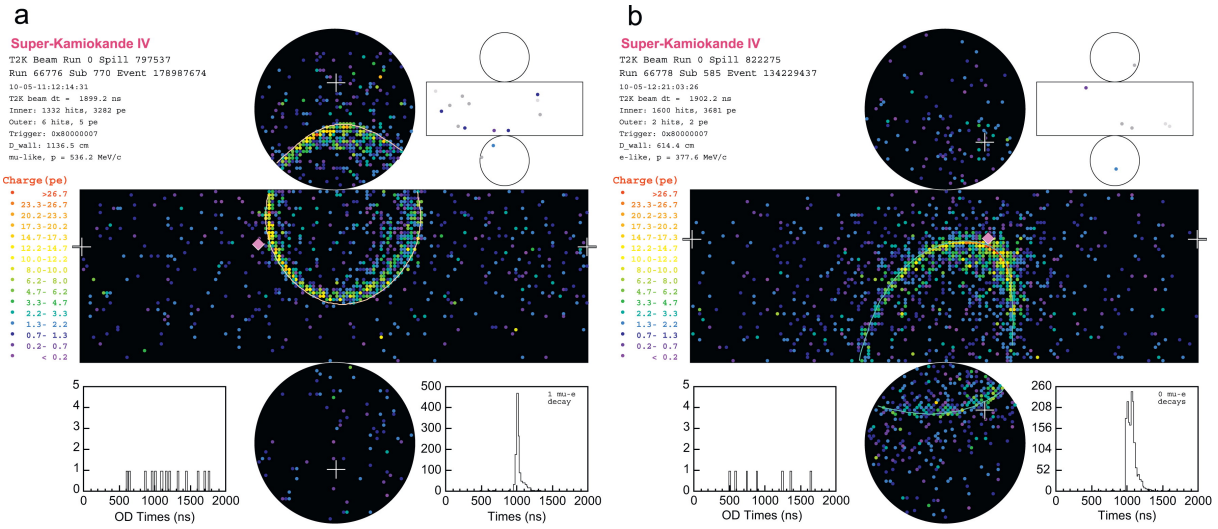


FIGURE 3.14: Super-K event display showing a muon-like and an electron-like neutrino event (on the left and right respectively). The PMT trace for the electron-like neutrino event is "fuzzy" compared to a muon-like neutrino event which produces a sharper ring [84].

approximately 42° in water ($n = 1.333$). This light is emitted whilst the particle energy (and thus its velocity) is above the Cherenkov threshold. However, once the energy of the particle decreases and falls below the Cherenkov threshold, Cherenkov light is no longer emitted. Thus, the emitted Cherenkov photons end up forming a ring-like profile on the wall of the tank instead of a solid circle. Since neutrino interactions are of interest, the outgoing charged leptons from charged current events are focused upon in detection. The Cherenkov photon ring patterns from these outgoing particles are detected by the PMTs lining the inner detector. This detection alongside the hit timing is used to reconstruct the direction, charge and momentum of the particle, alongside the position of the interaction vertex.

3.5 Software, Simulation and Reconstruction

The ND280 Software is used to simulate events within the experiment by creating a MC (Monte Carlo) data-set representing the real data. In order to do this, an initial neutrino flux prediction is required as an input. This requires specific software to be employed alongside a familiarity with the procedures to generate the MC data-set. The neutrino flux simulation and the ND280 detector simulation are mentioned in detail below. Finally, the simulation and reconstruction at Super-K is also discussed.

3.5.1 Neutrino Flux Simulation

Simulating a proton beam focusing on the T2K graphite target involves hadronic interactions which are hard to model accurately. The main uncertainties on the neutrino flux prediction are due to these hadronic interactions. However, the errors on this prediction can be significantly reduced using data from the NA61-SHINE [86] experiment which measures pion and kaon production as a function of momentum and angle for protons interacting with a carbon target.

A number of packages are also used to model interactions at the secondary beamline at T2K. FLUKA [87] is used for the in-target interactions and is passed to JNUBEAM [73], which uses GEANT4 [88] for propagation of products after impingement within the secondary beamline. JNUBEAM is a Monte Carlo simulation package that simulates the secondary beamline itself. JNUBEAM uses GCALOR [89] for hadronic interactions. Most importantly, in order to achieve an accurate neutrino flux prediction, the in-target simulation generated using FLUKA is tuned with data from NA61-SHINE, allowing for a more accurate flux prediction with better constrained uncertainties.

NA61-SHINE is a dedicated hadron production experiment located at CERN, Geneva, which has taken data using two different carbon targets: a carbon target at 4% of the nuclear interaction length; and a graphite target resembling the T2K target (1.9 interaction length) impinged on by a 31 GeV proton beam [90]. The latter case provides a good approximation to the T2K experiment beamline setup, and was specifically designed for that purpose.

3.5.2 ND280 Simulation and Reconstruction

NEUT [91] is used at the near detector to simulate neutrino interactions with nucleons and nuclei. The library was written specifically as a neutrino interaction generator for the T2K and Super-Kamiokande experiments. GEANT4 is used to simulate the ND280 and INGRID geometries alongside the resulting final-state particle trajectories from neutrino interactions. The results of this are output to ElecSim [71], a custom software package developed by the T2K collaboration.

ElecSim is used model the response of the ND280 to energy deposits from interacting particles. The package simulates the light due to deposited energy in FGD scintillator bars, alongside the transmission of light through the WLS fibres within the scintillator bars and the response of the MPPCs. With respect to the TPCs, the package simulates the ionised electrons (drift electrons) and the MicroMegas detectors. ElecSim outputs the result of this simulation as reconstructed data coupled with true event information.

Following a neutrino-nucleus interaction which occurs within the FGD, a number of resultant particles may propagate from the vertex into the neighbouring sub-detectors. Within a neighbouring TPC, a curved track may appear which passes through at high momentum; this may be a pion or a lepton. Whilst TPCs provide good tracking and identification of particles, short-ranged particles such as recoil protons are measured in the FGDs themselves [81].

The data collected from the FGDs and TPCs are used in conjunction with each other for particle identification (PID). Particle tracks are identified across the boundaries of the sub-detectors in reconstruction (Chapter 6, Section 6.3). Combining momentum information with energy loss as a function of distance (dE/dx) within the TPC allows for PID. Comparing dE/dx with known rates of energy loss per distance, for different particles produced in neutrino interactions, allows for particles to be distinguished from each other (Figure 3.15).

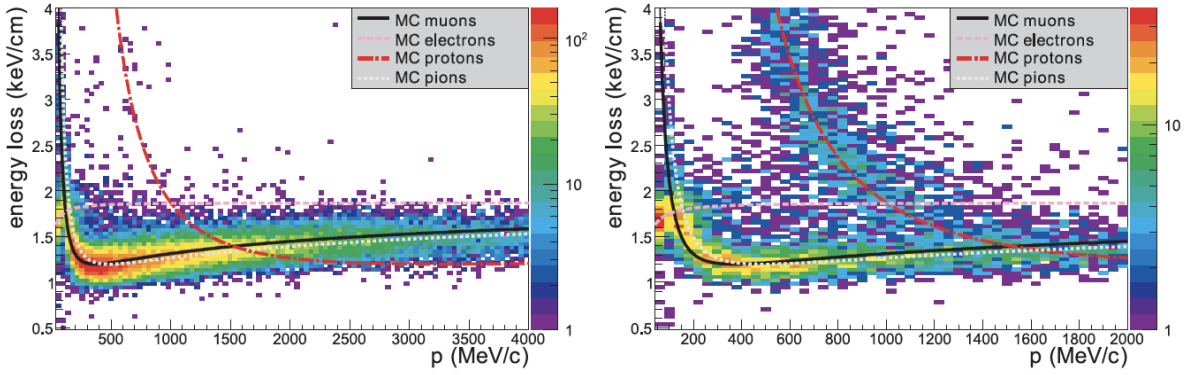


FIGURE 3.15: Energy loss plotted against momentum for different particles in the TPC [47]. The curves show theoretical predictions for the particles whilst the binned histogram shows data. Left: Negative Particles (muons with some low momentum electrons - no data matches the 'negative' protons prediction as expected). Right: Positive Particles (protons, pions, and positrons).

3.5.3 Super-K Simulation and Reconstruction

Super-K also serves as the far detector for the T2K experiment. With regards to the simulation of events: the NEUT event generator is used to model neutrino interactions in the Super-K tank. The SKDETSIM [84] software based on GEANT3 [88] simulates the produced particles as they move through the detector geometry. GCALOR [89] is used to simulate hadronic interactions.

Event reconstruction at Super-K is carried out by looking at the charge and timing information of PMTs which register light signatures from high energy particles. The end result is to identify particles along with their position and momentum. This reconstruction is carried out using fiTQun [92], an event reconstruction package for Super-K, which uses a likelihood function that is a function of particle track parameters such as momentum, particle ID, and vertex. This function can be expressed as a product over all PMTs. Maximising this function yields optical track parameters for every event [93].

Chapter 4

Trip-T Calibration and Electromagnetic Calorimeter Data Quality

This chapter will focus on my service task which consisted of assessing and reporting on the Electromagnetic Calorimeter (ECal) data quality within the Data Quality Group. The Data Quality group assesses the quality of the data being gathered at the ND280. I was the ECal data quality expert for Run 10 from November 2019 until February 2020. Every sub-detector of the ND280 has assessments carried out weekly, with their status at a particular time period being recorded by uploading a flag to the semi-offline database. Good flags (taking the value of 0) represent good data collection over a certain specified period, with any other positive number representing a bad status for the sub-detector. A value of -1 indicates that there is no data being recorded for that specified period. Following this, a global flag is then uploaded for the whole ND280 taking into account the flags of the sub-detectors. When all the sub-detectors (excluding the PØD) are flagged as good for data quality, the global flag is also good. This informs the analysers which run periods of the full data taking can be used.

4.1 ND280 Data Acquisition (DAQ) and Electronics

The sub-detectors in the near detector are made from different materials. The main material that allows for the detection of particles is scintillator. When a particle hits a scintillator bar, atoms in the bar excite and emit light which is transmitted to the end of the bar, the light is guided by wave-guides much like an optical fibre. This transmitted light reaches a sensor at the

end of the scintillator bar which can be a photodiode, photomultiplier or avalanche photodiode (APD) which converts photons to electrons.

The sub-detectors in ND280 are all Trip-T electronics based modules [94]. The Trip-T electronics system was developed by Fermilab and is the readout system used alongside MPPCs (Multi-Pixel Photon Counters) that records the data for ECal, INGRID, SMRD and PØD. The ND280 and INGRID detectors both have independent data acquisition (DAQ) systems, which have a similar architecture and are based on the MIDAS framework [95]. The electronics systems used for TPCs and FGDs, which are the more active parts of the ND280, are read out by MicroMegas (MM) modules and Multi-Pixel Photon Counters (MPPCs) respectively [96].

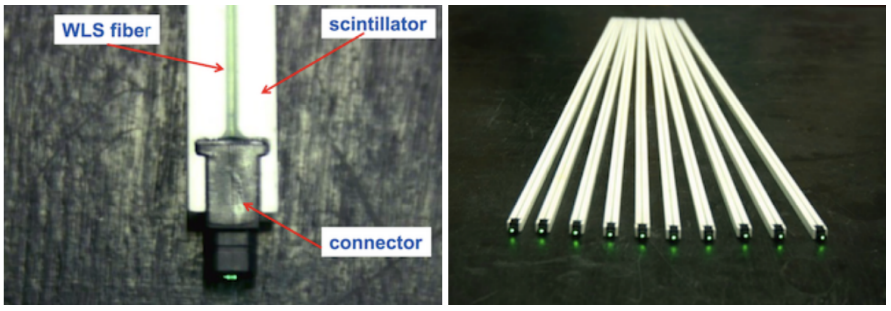


FIGURE 4.1: Photograph showing an example scintillator bar, WLS fibre and connector which would, in T2K, be linked to an MPPC for readout [97]. The image on the right from the perspective of the camera shows the scintillator bars in a spaced out Y configuration.

The MPPCs used in T2K consits of an array of 667 silicon avalanche photodiodes (APDs) as shown in Figure 4.2 below. These convert a light signal into an electric signal using the photoelectric effect. By observing the scintillation of scintillator planes, namely the electrical signals of scintillator bars through WLS fibres, the position of charged particles can be identified.

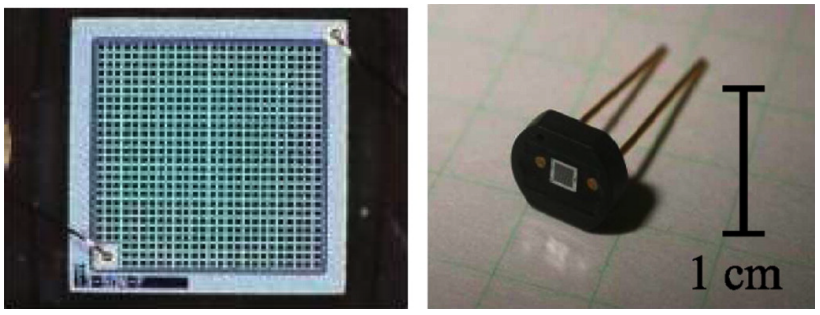


FIGURE 4.2: MPPC photograph and its pixel face. The 667 active APDs can be seen on an 26 x 26 grid of photodiodes (9 are lost due to the electrode on bottom left corner) [98].

Each scintillator bar in the sub-detector is read out by 1-2 MPPCs. The hierarchy of the electronics is as follows: A Trip-T Frontend Board (TFB) is able to read up to 4 Trip-T chips. A Trip-T chip is able to read from up to 16 MPPCs simultaneously. Therefore, each TFB can

read from upto 64 MPPCs (channels). A TFB has its data processed by a Readout Merger Module (RMM) with 48 TFBs [99] being read by an RMM. Because of the arrangement of these electronics, various sub-detectors can be isolated from the global DAQ in a standalone fashion (locally) by reference to certain RMMs which serve as the communication interface between a certain sub-detector and the ND280 DAQ. Figure 4.3 shows the electronic system layout in an ND280 sub-detector.

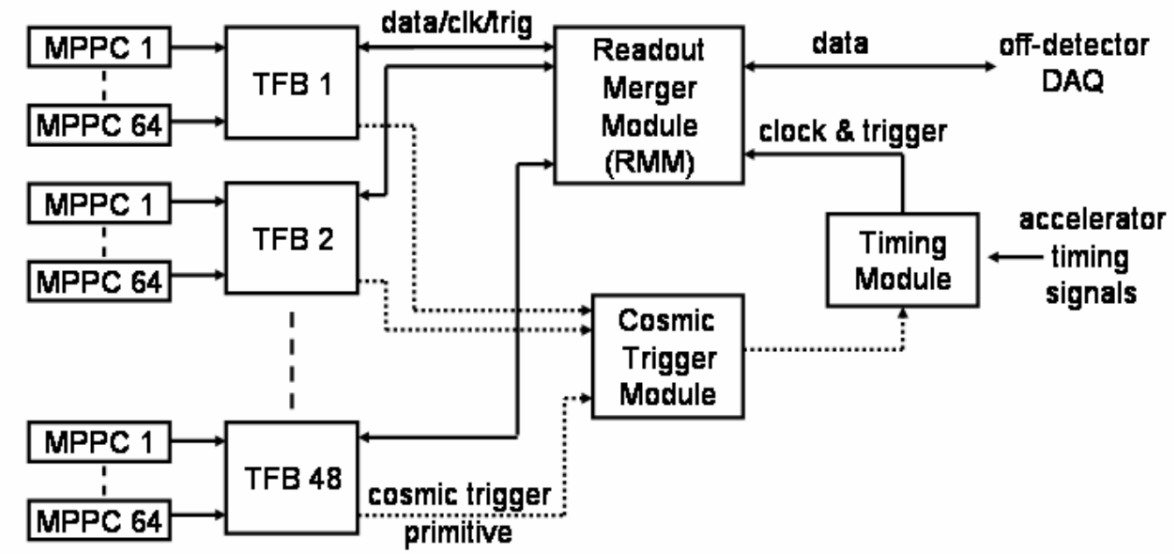


FIGURE 4.3: Trip-T based ND280 front end electronics system. Multiple RMMs are associated with each sub detector [100].

TFBs output measured ADC values. An analogue to digital converter (ADC) allows an analogue voltage to be converted into a digital number, this allows scintillator bar activity in sub-detectors to be quantified by certain values. Understanding what values constitute a signal, and what values constitute a background are essential in capturing data and ensuring data quality. To help in this, small intervals in which the detector gathers event data are controlled by certain clock modules. These clock modules allow for 'triggering', where a signal is sent to start a snapshot of data taking across the sub-detectors. Triggers are extremely important in order to isolate and separate beam interactions from cosmic interactions in the ND280. A Master Clock Module (MCM) issues triggers over the whole ND280 to the RMMs and to Slave Clock Modules (SCMs), these then trigger the sub-detectors. There is one SCM for every sub-detector which allows the option for a sub-detector to be configured to run standalone for calibration and debugging.

Three main triggers are registered across the ND280 DAQ: the beam trigger, the FGD cosmic trigger and the Trip-T cosmic trigger. When the accelerator timing signal from JPARC

is sent to the MCM, a beam trigger is issued. When there is no beam trigger and hits from the same trajectory are seen in both FGDs, an FGD cosmic trigger is issued. The final trigger is the Trip-T cosmic trigger which is issued when there is no beam trigger and hits are seen the opposite sides of the detector (PØD, ECal and SMRD). In the case of the FGD cosmic and Trip-T cosmic trigger, the events which are captured are mostly cosmic ray muons.

4.2 Electromagnetic Calorimeter (ECal) Data Quality

The data quality for the ECal consists of several weekly checks which include monitoring the voltage and current sensors on the ECal TFBs along with the temperature, the cooling system flow meters and raw data analysis. The summary of the data quality is given by a 12 bit summary flag with each bit representing an RMM. A value of 0 indicates good data quality, whereas any positive value indicates a problem with an RMM or multiple RMMs. Data not being written is indicated by -1 as a flag.

Any positive value for the summary flag can be broken down to show which constituent RMMs are an issue, as when multiple RMMs show problems, their values are summed together linearly to yield the summary flag. For example, a flag of 5 indicates a problem with RMM 0 and RMM 2 (as $2^0 + 2^2 = 5$). Likewise a flag of 530 indicates a problem with RMM 1, RMM 4 and RMM 9 (as $2^1 + 2^4 + 2^9 = 530$). A flag of 4095 indicates that there is a problem with all the RMMs, and therefore the whole ECal (as $2^0 + 2^1 + 2^2 + 2^3 + 2^4 + 2^5 + 2^6 + 2^7 + 2^8 + 2^9 + 2^{10} + 2^{11} = 4095$).

Summary flags are set for the duration of the data taking period and presented regularly at the Data Quality meeting. The information needed to produce a summary flag is extracted from an SQL database which stores all changes in state for the ECal sub-detector over its run period. An example ECal summary flag generated for the period of $13^{th} - 20^{th}$ January 2020 can be seen below in Figure 4.4.

The summary flag for the ECal is then uploaded to an SQL database to be used in generating a global flag. The ECal module and corresponding bit number, decimal value and binary value are shown below in Table 4.1.

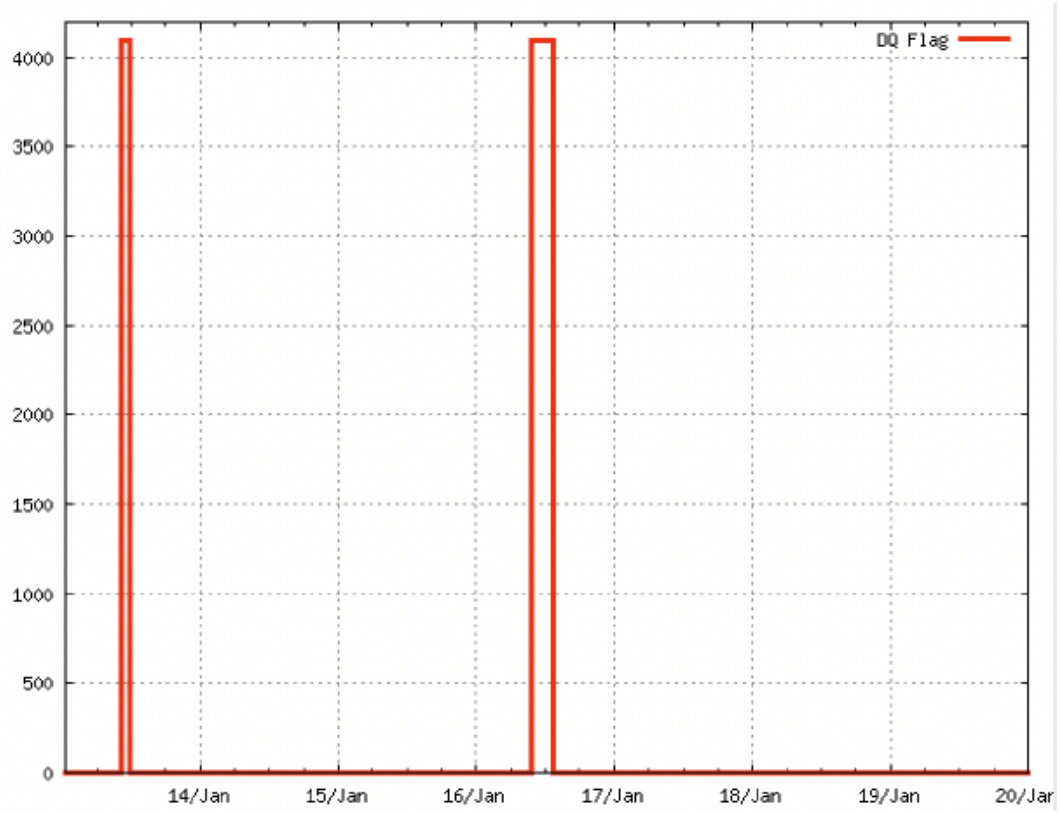


FIGURE 4.4: ECal data quality for the period of $13^{th} - 20^{th}$ January 2020. The flag of 4095 can be seen, signifying an issue with the ECal during that period.

TABLE 4.1: Table showing the bit number, decimal value and binary value for the corresponding ECal RMM.

Bit Number	Decimal	Binary Value	ECal module
1	1	2^0	RMM 0– Ds-ECal
2	2	2^1	RMM 1 - Ds-ECal
3	4	2^2	RMM 2– PØD-ECal South
4	8	2^3	RMM 3 - Barrel-ECal Top South
5	16	2^4	RMM 4 - Barrel-ECal Bottom South
6	32	2^5	RMM 5 - Barrel-ECal Side South
7	64	2^6	RMM 6 - Barrel-ECal Side South
8	128	2^7	RMM 7 - PØD-ECal North
9	256	2^8	RMM 8 - Barrel-ECal Side North
10	512	2^9	RMM 9 - Barrel-ECal Side North
11	1024	2^{10}	RMM 10 - Barrel-ECal Bottom North
12	2048	2^{11}	RMM 11 - Barrel-ECal Top North

4.2.1 MPPC Gains and Pedestals

Two main data checks were carried out; the observation of beam timing and the gain/pedestal monitoring. The MPPC gain and pedestal monitoring is presented in this section.

The pedestal values for the electronics along with the MPPC gains are monitored as large variations could indicate issues in the voltage supply which change the resolution and efficiency of the detector. Variations due to changes in temperature throughout the day/night are not usually troubling as these are quite small. These are visible on the plot for the period 3^{rd} - 9^{th} November 2019 seen below in Figure 4.5.

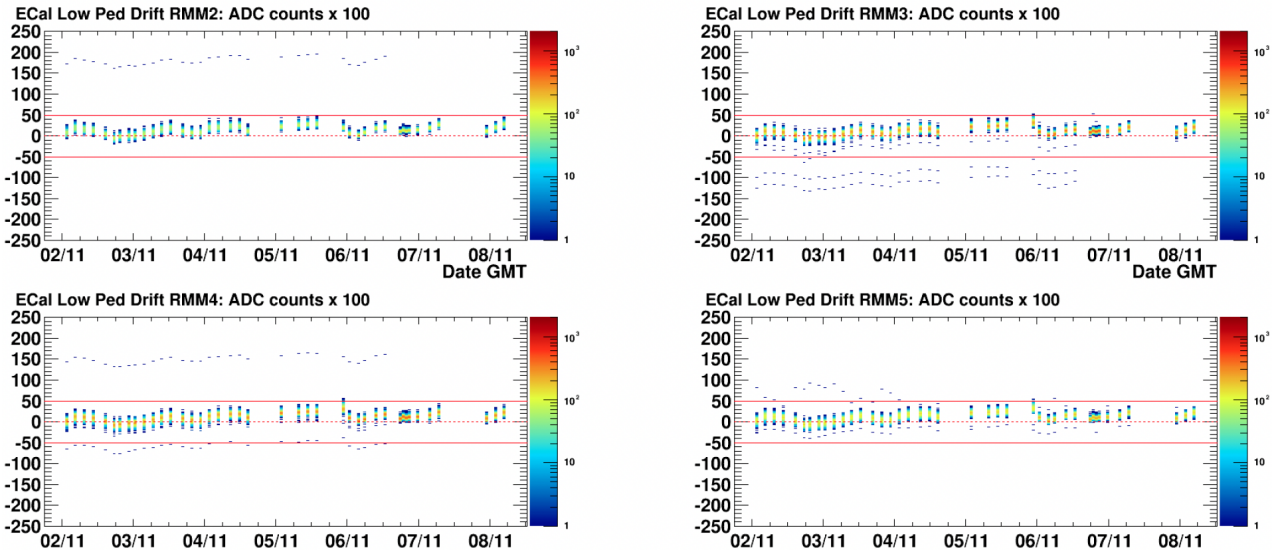


FIGURE 4.5: ECal low channel pedestal drift for multiple RMMs $2-5$. Red lines correspond to 0.5 ADC counts which indicate good performance. On the y-axis: ADC counts x 100.

The mean value recorded by the histogram of ADC counts, with no signal (only thermal noise) from the MPPC, is called the pedestal and corresponds to 0 pixels fired (see Figure 4.6 below).

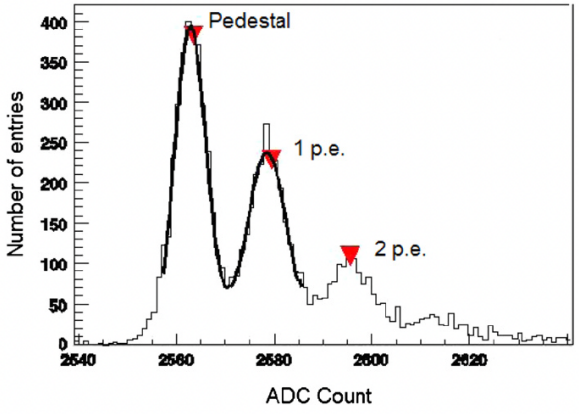


FIGURE 4.6: MPPC ADC spectrum from LED source with fitted Gaussian peaks shown [101].

Looking at a histogram of ADC counts showing only background noise (thermal noise) for each channel, the pedestal and gain can be calculated - the spectrum shows a large peak with equally spaced numerous peaks with decreasing amplitudes. The large initial fitted peak is called the pedestal and is essentially the effect of background noise (no MPPC activity).

The gain is defined as the spacing between the first (pedestal) peak and the second peak which corresponds to an ADC reading with no MPPC activity vs an ADC reading for one photo-electron event. Figure 4.7 below shows the ECal gain drift for RRMs $2-5$. Due to diurnal temperature variations in the detector cavern, the histograms display evaluations of the gain and pedestals 8 times in a day (24 hours), which corresponds to an evaluation every 3 hours.

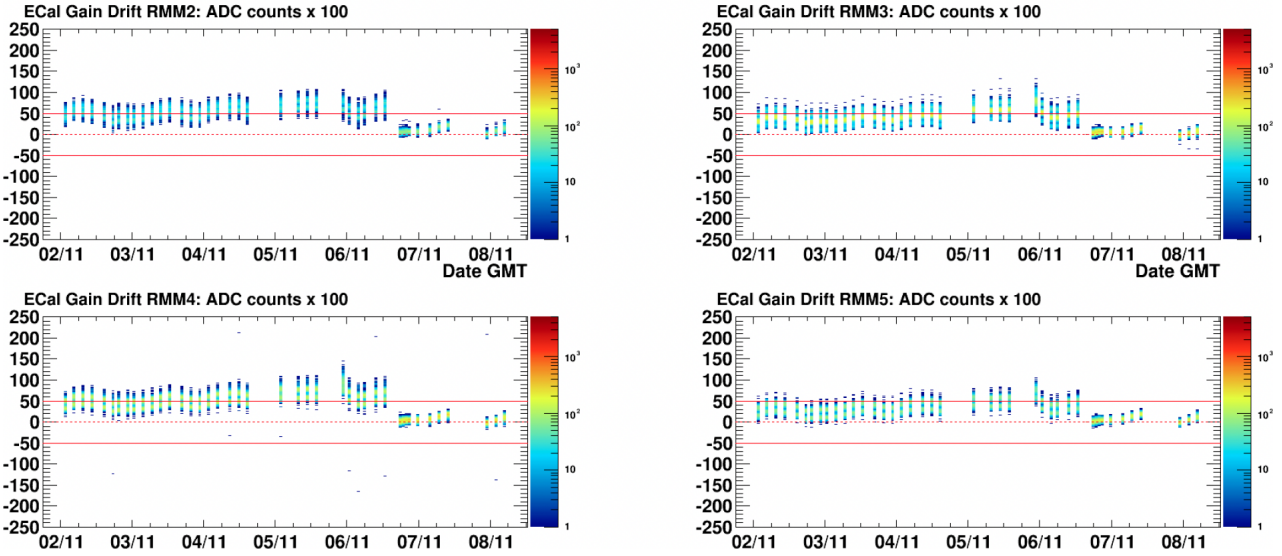


FIGURE 4.7: ECal gain drift for multiple RMMs $2-5$. Red lines correspond to 0.5 ADC counts which indicate good performance.

Furthermore, the readout system at T2K has to be able to resolve all time signals, ranging from MPPC signal sizes to the time signals of a few photo-electrons. A single channel reading an input signal is liable to voltage and noise dispersion and such disturbances could affect its resolution of a couple of photo-electrons. In order to account for this dispersion, the MPPC input signals are split into high and low gain channels, with pedestal values for each channel being evaluated separately.

Using the pedestal and gain values allows one to convert from a measured ADC value into a number of photo-electrons producing the signal. This is highly important because the MPPCs output an ADC value similar to most sensors used in experiments, in order to relate this to underlying physics processes it is imperative that the number of photo-electrons that generated this response be known. This number of photo-electrons (N_{pe}) is given by the ADC value for

that MPPC (ADC) subtracted by the pedestal (P), all divided by the gain (G), as shown in Equation 6.1 for the high-gain channel below:

$$N_{pe} = \frac{(ADC - P)}{G} \,. \tag{4.1}$$

This equation does not include corrections for nonlinear electronics and different scaling on the low and high gain channel readouts. However, essentially, it can be seen that the pedestal value is an offset, with the gain being the factor that relates the ADC value to the number of photo-electrons. It is therefore essential that these values are calculated accurately. This depends on the electronics being calibrated correctly due to shifts in the pedestals and gains; these may be due to changes in detector temperature (mentioned prior), or underlying voltage supply issues.

The gain and pedestal constants for special runs (cosmics) were processed and uploaded once a week. During these runs there was no beam present, allowing for the offline population of ADC spectra. Processing involved locating the special runs (cosmics) and submitting a job on the semi-offline cluster to populate the ADC histograms. The resultant files showing the gain and pedestals were downloaded to check they were generated correctly (within ~ 1 ADC of normal sub-detector values), if this was the case, these were uploaded. For normal runs, gain and pedestal constants were uploaded automatically.

4.2.2 Beam Timing

Reconstruction of tracks within the ND280 detector involves looking at the time correlation between signals across sub-detectors. This is essential in being able to filter out (veto) background signals, alongside determining the track direction. When beam-line neutrinos are expected, a beam trigger is issued, and a small window on the order of nanoseconds in which detector hits occur are recorded. This hit time in relation to the beam trigger is recorded, by looking at these offsets across all the readout electronics; the bunch structure can be predicted from the TFB readouts.

Figure 4.8 below shows the hit timing relative to the triggers. It can be seen that the bunch timing is relatively stable over time. A certain level of jitter across the bunches is present, but this is not flagged as problematic as fluctuations up to 100 ns are expected.

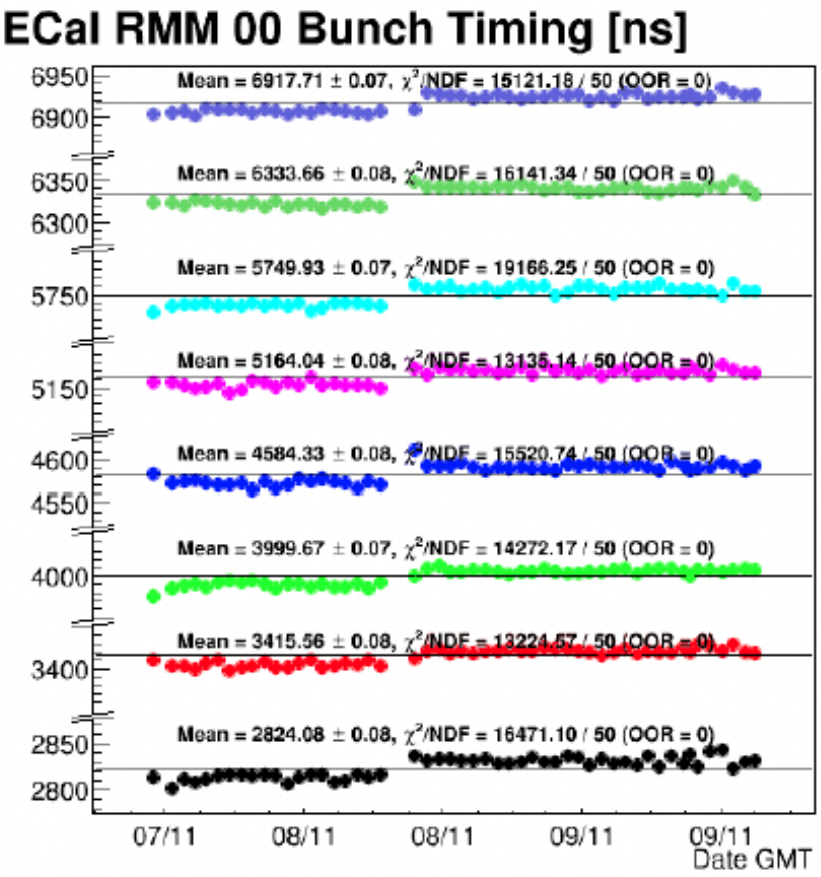


FIGURE 4.8: ECal hit time relative to the beam triggers for 8th - 9th Nov 2019 (from Run 10). The 100 ns reset window is shown as double lines on the y-axis between the 480 ns readout cycles.

4.3 Run 10 Data Quality Summary

The flags for all the sub-detectors are combined in order to generate a global data quality flag for Run 10 seen below in Figure 4.9. The ND280 detector magnet was broken during the majority of Run 10 and this is reflected in the bad DQ flag. Data quality was still carried out for all the sub-detectors that required a magnetic field to complete their function.

Ultimately, due to the broken magnet, the data over the total Run 10 period for these sub-detectors (including the ECal) was marked as bad and deemed as incomplete.

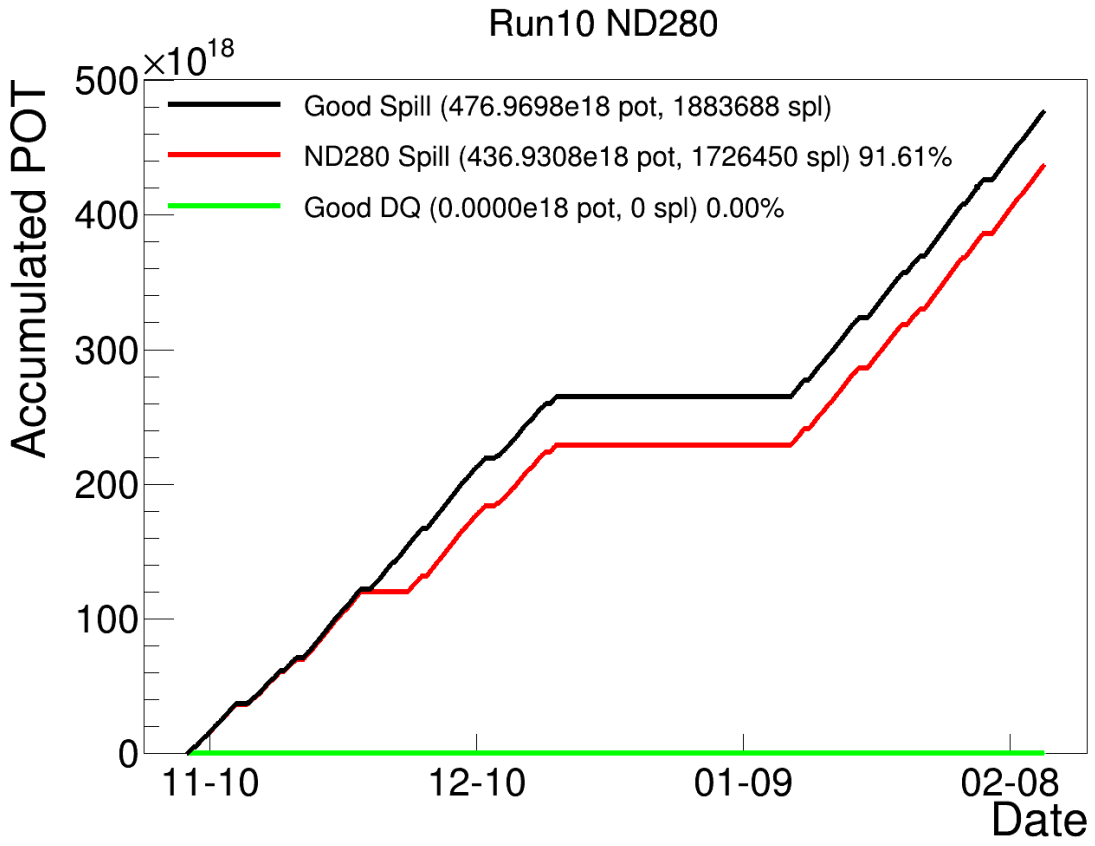


FIGURE 4.9: Accumulated Protons on Target (POT) during Run 10 (Nov 2019 - Feb 2020), showing good spill efficiency and a bad DQ flag (T2K Internal).
Good Spill: the good beam spill delivered by JPARC - the total beam delivered (in POT).
ND280 Spill: good beam spill received when ND280 DAQ was on and recording data (in POT).

Chapter 5

The Oscillation Analysis at T2K

In this chapter the oscillation analysis at T2K will be discussed, alongside the inputs to the oscillation, how to generate a prediction, the fit mechanics and, finally, extracting results from the fit. Measurements relating to neutrino oscillations are determined from observing a neutrino beam as it travels from one point to another. The neutrino oscillation probability is a function of the energy and the distance travelled (Section 2.2, Figure 2.3); therefore, oscillation measurements consist of measuring the neutrino events of a given flavour at a certain distance and comparing this to a prediction.

The neutrino beam is measured at two points in T2K, once at ND280, which serves as the near detector, and once at Super-K, which serves as the far detector. Close to the beamline at ND280, the neutrino beam is "un-oscillated", meaning that neutrinos have not been able to travel far enough to have undergone a flavour oscillation. At Super-K the neutrino beam is considered to have "oscillated", meaning that the neutrinos have been able to travel far enough to undergo a flavour oscillation. Essentially, by comparing ND280 and Super-K measurements, a neutrino deficit or surplus at the far detector is seen as evidence of neutrino oscillation.

5.1 The T2K Oscillation Analysis

The first step in the oscillation analysis at T2K is to generate a near detector event rate. The predicted near detector (ND280) event rate can be described as shown below in Equation 5.1:

$$N_{ND280}\left(p_{\text{reco}}\right) = \sum_{i} \Phi_{\alpha}\left(E_{\text{true}}\right) \times \sigma_{\alpha}^{i}\left(p_{\text{true}}\right) \times \epsilon_{\alpha}\left(p_{\text{true}}\right). \tag{5.1}$$

$N_{ND280}(p_{\text{reco}})$ represents the predicted ND280 event rate as a function of the outgoing reconstructed lepton momentum (which is a 4-vector), i represents the interaction type – CCQE, CCRES, or CCDIS for example, which is the index that is summed over. $\Phi_\alpha(E_{\text{true}})$ represents the neutrino flux as a function of the true energy. $\sigma_\alpha^i(p_{\text{true}})$ represents the neutrino interaction cross section (detailed in Chapter 6, Section 6.7.4) as a function of momentum for interaction type i and neutrino flavour α which, when convoluted with the neutrino flux, provides the base number of neutrinos that interact in the detector. To account for the fact that only a percentage of these events are detected, $\epsilon_\alpha(p_{\text{true}})$ represents the efficiency of the detector.

The far detector, Super-K, which is 295 km away from the source shows a difference in detected events due to noticeable neutrino oscillation over this distance. To account for this in the far detector event prediction, the un-oscillated rate is updated by including an additional factor describing neutrino oscillations (Equation 5.2 below).

$$N_{SK}^{\alpha\to\beta}(p_{\text{reco}}) = \sum_i \Phi_\alpha(E_{\text{true}}) \times \sigma_\beta^i(p_{\text{true}}) \times \epsilon_\beta(p_{\text{true}}) \times P_{\alpha\beta}(E_{\text{true}}). \tag{5.2}$$

$N_{SK}^{\alpha\to\beta}(p_{\text{reco}})$ represents the Super-K predicted event rate as a function of lepton kinematics. The remainder of the terms are similar to the ND280 event prediction in Equation 5.1, except for the inclusion of the probability of oscillation, $P_{\alpha\beta}(E_{\text{true}})$. This term contains the parameters describing neutrino oscillation and serve as the signal parameters in the oscillation analysis, with all the other parameters becoming systematic uncertainties in relation to them.

Equation 5.2 describes the event rate for an 'appearance' experiment, where a change of neutrino flavour (from α to β), namely from muon to electron neutrino, is observed at the far detector. $N_{SK}^{\alpha\to\alpha}$ looks at a 'disappearance' experiment where the same neutrino flavour (from α to α) is measured at the near and far detector to observe any change in event rate, giving information about different parameters. Figure 5.1 below shows the neutrino oscillation analysis pathway leading to a detector prediction.

It can be seen that in order to get an accurate detector prediction it is vital to understand the cross-section parameters, specifically the hadron interaction model, as this forms one of the greatest systematic uncertainties [102]. Therefore, getting an accurate detector prediction is dependent on making sure that all parameters in the model are constrained in order to reduce the total systematic uncertainty.

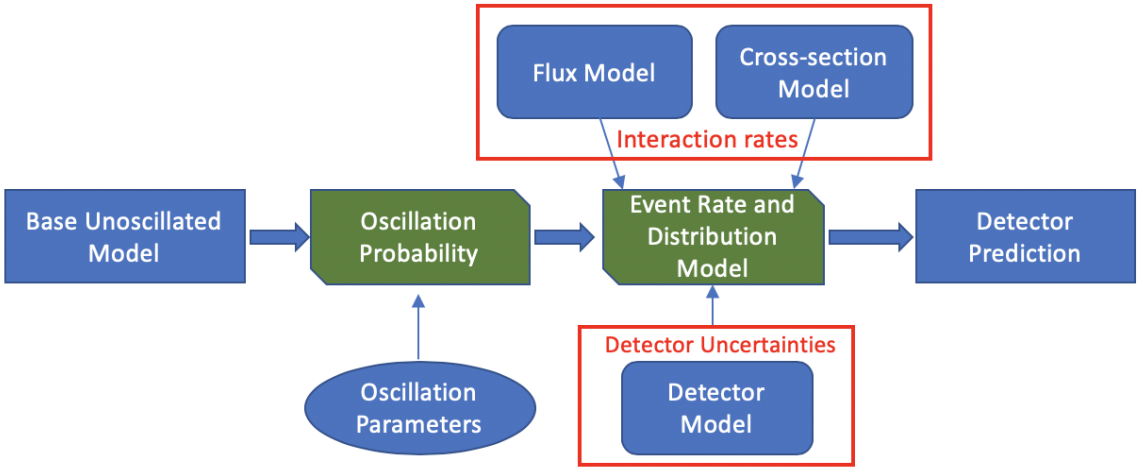


FIGURE 5.1: Neutrino oscillation analysis pathway.

5.2 The Near Detector Fit

Fitting is a procedure in which observed data is compared to predicted data from a model. This is in order to determine its goodness of fit and concurrence with an assumed hypothesis. Predicted data is generated through parameters of the model which are used to describe the physics. If this comparison can be improved, then this predicated data is 'regenerated' by adjusting the underlying parameters which describe the physics to generate a new prediction.

A metric is used to describe how accurately the observed events fit the predicted events. This metric, called the likelihood, is calculated after successive predictions in the process of fitting until maximised to the fullest extent it can be. The conditions for having reached this maximised extent depend on the fitter method used (mentioned below). Regardless, at this point the fit is said to have converged. The underlying physics parameters making up the final prediction are then extracted, the parameters which are not sought after removed from the final result, and the sought after parameters such as δ_{CP} are stated with an associated error.

Fitting is done in final state lepton kinematics with momentum p and direction $\cos\theta$, where θ is the angle that the outgoing lepton makes to the incoming neutrino axis. The prediction is generated in these kinematic variables using Monte Carlo methods, as this lepton final state is common to all charged current neutrino interactions. There are two fitting groups at the near detector; each of the two groups are differentiated by their adoption of either one of the two fundamental views of statistics: the first group is called BANFF (Beam And ND280 Flux extrapolation task Force) who employ a Frequentist based framework which uses a gradient descent approach called MINUIT [103] to maximise likelihood at the near detector (which is

run again for the far detector). The second group is called MaCh3 (Markov Chain for 3 flavour oscillation fitting) who employ a Bayesian approach which looks at the probability calculation in terms of a prior and a posterior. The prior contains all the initial parameters, which, when convoluted with the likelihood, yields an updated posterior. The BANFF Frequentist framework uses a covariance matrix based approach to yield sought parameters with confidence intervals, whereas the MaCh3 Bayesian framework uses MCMC [104] (Markov Chain Monte Carlo) to sample the posterior and yield sought parameters with credible intervals. In this thesis we will look at the latter Bayesian approach.

5.3 T2K Likelihood and Bayes' Theorem

In Bayes' Theorem, prior information about parameters, $P(\vec{\theta})$, is convoluted with the likelihood, which is the probability of the observed data given the model, $P(D|\vec{\theta})$. This results in a posterior distribution of the probability of the model given the data, $P(\vec{\theta}|D)$ which contains updated information about all parameters within the model taking into account the data presented. Parameters of interest can then be extracted by sampling this posterior as mentioned above, with unwanted parameters marginalised through integration (Section 5.4.2). Thus, a hypothesis or model is tested in Bayesian statistics through Bayes' Theorem which is given below in Equation 5.3:

$$P(\vec{\theta}|D) = \frac{P(D|\vec{\theta})P(\vec{\theta})}{\int P(D|\vec{\theta})P(\vec{\theta})d\vec{\theta}}. \tag{5.3}$$

It can be seen here that $P(D|\vec{\theta})$ is the probability of measuring data (D) given the model parameter values (θ). This is evaluated by comparing between the observed (data) events and the predicted events. This likelihood is calculated using a Poisson probability distribution comprising of individual bin likelihoods. The likelihood for one bin is given in Equation 5.4. The parameter of the Poisson distribution is the difference between actual observed data n within a bin and the model Monte Carlo simulation $\lambda(\theta)$ within that same bin. For a single bin (b) the likelihood ratio (referred to as the likelihood) is:

$$\mathcal{L}_b = \frac{\lambda(\vec{\theta})^n e^{-\lambda}}{n!} / \frac{n(\vec{\theta})^n e^{-n}}{n!}. \tag{5.4}$$

The Monte Carlo simulation data is produced following the procedure highlighted in Chapter 3, Section 3.5. To make the likelihood in Equation 5.4 easier to deal with computationally, the natural log of both sides is taken. When summing over all bins for the sample, the sample log-likelihood becomes:

$$-\log \mathcal{L}_{\text{Sample}} = \sum_{\text{All bins}} \left[\lambda(\vec{\theta}) - n + n \log \frac{n}{\lambda(\vec{\theta})} \right] . \tag{5.5}$$

This log-likelihood is modified so that it contains extra information which captures Monte Carlo (MC) statistical error. This error is due to the fact that an unlimited amount of MC is not generated; in particle physics experiments a sample of MC events are taken out of a potentially infinite number to represent the prediction. The sample log-likelihood is then modified as below:

$$-\log \mathcal{L}_{\text{Sample}} = \sum_{\text{All bins}} \left[\beta\lambda(\vec{\theta}) - n + n \log \frac{n}{\beta\lambda(\vec{\theta})} + \frac{(\beta-1)^2}{2\sigma_\beta^2} \right] , \tag{5.6}$$

the additional term and multiplicative factor β accurately represents the MC statistical error within the likelihood. This method of looking at the likelihood error is called the Barlow-Beeston method [105]. The term σ_β is the error on β. Maximising the likelihood with respect to β, which would mean minimising the $-logL$ in Equation 5.6, gives a quadratic equation which can be solved. Therefore, the parameter β can be determined for each bin analytically [106].

The total likelihood for the T2K experiment is a combination of the sample likelihood and the systematic likelihood. The systematic likelihood is due to the errors on the model parameters themselves, some of which are Gaussian distributed and others which are flat in distribution. The model parameters are broadly grouped into 4 categories: flux, detector (ND280 and Super-K), cross-section, and PMNS (oscillation) parameters. Each model parameter also has a correlation with other model parameters; this information is captured by the covariance matrix. Some priors of the covariance matrix come from inputs from previous experiments (simulated or real), this is then captured by a Gaussian prior. However, the rest of the parameters do not have prior data, therefore the priors for these are flat. The log likelihood for the systematic parameters yields:

$$-\log \mathcal{L}_{\text{Syst}} = \sum_{\text{Parameters}} \frac{1}{2} \left[(\theta_i - \mu_i) \left(\mathbf{V}\right)^{-1}_{ij} (\theta_j - \mu_j) \right] , \tag{5.7}$$

Where the value of parameter i is given by θ_i that has a central value μ_i, $\mathbf{V_{ij}}$ is the covariance matrix element i, j. This matrix describes the uncertainties between all systematic parameters i and j (discussed in Chapter 6, Section 6.7). The fitting process involves determining what values of these parameters accurately reproduce the experimental results, the error on these parameters, and how much shifting the values of these would impact others. It can therefore be seen that the total likelihood is the sample likelihood convoluted with the systematic likelihood, which when taken to a log likelihood yields:

$$-\log \mathcal{L}_{\text{Total}} = \sum_{\text{All bins}} \left[\lambda(\vec{\theta}) - n + n \log \frac{n}{\lambda(\vec{\theta})} + \frac{(\beta - 1)^2}{2\sigma_\beta^2} \right] + \sum_{\text{Parameters}} \frac{1}{2} \left[(\theta_i - \mu_i) \left(\mathbf{V} \right)_{ij}^{-1} (\theta_j - \mu_j) \right]. \tag{5.8}$$

In Equation 5.8, 'All bins' refer to near detector and far detector binning. 'Parameters' refer to all the categories of parameters: flux, cross-section, detector, and oscillation parameters. The total number of parameters (across both the near and far detector) are $\sim$700.

5.4 Markov Chain Monte Carlo (MCMC)

The MaCh3 fitter maps out the likelihood as a probability density and conducts a semi-random walk using Markov Chain Monte Carlo (MCMC). This map of the likelihood gives the posterior distribution (within a normalisation factor) which allows MaCh3 to extract final sought parameter values by marginalising and integrating out all the systematic parameters.

As seen in Equation 5.2, the prior contains initial values of all the parameters given by previous experimental results or estimates. This prior is multidimensional alongside the likelihood and posterior which has dimensions corresponding to the number of parameters. Therefore, this posterior contains information about the PMNS parameters of interest such as θ_{ij} and δ_{CP}. This posterior distribution is not generally analytically solvable, so it is sampled using a Markov Chain Monte Carlo (MCMC) method to get estimates for the parameters. MCMC is used due to the high number of parameters within the model; it is computationally much more efficient than random sampling as less time is spent sampling low probability areas.

A Markov chain is a stochastic (random) model which describes a sequence of events in which any given event only depends on what was immediately before it. In algorithmic terms, a generated point, x_n only depends on x_{n-1}. This means that compared to other points which do not subsequently follow on from one another, there is true independence between points. The history of a walk is irrelevant in order to take the next step.

MCMC methods generate a Markov chain with a distribution of the posterior after stepping within the parameter space. The result of the MCMC algorithm is a series of steps, ideally a Markov chain (conditions of which are mentioned below), in N dimensions that is extracted past a 'burn-in' point. A burn-in point is the point after which the MCMC ceases jumping around the parameter space, settling in a position and finding a preferred region of minima. The series of steps after this point, which have converged, are stable in their distribution and so best represent the posterior distribution $P(\vec{\theta}|D)$ in the multidimensional parameter space. As additional steps are added, the distribution more accurately represents the posterior (Figure 5.2).

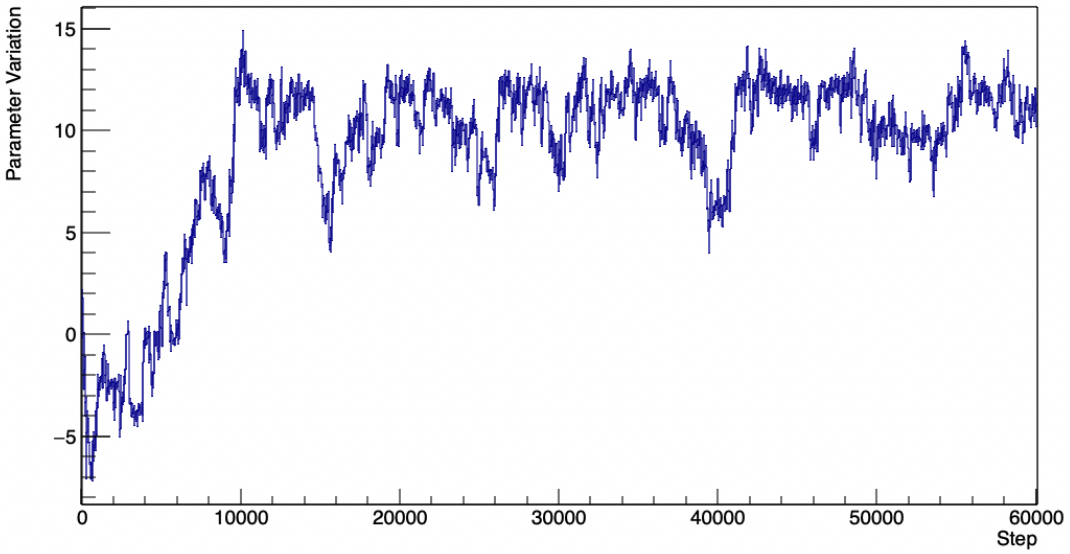


FIGURE 5.2: The MCMC step trace for a CCQE binding energy parameter (detailed in Chapter 6, Section 6.7.4). The first 60,000 steps are shown, the initial 10,000 steps show the chain as it attempts to reach a stationary state. The central value and uncertainty of the parameter is evaluated by considering the density of the steps post burn-in.

A Markov chain is said to have converged when all of the subsequent steps x_{n+1} cohere with x_n and there appears to be no significant displacement from this point after subsequent steps. A sequence of steps will become a true Markov chain that converges if they are:

1. **Recurrent** - All steps beyond the burn-in point end up sampling from the stable distribution that was reached. Additional steps now end up giving a better representation of this stable distribution.

2. **Aperiodic** - The steps must not form a closed loop such that the same steps end up being traversed periodically.

3. **Irreducible** - To avoid being stuck in local minima, every point within the probability space has a non-zero probability of being selected.

The main challenge with a Markov chain is reaching a steady distribution from a random walk. In MaCh3, the walk is directed to some extent and is semi-random. This is achieved by using the Metropolis-Hastings algorithm to yield a Markov chain which converges to give a stable distribution.

5.4.1 Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm [107] (as known now) was originally proposed in a paper by a team of five researchers in 1953. The discovery is usually attributed to N. Metropolis, however

recent findings [108] indicate that it was another one of the sub authors, Marshall Rosenbluth, and his wife, Arianna Rosenbluth, who did the majority of the work on the algorithm. The algorithm was extended by W. K. Hastings in 1970 for more general cases, and since then has become widely used.

The algorithm is used to help a Markov chain reach a stationary distribution and converge. The algorithm directs a semi-random walk in parameter space, by looking at the surroundings of the parameter space and proposing a next step which is then evaluated. It achieves this by following the steps below:

1. **Initialising** - Each parameter is initialised to its current value. If starting, the initial seed value is the prior value of the parameter.

2. **Proposing** - A proposal is made for each parameter based on a proposal function which is a multivariate Gaussian with the parameter value at the initial step as the central value, and prior uncertainties as the width multiplied by a scaling factor. This scaling factor is optimised for each parameter.

3. **Accepting** - Instead of generating a new MC prediction and total likelihood for the proposed step, which would be computationally too intensive, the Monte Carlo prediction is 're-weighted' and this is used to calculate a sample likelihood. Re-weighting is the process by which the effect of each parameter on the prediction is captured by a spline response function. Changing the parameter itself by some amount changes the prediction by some amount. This allows one to scale the MC prediction for multiple parameter values instead of generating it again from scratch.

   The acceptance probability for the proposed step is the ratio of the posterior probabilities for the initial and proposed step, this is the case because the proposal functions are symmetric. The probability of acceptance, α, for the next step is given by:

   $$\alpha = \min\left(1, \frac{P\left(\vec{x}_p \mid D\right)}{P\left(\vec{x}_i \mid D\right)}\right), \quad (5.9)$$

   where x_i is the initial step for all the parameters and x_p is the proposed step for all the parameters. A random number r is generated from the distribution [0, 1] which is uniform. This is then compared to α. For the proposed step to be accepted, the condition is that r $\geq \alpha$. If this holds, the proposed step is accepted and the process moves to Step 4. If it does not hold, the proposed step is rejected, the initial step is kept, and the process starts again from Initialising.

4. **Repeat** - The process executes Step 2 and 3 again. This is iterated N times.

When a next step has an improved likelihood, the evaluation is to always accept. If the next step does not have an improved likelihood, then it is less likely to be accepted. However, a non-zero probability always remains for acceptance as this ensures that steps do not become trapped in local minima.

By analysing the high-dimensional posterior and using MCMC methods to step in the multidimensional parameter space, the parameters of interest, such as δ_{CP} and the mixing angles can be determined by taking the results of the MCMC walk (also called 'chain') in the parameter space. As mentioned before, the initial steps of the walk (called the burn-in) are discarded as they may not provide an accurate representation of the posterior. After a large number of steps, usually in the millions, it is assumed that the chain has sufficiently sampled the posterior distribution, such that the distribution of points in the chain are proportional to the probability distribution of the posterior. Figure 5.3 below shows the above steps in one image to demonstrate how the MCMC and specifically the Metropolis-Hastings algorithm works.

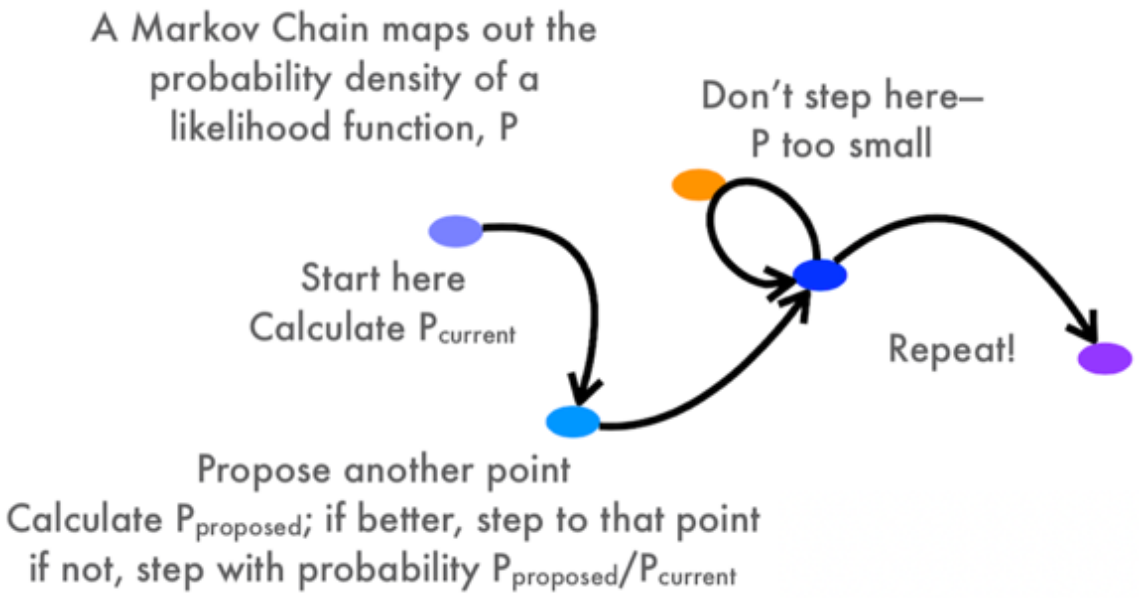


FIGURE 5.3: An explanation of the MCMC Method (T2K Internal, Image by A. Kaboth).

The parameter space that MCMC methods are employed in consist of ∼700 model parameters in total, this includes the flux, detector (both near and far), cross-section and oscillation parameters. A post-fit covariance matrix is used to capture all of the correlations between these parameters. By projecting the multidimensional posterior onto 2 dimensions and creating a matrix, the relationship between the different parameters can be seen (see Figure 5.4 below). Furthermore, the relationship between the ND280 and the Super-K detector allows for some of

the cross-section parameters Super-K shares to be constrained by using the posterior from the fit at ND280.

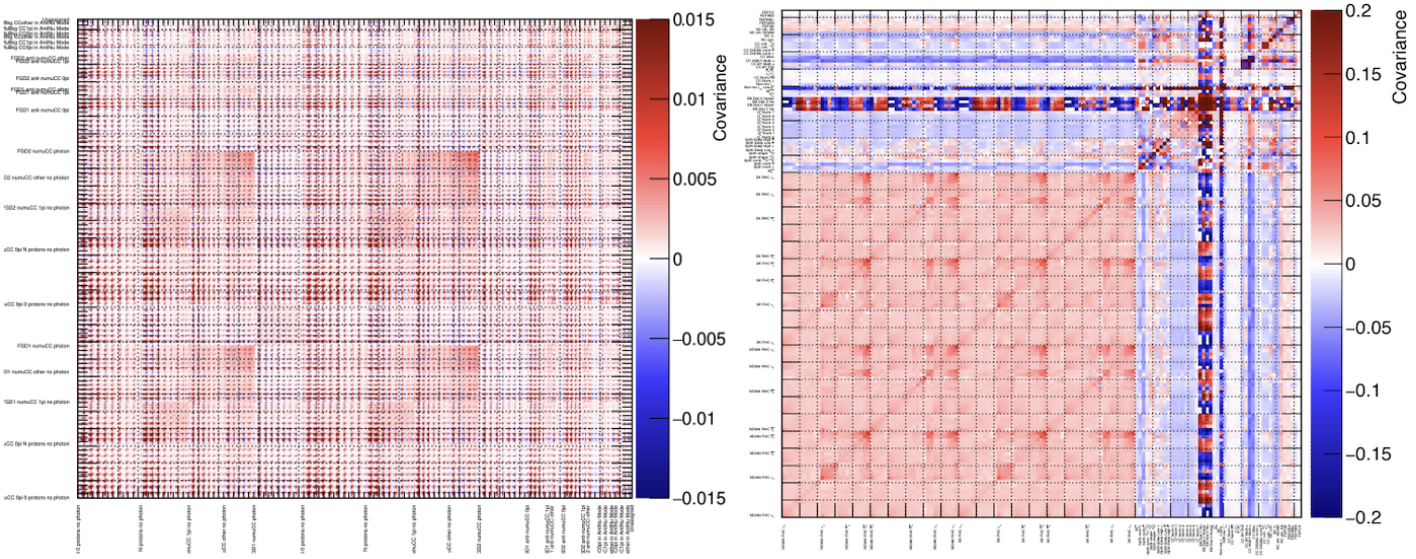


FIGURE 5.4: Left: The oscillation analysis 2021 near detector covariance matrix showing the covariance between ND280 near detector parameters (sample bins) [109].
Right: A post-fit covariance matrix showing the Super-K far detector parameters (in red, bottom left corner), and the cross-section parameters (top left). [T2K Internal]
The two matrices show the majority of the non-oscillation parameters in a fit (flux parameters, $\sim$100 are not shown).

The fitting process mentioned above is carried out in order to determine the best estimate for the oscillation parameters. By marginalising the sampled posterior and integrating over the many non-oscillation parameters, one can extract the distributions for sought after oscillation parameters such as $\sin^2\theta_{13}$, $\sin^2\theta_{23}$, Δm^2_{32}, and δ_{CP}.

5.4.2 Marginalisation

In MaCh3, the near detector and the far detector are fit simultaneously using MCMC methods. The end result is a multidimensional posterior distribution of all the oscillation and non-oscillation parameters. The non-oscillation parameters are called nuisance parameters, and include near and far detector parameters such as sample bins, flux parameters, and cross section parameters (detailed in Chapter 6, Section 6.7.4). The multidimensional posterior is then marginalised so that the oscillation parameters can be found.

Marginalisation can be understood as the method by which to extract the central value and width of parameters from the posterior, allowing us to better understand their behaviour. Given the fact that the distribution is multidimensional, all the other dimensions - except the dimensions (parameters) to be extracted - are integrated over their whole space. This means that only the

parameters of interest remain as a distribution from the posterior. For a parameter p from the model θ, the marginalisation integral given the data D is expressed as:

$$P(\theta \mid D) = \int P\left(\vec{\theta'}, p \mid D\right) d\vec{\theta'}\,. \tag{5.10}$$

$\vec{\theta'}$ are all the model parameters except p. Figures below show the 1-dimensional plots for two example parameters showing their central values and widths calculated using different methods in Figure 5.5 below.

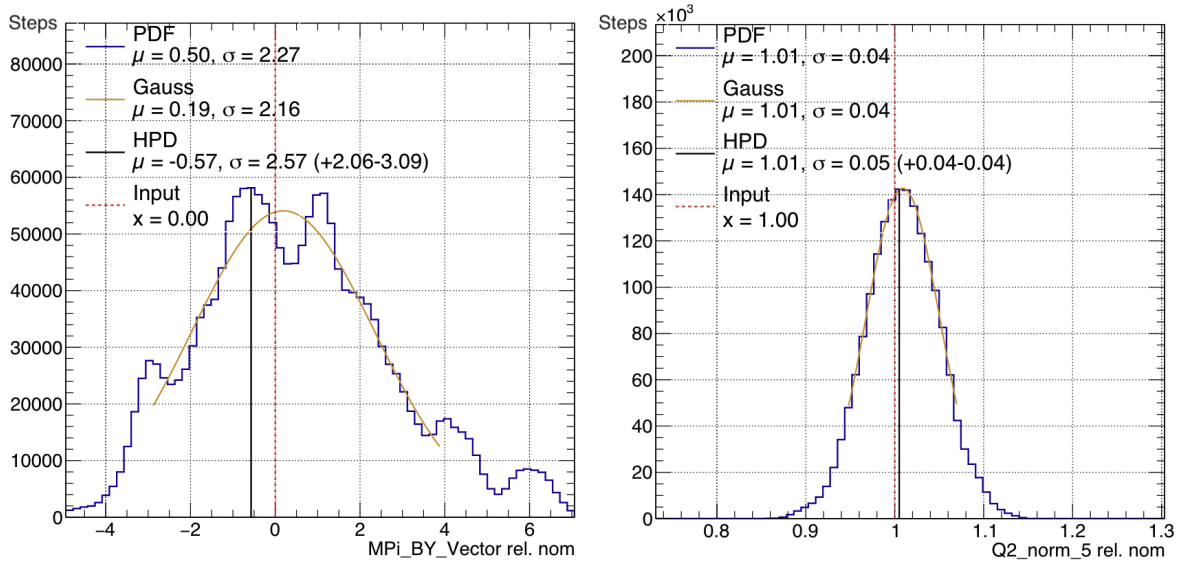


FIGURE 5.5: An example of two cross-section parameters (post-fit) from the marginalised posterior. The red dotted line shows the prior value, whereas the golden line shows the fitted Gaussian distribution and the black line shows the highest posterior density.

The three methods for working out the central value and width for the parameters are: 1) the mean and RMS value, 2) a Gaussian fitting method where the central value is the mean and the uncertainty is the width, 3) HPD (Highest Posterior Density point) which is the mode of the histogram. The three methods are usually all similar for Gaussian like distributions as seen in Figure 5.5 (right).

For non-Gaussian distributions the three methods give differing outputs as shown on the left-hand side of Figure 5.5. This may seem counter-intuitive at first glance, as some methods (such as the Gaussian fitting method) give the most probable value of the parameter in a region where there is non-maximal probability as seen in Figure 5.5 (left). However, this is not problematic in itself as it is something true about the parameter as it is. The important point is that this non-Gaussian nature must be taken into account when interpreting other parameters. Furthermore, by understanding that only the full ~ 700 parameter posterior is what accurately

represents the results, and that the 1-dimensional plots are only a partial picture of what is happening, it is not a cause for concern as the full posterior is what is used in the fit.

These 1-dimension parameter post-fit plots are also useful because they can be compared to results from other fitting groups within T2K, such as BANFF. Validation is carried out between the two fitting groups to confirm results and study anomalies if they arise. Some differences may be expected due to the fit mechanics and the marginalisation techniques of MaCh3. Outside of T2K however, a comparison is hard as different experiments may parameterise their likelihood in a different manner, therefore reducing the possibility for correct benchmarking.

5.5 Values and Uncertainties for Oscillation Parameters

In order to find a best fit for the oscillation parameters from the multidimensional posterior, the mode of the posterior distribution is taken as this corresponds to the most probable point. In MaCh3 this is done by producing a series of binned intervals across the parameter space which capture information about the posterior density. The bin which contains the greatest number of points is therefore the mode. However, to ensure that there are a sufficient number of points within each bin to yield a mode, the steps within the MCMC chain must be significantly large.

There are 4 oscillation parameters T2K is sensitive to, and when the posterior is marginalised for the parameters of interest, a 4 dimension posterior remains. Finding the mode of a 4D posterior using the above method can become difficult. The number of bins needed in a 4 dimensional histogram increases significantly. Therefore, a kernel method is used to turn these series of points within bins into a smooth distribution. The maximum of this distribution can then be found using gradient-descent algorithms such as MINUIT. This yields a best fit point in 2D space (for example $\sin^2\theta_{23}$-Δm^2_{32}), when marginalising further to 1D it yields a central value.

An uncertainty estimate must also be provided, this is carried out by producing credible intervals within the posterior distribution. This is seen on a 'contour' plot in MaCh3 when looking at oscillation parameters in 2D (see Chapter 7, Section 7.5). A 90% credible interval, for example, is defined as the interval across which the integral of the marginalised posterior yields 90% of the whole distribution. This can be written as the integral of the marginalised posterior distribution which gives a probability of 0.9. There are naturally many ways to define this interval in a 4 dimensional space for a 90% credible interval, and each can be valid. However, in this analysis the highest posterior density (HPD) is used to determine the credible interval.

Chapter 6

The Near Detector Selections and Systematic Sources

The Monte Carlo prediction that is generated at the near detector alongside the observed data is categorised into certain selections. This is done by applying cuts on MC and data events, specifically on particle information, co-ordinate locations, and sub-detectors such that events are apportioned into selections of interest (or excluded). These selections of interest (also known as samples) are then binned in lepton kinematics (p, $\cos\theta$), where p is the magnitude of the momentum of the lepton, and θ is the angle that the outgoing lepton makes to the incoming neutrino axis.

These samples are then fit using the Markov Chain Monte Carlo process as described in Chapter 5 previously. After each sample is fit to data, the optimal model parameters which were varied from their input values to result in the final fit are then extracted. These samples and their definitions will be further detailed in this chapter, along with the model parameter inputs and their centrality to the near detector fit.

6.1 Motivation

As mentioned prior, the MC prediction and the data are all binned in lepton kinematics. This is where the mechanics of fitting takes place, as the binned prediction can be compared to the binned data. The Oscillation Analysis group generates a model prediction and uses data to constrain it. The log likelihood (mentioned prior) is then used as a test statistic to determine the goodness of fit of the prediction to the data for all the samples.

The near detector constraint allows T2K to make world renowned measurements of the neutrino oscillation parameters. In order to achieve this there are inputs to the model which are considered alongside further detector details. These can be understood by looking at the prediction for the number of events at the far detector given by Equation 6.1 (mentioned prior in Chapter 5, Section 5.1):

$$N_{SK}^{\alpha\to\beta}\left(p_{\text{reco}}\right) = \sum_{i} \Phi_{\alpha}\left(E_{\text{true}}\right) \times \sigma_{\beta}^{i}\left(p_{\text{true}}\right) \times \epsilon_{\beta}\left(p_{\text{true}}\right) \times P_{\alpha\beta}\left(E_{\text{true}}\right). \tag{6.1}$$

$N_{SK}^{\alpha\to\beta}\left(p_{\text{reco}}\right)$ represents the Super-K predicted event rate as a function of lepton kinematics. i represents the interaction type which is the index that is summed over. $\Phi_{\alpha}\left(E_{\text{true}}\right)$ represents the neutrino flux as a function of the true energy. $\sigma_{\alpha}^{i}\left(p_{\text{true}}\right)$ represents the neutrino cross section as a function of momentum for different interaction type i and neutrino flavour α which, when convoluted with the neutrino flux, provides the base number of neutrinos that interact in the detector. To account for the fact that only a percentage of these events are detected, $\epsilon_{\alpha}\left(p_{\text{true}}\right)$ represents the efficiency of the detector.

All of these parameters are interrelated with each other as shown by the covariance matrix (displayed in Chapter 5, Figure 5.4). Therefore, a variation in one of the systematic parameters affects the parameters of interest. This means that systematic parameters must be constrained well, without the parameters of interest (oscillation parameters) being influenced; this is one of the reasons for a near detector at T2K in order to individually constrain systematic sources.

These parameters affect the log-likelihood as mentioned and consist of beam, near detector, and cross section parameters which come from T2K working groups: the Systematics, Selections and Validations (SSV) group provides the ND280 detector model parameters, with the beam group providing the beam parameters, and the Neutrino Interaction Working Group (NIWG) providing the cross-section parameters from their parameterised model of neutrino interactions. A far detector prediction requires the Super-K detector model parameters. Since T2K and Super-K (SK) are separate collaborations, the T2K-SK working group is assigned to provide these model parameters to T2K along with the selections.

Certain selections are specifically picked and developed in order to constrain the beam and the cross section parameters. These selections and their definitions will be discussed below.

6.2 Selections

It is important that the selections which most constrain the systematic parameters such as the beam and the cross section are used. This is to ensure that the oscillation parameters are

better constrained, which is the objective of the fit. Given this approach, the Charged Current (CC) neutrino interactions become the interactions of interest, as they are the easiest to model alongside capturing information about the flavour of the neutrino. Alongside this, the CC interactions leave a clean signal of a muon track within a TPC sub-detector which neighbours FGD1 or FGD2. The cuts made to the reconstructed data and prediction are intentionally designed to capture as many charged current interactions as possible.

There are 3 key interactions that fall under charged current (CC) interactions. Going from the easiest to model to the hardest to model, these are: charged current quasi-elastic (CCQE), charged current resonant single pion production (CCRES), and charged current deep inelastic scattering (CCDIS). As mentioned, all of these interactions contain the important signal of a charged muon in the final state. However, the three interactions may be distinguished by the number of emitted pions (0 to N) that are found in the final state. Therefore, considering the underlying interaction theory, splitting event types by the number of pions in the final state yields an approach at the experimental level to categorise neutrino interaction events under event topologies: with CC0Pi (0 pions in final state), CC1Pi (1 pion in final state), and CCOther (multiple pions in final state) being mapped. This way of categorising by topology is used globally within neutrino physics experiments and allows for selections to be developed, producing 3 samples that are enhanced in CCQE, CCRES and CCDIS interactions in a model independent way respectively.

However, this above mapping of topology to interaction is not 100 % accurate. For example, there maybe a small number of events which end up being reconstructed as CC0Pi but had a final state pion which was absorbed before it could be emitted. This is due to factors such as Final State interactions which must be considered and why parameters incorporating these interactions are included within the model (see Section 6.7.4).

The different topologies also categorise samples by a presence of a photon. If a photon is not present, then samples are categorised by the number of pions in the final state. The CC0Pi sample (with no final state photon) has further been split by considering if protons are present in the final state. This is all in effect to populate a kinematic space in which underlying interactions can be better understood, and a space in which these interactions can easily, and accurately, be modelled by the MC prediction. The purer the sample, the better the modelling and application of systematics will be.

The near detector selections are broadly split by neutrino (FHC) mode and antineutrino (RHC) mode. FHC and RHC modes have slightly different systematic parameters associated with them. They are further split by considering the sub-detector that lepton kinematics are recorded in; FGD1 and FGD2. This is because FGD1 is made up of solid scintillator plates, whereas FDG2

has layers of water in between scintillator plates. This means that FGD2 additionally looks at interactions on oxygen, and therefore the two different FGDs have different event reconstruction methods and are treated by different systematic parameters.

The FHC selection and RHC selection hierarchy is shown below in Figure 6.1 and Figure 6.2, yielding a total of 22 selections. There are 10 selections for FHC and 12 selections for RHC mode.

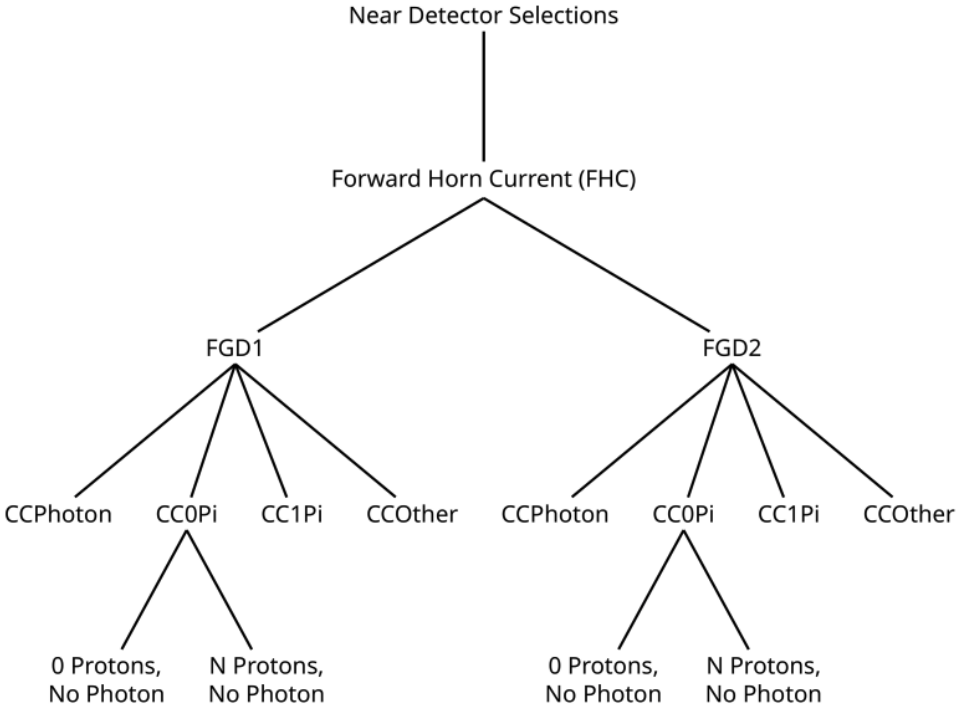


FIGURE 6.1: Selection map for FHC showing the different near detector samples split by pion and photon final states: CCPhoton, CC0Pi, CC1Pi and CCOther.

The FHC samples are split by photon presence; this captures more of the hard-to-model CCDIS and CC1Pi background π^0 events, therefore providing the no photon samples with a higher purity of π^+ final state events from true CCQE and CCRES interactions. Furthermore, separating by protons helps to improve the CCQE and 2p2h interaction predictions in FHC [110]. Therefore, photon proton samples significantly help to constrain the far detector multi-ring sample (see Section 6.6).

It is important to evaluate the wrong-sign contamination (i.e $\bar{\nu}$ in FHC, ν in RHC) at ND280 as there is no magnetic field at the far detector (Super-K) to determine if the neutrinos arriving are neutrinos or antineutrinos. Evaluating this contribution is important as it allows for a prediction at the far detector of the number of rings formed by wrong-sign neutrinos for FHC and RHC modes respectively.

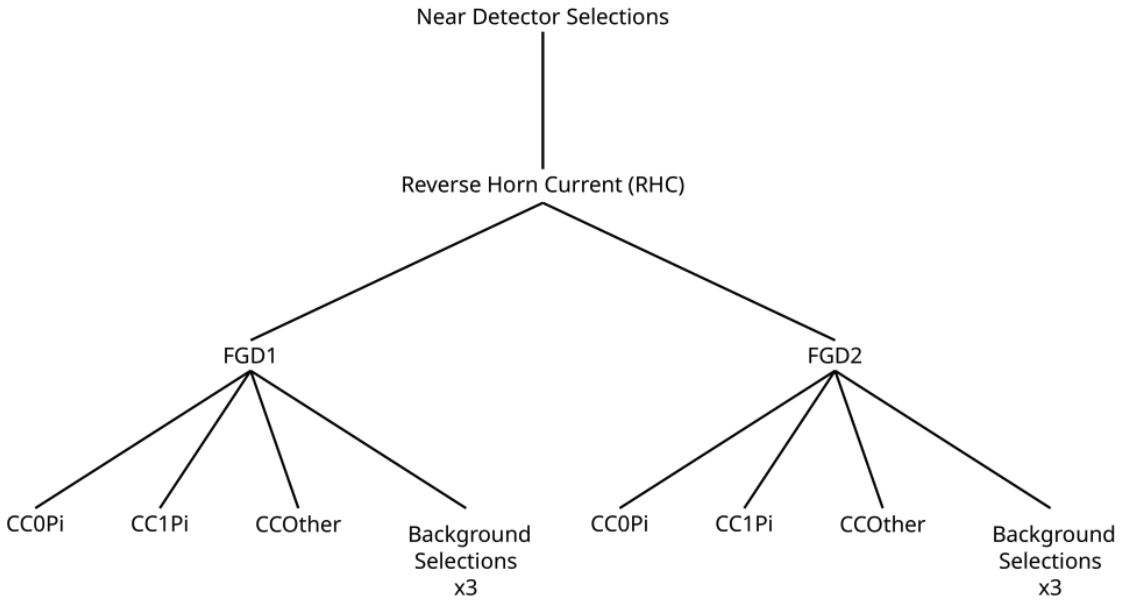


FIGURE 6.2: Selection map for RHC showing the different near detector samples split by pion final states: CC0Pi, CC1Pi and CCOther. There are additional background selections which consider the beam mix with a sizeable percentage of muon neutrinos in RHC anti-neutrino mode that also interact. These are also split by pion topology.

In this analysis, runs 2 - 9 were used, with the FHC runs being focused on as shown in Table 6.1.

TABLE 6.1: Table showing run number and type, along with the protons on target (POT) for Data, MC and MC Sand [109].

T2K Run	Data POT/10^{20}	MC POT/10^{20}	MC Sand POT/10^{20}
Run 2 Air	0.360	16.802	10.730
Run 2 Water	0.434	12.038	10.730
Run 3 Air	1.593	30.780	10.730
Run 4 Air	1.789	36.122	10.730
Run 4 Water	1.695	36.122	10.730
Run 5 Water	0.445	22.110	12.025
Run 6	3.422	34.698	12.025
Run 7	2.439	33.310	12.025
Run 8 Air	4.150	44.532	10.730
Run 8 Water	1.581	27.168	10.730
Run 9	2.303	26.607	12.025
Total	20.211	320.289	123.210

The MC Sand rate is from neutrino interactions with the sand in the pit outside ND280, the results of which end up being detected within the ND280 (fiducial volume). This forms an undesirable background (sand muons [111]) when analysing events.

6.3 FHC ν_μ Selection Criteria

The selection criteria are intentionally designed to capture as many charged current (CC) interaction events as possible. There are 5 different samples under the FHC selection for both FGD1 and FGD2. The events must occur in either of the FGDs, with a track identified as a forward going negatively charged muon which then crosses the neighbouring TPC. These events also have a quality threshold they must meet in order to be considered, alongside the highest momentum negative track from the event being reconstructed as a muon. The corresponding selection cuts are as follows:

Event Quality - The event must be part of a beam spill under a time period that is flagged as good in terms of data quality. This is only the case for data events as MC events are all assumed to have good data quality. The data quality flag for the ND280 sub-detectors is mentioned in detail in Chapter 4.

Total Multiplicity - The event must have at least one reconstructed track that crosses a TPC. The event reconstruction algorithm works first by clustering all the hits within the ND280 sub-detectors and combining these to form tracks. This information consisting of tracks and hits are then fitted to yield certain particles and interactions in that event.

Fiducial Volume - Reconstructed tracks must originate within the fiducial volume (FV) of FGD1 or FGD2. The fiducial volume is the volume within the FGD where the most accurate reconstruction occurs. Furthermore, one reconstructed track must have a segment present across at least one FGD and TPC. The FGD1 and FGD2 fiducial volumes are defined below:

$$\text{FGD1} => \ |x| < 874.51 \text{ mm}; -819.51 < y < 929.51 \text{ mm}; 136.875 < z < 446.955 \text{ mm}\,,$$

$$\text{FGD2} => \ |x| < 874.51 \text{ mm}; -819.51 < y < 929.51 \text{ mm}; 1481.45 < z < 1807.05 \text{ mm}\,,$$

The cuts in the XY plane serve to remove the edges of the sub-detector within 5 bars where sand muons (due to the interaction between the ND280 pit and neutrinos) are detected; this is also where the reconstruction is at its weakest. The Z axis cuts make sure that the FGD of interest is selected, alongside removing the first upstream module of each FGD.

Upstream Background Veto - Events with tracks that start outside of one of the FGD fiducial volumes may be reconstructed improperly to seem as though they are within the FGD, this cut starts by looping over all TPC tracks and finding the highest momentum secondary track (which may be a pion for example). For FGD2, events with a secondary track starting in FGD1 or more than 150 mm upstream from the muon candidate (the primary track) are then rejected.

Broken Track - When a track for the muon candidate appears in the last few layers of an FGD before appearing in the neighbouring TPC, the track can be mis-reconstructed as two separate tracks and thus is termed 'Broken'. This first part of the track is disregarded, with the second part considered the muon candidate. Any event with a muon candidate track that starts 425 mm within the FGD upstream edge is rejected given that there is also one contained FGD track present.

Muon Particle Identification - The particle identification procedure for the muon within the TPC is mentioned in Chapter 3.5.2, where the dE/dx value is compared to what would be expected for different particle hypotheses. The likelihood hypothesis for the particles is a function of the *Pull*, for particle p it is given by:

$$Pull_p = \frac{\mathrm{d}E/\mathrm{d}x_{\text{measured}} - \mathrm{d}E/\mathrm{d}x_{\text{expected},p}}{\sigma_{\left(\mathrm{d}E/\mathrm{d}x_{\text{measured}} - \mathrm{d}E/\mathrm{d}x_{\text{expected},p}\right)}}. \tag{6.2}$$

Where σ is the standard deviation, p represents a candidate particle such as a muon, pion, electron or a proton. These four particles are the most relevant particle hypotheses that are considered. This set of candidate particles can be called *cp*. The likelihood is then given by:

$$L_p = \frac{e^{-Pull_p^2}}{\sum_{cp} e^{Pull_{cp}^2}}. \tag{6.3}$$

Only reconstructed tracks which have a momentum of less than 500 MeV/C are considered. Electrons are then cut out by using:

$$\frac{L_\mu + L_\pi}{1 - L_{pr}} > 0.8. \tag{6.4}$$

Protons and pions are then discarded by looking at the likelihood for muons and setting the cut at $\mathrm{L}_\mu > 0.05$.

The FHC samples are further divided by photon presence, where a cut is made on variables in order to determine whether a photon is present in the interaction. This is done using the TPC π^0 tag in conjunction with the ECal which detects photons that convert from signal channels [112]. If there is no photon present, then the sample is further divided by pion multiplicity. Positively charged pions are seen for FHC runs, negatively charged pions are seen for RHC runs.

Pion PID - The number of secondary tracks of a pion candidate identified in a TPC determines the amount of charged pions. Pion and electron hypotheses are tested for negatively charged tracks, whereas the proton, pion and positron hypotheses are tested for positively charged tracks. These are computed using the same method with pulls as mentioned earlier. The cut for rejecting protons from the sample is:

$$\frac{L_\mu + L_\pi}{1 - L_{pr}} > 0.8, \tag{6.5}$$

for reconstructed tracks which have a momentum of less than 500 MeV/c. Muons are discarded by setting the likelihood for pions at $L_\pi > 0.3$. Equation 6.4 mentioned prior has the same expression, however the difference is that specific particles are cut by restricting the value of a certain particle likelihood in the expression (such as L_μ or L_π).

6.3.1 The CCPhoton Sample

The CCPhoton sample apportions a majority of events with π^0 particles decaying to photons. π^0 particles mostly decay to two photons, the photons either form a pair of particles that are detected in the TPC, or they can interact within the ECal. Looking at the ECal and the TPC activity following this decay allows for this sample to be cut and selected. The main benefit of the sample is that being π^0 rich, it improves the purity of the remaining samples allowing the CCPi samples to focus on negative and positive pions. This is greatly advantageous at the Super-K far detector for studying multi ring (MR) samples related to charged pions.

6.3.2 The CC0Pi Sample

The first pion multiplicity related sample as seen in Figure 6.1 is the CC0Pi no photon sample. From the FHC selection, those events which pass the cuts are then cut by Pion PID. CC0Pi no photon events have a single muon track but no pions, positrons or electrons identified within the TPC or within the FGDs. The majority of this sample contains CCQE signal events which are important as the neutrino energy can be reconstructed through lepton kinematics.

The CC0Pi no photon sample is then split by the multiplicity of protons an event contains (either 0 or N protons). This selection was developed for the 2021 oscillation analysis (OA2021) in order to probe the 2p2h and CCQE interactions. Proton candidates are found by looking at tracks which are not categorised as a muon or a pion by the cuts above, and where $L_{pr} > 0.5$. If there is no reconstructed proton, then the event is classified in the CC0Pi 0 protons no photon, otherwise it is classified as CC0Pi N protons no photon as seen above.

6.3.3 The CC1Pi Sample - Split by Pion Tag for FHC

The CC1Pi no photon sample contains one pion in the final state along with the muon signal. The new final state selection developed in this thesis takes advantage of various methods to tag pions within the CC1Pi sample. By using these tags, the muons themselves can be tagged by sub-detector type by looking at the pion in the final state. This tag can serve as a proxy for pion kinematics. Pions are tagged by ND280 sub-detector type, but this had not been studied prior, the 3 tags are: TPC pions, FGD pions, and Michel pions.

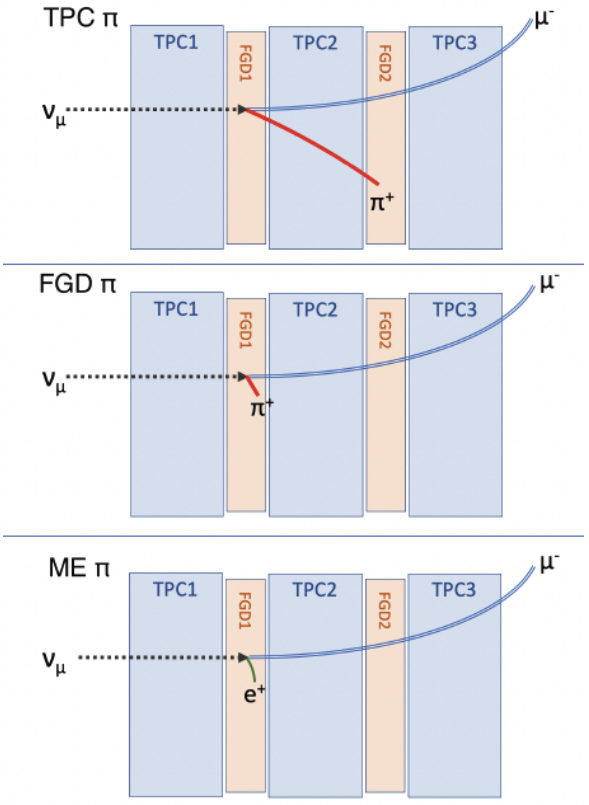


FIGURE 6.3: The different pion tag definitions illustrated within the sub-detectors. The highest momentum negative track of the muon can be seen (in blue) with a secondary track associated with it (in red, green). (T2K Internal, Image by J. Walsh).

Figure 6.3 shows that the TPC tagged pion will make it out of the FGD as it has a higher momentum. The last two tagged pions; the FGD and the Michel pion are much lower momentum and remain in the FGD. Two methods are used in order to distinguish between the types:

FGD Reconstruction - Self-contained FGD tracks are sought for a FGD pion candidate or Michel pion candidate. Therefore, tracks which do not meet this condition are rejected; this includes the highest momentum negative track starting in a different FGD. The candidate tracks must also be in the same bunch window as the muons in order to be considered as pion candidates. Protons are distinguished from pions by looking at the energy deposited in the

FGD. A pion has a distinct energy deposited per track length profile which peaks early and falls rapidly, whereas the proton profile peaks later, further down in track length, and falls steadily.

Michel Electron - The low momentum pions that do not leave a sizeable track within the FGD are identified by a secondary electron which is detected. This electron corresponds to a Michel electron which the low momentum pion (Michel pion) decays to after 2.19 μs. The detector classifies delayed activity as activity after 100 ns past the first neutrino interaction. The Michel electron candidate must meet a certain requirement of hits within the FGD, alongside being outside the beam time window (due the delay in Michel pion decay).

Figure 6.4 below shows the distribution of different tagged pions within the CC1Pi sample.

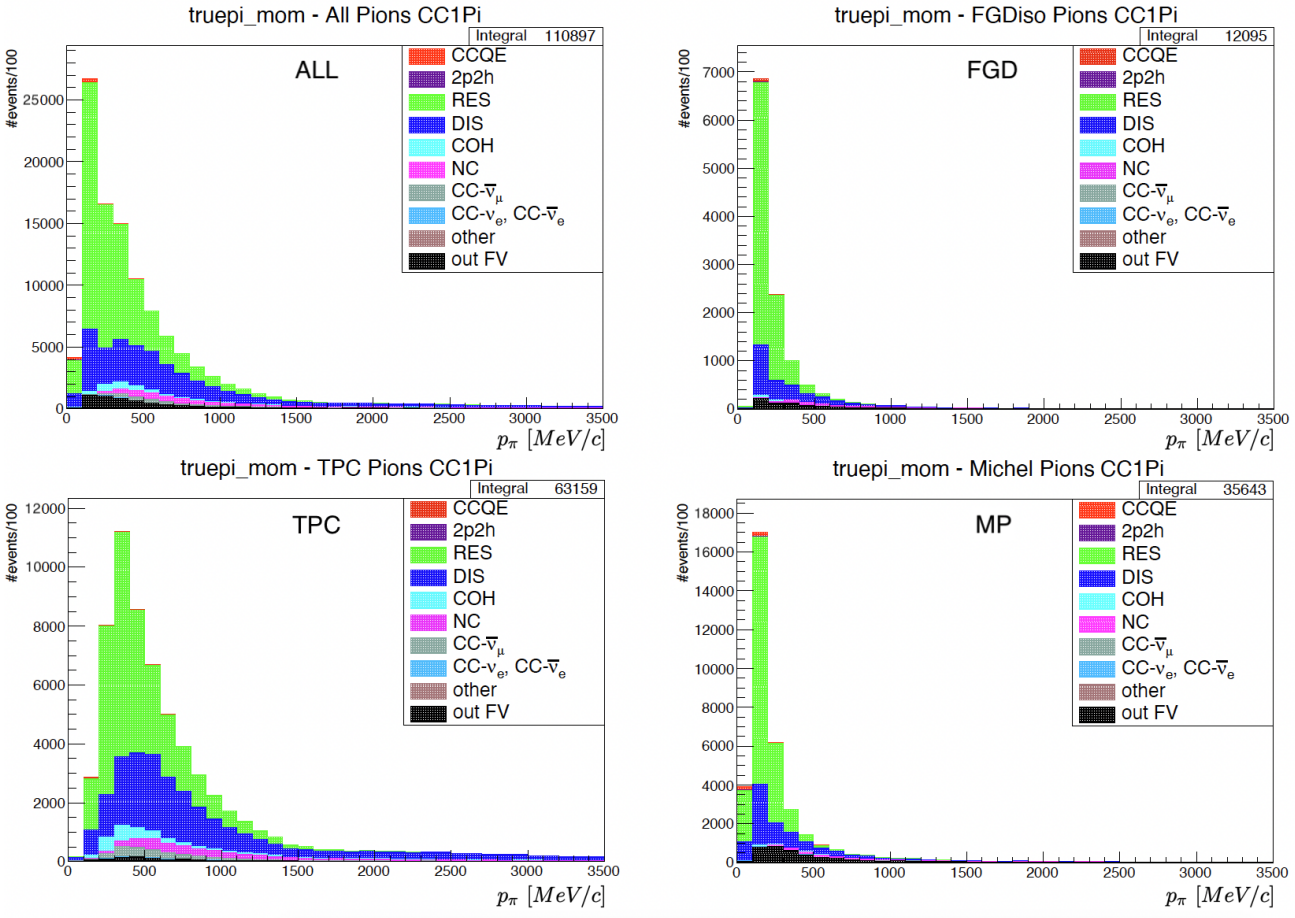


FIGURE 6.4: Histograms of the CC1Pi selection split by pion tag type showing true pion momentum from 0 MeV to 3500 MeV. Legend text – COH: CC Coherent, NC: Neutral Current, CC-$\bar{\nu}_\mu$: $\bar{\nu}_\mu$ scattering, CC-ν_e: ν_e scattering, CC-$\bar{\nu}_e$: $\bar{\nu}_e$ scattering, out FV - out of fiducial volume.

In Figure 6.4 many interesting features can be seen of the pion distribution, TPC pions peak near 400 MeV whereas the FGD and Michel pions have a peak near 200 MeV; this is due to their lower momentum. Michel pions dominate at momentum under 200 MeV. At the peak energy, the dominant interaction is the CC resonant pion production as expected.

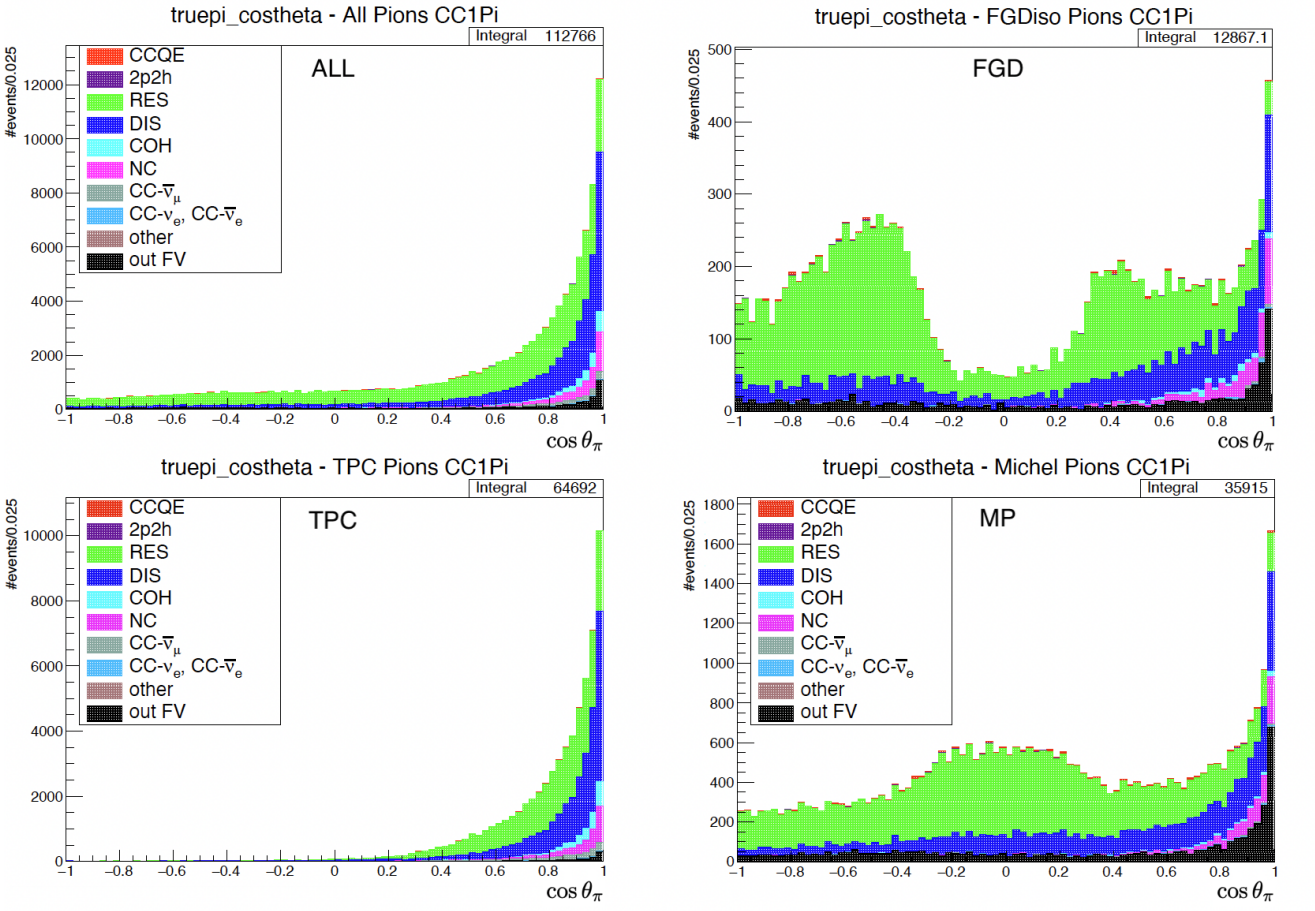


FIGURE 6.5: CC1Pi selection split by pion tag type showing events per $\cos\theta$ bin from -1 to 1.

Looking at the $\cos\theta$ distribution in Figure 6.5, it can be seen that there is a bimodal distribution for FGD pions at -0.3 to 0.3 with an region of minimal events due to the poor reconstruction as near vertical tracks do not leave many hits in the FGD (considering both XY and Z planes). This is similarly the case for vertex reconstruction for Michel pions in determining an accurate value for $\cos\theta$ [113]. The TPC pions become more dominant at lower angles (forward going). Ultimately, the majority of pions are forward going although there are a significant number of backward going pions ($\sim$ 18%). For each tag, the highest number of events are in the highest angle bin with the hard-to-model DIS most interaction dominant there also.

Implementing pion kinematics into the fit is of importance due to new far detector multi-ring samples with CCRES target interactions that have a strong dependence on pion kinematics (Section 6.6). In order to start to leverage this new sample, there must be an intermediate step taken of determining how lepton kinematics change by pion tag type. In this thesis, the pion tag information is incorporated by looking at the associated leptons in the event and relating them to their pion tag. This is achieved by 'stacking' the CC1Pi histogram in lepton kinematics by offsets in $\cos\theta$, with different regions of phase space representing TPC pion-tagged leptons, FGD pion-tagged leptons, and Michel pion-tagged leptons. The selection map is shown in Figure 6.6:

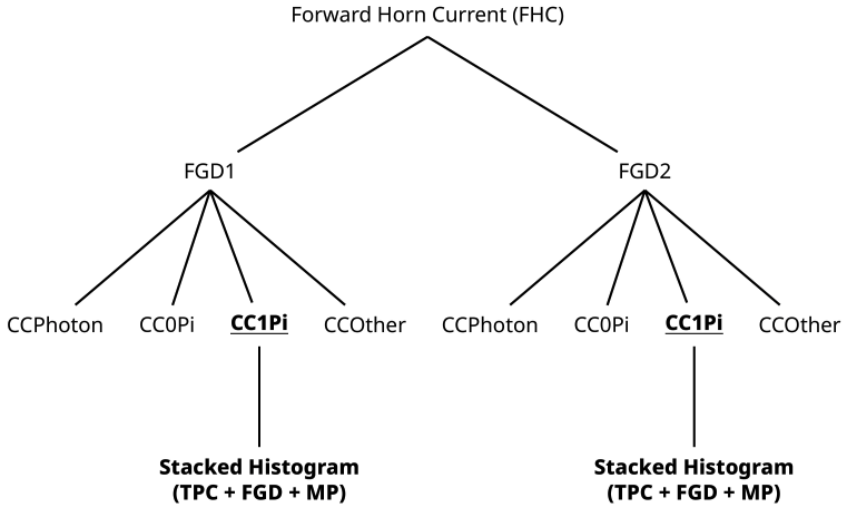


FIGURE 6.6: Selection map for FHC showing the different near detector samples. CC1Pi is shown in detail (bold), now representing a histogram for lepton kinematics split by pion tag.

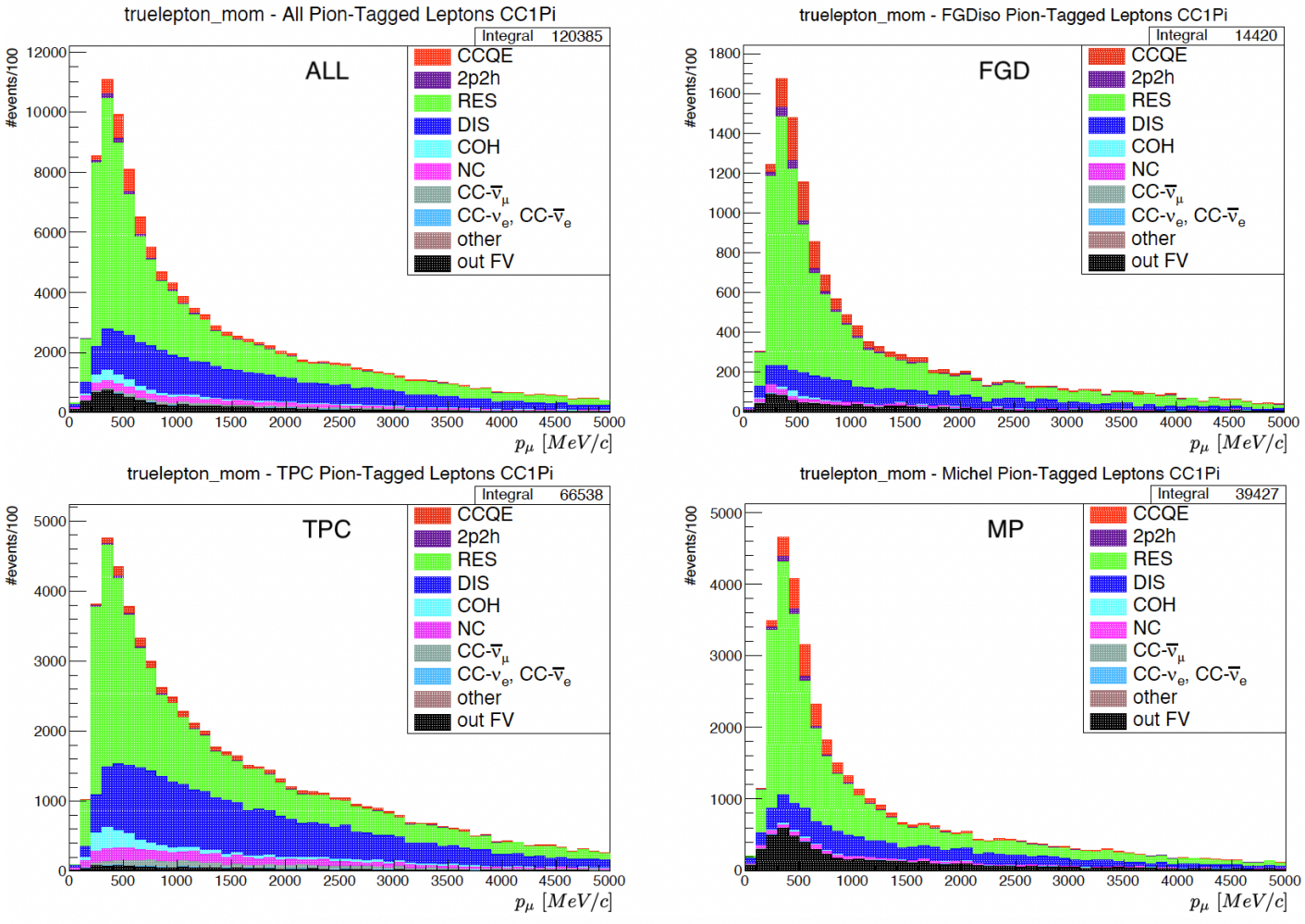


FIGURE 6.7: Histograms of the CC1Pi selection split by pion tag type.

Figure 6.7 shows the individual histograms for number of events per 100 MeV bin in muon momentum for events with the 3 types of pion tag. Events are broken down by underlying interaction type. The biggest contribution to the total number as the peak T2K neutrino energy for CC1Pi is the CCRES interaction, with the CCDIS being an equal contributor at energies greater than 1 GeV.

Looking at the TPC pion-tagged leptons, it can be seen that the CC Coherent interaction events also play a role at the T2K energy; this is not seen as significantly throughout the FGD or the MP distribution. Around the peak T2K neutrino energy there are also a number of CCQE interactions, which are less prominent for the high kinematic TPC pion-tagged leptons, but are more prominent for the FGD and Michel tagged-pions. CCQE interactions are expected to be the dominant interaction mode for CC0Pi, but still show up in CC1Pi due to final state interactions in the model. The Michel pion-tagged leptons show a large proportion of out of fiducial volume events around 0.6 GeV which is expected due to their having low momentum.

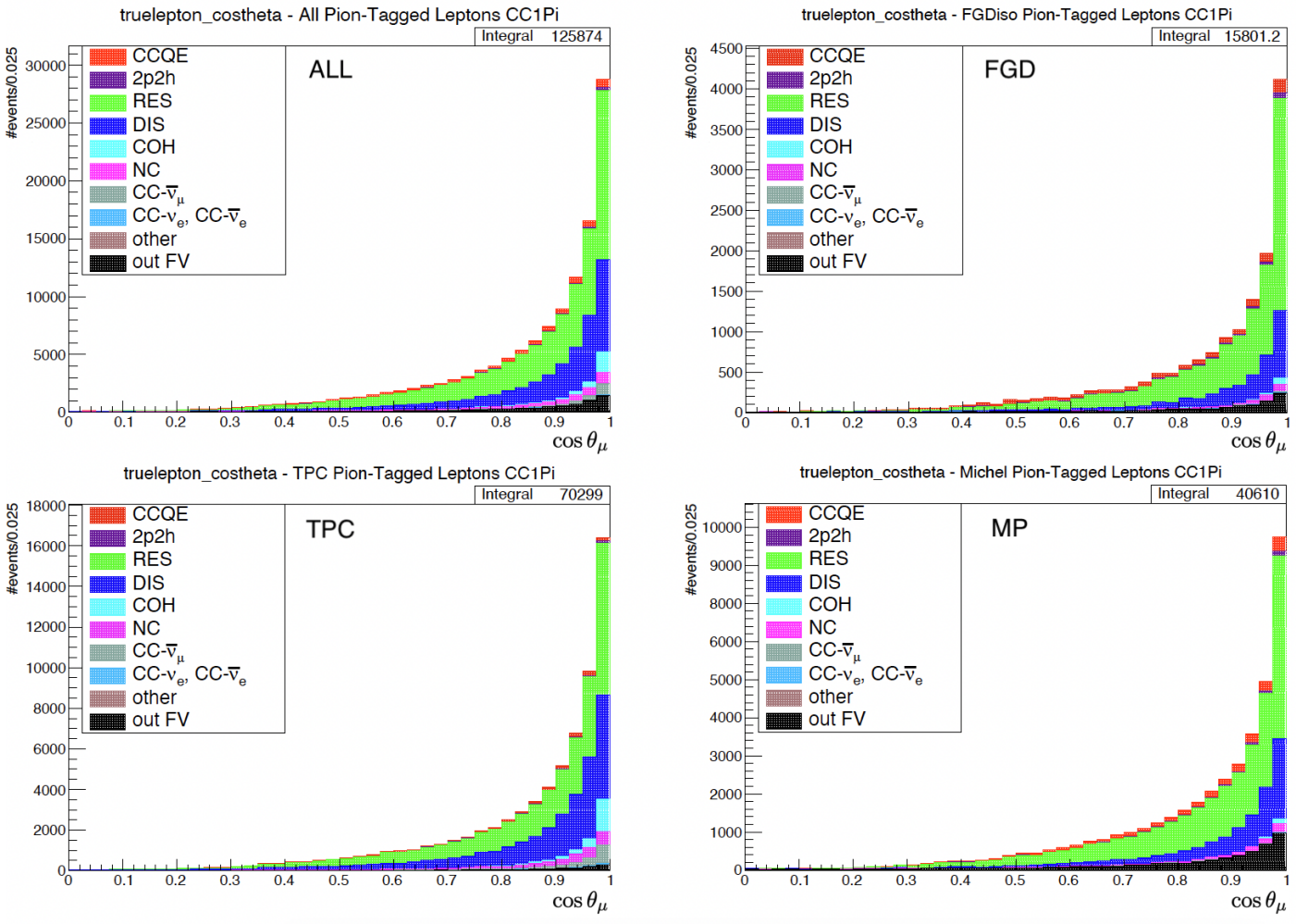


FIGURE 6.8: Histograms of the CC1Pi selection split by pion tag type.

The histogram integrals give the total event numbers in each pion tag sample up to 5 GeV. In Figure 6.7, out of the total 120,385 MC events (100%), 66,538 events are TPC pion-tagged (55%), 39,427 events are Michel pion-tagged (33%) and 14,420 events are FGD pion-tagged (12%). It can be seen that the majority of events making up the CC1Pi sample, 88%, are TPC or Michel pion-tagged events.

Figure 6.8 shows the individual histograms for number of events per 0.025 bin width in $\cos\theta_\mu$. Many interesting features can be seen, the most important is that statistically all of muons are forward going, with $\sim$ 75% deviating less than 45° away from the forward going neutrino axis. TPC, FGD and Michel pion tagged events have $\sim$ 50% of their muons within the $\cos\theta_\mu$ bin 0.9 to 1, with the majority of events of interaction type CCRES.

.FGD and Michel pion tagged events have a greater proportion of muons within the $cos(\theta_\mu)$ bin 0.975 to 1 under the CCRES interaction. It can be seen that this dominates the proportion of CCDIS events compared to the same bin within TPC pion tagged leptons. Ultimately, by studying the distributions shown in Figures 6.4 – 6.8, it can be seen that splitting by pion tag in lepton kinematics serves as an effective proxy for pion kinematics.

6.3.3.1 Data-MC Comparisons for CC1Pi - Split by Pion Tag for FHC

To check that the data and Monte Carlo (MC) showed good agreement for the new sample, plots were generated showing Data-MC comparisons. These plots are shown below in reconstructed momentum and $\cos\theta_\mu$ for different pion tags in Figure 6.9.

It can be seen below in Figure 6.9 that the Data-MC comparisons are in good agreement at the T2K peak energy and beyond. The MC is scaled to data as more MC events are simulated than in the actual data. Therefore, when plotted, the MC event rate needs to be normalised for the amount of POT (Protons on Target). The integral of the data evaluates to a whole number, whereas the MC (having being scaled to data) evaluates to a decimal.

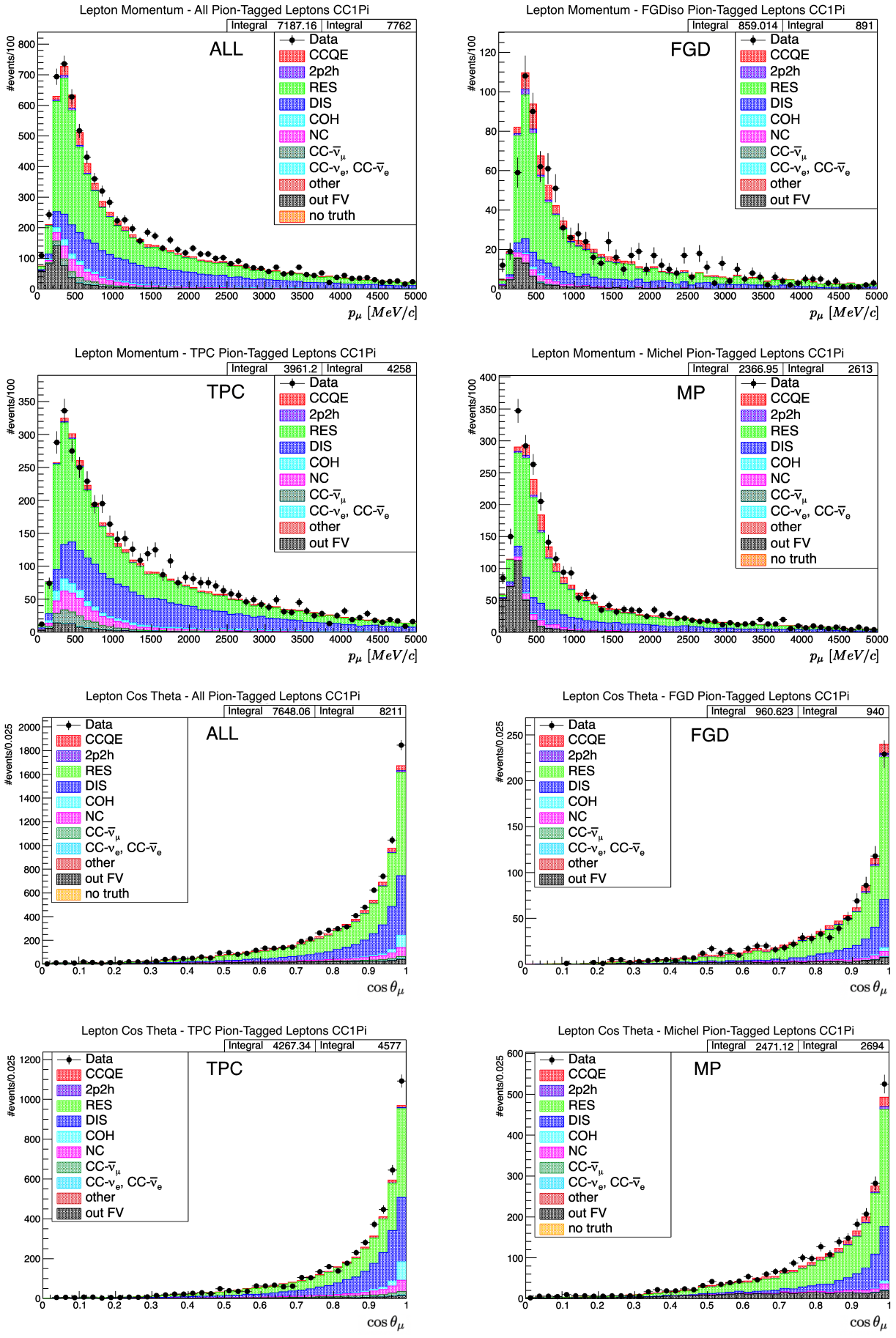


FIGURE 6.9: Data-MC comparison of the CC1Pi selection split by pion tag type.

6.3.4 The CCOther Sample

All non photon events that do not fall into CC1Pi no photon or CC0Pi no photon are apportioned into the CCOther selection which contains events with one or multiple pions and a negative muon candidate track. These pion(s) can be either be negative, neutral or positive reconstructed pions.

6.3.5 Purities and Efficiencies

In order to understand how clean a sample is, there are two measures which can be looked at, one is the purity of the sample, another is the efficiency.

The two metrics are defined below:

$$\text{Efficiency} = \frac{\text{No. of selected signal events}}{\text{No. of generated signal events}}, \tag{6.6}$$

$$\text{Purity} = \frac{\text{No. of selected signal events}}{\text{No. of selected events}}. \tag{6.7}$$

Table 6.2 below shows the purities and efficiencies of the FHC samples.

TABLE 6.2: Purities and efficiencies for the different FHC samples in FGD1 and FGD2 [112].

Sample	Efficiency(%)		Purity(%)	
	FGD1	FGD2	FGD1	FGD2
CC0π	46.86 ± 0.05	47.47 ± 0.05	76.3 ± 0.1	72.8 ± 0.1
CC1π^+	27.06 ± 0.08	23.03 ± 0.08	60.2 ± 0.2	58.5 ± 0.2
CC-Photon	43.0 ± 0.1	43.9 ± 0.1	53.9 ± 0.2	54.2 ± 0.2
CC-Other	21.1 ± 0.1	20.6 ± 0.1	52.2 ± 0.4	50.3 ± 0.4

6.4 RHC Selection Criteria

6.4.1 RHC $\bar{\nu}_\mu$

The RHC selections are selected based on the same criteria as the FHC selection, however for RHC $\bar{\nu}_\mu$, the highest momentum track that is considered is positively charged. As mentioned in Chapter 3, the event rate in RHC has a higher ν contamination than $\bar{\nu}$ in FHC. This is due to the cross section of ν being roughly double that of $\bar{\nu}$. Given this, the cuts for RHC are slightly different to FHC cuts.

Highest Momentum Track - The highest momentum positive track is required to be the highest momentum track in order to ensure that π^+ pions mis-identified as μ^+ muons are reduced.

Upstream Background Veto - Events entering back stream into the FGD1 upstream fiducial volume are rejected, alongside events coming from the PØD or the magnet.

Muon PID - As in the FHC ν_μ selection, the likelihood for the hypothesis of different candidate particles is calculated. These are calculated the same way using dE/dx, the energy deposited per distance within the TPC. However, the value of the cut is set accordingly for the different likelihoods, with the proton rejection cut as below:

$$\frac{L_\mu + L_\pi}{1 - L_{pr}} > 0.9. \tag{6.8}$$

Only reconstructed tracks which have a momentum of less than 500 MeV/c are considered. Muons are then cut by considering that $L_\mu > 0.1$.

As seen in Figure 6.2 the selection is then divided based on pion multiplicity, the procedure for this is equivalent to the procedure used in the FHC ν_μ selection.

6.4.2 RHC ν_μ

RHC ν_μ selection cuts are similar to the RHC $\bar{\nu}_\mu$ and FHC cuts whilst including some additional cuts. Due to the higher interaction cross section of ν_μ on matter, these neutrinos are significant in event number to be accounted for – despite being a background to the RHC selection. The following additional cuts are made:

Highest Momentum Track - The highest momentum negative track is required to be the highest momentum track.

Upstream Background Veto - Events entering back stream into the FGD1 upstream fiducial volume are rejected, alongside events coming from the PØD or the magnet.

Muon PID - As with the RHC $\bar{\nu}_\mu$ selection, the likelihood for the hypothesis of different candidate particles is calculated. These are calculated the same way using the energy deposited per distance within the TPC. However, the value of the cut is set accordingly for the different likelihoods, with the proton rejection cut as below:

$$\frac{L_\mu + L_\pi}{1 - L_{pr}} > 0.7. \tag{6.9}$$

Only reconstructed tracks which have a momentum of less than 500 MeV/c are considered. Muons are then cut by considering that $L_\mu > 0.1$.

This selection is split by pion multiplicity and forms the 'Background Selections' as seen in Figure 6.2.

6.5 Binning

The selections are then binned in lepton kinematics, p (reconstructed momentum) and $\cos(\theta)$, where θ is the angle the outgoing lepton makes to the incoming neutrino axis. The binning is decided based on MC and data event thresholds, with $>$ 20–25 minimum MC events and 1–2 data events. The binning must be detailed and precise enough to describe the regions of interest, whilst being big enough to meet the minimal thresholds for data and MC events.

FHC CC0Pi 0 Protons No Photon
p (MeV/c): 0, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1500, 1750, 2000, 2500, 3000, 5000, 30000.
cos θ: -1.0, 0.5, 0.6, 0.68, 0.72, 0.76, 0.82, 0.84, 0.86, 0.88, 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.955, 0.96, 0.965, 0.97, 0.975, 0.98, 0.985, 0.99, 0.995, 1.0.

FHC CC0Pi N Protons No Photon
p (MeV/c): 0, 250, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1000, 1100, 1200, 1300, 1500, 1600, 1750, 2000, 2750, 5000, 30000.
cos θ: -1.0, 0.55, 0.65, 0.75, 0.8, 0.85, 0.88, 0.9, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.988, 1.0.

FHC CC1Pi No Photon (Stacked Histogram)
p (MeV/c): 0, 300, 350, 400, 500, 600, 650, 700, 750, 1000, 1200, 1500, 2000, 3000, 30000.
TPC tag; cos θ: -1.0, 0.6, 0.75, 0.8, 0.85, 0.88, 0.9, 0.92, 0.93, 0.95, 0.97, 0.99, 1.0.

FGD tag; cos θ: 1.0, 2.83, 2.92, 3.0.

Michel tag; cos θ: 3.0, 4.8, 4.9, 4.93, 4.96, 4.98, 5.0.

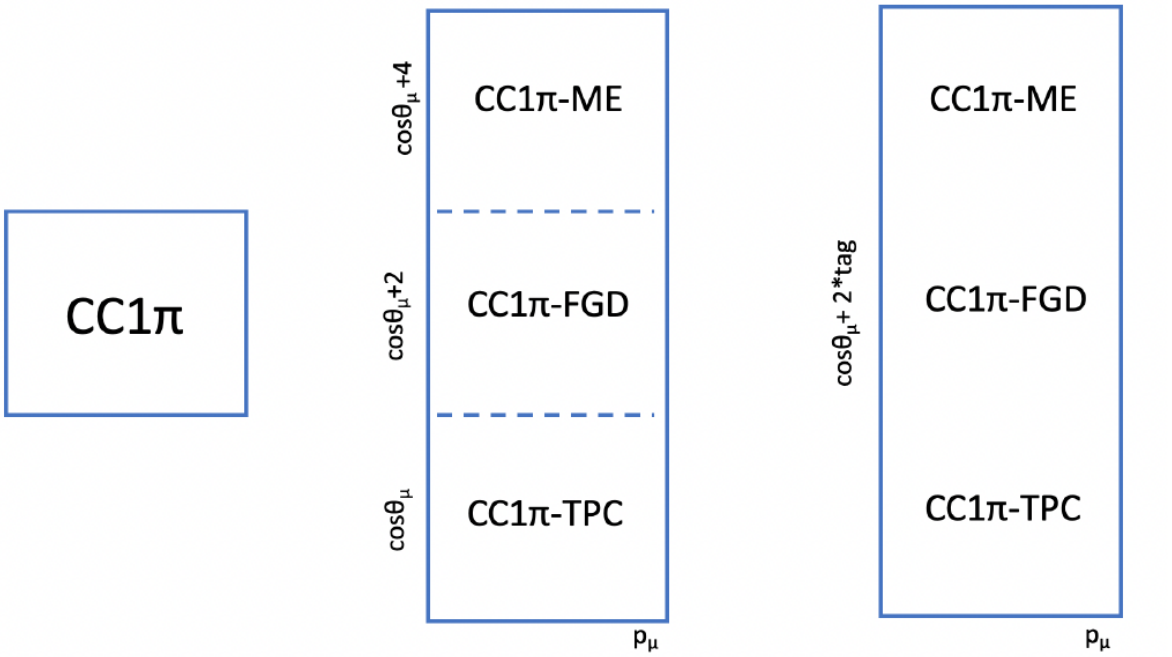


FIGURE 6.10: Histogram of the CC1Pi selection split by pion tag type in cos θ. The same momentum binning is applied across all tag samples. (T2K Internal, Image by J. Walsh).

I studied, modified, and implemented changes to the sample in order to fill a stacked histogram. This ensured that the fit mechanics seamlessly incorporated the separation of tagged events into their respective phase space.

FHC CCOther No Photon

p (MeV/c): 0, 350, 500, 600, 650, 700, 800, 900, 1000, 1150, 1250, 1500, 2000, 5000, 30000.

cos θ: -1.0, 0.6, 0.7, 0.8, 0.85, 0.88, 0.9, 0.92, 0.94, 0.96, 0.98, 1.

FHC CCPhoton

p (MeV/c): 0, 300, 400, 500, 600, 650, 700, 750, 800, 900, 1000, 1100, 1250, 1500, 1600, 1750, 2000, 2500, 3000, 5000, 30000.

cos θ: -1.0, 0.6, 0.7, 0.76, 0.8, 0.84, 0.86, 0.88, 0.89, 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1.

RHC CC0Pi

p (MeV/c): 0, 300, 400, 500, 550, 600, 650, 700, 750, 800, 900, 1000, 1100, 1200, 1500, 2000, 4000, 30000.

cos θ: -1, 0.6, 0.7, 0.8, 0.85, 0.9, 0.92, 0.93, 0.94, 0.95, 0.96, 0.965, 0.97, 0.975, 0.98, 0.985, 0.99, 0.995, 1.

RHC CC1Pi

p (MeV/c): 0, 500, 700, 900, 1300, 2500, 30000.

cos θ: -1, 0.7, 0.8, 0.9, 0.94, 0.96, 0.98, 0.99, 1.

RHC CCOther

p (MeV/c): 0, 600, 800, 1000, 1250, 1500, 2000, 4000, 30000.

cos θ: -1, 0.7, 0.8, 0.85, 0.9, 0.93, 0.95, 0.97, 0.98, 0.99, 1.

RHC BKG CC0Pi

p (MeV/c): 0, 300, 500, 700, 800, 900, 1250, 1500, 2000, 4000, 30000.

cos θ: -1, 0.7, 0.8, 0.85, 0.88, 0.9, 0.92, 0.94, 0.96, 0.97, 0.98, 0.99, 1.

RHC BKG CC1Pi

p (MeV/c): 0, 600, 800, 1500, 30000.

cos θ: -1, 0.7, 0.8, 0.86, 0.9, 0.94, 0.96, 0.97, 0.98, 0.99, 1.

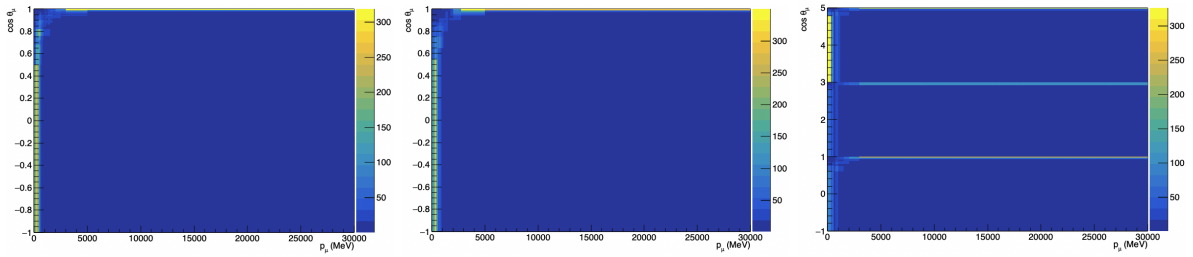
RHC BKG CCOther

p (MeV/c): 0, 600, 1000, 1250, 2000, 4000, 30000.

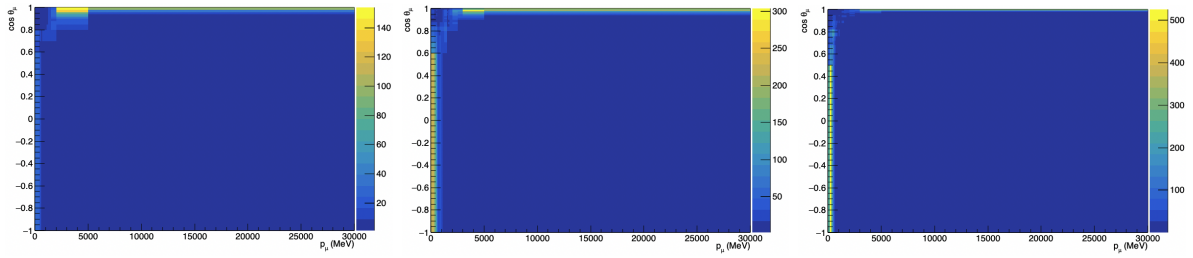
cos θ: -1, 0.7, 0.8, 0.86, 0.9, 0.93, 0.95, 0.97, 0.99, 1.

The binning for all the samples above are shown in Figure 6.11 below. It can be seen that the first bin in $\cos\theta$ bin is the largest for all samples. For CC1Pi No Photon the first bin range is from -1 to 0.6. Alongside this, the first momentum bin is 0–300 MeV; this is selected in order to ensure that all bins meet the data and MC requirements throughout the histogram, as the momentum bin width is imposed throughout the whole $\cos\theta$ phase space (-1 to 1).

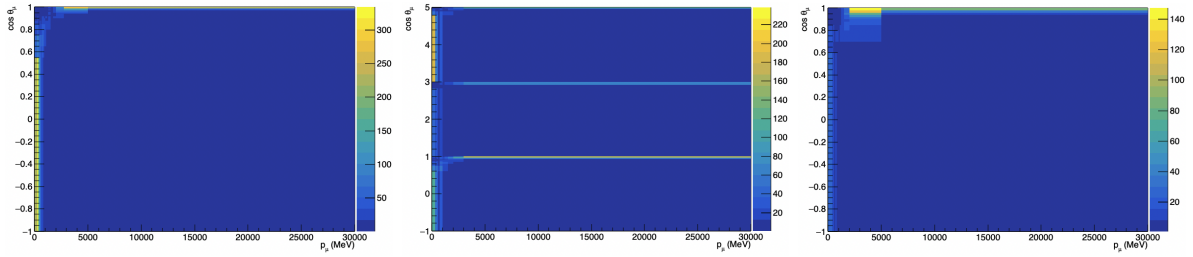
This is especially seen throughout the CC1Pi No Photon sample, as only $\cos\theta$ bins change throughout the stacked histogram. Moreover, this means that lower momentum bins with a wide cosine angle will end up with significantly higher events. This is similar for bins at higher momentum where the bin width is big at 25000 MeV, but the $\cos\theta$ bin is only 0.01 in width.



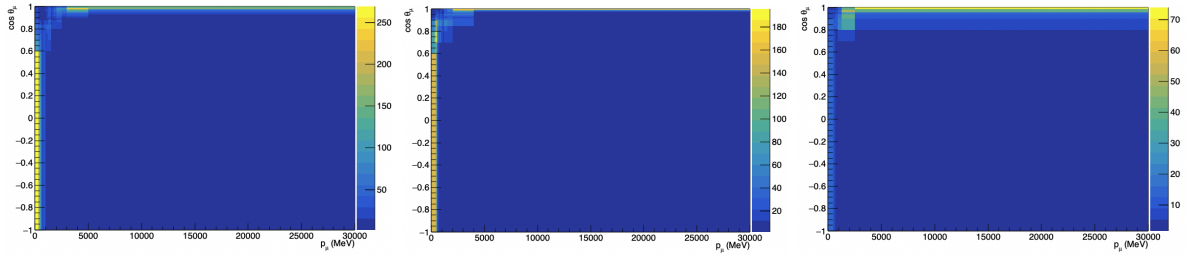
(a) FGD1 CC0Pi 0ProtonsNoPhoton (b) FGD1 CC0Pi NProtonsNoPhoton (c) **FGD1 CC1Pi NoPhoton**



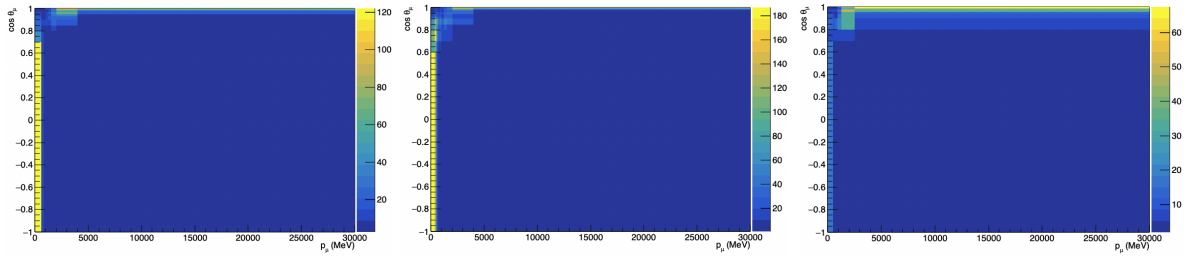
(d) FGD1 CCOther NoPhoton (e) FGD1 CCPhoton (f) FGD1 CC0Pi 0ProtonsNoPhoton



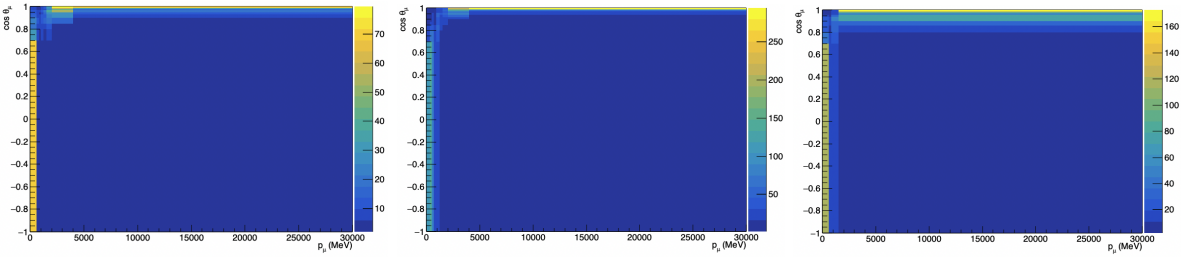
(g) FGD2 CC0Pi NProtonsNoPhoton (h) **FGD2 CC1Pi NoPhoton** (i) FGD2 CCOther NoPhoton



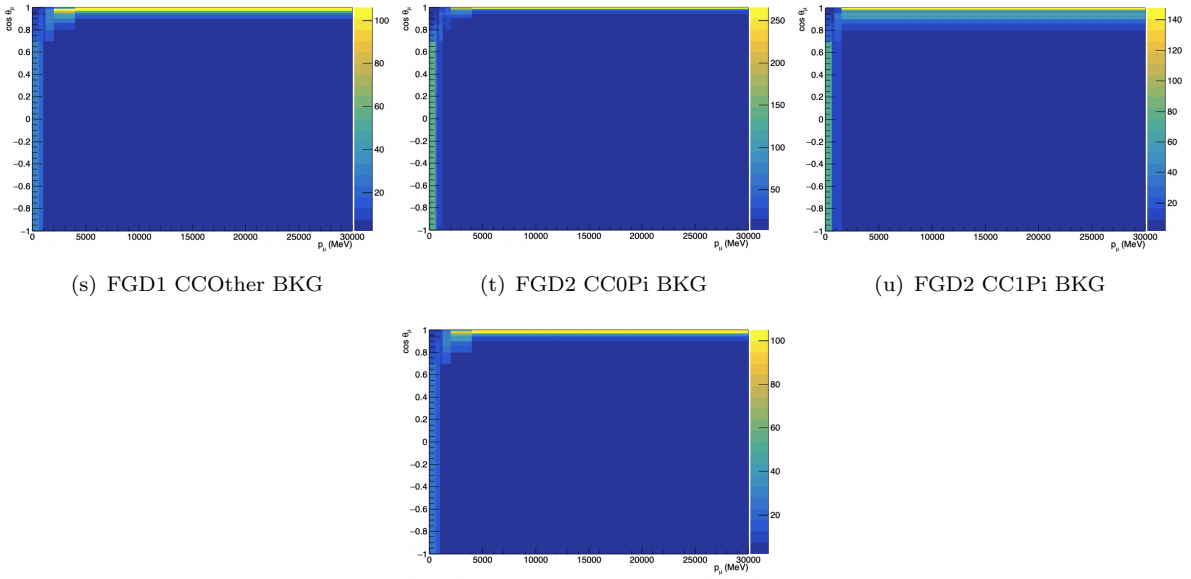
(j) FGD2 CCPhoton (k) FGD1 CC0Pi RHC (l) FGD1 CC1Pi RHC



(m) FGD1 CCOther RHC (n) FGD2 CC0Pi RHC (o) FGD2 CC1Pi RHC



(p) FGD2 CCOther RHC (q) FGD1 CC0Pi BKG (r) FGD1 CC1Pi BKG



(s) FGD1 CCOther BKG (t) FGD2 CC0Pi BKG (u) FGD2 CC1Pi BKG

(v) FGD2 CCOther BKG

FIGURE 6.11: Binning of 22 FHC and RHC samples from runs 2 - 9. Samples labelled in bold, (c) and (h), illustrate where the pion tag feature in this thesis was implemented.

6.6 Far Detector (Super-K) Selections

The Super-K selections which are used in the oscillation analysis are briefly mentioned here. The full selection criteria can be found in [114] and [115].

FHC $1R\mu$ **(Single Muon Ring):** This sample bins the number of single cherenkov ring events in FHC mode from leptons reconstructed as muons in charged current interactions.

RHC $1R\mu$ **(Single Muon Ring):** This sample bins the number of single cherenkov ring events in RHC mode from leptons reconstructed as muons in charged current interactions.

FHC $1Re$ **(Single Electron Ring):** This sample bins the number of single cherenkov ring events in FHC mode from leptons reconstructed as electrons in charged current interactions.

RHC $1Re$ **(Single Electron Ring):** This sample bins the number of single cherenkov ring events in RHC mode from leptons reconstructed as electrons in charged current interactions.

FHC $1Re\ 1de$ **(Single Electron Ring, 1 Delayed Electron):** This sample bins, in FHC mode, the number of single cherenkov ring events from leptons reconstructed as electrons alongside 1 decay electron in charged current interactions which are the result of Michel pion decay.

FHC $\nu_\mu\ CC1Pi$ **(Multi-Ring – 1 Muon Ring and 2 Delayed Electrons; Or: 2 Muon Rings and 1-2 Delayed Electrons):** This sample contains events which have 2 reconstructed

muon-like rings with 1 or 2 delayed electron rings, or 1 reconstructed muon ring with 2 delayed electron rings.

At T2K, only leptons and pions cross the Cherenkov threshold to become visible at Super-K. The threshold energies are: 5 MeV for $e^{\pm}$, 120 MeV for $\mu^{\pm}$ and 159 MeV for pions [116]. The single ring samples ($1R\mu$ and $1Re$) aim to apportion CCQE interactions. Samples which include delayed electrons such as FHC $1Re$ $1de$ and FHC ν_μ $CC1Pi$ aim to apportion CCRES interactions.

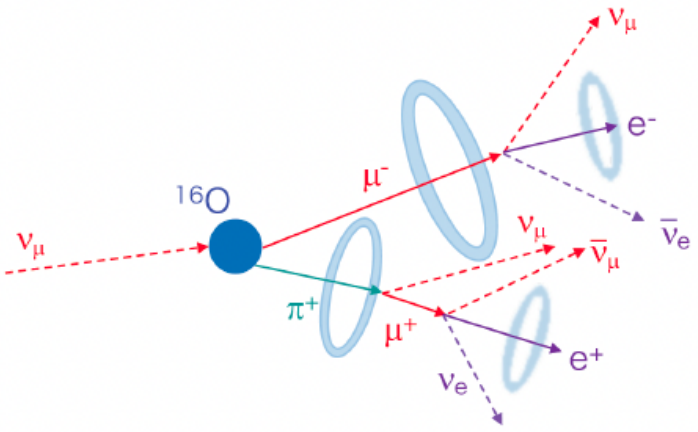


FIGURE 6.12: A schematic of a FHC ν_μ $CC1Pi$ event. The pion can be seen in this sample along with the multiple decay rings. (T2K Internal).

In OA2021, the two muon multi-ring samples (one example seen in Figure 6.12 above) were merged into one FHC ν_μ $CC1Pi$ sample. This was done in order to reduce dependence on pion kinematics. For future analyses, incorporating pion kinematics and splitting the CC1Pi sample by pion tag type at the near detector, especially by Michel tag, allows for better fitting of the multi-ring sample leading to better constraints on the oscillation parameters.

6.7 ND280 Systematic Parameters in the Near Detector Fit

Many different sources of uncertainty exist at the ND280 detector. These can be parameterised into 3 groups: flux, detector, and cross-section systematic parameters. Flux systematic parameters capture the uncertainties associated with the neutrino beam. Detector systematic parameters capture the uncertainties due to the reconstruction and detector response. Finally, cross-section systematic parameters capture the uncertainties associated with the neutrino interaction model [109]. These are listed and described in the section below.

A covariance matrix brings these uncertainties at the near detector into the fit as seen in Equation 5.7, Chapter 5. The inputs to this covariance may be from prior experimental data or not. If the prior input is known from experimental data, then the terms in the likelihood are Gaussian, otherwise they are constant (flat prior). This ensures that parameters in the fit are allowed to float according to prior knowledge.

6.7.1 Flux Systematic Parameters

As mentioned in Section 3.5.1, the neutrino flux simulation has two sources of systematic parameters: 1) Hadronic interaction systematic parameters and 2) Non-interaction systematic parameters, consisting of: alignment, modelling, and beam condition systematic parameters.

Figure 6.13 and 6.14 below show graphs of the fractional error from these main uncertainties across neutrino energy for ND280 and Super-K. The graph shows the hadronic interaction systematic that encapsulates the uncertainties from the interaction of the hadrons and the proton beam within the target. The non-interaction uncertainties in the graphs are: the proton beam profile and off-axis angle (alignment of proton beam), the horn current for the magnetic horns and their magnetic field (horn current & field), alignment of the focusing horns and the target (horn & target alignment), horn and decay volume materials (material modelling), and protons on target (number of protons). As seen in the graphs below, the hadronic interaction is the most dominant contribution over the energy phase space for both the far and near detectors.

Furthermore, the improvements to the flux uncertainty using a new version of the model (21bv2) can be seen across both detectors, especially in high neutrino energy regions. The total flux uncertainty for both detector around the peak neutrino energy (600 MeV) can be seen to be less than 10% in many cases, ranging from 6% – 10% across different neutrino modes.

The flux model is parameterised according to 100 neutrino energy bins with 50 bins for ND280 and 50 bins for Super-K which form a covariance matrix. The systematic parameters all have a central value of 1.0, with a Gaussian width corresponding to the flux covariance matrix. The bins in neutrino energy differ depending on the detector (near or far), the neutrino type, and the horn current mode. The neutrino energy bins (in GeV) are listed below:

ND280 and SK, FHC ν_μ and RHC $\bar{\nu}_\mu$: E_ν^{true}: 0, 0.4, 0.5, 0.6, 0.7, 1.0, 1.5, 2.5, 3.5, 7.0, 30

ND280 and SK, FHC $\bar{\nu}_\mu$ and RHC ν_μ : E_ν^{true}: 0, 0.7, 1.0, 1.5, 2.5, 30

ND280 and SK, FHC ν_e and RHC $\bar{\nu}_e$: E_ν^{true}: 0, 0.5, 0.7, 0.8, 1.5, 2.5, 4.0, 30

ND280 and SK, FHC $\bar{\nu}_e$ and RHC ν_e : E_ν^{true}: 0, 2.5, 30

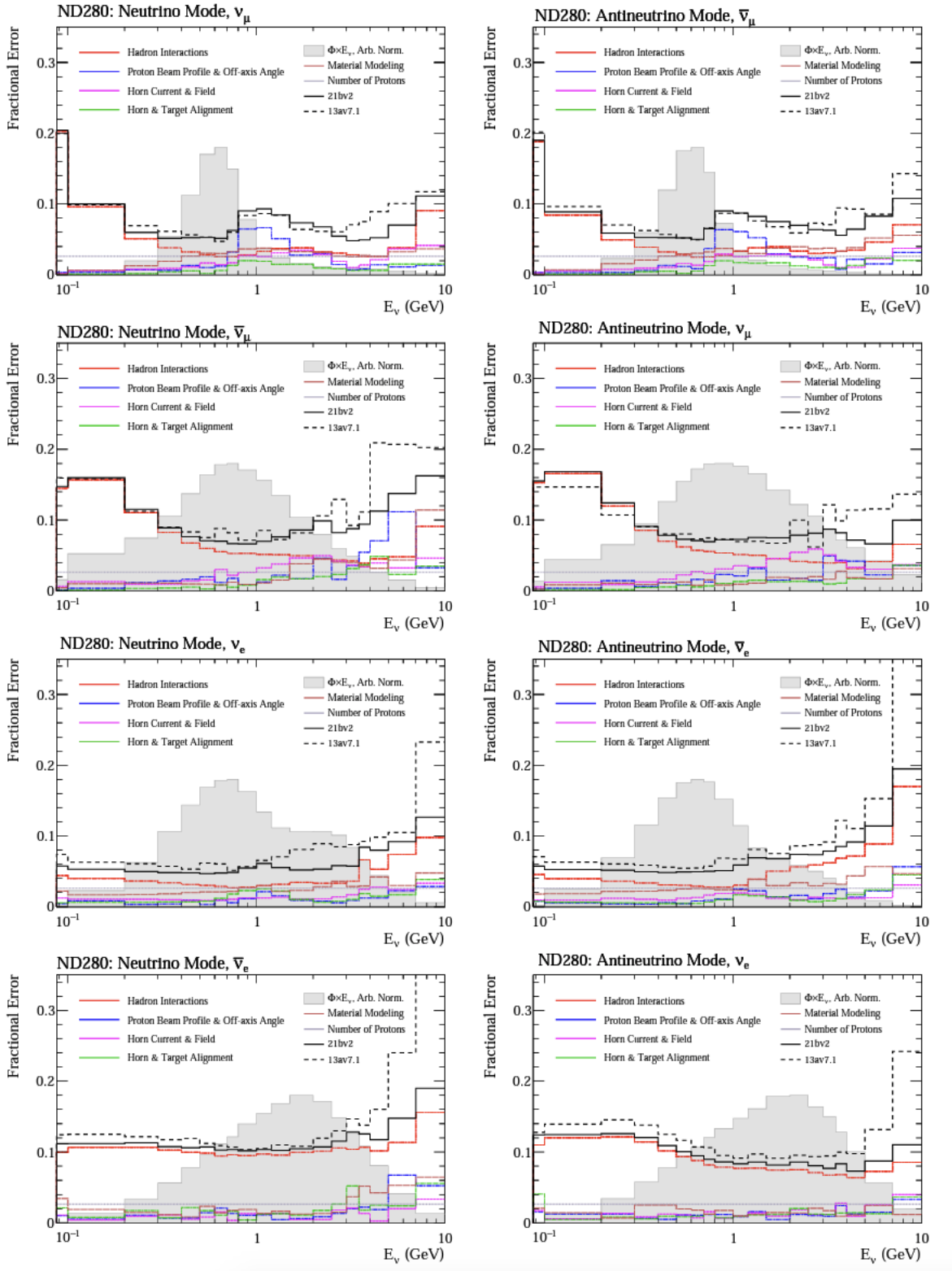


FIGURE 6.13: Fractional errors for different flux uncertainties across neutrino energy for ND280. The dotted line shows the total error of the last flux model (13av7.1), whereas the solid line shows the total error of the current flux model used in this analysis (21bv2) [117].

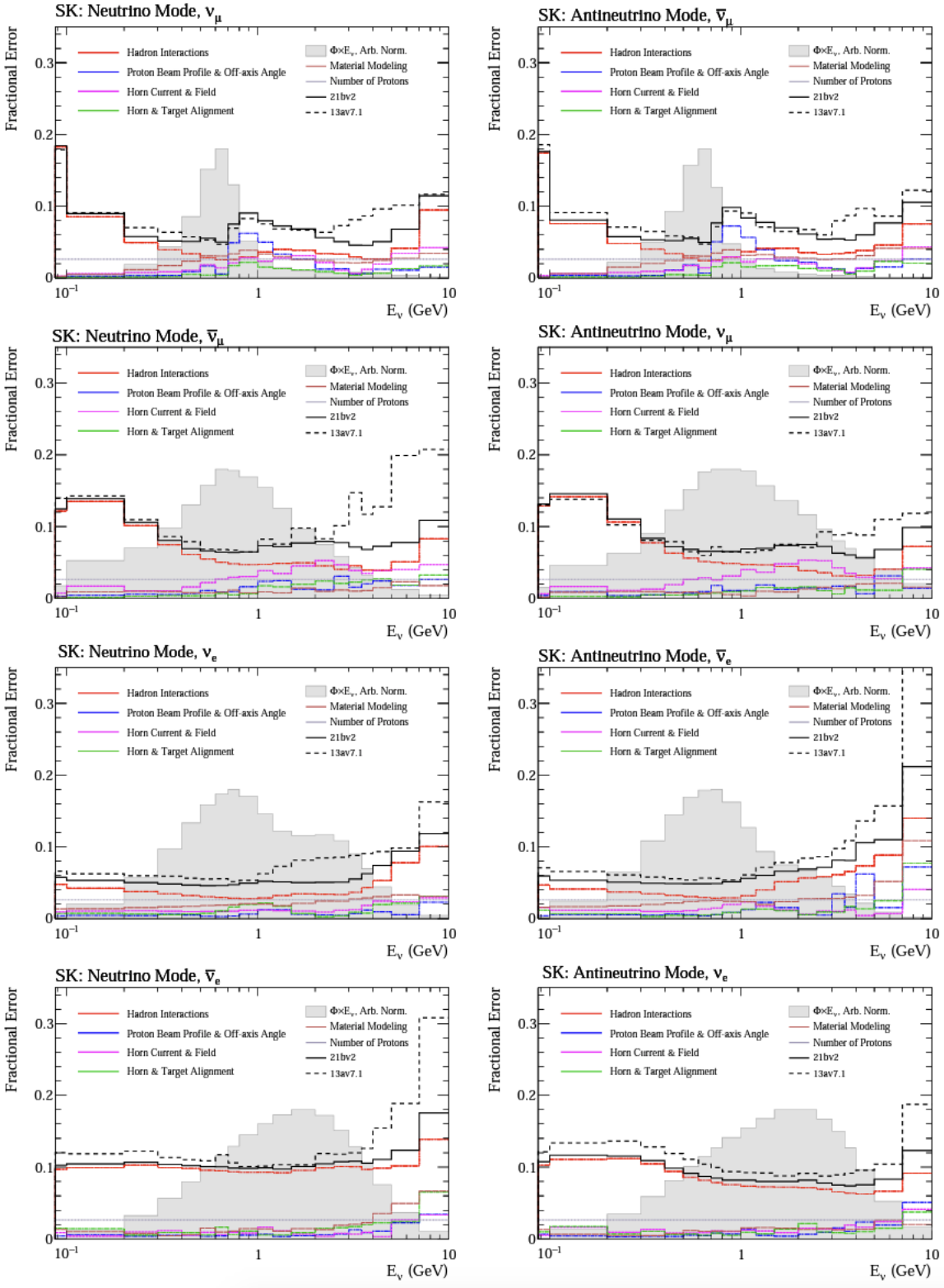


FIGURE 6.14: Fractional errors for different flux uncertainties across neutrino energy for Super-K. The dotted line shows the total error of the last flux model (13av7.1), whereas the solid line shows the total error of the current flux model used in this analysis (21bv2) [117].

6.7.2 Detector Systematic Parameters

Due to the imperfections in modelling detectors, the ND280 detector systematic parameters are key in determining how reconstructed events end up being categorised into different samples. These systematic parameters are then applied in one of the 3 different ways as seen below:

- **Efficiency Systematic Parameters (E):** These are applied as a weight to individual events that have been selected. The uncertainty originates from the reconstruction and detection of particles.
- **Observable Variable Systematic Parameters (O):** Reconstructed variables are smeared; which is when kinematics are modified according to detector performance. These smeared variables are then run through the selection algorithm again, this can result in events changing their highest momentum track, their selection, and topology.
- **Normalisation Systematic Parameters (N):** These scale event rates up or down through applying a normalisation to the bins.

Detector systematic parameters can be grouped under 17 different sources [118]. The letter indicated in brackets ([E], [O], [N]) identifies the way in which the systematic parameter is applied as detailed above:

TPC Track Efficiency [E] - The efficiency of reconstructing TPC tracks vary, this parameter takes this into account by measuring the efficiency of the TPC reconstruction for control samples from beam and cosmic events. The difference in reconstruction efficiency between beam and cosmic events is used as a systematic source.

TPC PID [O] - As mentioned prior, the particle identification in the TPC relies on the energy lost within it, this is captured by the *Pull*. The difference between the pulls for the data and the MC is used as a systematic parameter.

TPC-FGD Matching [E] - The efficiency of matching TPC tracks with FGD tracks at the boundary between the two sub-detectors is captured by this systematic parameter.

TPC Cluster Efficiency [E] - Hits are clustered within the TPC, the difference between the efficiency of the MC and data clusters from beam and cosmic control samples forms this systematic parameter.

TPC Momentum Scale [O] - There is uncertainty in the momentum measured by a TPC due to magnetic field uncertainties. The magnetic field is measured by probes (using the Hall

effect) within ND280, the results of these are used to give a scaling factor that can be applied to the MC.

TPC Momentum Resolution [O] - TPC tracks from the MC are smeared due to a discrepancy between the momentum resolutions in data and MC. The discrepancy is accounted for by smearing the TPC MC tracks such that the resolution is matched by data.

Michel Electron Efficiency [E] - A certain level of hits are required within the FGD by an electron in order to pass the selection cuts and thus for a Michel electron to be detected. This systematic is an efficiency that is measured from cosmic events, serving as a control sample, that is defined as the probability that when a muon is stopped in the FGD, an expected Michel electron is detected.

FGD PID [O] - The particle identification in the FGD relies on the energy lost within it, this is captured by the *Pull*. The difference between the pulls for the data and the MC is used as a systematic parameter.

FGD Mass [N] - The uncertainty on the FGD mass changes the target nuclei within it, thus changing the event rate. The difference between measured FGD mass, and simulated FGD mass is applied as a systematic parameter.

FGD Hybrid Track Efficiency [E] - FGD only tracks which are reconstructed correctly in the vicinity of FGD-TPC tracks are calculated as a proportion within a sample of GEANT4 particle simulations in events with one or two reconstructed tracks entering a TPC as a control sample. This efficiency is calculated for both data and MC, where the difference between them is used as a systematic parameter.

TPC Charge ID [E] - The reconstruction's ability to distinguish between a track in the TPC based on its curvature, using the MC and data, carries a level of uncertainty which is applied as a systematic parameter.

Pion Secondary Interactions [N] - Pions in the final state from neutrino interactions can further interact in the detector, this leads to inefficiencies in pion detection. The difference between the pion secondary interaction cross section for data and MC is used to account for this effect and is applied as a systematic parameter.

Proton Secondary Interactions [N] - Protons from neutrino interactions can interact in the detector as a secondary interaction, this leads to inefficiencies in proton detection. The difference between the proton secondary interaction cross section for data and MC is used to account for this effect and is applied as a systematic parameter.

Sand Muon Flux [N] - The sand in the pit outside ND280 can serve as a target for some neutrino interactions, the results of which end up being detected within the ND280 (fiducial volume). This forms an undesirable background (sand muons) when analysing events. A systematic is introduced which captures the uncertainty of such interactions by looking at an apt simulation and comparing that to data.

Out Of Fuducial Volume (OOFV) [N] - There are a number of events entering the fiducial volume of the detector due to a wrong reconstruction or an out of fiducial volume event. This background is calculated for events which occur within the beam trigger, the difference between data and MC is then used as a systematic parameter.

Pile-up Systematic [N] - OOFV events (from sand muons) that occur coincident with events within the fiducial volume may end up getting rejected due to the veto cut as described prior in the selection cuts. The number of these events occurring in data and MC are subtracted from each other in order to find the difference; this is then used as a systematic parameter.

B-Field [O] - This uncertainty comes from the fact that the magnetic field is not uniform through the detector. Therefore, there is an uncertainty in modelling this inhomogeneity within the TPCs.

6.7.3 Detector Binning

To apply all of these systematic effects to each individual event would be too computationally intensive. Therefore, the MC is re-weighted by a weight calculated for each systematic source. The weights for each systematic are worked out for each bin by 'throwing toys', which is achieved through populating a histogram in muon kinematics using randomly generated values for the systematic parameters (toy experiments), and then repeating this process (throwing). This process is repeated 2000 times. Following this, the systematic uncertainties are found for each bin by calculating the relative error (RMS / mean). All of the systematic parameters have a Gaussian prior except the B-Field systematic which is the only flat prior systematic.

Due to the fact that the number of bins within the sample binning are high, the computational requirements become intensive. Therefore, in order to speed up the process sufficiently for oscillation analyses, the number of parameters (bins) must be reduced. To do this, I merged adjacent bins with similar systematic uncertainty responses. The procedure was as follows:

1. Set the initial detector binning equal to the sample binning.
2. Pass the binning to the bin merging algorithm to see if adjacent bins are within the tolerance specified. The tolerance used is 7 %, which is the difference in correlation between

the two bins. Potential bins to merge which fall within the tolerance (under 7 %) are outputted and identified to be merged.

3. Merge the recommended bins and update the detector binning.

4. Repeat steps 2) and 3) a maximum of 2 iterations until a final binning is decided. The criteria to stop is determined by prior experience, which is balancing the need to reduce parameters in the fit vs the need to not merge excessively as to wash out important effects.

In this thesis, when implementing the new FHC FGD1 and FGD2 CC1Pi stacked histogram I followed the above process, and the total bin number was reduced from an initial 1070 detector bins (in step 1) to a final 754 detector bins (after repeated merging). This ended up meeting the requirement for detector binning of at least 1 data event and $\geq$ 20 MC events within all bins. Compared to the non-split by tag CC1Pi sample used in OA2021, this updated binning for pion-tagged leptons is a 202 bin increase, where the previous covariance matrix used in the 2021 Oscillation Analysis had 552 bins with the original non-split by tag CC1Pi sample. Histograms of event number distributions from 3 bins in different pion-tag regions of the CC1Pi stacked histogram are shown below in Figure 6.15:

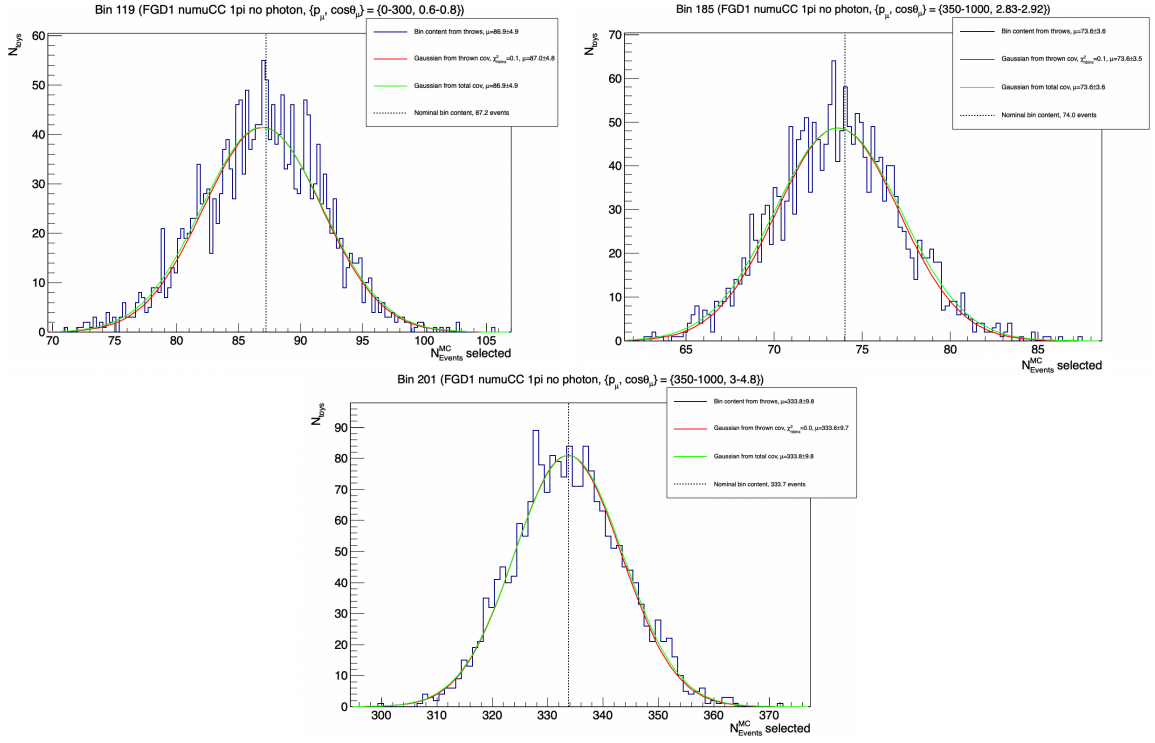


FIGURE 6.15: The result of throwing detector systematics 2000 times: the distribution of the number of events in 3 detector covariance bins, each spanning an area of phase space (TPC, Michel, and FGD) within the new CC1Pi sample split by pion tag. The blue line shows the bin content from throws, the green line shows the Gaussian fit with MC stats, the red line shows the Gaussian without MC Stats, and the dotted black line shows the number of nominal events.

The final detector binning following the merging process is as follows (identical binning for FGD1 and FGD2):

FHC CC0Pi 0 Protons No Photon

p (MeV/c): 0, 200, 300, 800, 850, 2000, 5000, 30000.

cos θ: -1, 0.5, 0.6, 0.72, 0.82, 0.86, 0.9, 0.95, 0.96, 1.

FHC CC0Pi N Protons No Photon

p (MeV/c): 0, 250, 1200, 1300, 2750, 5000, 30000.

cos θ: -1, 0.55, 0.65, 0.8, 0.85, 0.97, 0.98, 0.988, 1.

FHC CC1Pi No Photon (Stacked in $cos(\theta)$**: -1–1, TPC; 1–3, FGD; 3–5, Michel])**

p (MeV/c): 0, 300, 350, 1000, 1200, 1500, 2000, 3000, 30000.

cos θ: -1, 0.6, 0.8, 0.85, 0.92, 0.93, 0.97, 0.99, 1.0, 2.83, 2.92, 3.0, 4.8, 4.9, 4.93, 4.96, 4.98, 5.

FHC CCOther No Photon

p (MeV/c): 0, 350, 1250, 2000, 5000, 30000.

cos θ: -1, 0.6, 0.8, 0.92, 0.94, 0.98, 1.

FHC CCPhoton

p (MeV/c): 0, 300, 500, 1500, 3000, 5000, 30000.

cos θ: -1, 0.7, 0.76, 0.8, 0.84, 0.88, 0.89, 0.96, 1.

RHC CC0Pi

p (MeV/c): 0, 300, 2000, 4000, 30000.

cos θ: -1, 0.6, 0.8, 0.9, 0.96, 1.

RHC CC1Pi

p (MeV/c): 0, 500, 30000.

cos θ: -1, 0.7, 1.

RHC CCOther

p (MeV/c): 0, 600, 800, 30000.

cos θ: -1, 0.7, 0.95, 0.97, 1.

RHC BKG CC0Pi

p (MeV/c): 0, 300, 1500, 30000.

cos θ: -1, 0.7, 1.

RHC BKG CC1Pi

p (MeV/c): 0, 600, 800, 30000.

cos θ: -1, 0.7, 1.

RHC BKG CCOther

p (MeV/c): 0, 600, 30000.

cos θ: -1, 0.7, 1.

The full merged detector covariance is shown below in Figure 6.16.

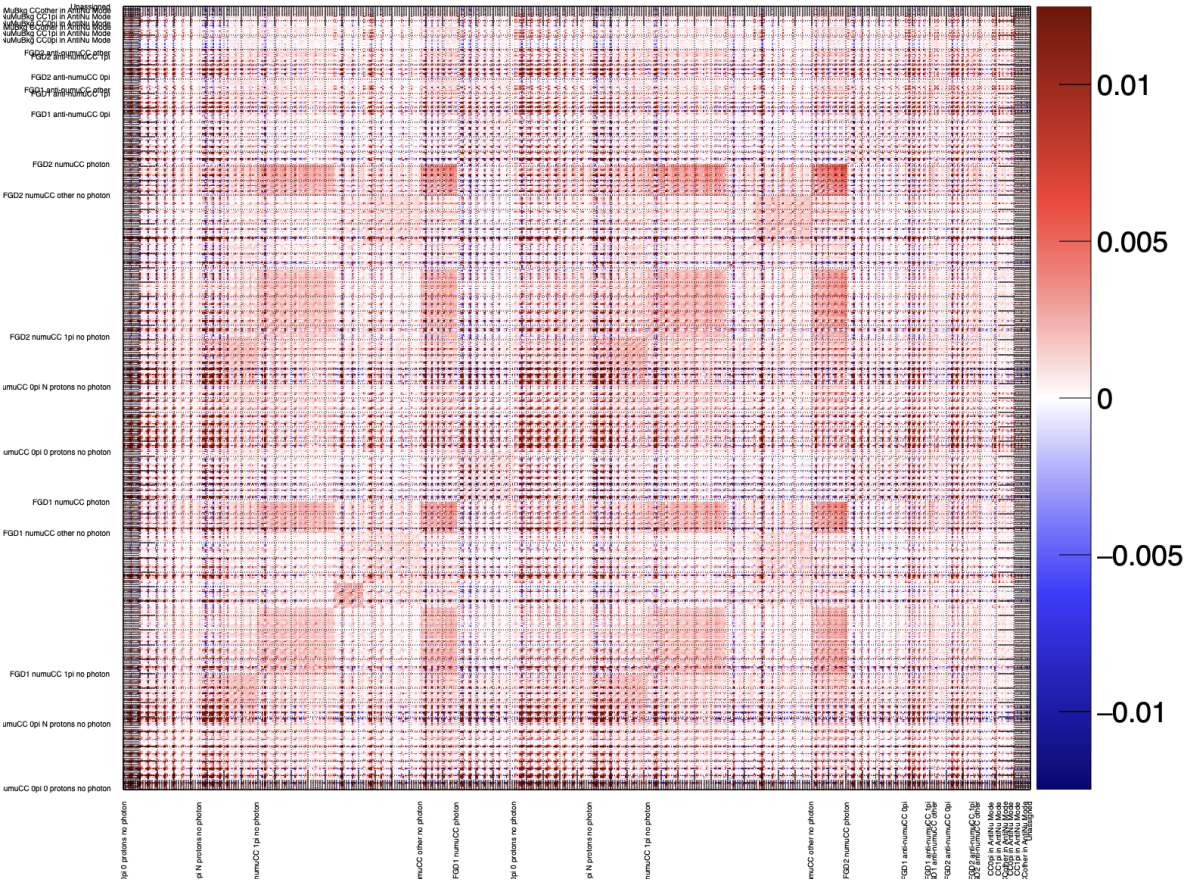


FIGURE 6.16: Final detector covariance matrix as the result of merging (754 bins).

The new FHC CC1Pi split by pion tag sample compared to the non-split FHC CC1Pi sample requires additional detector systematic parameters to account for the increased binning. This allows for a better consideration of the uncertainties. The new split sample does not necessitate changing the flux or cross-section parameters as the previous non-split CC1Pi sample treatment can remain the same (the events themselves are unchanged). In general, however, developing the cross-section model to better account for single pion production (SPP) is important to better constrain the oscillation parameters. This thesis work provides an important base for developing such changes in the future.

6.7.4 Cross Section Systematic Parameters

The cross section systematic parameters in T2K are determined from the interaction model used in NEUT which is tuned with external data and additional theory models. As mentioned prior, NEUT is the interaction generator used in T2K to simulate neutrino-nucleus interactions, with neutrinos of energy ranging from 100 MeV to the order of a couple of TeV. The observable kinematics are predicted by NEUT over this wide range of energies using the most suitable model for the interaction type with the neutrino energy. For the T2K setup, this energy ranges from 100 MeV up until $\sim$30 GeV. Alongside this, the phenomena of the re-scattering of hadrons is modelled, allowing final state interactions with pions and nucleons to be accounted for [119].

These systematic parameters are applied on individual events depending on their interaction type. There are two ways these systematic parameters are applied, either through a normalisation or a shape uncertainty. Normalisation uncertainties apply a weight to events which are exactly equal to the systematic value. Whereas shape systematics apply a weight that depends on the kinematics of an event; this is captured by a spline response function. These response functions are made by working out the weight for certain values of a parameter at evenly spaced intervals and interpolating through the points (in a 1-to-1 fashion). This means that when a parameter is changed, the new weight is found by reading the spline graph at that parameter value and applying the corresponding weight.

The systematic parameters of some of the common interaction types are detailed below, using NEUT 5.4.2:

Charged Current Quasi-elastic (CCQE): A spectral function (SF) is used to simulate this interaction type with the axial mass, M_A^{QE}, set at 1.21. The M_A^{QE} prior is determined through previous bubble chamber data, and is a shape parameter with an associated spline function as mentioned. There are 3 normalisation parameters for bins within Q^2, the prior uncertainty and central value for which are determined using MINERvA data [120]. These parameters are in the regions: 0.25 – 0.5 GeV^2, 0.5 – 1 GeV^2, and $>$ 1.00 GeV^2. 13 new shape parameters have been added, which allow for alterations to the 2D SF (P and S shells), with 7 parameters for oxygen and 5 for carbon [121].

As seen below in Figure 6.17, there are 4 E_b parameters and 1 $\alpha(q3)$ parameter which are needed in this model to fit the binding energy of the targeted nuclei (carbon and oxygen); this is important for reconstructing the neutrino energy accurately. Additionally 6 shape parameters have been added to account for Pauli blocking and optical potential for both carbon and oxygen nuclei.

Charged Current Resonance (CCRES): The resonant interactions are modelled in NEUT by application of the Rein-Sehgal model that describes resonant pion production [122]. As seen below in Figure 6.17, there are 11 parameters associated with this model. 2 are normalisation parameters due to single pion production. The remaining 9 shape parameters are: M_A^{RES}, the resonance axial mass, C_5^A, the axial form factor normalisation, $I-\frac{1}{2}$ background at low and high pion momenta to resonant pion production. These 4 parameters have specific prior uncertainties and central values determined from bubble chamber data. The remaining 5 shape parameters are resonant binding energy parameters (4) for carbon and oxygen targeted by neutrinos or antineutrinos, alongside the resonant delta decay (RS Delta Decay).

Two-Particle-Two-Hole (2p2h): Using the Nieves et al. model [123], 2p2h interactions are generated. Modelling these requires 12 parameters in total, with 4 shape parameters coming from the fact that at one extreme of energy and momentum transfer pion-less decay in the 2p2h model accords with the Nieves et al. model (2 parameters), whilst at the other extreme it accords with the Martini model [124] (2 parameters). There are 4 shape parameters due to 2p2h events on oxygen (2) or carbon (2). There is a final shape parameter to account for pair type (PN-NN shape). These 2p2h parameters cannot be fully constrained by near detector only fits and so these parameters are fixed. With a far detector fit, or a joint fit with more than ND only data available, these parameters become free floating. There are 3 normalisation parameters due to neutrino and antineutrino events (2) alongside a scaling for carbon to oxygen.

Charged Current Coherent (CCCOH): The new Berger-Seghal model [125] is used to simulate coherent interactions where the neutrino interacts as a whole with the nucleus. There are two normalisation parameters: one for coherent interactions on oxygen and one for coherent interactions on carbon.

Charged Current Deep Inelastic Scattering (CCDIS): In CCDIS interactions, the neutrino interacts with the quarks of a nucleon directly. Due to this, the cross sections are calculated using Parton Distribution Functions (PDF) which describe the probability of finding a quark with a certain fraction of momentum from the nucleon. These interactions produce a multiplicity of pions and CCDIS interactions are modelled in NEUT using a model with corrections from Bodek and Yang [126]. In total, there are 7 parameters: 2 are normalisation parameters of multi pion production from DIS for neutrinos and antineutrinos. The remaining 5 shape parameters consist of 3 parameters describing Bodek-Yang corrections to the model and 2 describing the pion multiplicity cross section and shape.

Final State Interactions (FSI): For pion final state interactions, a NEUT cascade model is applied. The pions which are produced within the nucleus can undergo secondary interactions which can change event topology. In the model there are 7 final state interaction parameters

which are all shape parameters. FSI QE is concerning low pion momentum charge exchange, FSI QE High deals with pion charge exchange at high momenta. The other parameters deal with pions, in particular hadron production, absorption, and pion charge exchange. The final shape parameter deals with the final state interaction of the nucleon itself.

Parameter	Validity Range	generated value	Prior mean	Prior Error	Type	Unit
M_A^{QE}	all	1.21	1.03	0.06	shape	GeV/c^2
Q^2 norm 0.25-0.50 GeV^2	all	1	1	0.11	norm.	—
Q^2 norm 0.50-1.00 GeV^2	all	1	1	0.18	norm.	—
Q^2 norm $>$ 1.00 GeV^2	all	1	1	0.4	norm.	—
PShell MF Norm C	all	0	0	0.2	shape	—
SShell MF Norm C	all	0	0	0.35	shape	—
SRC Norm C	all	1	1	2	shape	—
PShell MF PMiss C	all	0	0	1	shape	—
SShell MF PMiss C	all	0	0	1	shape	—
P1 2Shell MF Norm O	all	0	0	0.2	shape	—
P3 2Shell MF Norm O	all	0	0	0.45	shape	—
SShell MF Norm O	all	0	0	0.75	shape	—
SRC Norm O	all	1	1	2	shape	—
P1 2Shell MF PMiss O	all	0	0	1	shape	—
P3 2Shell MF PMiss O	all	0	0	1	shape	—
SShell MF PMiss O	all	0	0	1	shape	—
Pauli Blocking C ν	all	0	0	1	shape	—
Pauli Blocking O ν	all	0	0	1	shape	—
Pauli Blocking C $\bar{\nu}$	all	0	0	1	shape	—
Pauli Blocking O $\bar{\nu}$	all	0	0	1	shape	—
Optical Potential C	0–1	0	0	1	shape	—
Optical Potential O	0–1	0	0	1	shape	—
2p2h norm ν	>0	1	1	1	norm.	—
2p2h norm $\bar{\nu}$	>0	1	1	1	norm.	—
2p2h norm CtoO	>0	1	1	0.2	norm.	—
2p2h Edep low E_ν	0–1	1	1	1	shape	—
2p2h Edep high E_ν	0–1	1	1	1	shape	—
2p2h Edep low $E_{\bar{\nu}}$	0–1	1	1	1	shape	—
2p2h Edep high $E_{\bar{\nu}}$	0–1	1	1	1	shape	—
PN-NN Shape	-1–1	0	0	0.33	shape	—
2p2h shape C np	-1–1	0	0	1	shape	—
2p2h shape C NN	-1–1	0	0	1	shape	—
2p2h shape O np	-1–1	0	0	1	shape	—
2p2h shape O NN	-1–1	0	0	1	shape	—
C_A^5	all	1.01	0.96	0.15	shape	—
M_A^{RES}	all	0.95	1.07	0.15	shape	—
I-$\frac{1}{2}$ bkg. low p_π	all	1.3	0.96	1.3	shape	—
I-$\frac{1}{2}$ bkg.	all	1.3	0.96	0.27	shape	—
RES Eb C ν_μ	0–25	0	25	25	shape	—
RES Eb 0 ν_μ	0–25	0	25	25	shape	—
RES Eb C $\bar{\nu}_\mu$	0–25	0	25	25	shape	—
RES Eb O $\bar{\nu}_\mu$	0–25	0	25	25	shape	—
RS Delta Decay	0–1	1	1	1	shape	—
SPP Pi0 numu	all	1	1	0.3	norm	—
SPP Pi0 numubar	all	1	1	0.3	norm	—
π-FSI QE	all	1.069	1.069	0.313	shape	—
π-FSI QE High	all	1.824	1.824	0.859	shape	—
π-FSI Hadron production	all	1.002	1.002	1.101	shape	—
π-FSI Absoprtion	all	1.404	1.404	0.432	shape	—
π-FSI Charge Exchange	all	0.697	0.697	0.305	shape	—
π-FSI Pion Charge Exchange	all	1.8	1.8	0.288	shape	—
Nucleon FSI	-1–1	0	0	0.3	shape	—
CC Coh. C	>0	1	1	0.3	norm.	—
CC Coh. O	>0	1	1	0.3	norm.	—
MPi Multi TotXsec	all	0	0	1	shape	—
MPi BY Vector	all	0	0	1	shape	—
MPi BY Axial	all	0	0	1	shape	—
MPi Multiplicity shape	all	0	0	1	shape	—
CC BY DIS	all	0	0	1	shape	—
CC DIS MultiPi ν	>0	1	1	0.035	norm	—
CC DIS MultiPi $\bar{\nu}$	>0	1	1	0.065	norm	—
CC Misc.	>0	1	1	1	norm.	—
NC Coh.	>0	1	1	0.3	norm.	—
NC 1γ	>0	1	1	1	norm.	—
NC other near	>0	1	1	0.3	norm.	—
NC other far	>0	1	1	0.3	norm.	—
CC norm ν	>0	1	1	0.02	norm.	—
CC norm $\bar{\nu}$	>0	1	1	0.01	norm.	—
ν_e ν_μ	>0	1	1	0.0282843	norm.	—
$\bar{\nu}_e$ $\bar{\nu}_\mu$	>0	1	1	0.0282843	norm.	—
E_b C ν	-10–15	0	2	6	function	MeV
E_b C $\bar{\nu}$	-10–15	0	0	6	function	MeV
E_b O ν	-10–15	0	4	6	function	MeV
E_b O $\bar{\nu}$	-10–15	0	0	6	function	MeV
α q3	0–1	0	0	1	function	—

FIGURE 6.17: Table showing all the cross-section parameters used in the analysis. Table from [121].

Chapter 7

Results

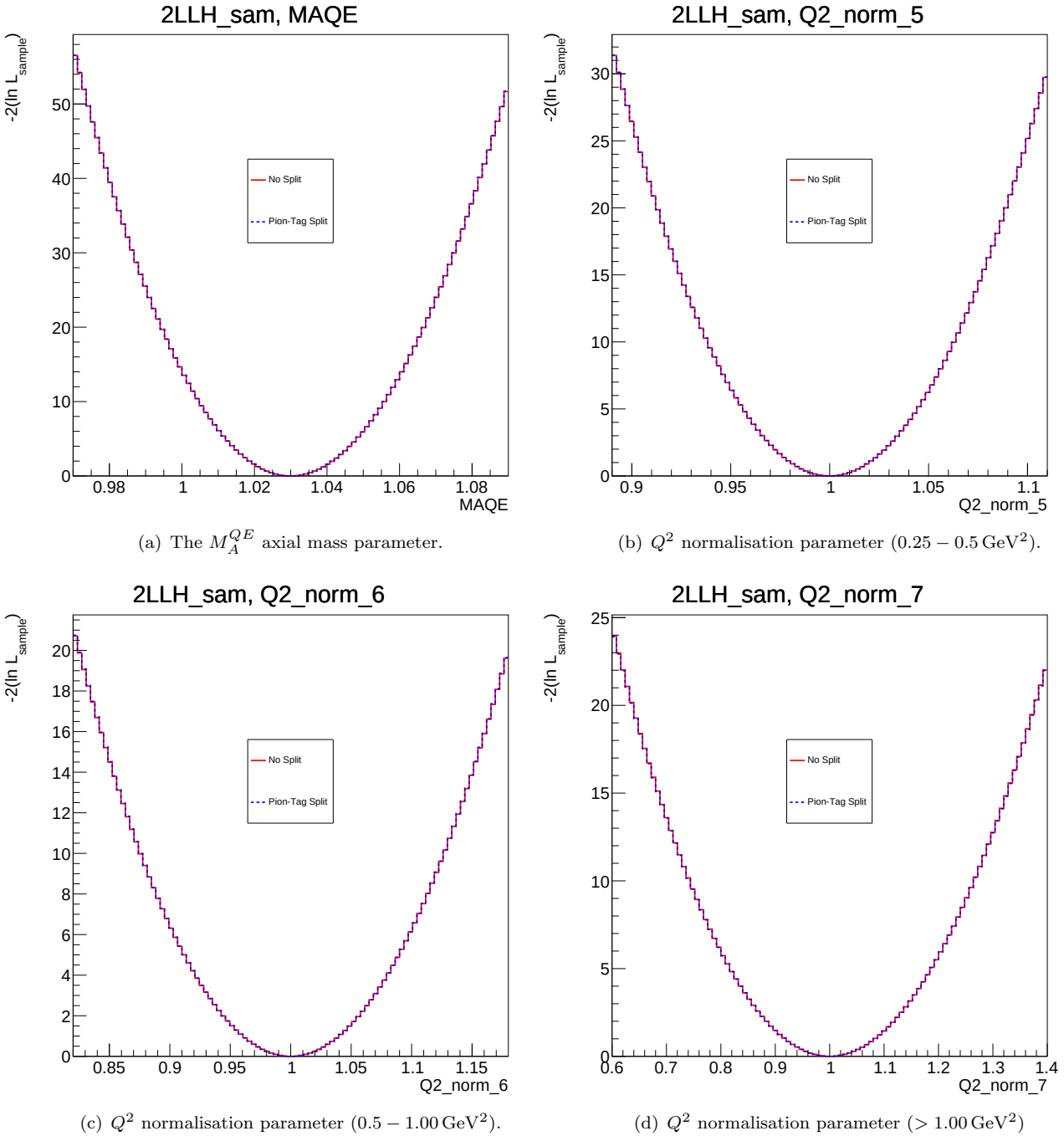
The impacts from the new feature implementation in this thesis, the FHC CC1Pi pion tag split in lepton kinematics, were quantified by performing Asimov fits [127] and comparing this new stacked histogram feature (known as the split case) with the original FHC CC1Pi non-split feature (known as the no-split case). Cross-section parameters which had stronger constraints are identified, and a log-likelihood scan is performed to determine the constraints in the split vs no-split case. Comparing both results, improved constraints in some parameters can be seen clearly (these scans are confirmed with the other T2K fitting framework BANFF). Near detector data fits were also run, alongside the production of far detector spectra. This study serves as a necessary intermediate step on the way to implementing pion kinematics within the MaCh3 fit.

7.1 Log Likelihood Scans

Log likelihood (LLH) scans for the cross-section parameters are shown below. A log-likelihood scan takes the MC simulation and varies one parameter whilst fixing the others. The scan steps through 150 equally spaced points in the domain -1σ to $+1\sigma$. For each parameter value, the MC is re-weighted for the parameter at that point with the corresponding χ^2 value being evaluated at each step. The result of this scan forms a parabola-like trace which has a symmetrical effect on the bins in most cases. The minimum of the scan is at the initial parameter value, as there cannot be a better fit to the nominal MC at any other parameter value except at the initial parameter value the nominal MC was generated with. After the parameter has been varied over the full domain, the parameter is reset and the process is repeated again for the next parameter. Only the diagonals of the covariance matrix are used for penalty terms in the LLH calculation, correlations are ignored when varying one parameter as this could severely impact the scan. Therefore, the LLH scans serve as a partial description of the sensitivity to the fit and a

validation of the framework. The cross-section parameters are listed in Chapter 6, Section 6.7.4, Figure 6.17. All cross-section parameter scans below are grouped according to interaction type.

7.1.1 CCQE Parameters.



(a) The M_A^{QE} axial mass parameter.

(b) Q^2 normalisation parameter $(0.25 - 0.5\,\mathrm{GeV}^2)$.

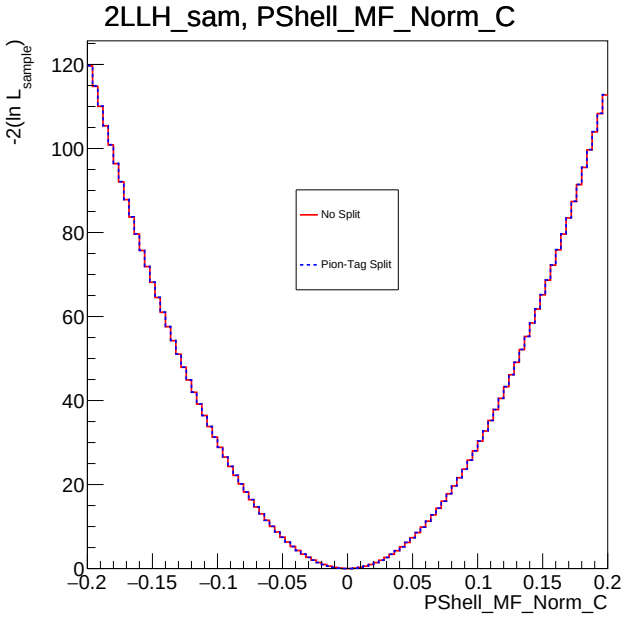
(c) Q^2 normalisation parameter $(0.5 - 1.00\,\mathrm{GeV}^2)$.

(d) Q^2 normalisation parameter $(> 1.00\,\mathrm{GeV}^2)$

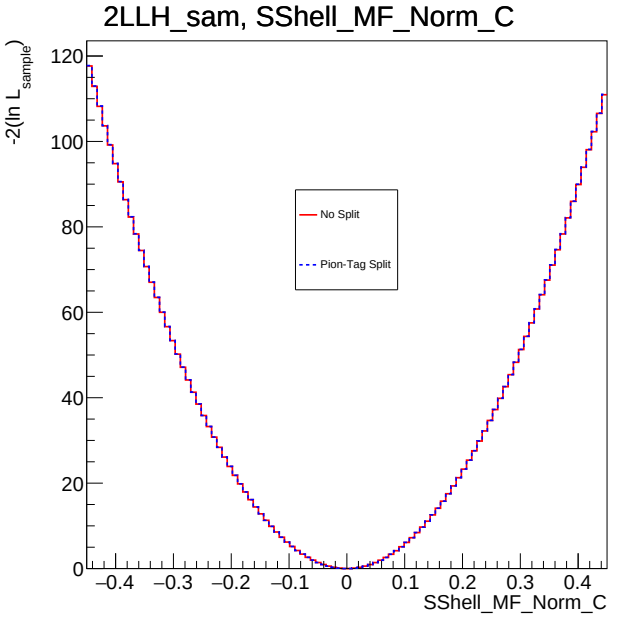
FIGURE 7.1: CCQE Parameters.

The parameters in Figure 7.1 are related to the Spectral Function used to model CCQE interactions. The Q^2 squared parameters cover physics beyond the dipole factorisation of the form

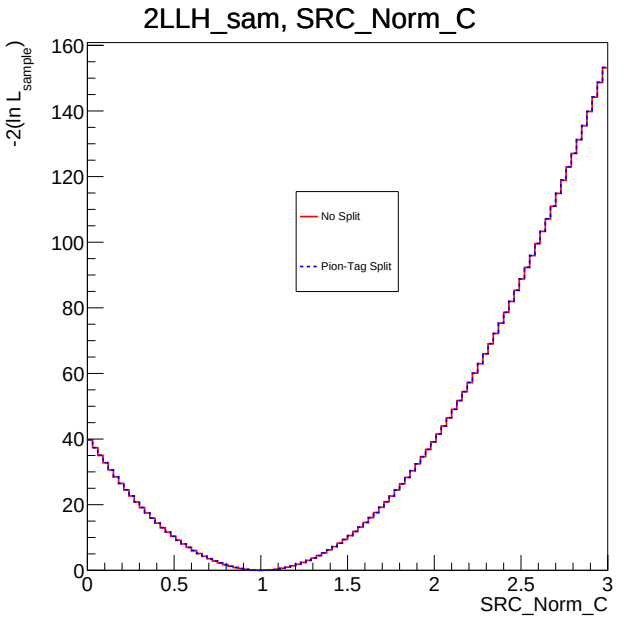
factor. The same constraint is seen here in the no-split feature compared to the split feature. This is expected as pion tag splitting is not in the CC0Pi sample which is CCQE interaction rich.



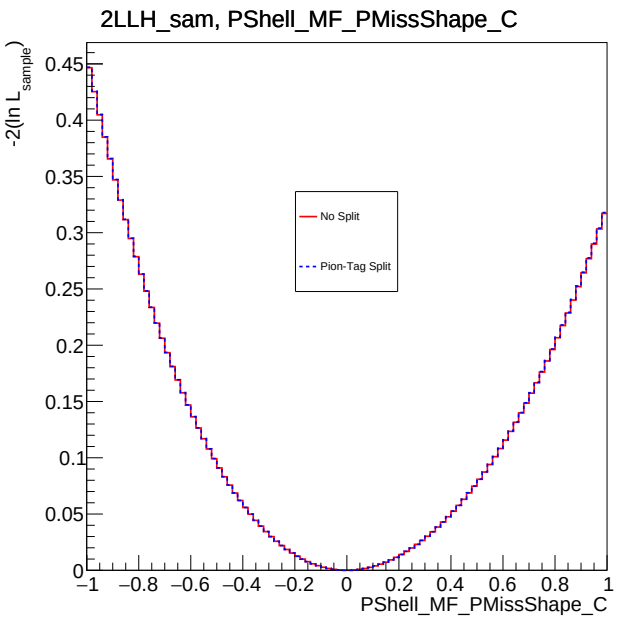
(a) P Shell Mean Field Normalisation for Carbon



(b) S Shell Mean Field Normalisation for Carbon



(c) Short Range Correlations Normalisation for Carbon

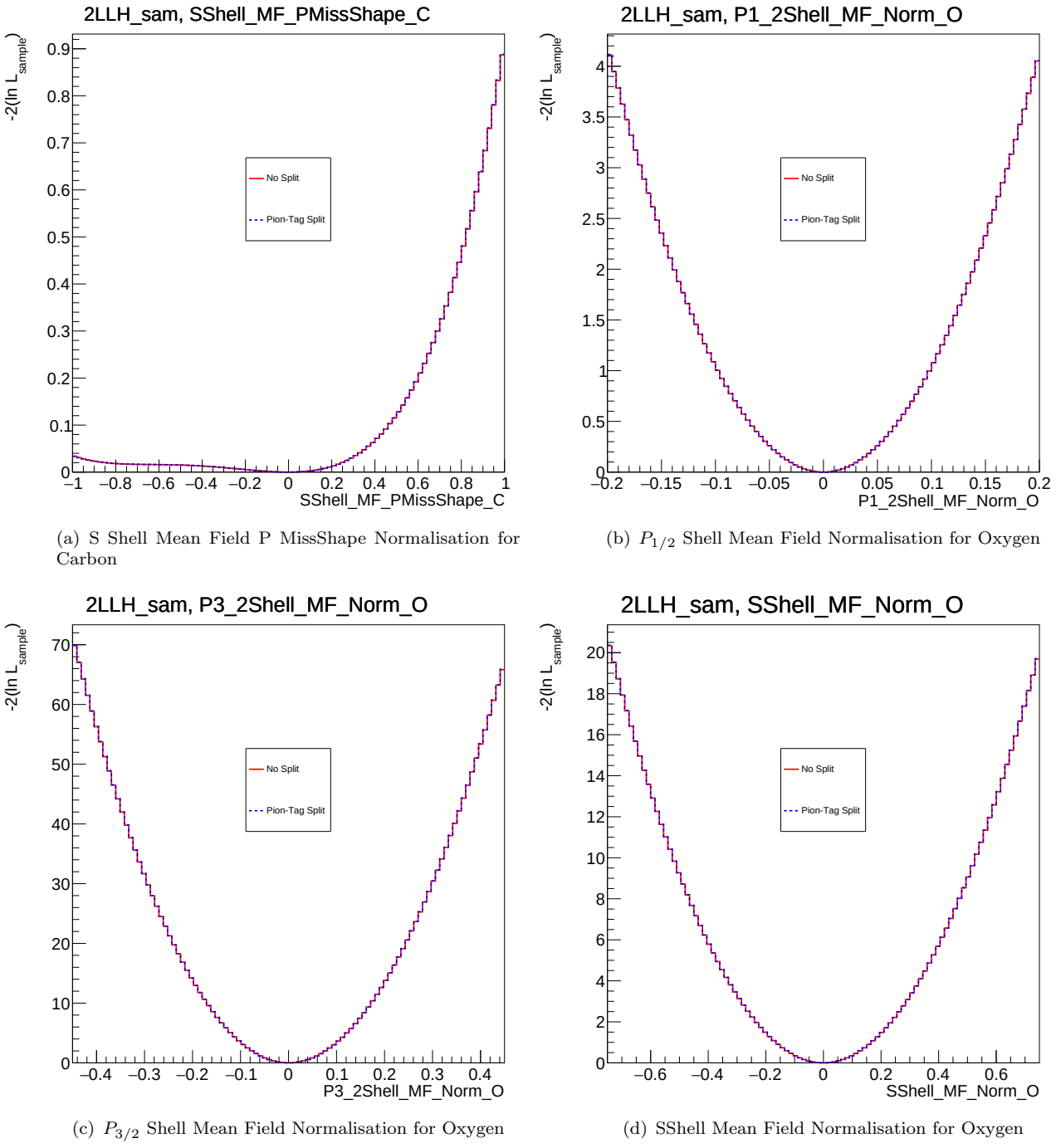


(d) P Shell Mean Field P MissShape Normalisation for Carbon

FIGURE 7.2: CCQE Parameters.

The mean field normalisation and short range correlation parameters in Figure 7.2 describe the shell model of the nucleus for Carbon with P and S shells. These are used to model CCQE

interactions, and therefore the same constraints are seen between the two features. This is expected as the split feature does not affect the CC0Pi sample which is CCQE interaction rich.



(a) S Shell Mean Field P MissShape Normalisation for Carbon

(b) $P_{1/2}$ Shell Mean Field Normalisation for Oxygen

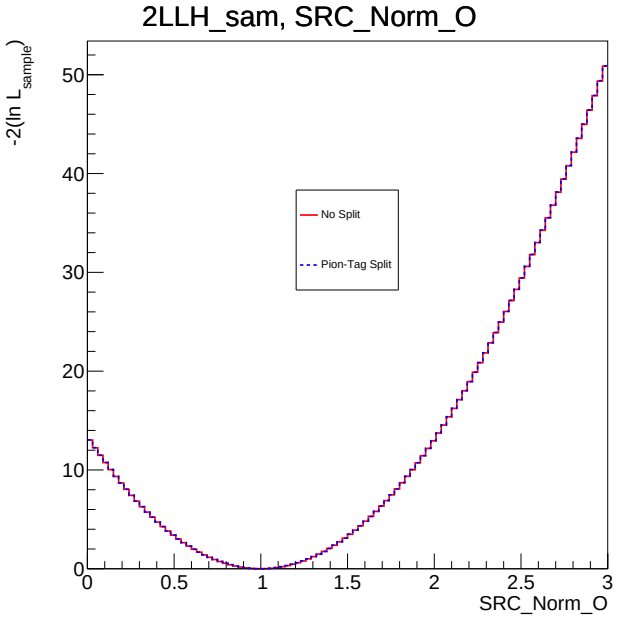
(c) $P_{3/2}$ Shell Mean Field Normalisation for Oxygen

(d) SShell Mean Field Normalisation for Oxygen

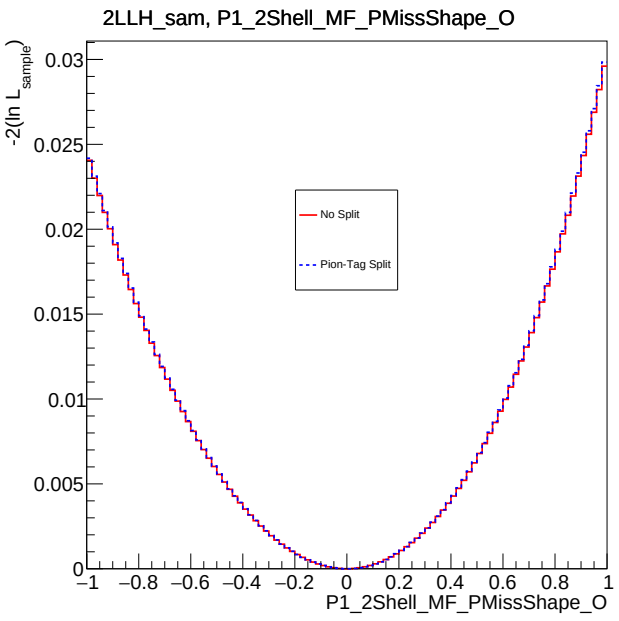
FIGURE 7.3: CCQE Parameters.

The mean field normalisation parameters for Carbon and Oxygen form a set of shape parameters for CCQE interactions in Figure 7.3. The oxygen parameters describe the shell model of the nucleus for oxygen with P and S shells, which are used to model CCQE interactions. The same

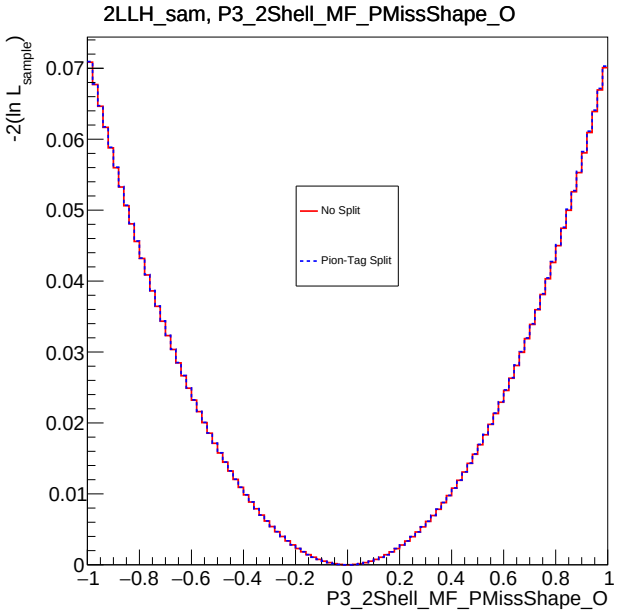
constraints are seen between the two features. This is an expected response as the split does not affect the CC0Pi sample which is CCQE interaction rich.



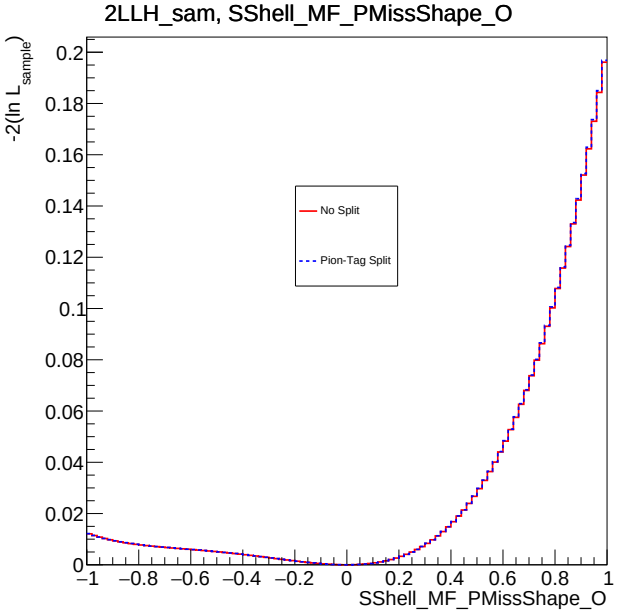
(a) Short Range Correlations Normalisation for Oxygen



(b) $P_{1/2}$ Shell Mean Field P MissShape Normalisation for Oxygen



(c) $P_{3/2}$ Shell Mean Field P MissShape Normalisation for Oxygen

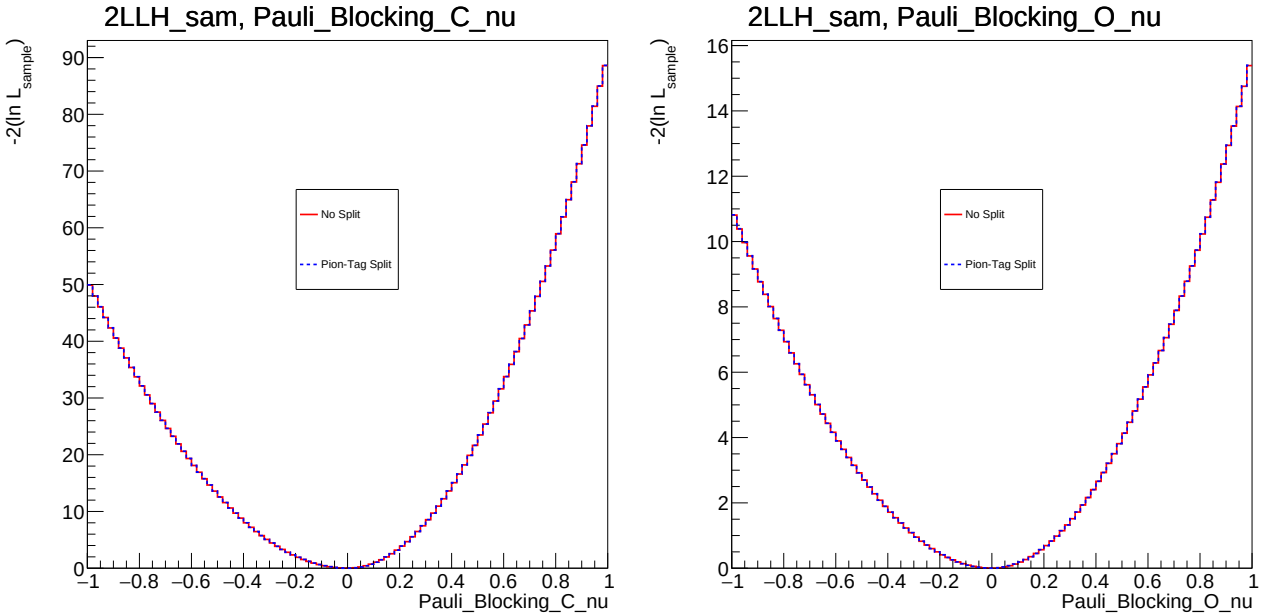


(d) S Shell Mean Field P MissShape for Oxygen

FIGURE 7.4: CCQE Parameters.

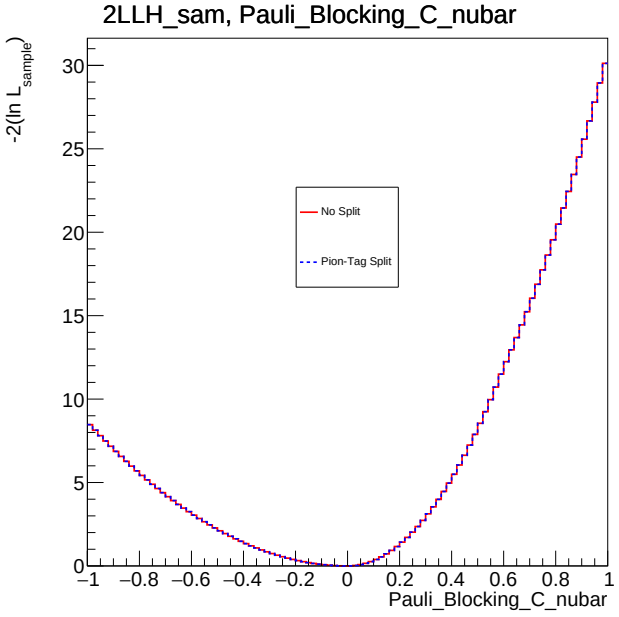
The short range correlation parameter and the mean field shape parameters for oxygen are shown in Figure 7.4, which play a role in modelling FGD2 CCQE interactions (due to the presence of water in FGD2) and interactions at the far detector. The mean field normalisation

parameters for oxygen form a set of shape parameters for CCQE interactions describing the shell model of the nucleus for oxygen with P and S shells. The same constraints are seen between the two features which is expected as the split does not affect the CC0Pi sample which is CCQE rich.

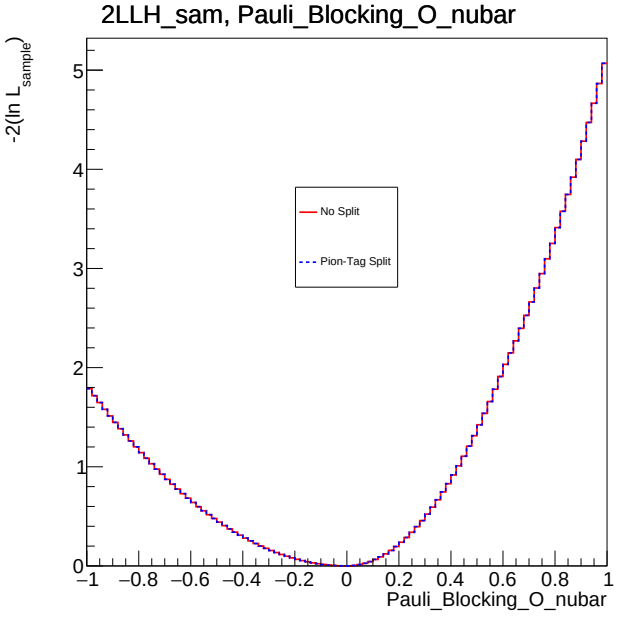


(a) Pauli Blocking ^{12}C ν

(b) Pauli Blocking ^{16}O ν



(c) Pauli Blocking ^{12}C $\bar{\nu}$

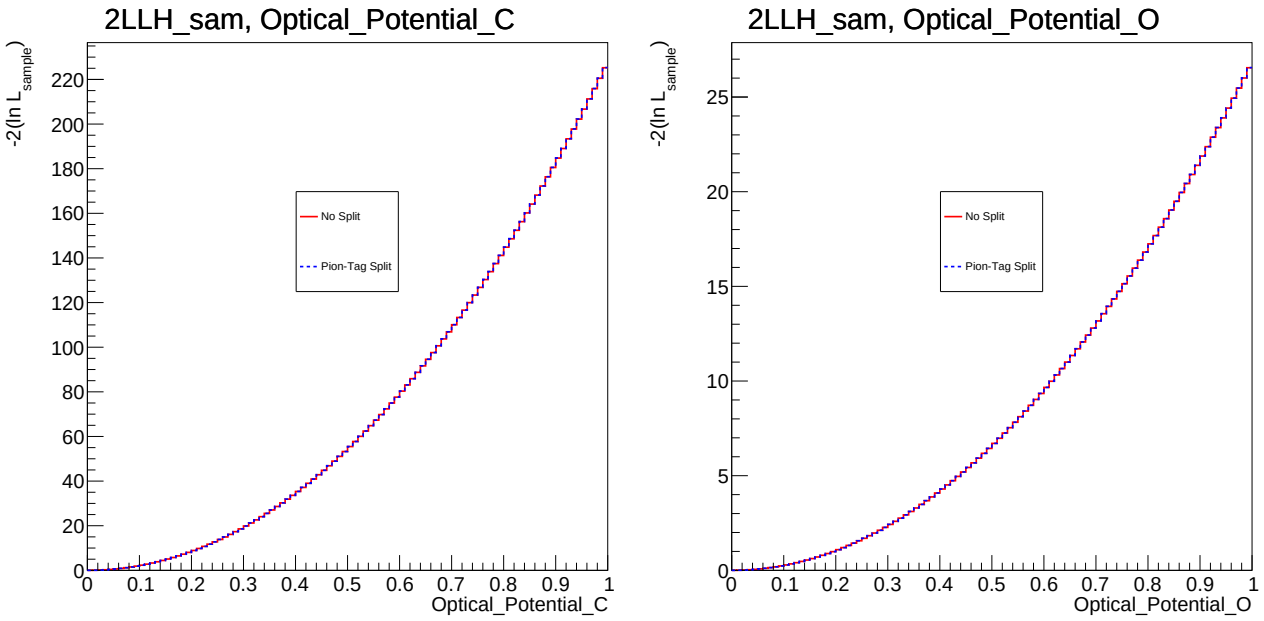


(d) Pauli Blocking ^{16}O $\bar{\nu}$

FIGURE 7.5: CCQE Parameters.

The Pauli Blocking parameters are shown for neutrino and antineutrinos on Oxygen and Carbon in Figure 7.5. The nucleons in the target are not stationary in the ground state, and move with a certain momentum. In the ground state of the nucleus, energy levels are occupied up until the

Fermi momentum. For interactions to occur, the incoming neutrino must excite these nucleons to beyond this momentum. This is known as Pauli blocking. The parameters above affect interactions with nucleons within a CCQE interaction. The same constraints are seen between the two features which is expected as the split does not affect the CC0Pi sample which is CCQE rich.



(a) Optical Potential for Carbon

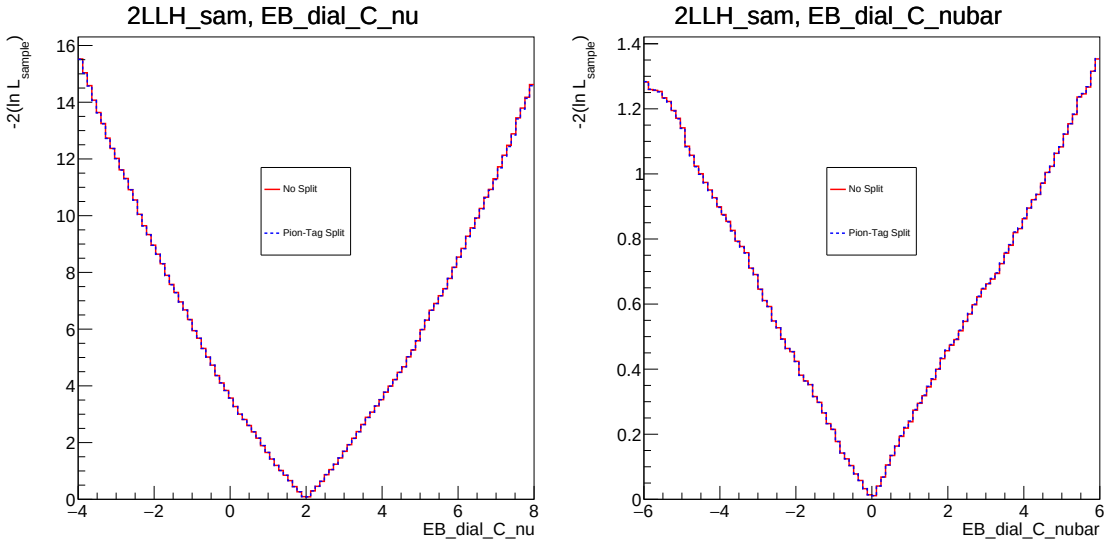
(b) Optical Potential for Oxygen

FIGURE 7.6: CCQE Parameters.

For systems with many degrees of freedom, such as the nucleus of carbon or oxygen, a mean field approximation has to be made. The mean field approximation looks at each nucleon of the system moving under the effect of the rest. The optical potential parameter for Carbon and Oxygen, in Figure 7.6 above, describes final state nucleon interactions for CCQE interactions. This parameter has a flat prior and the effect is not symmetric in all bins, therefore there it has a different shape. However, there is still no difference seen between the two features in this scan, as expected.

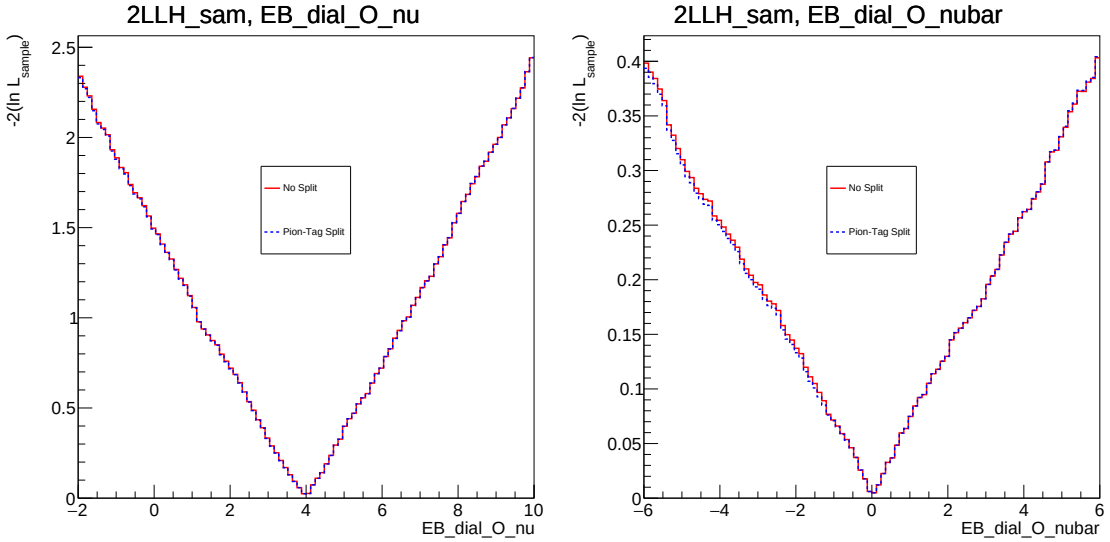
The CCQE model parameters alter both nucleon and lepton kinematics within a CCQE interaction. Carbon and Oxygen uncertainties are applied when required per interaction. Ultimately, for the 22 CCQE parameters above, it can be seen that there are no significant difference in the LLH scans. The dotted blue line represents the split feature, whereas the red line represents the chi squared value for the no-split feature.

It can be seen in Figures 7.1 - 7.6 that there is no additional constraint between the split and no-split histograms above. This is expected for the CCQE parameters as they are linked to the majority of events in the CC0Pi histograms which are identical between the two features.



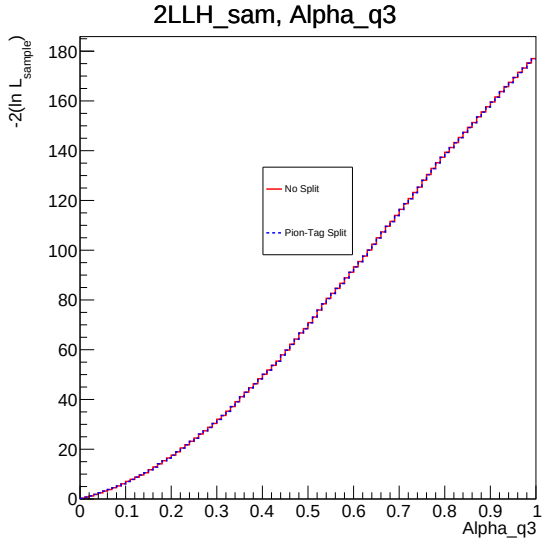
(a) CCQE E_b $^{12}C\,\nu$

(b) CCQE E_b $^{12}C\,\bar{\nu}$



(c) CCQE E_b $^{16}O\,\bar{\nu}$

(d) CCQE E_b $^{16}O\,\bar{\nu}$

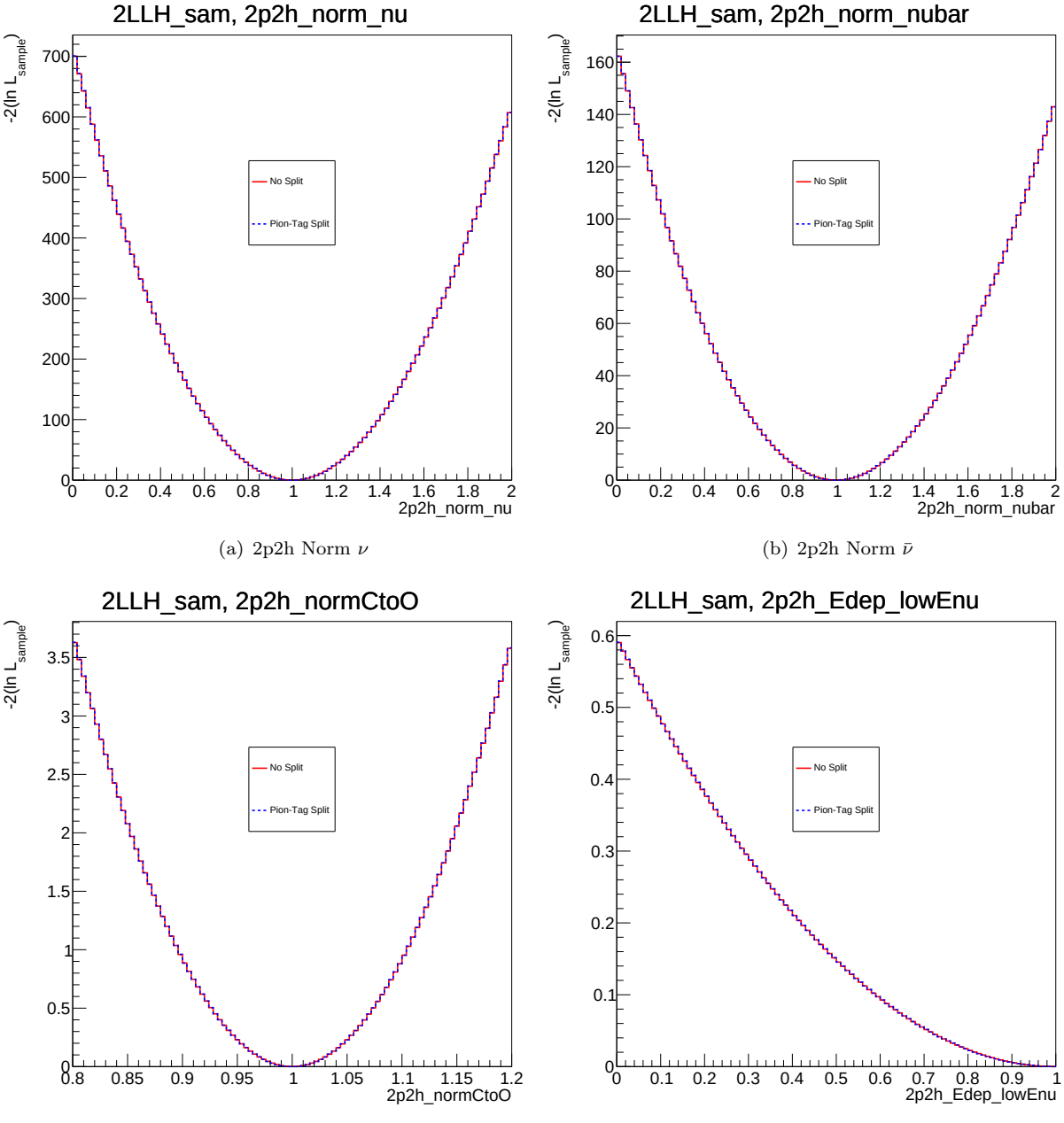


(e) α correction (q_3)

FIGURE 7.7: CCQE Binding Energy Parameters. The α correction (q_3) has a flat prior.

The CCQE Binding Energy parameters are shown above in Figure 7.18. These parameters allow for the energy shifts needed to remove nucleons from the nucleus. In MaCh3, the result of these shifts are implemented directly on lepton kinematics, leading to event migration from one bin to another. Additionally, the α correction is included in the binding energy calculation as a result of its better describing data from electron scattering experiments [128].

7.1.2 2p2h Parameters



(a) 2p2h Norm ν

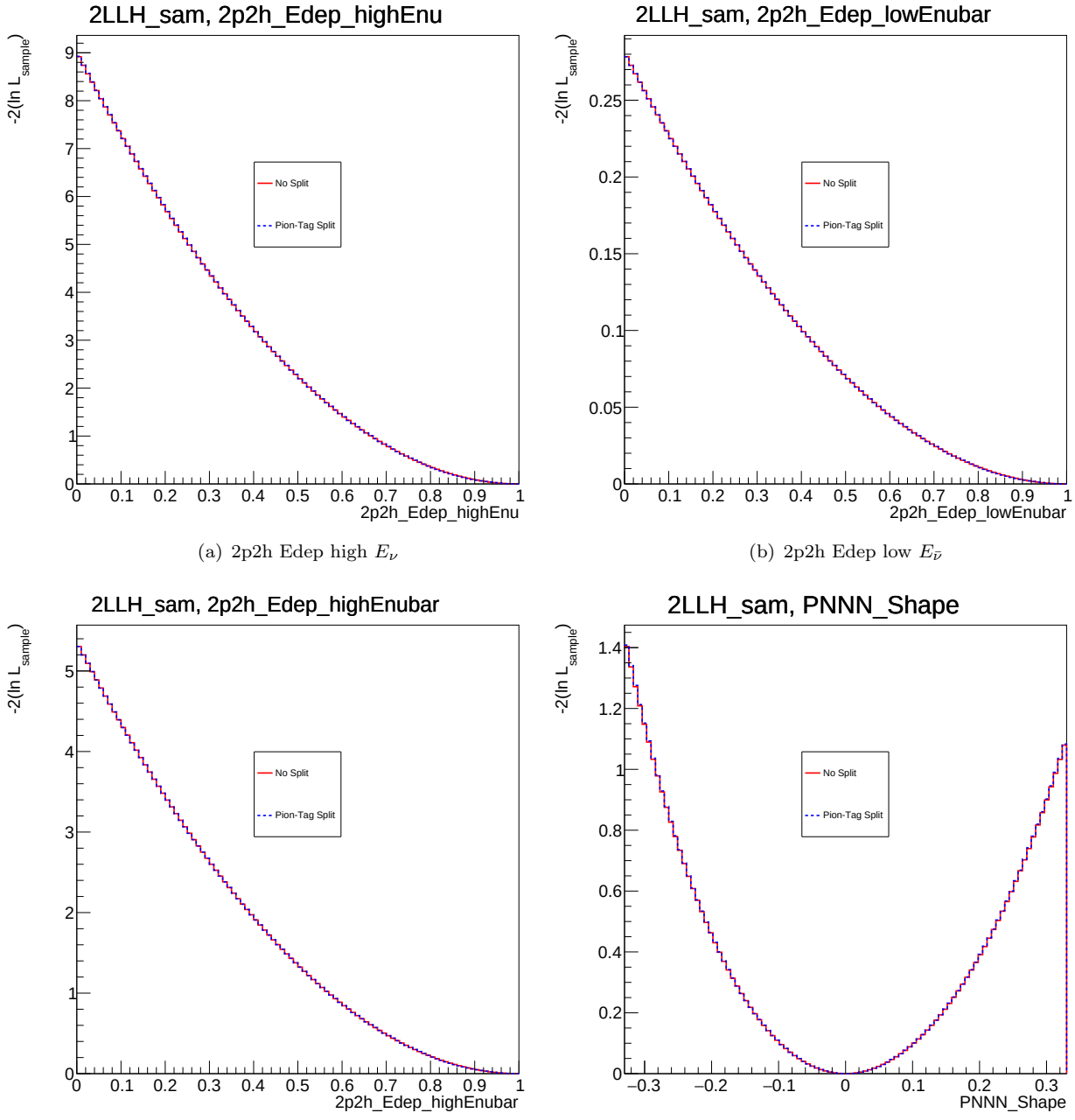
(b) 2p2h Norm $\bar{\nu}$

(c) 2p2h Norm $^{12}C/^{16}O$

(d) 2p2h Edep low E_ν

FIGURE 7.8: 2p2h Parameters.

Three normalisation parameters are shown above in Figure 7.8 for neutrino and antineutrino 2p2h interactions. No additional constraints are seen in these parameters in Figure 7.8 above. The 2p2h Norm $^{12}C/^{16}O$ parameter allows the scaling of the 2p2h cross-section from carbon to oxygen which is needed at the far detector. The 2p2h energy dependence low E_ν shape parameter allows for the modelling of the energy dependence of the 2p2h interactions. There is no change seen in constraint between the two features which is expected as the majority of 2p2h interactions are captured in the CC0Pi histogram which has not been split by pion tag.



(a) 2p2h Edep high E_ν

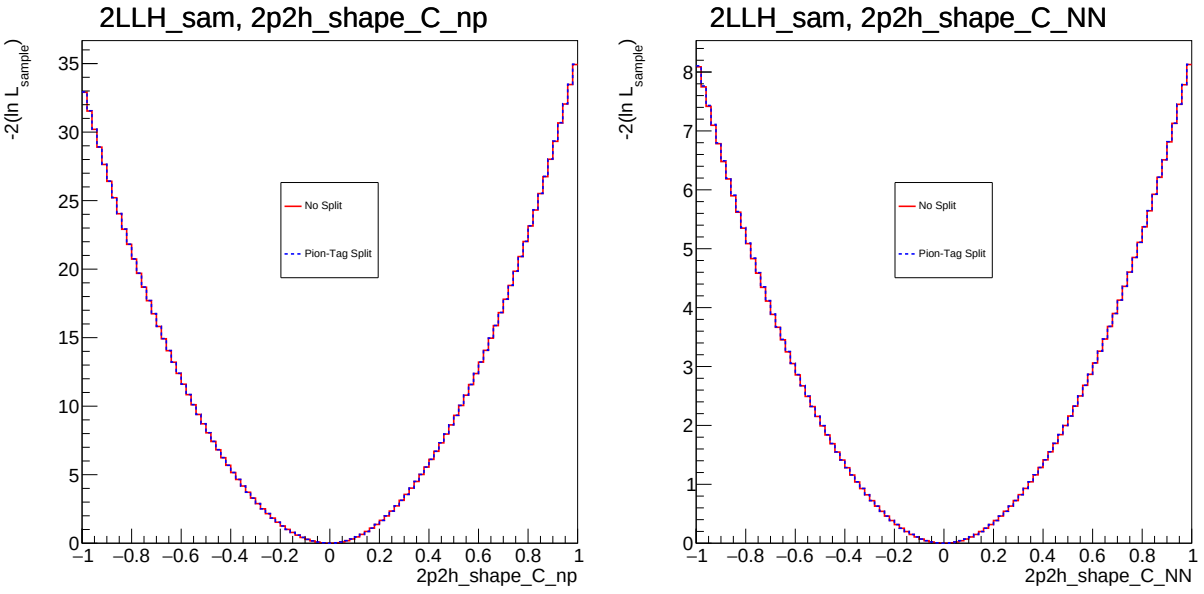
(b) 2p2h Edep low $E_{\bar{\nu}}$

(c) 2p2h Edep high $E_{\bar{\nu}}$

(d) PN-NN Shape

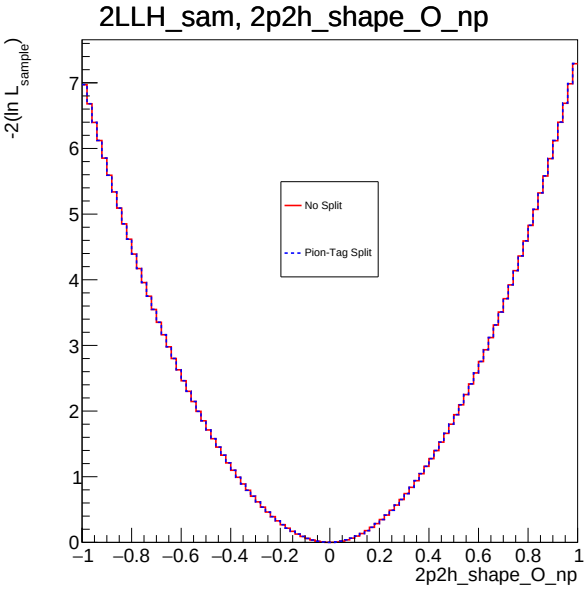
FIGURE 7.9: 2p2h Parameters.

In Figure 7.9, the 2p2h energy dependence E_ν and $E_{\bar{\nu}}$ shape parameters allows for the modelling of the energy dependence of the 2p2h interactions. These all have a flat prior. The PNNN shape parameter accounts for the fact that neutrinos and antineutrinos interact with different nucleon pairs. In a 2p2h interaction, neutrinos interact with *pn* or *nn* nucleon pairs, whereas antineutrinos interact with *pn* or *pp*. This is important in populating the CC0pi 0 Proton and N Proton (no photon) samples correctly. There is no change seen between the two scans which is expected as the majority of 2p2h interactions are captured in the CC0Pi histogram which is the same between the two features.

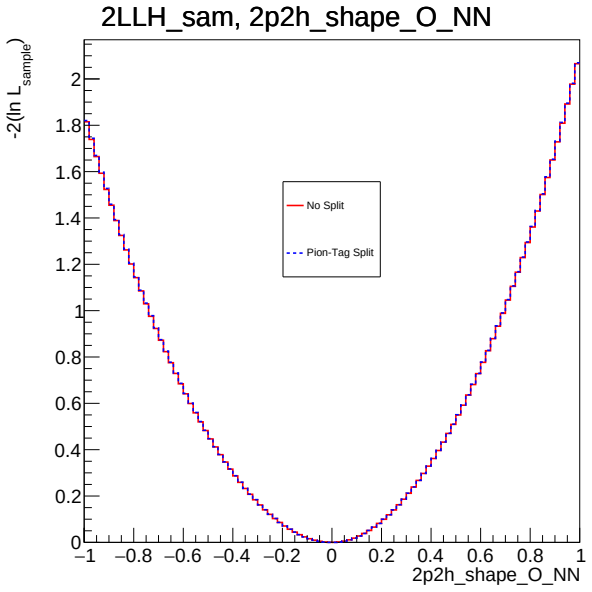


(a) 2*p*2*h* shape ^{12}C np

(b) 2*p*2*h* shape ^{12}C nn



(c) 2*p*2*h* shape ^{16}O np

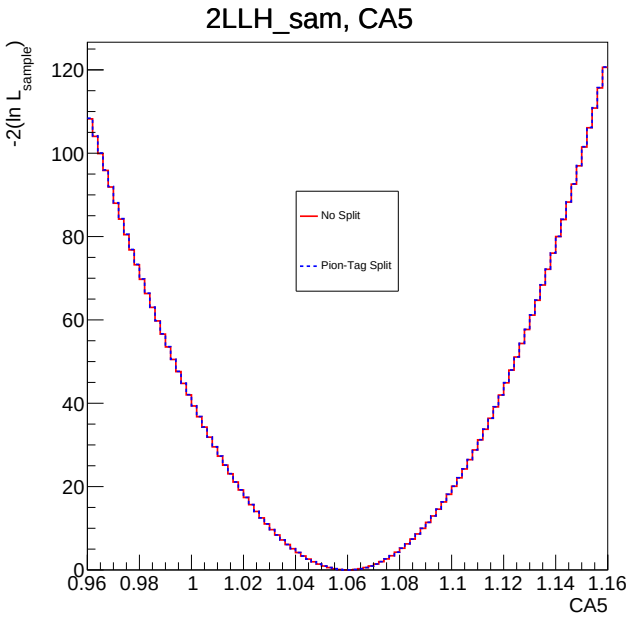


(d) 2*p*2*h* shape ^{16}O nn

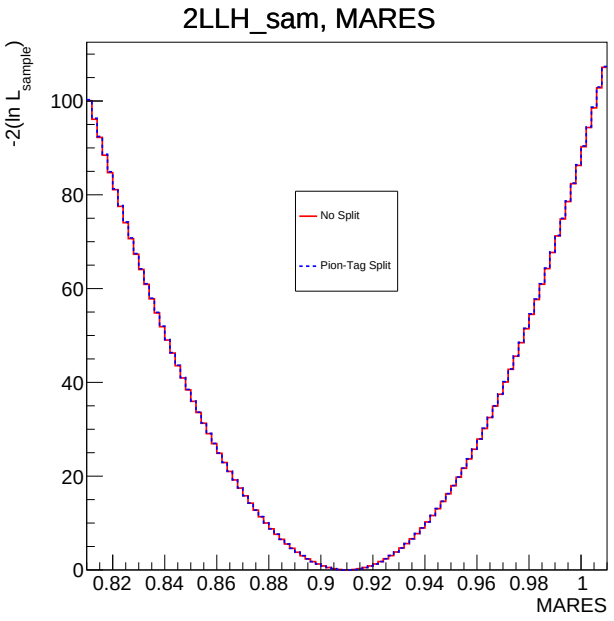
FIGURE 7.10: 2p2h Parameters.

The parameters shown in Figure 7.10 above are nucleon pair shape parameters which accounts for different ends of the 2p2h interaction energy spectrum. These change the momentum and energy distribution; at one end of the parameter, the interactions are from pion rich delta decays as the Nieves model predicts. Whereas at the other end, pion-less delta decay occurs in accordance with the Martini model.

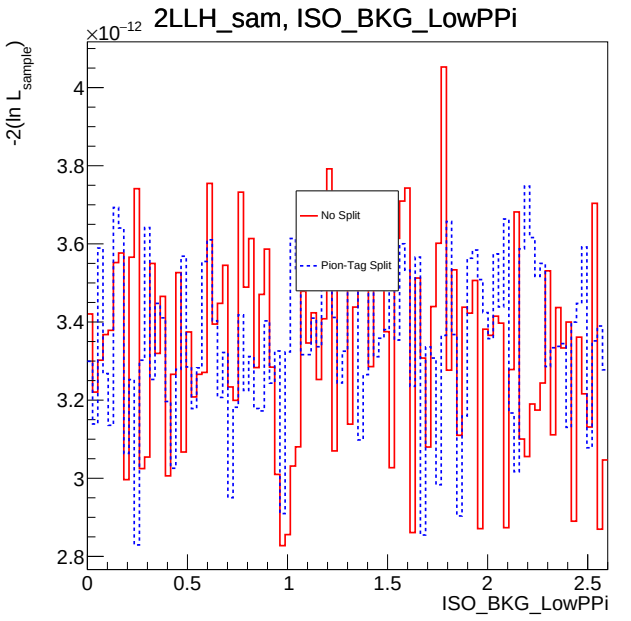
7.1.3 CCRES Single Pion Production (SPP)



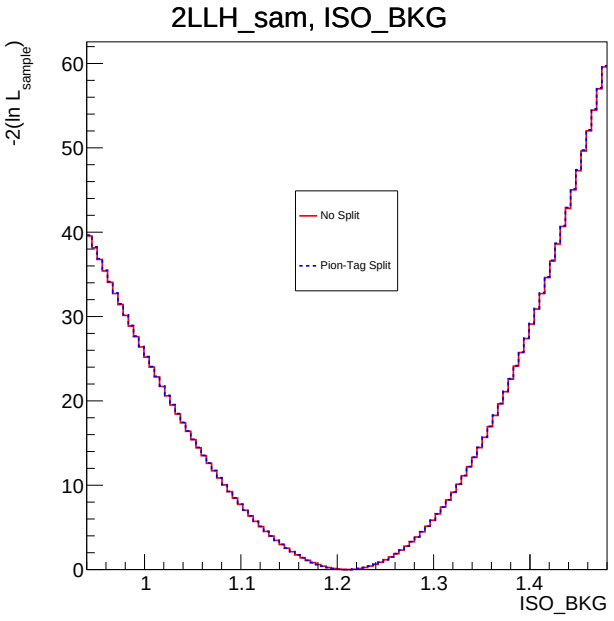
(a) C_5^A, the axial form factor for CCRES interactions



(b) M_A^{RES}, the axial mass for CCRES interactions



(c) Non-res $I_{1/2}$ Low p_π, y-axis scale negligible at 10^{-12}

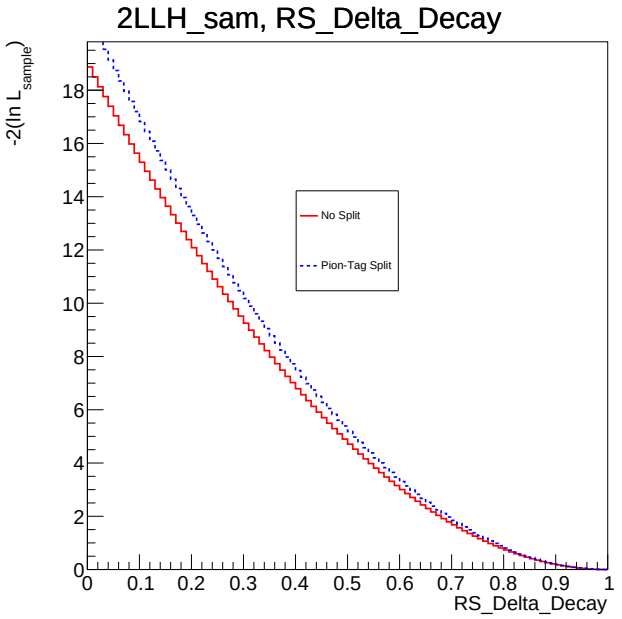


(d) Non-res $I_{1/2}$

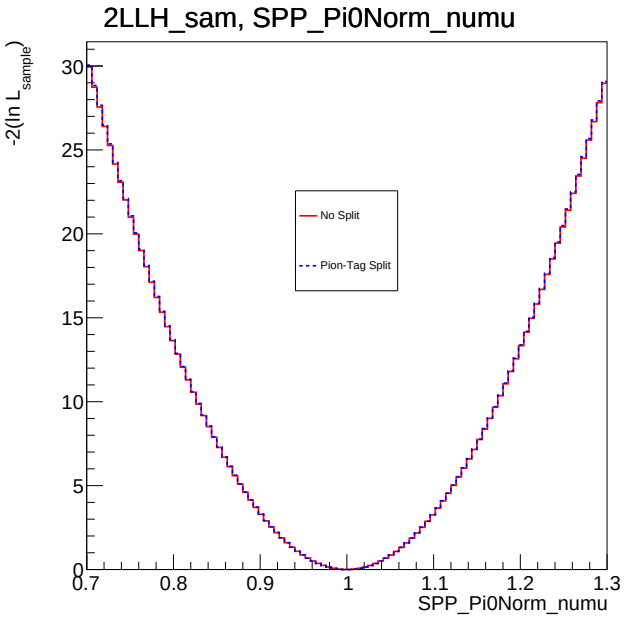
FIGURE 7.11: SPP Parameters.

The Non-res $I_{1/2}$ Low p_π parameter (Figure 7.11) is exclusive to the far detector, so LLH scans at the near detector only show statistical fluctuations. Furthermore, in Figure 7.11, M_A^{RES}, the resonance axial mass, C_5^A, the axial form factor normalisation, and Non-res $I_{1/2}$ are non-CCQE parameters which show a correspondence between the two features as well. The Non-res $I_{1/2}$ parameters (c) and (d) describe the backgrounds to the pion production in CCRES due to

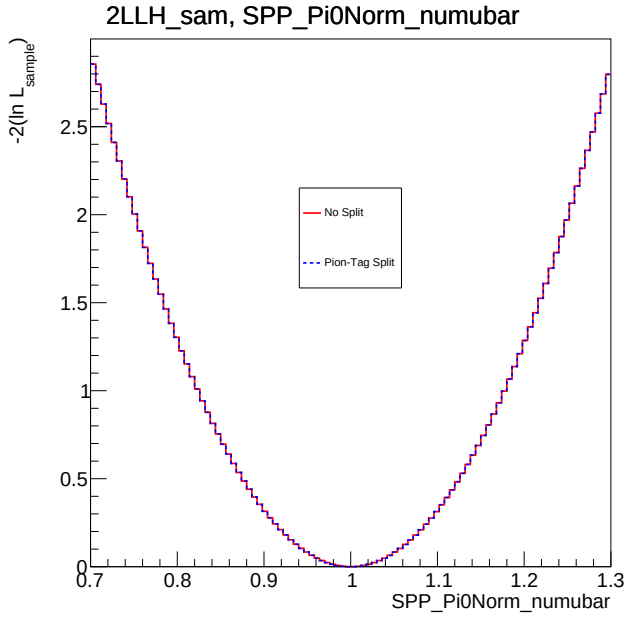
isospin interaction between neutrons and protons. It may be seen that separating the sample by pion tag does not add any additional constraints in these parameters.



(a) Rein-Seghal Delta Decay



(b) SPP π^0 Norm ν_μ

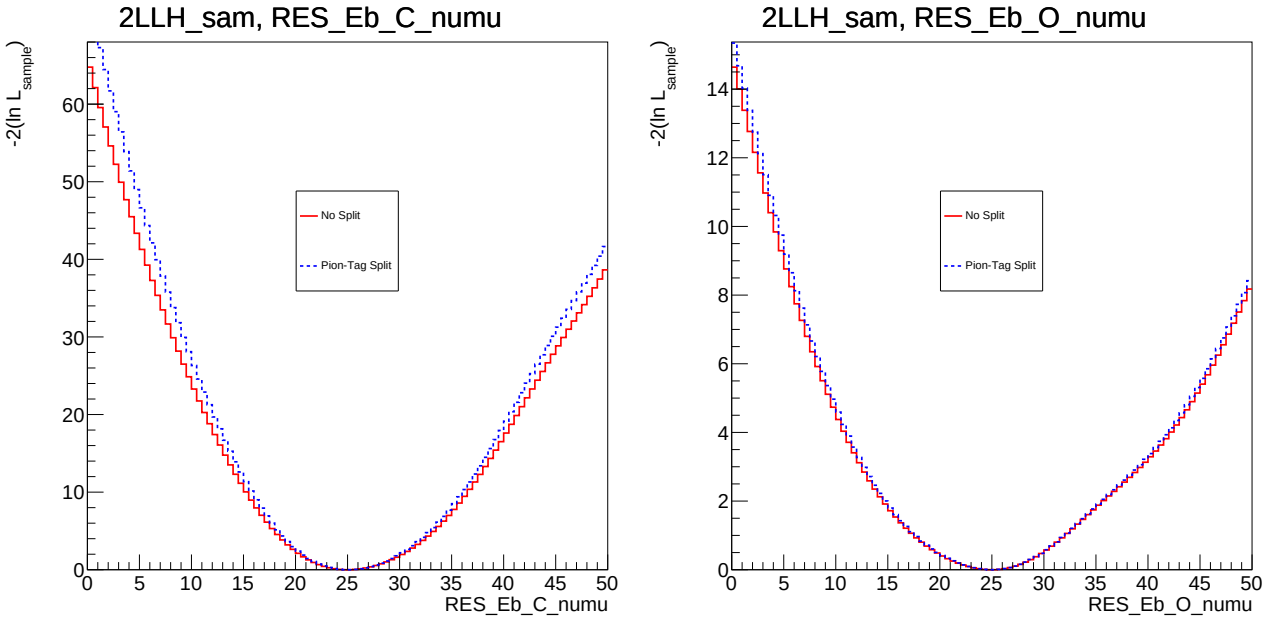


(c) SPP π^0 Norm $\bar{\nu}_\mu$

FIGURE 7.12: SPP Parameters.

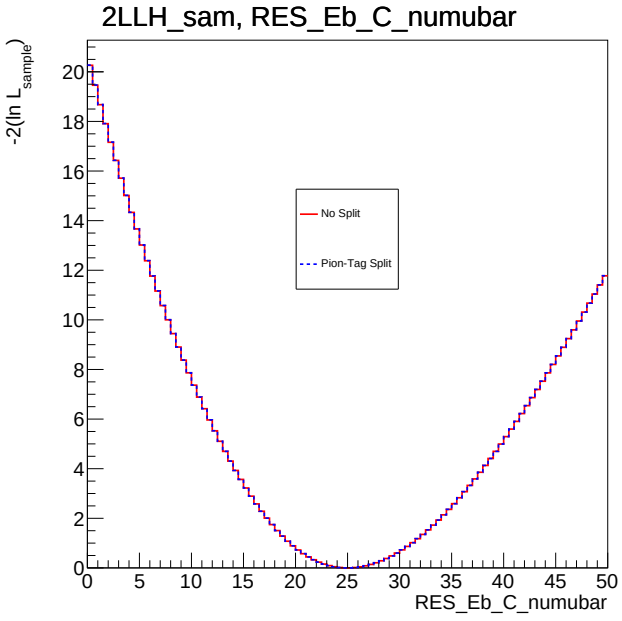
SPP parameters are shown above in Figure 7.12. The Rein-Seghal delta decay spline parameter has a flat prior which affects the way that pions are ejected from the nucleus. At one end of the parameter, a pion is ejected in an isotropic fashion with the nucleon in the rest frame. At the other end, the delta resonance transition is calculated using elements of the matrix and

contracting it with spherical harmonics. Splitting the CC1Pi sample by pion-tag type leads to a better constraint in this parameter. The following two SPP parameters are normalisation parameters for CCRES interactions that produce one π_0 (for neutrinos and anti neutrinos interactions respectively).

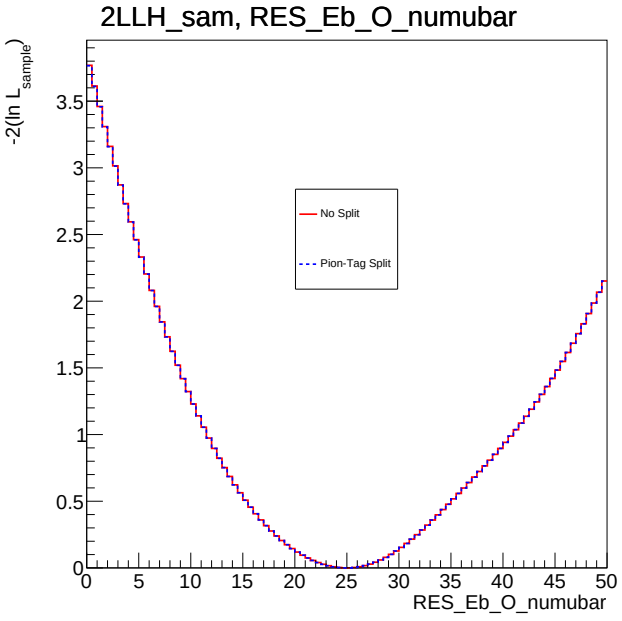


(a) RES E_b ^{12}C ν_μ

(b) RES E_b ^{16}O ν_μ



(c) RES E_b ^{12}C $\bar{\nu}_\mu$



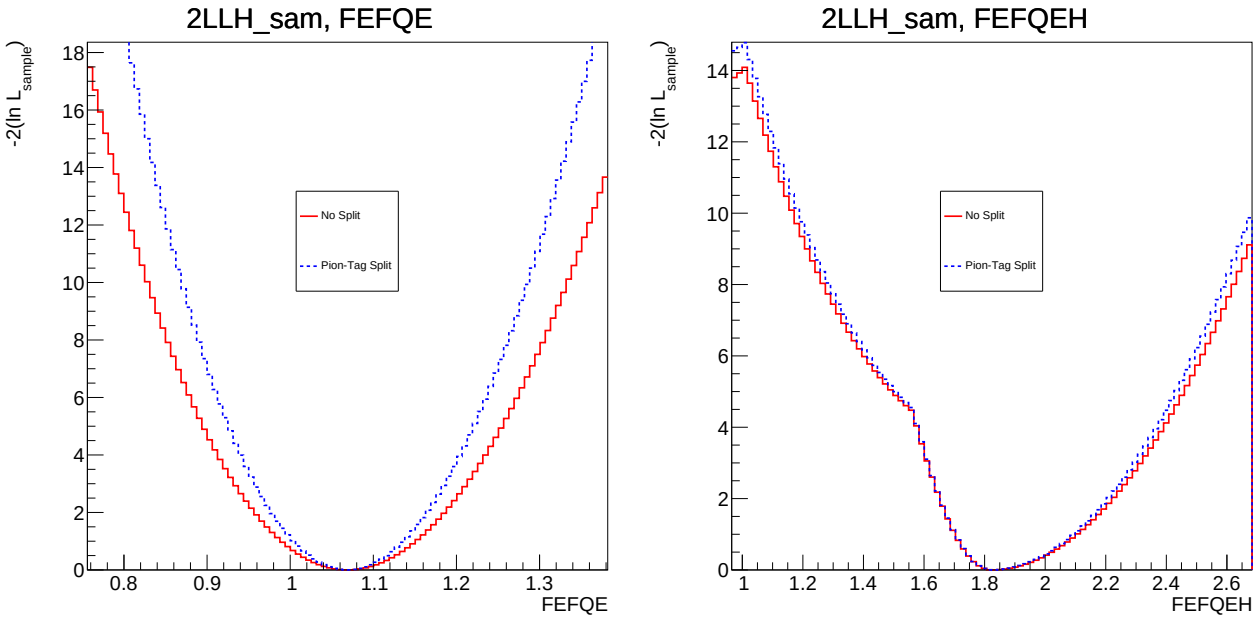
(d) RES E_b ^{16}O $\bar{\nu}_\mu$

FIGURE 7.13: CCRES Parameters.

Additional constraints are seen in the CCRES binding energy parameters, (a) and (b) with the split feature (Figure 7.13). Unlike CCQE binding energy parameters which are shifts applied

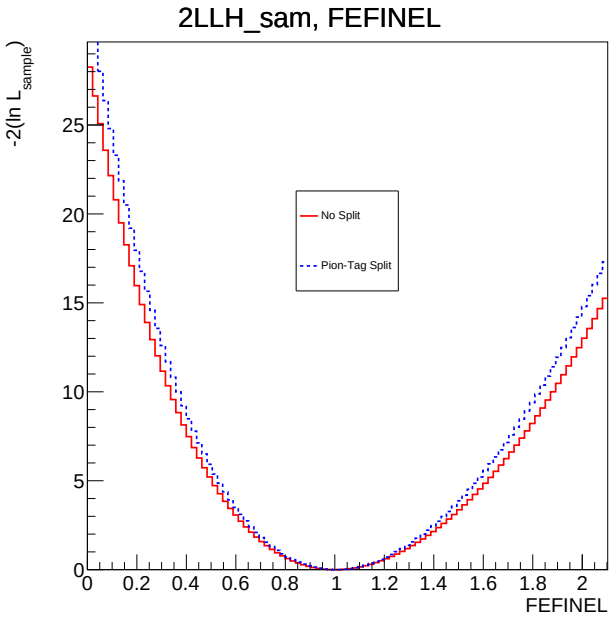
directly to lepton kinematics, the CCRES binding energy parameters are applied during the re-weighting process. There are four parameters in total: one for neutrinos with carbon, one for antineutrinos with carbon, one for neutrinos with oxygen, one for anitneutrinos with oxygen. Better constraints are seen in the split feature for carbon and oxygen binding energy parameters with neutrinos.

7.1.4 Final State Interaction (FSI) Parameters

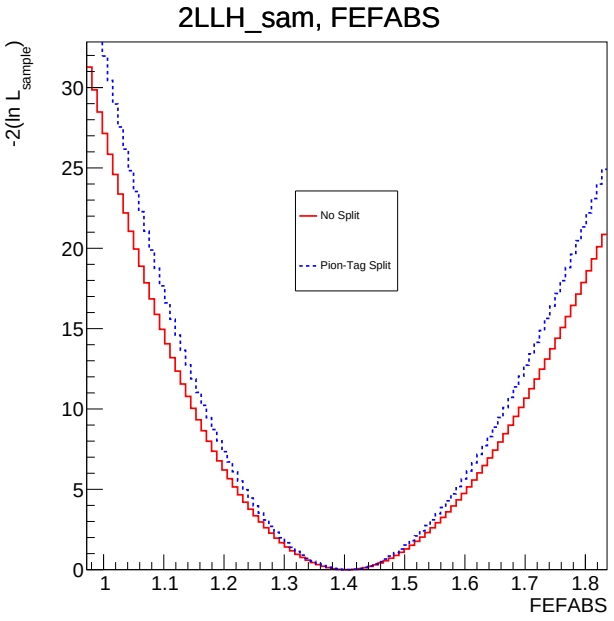


(a) FSI Quasi-elastic (low energy) parameter

(b) FSI Quasi-elastic (high energy) parameter



(c) FSI Inelastic (pion production) parameter

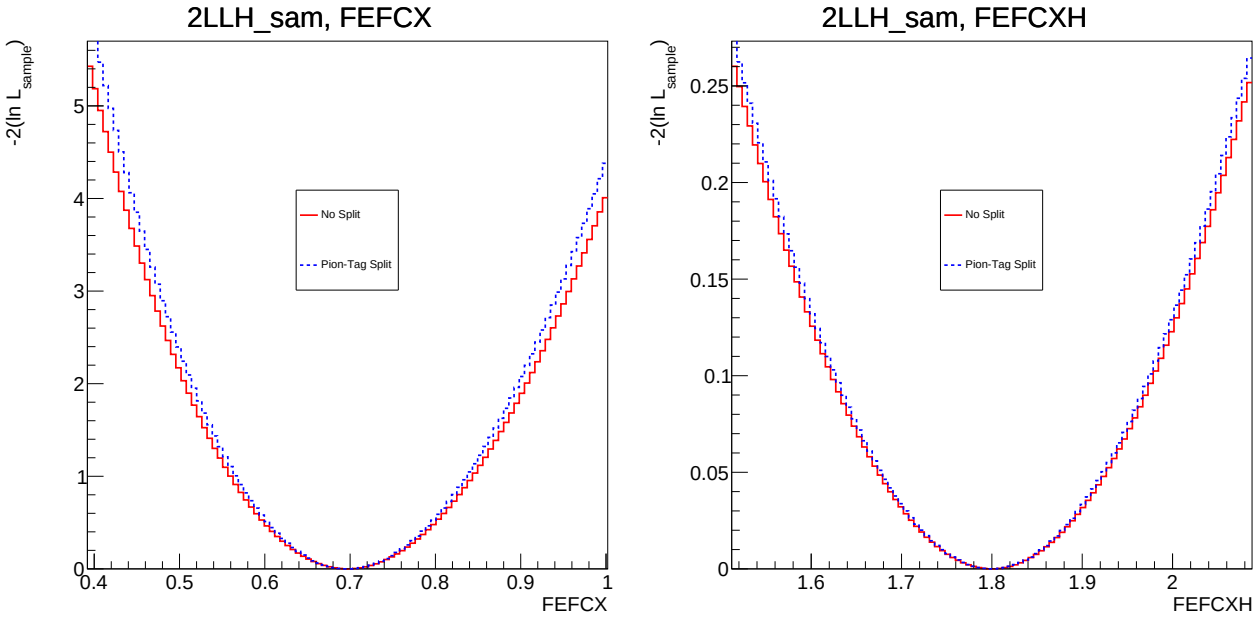


(d) FSI Absorption (pion absorption) parameter

FIGURE 7.14: FSI Parameters.

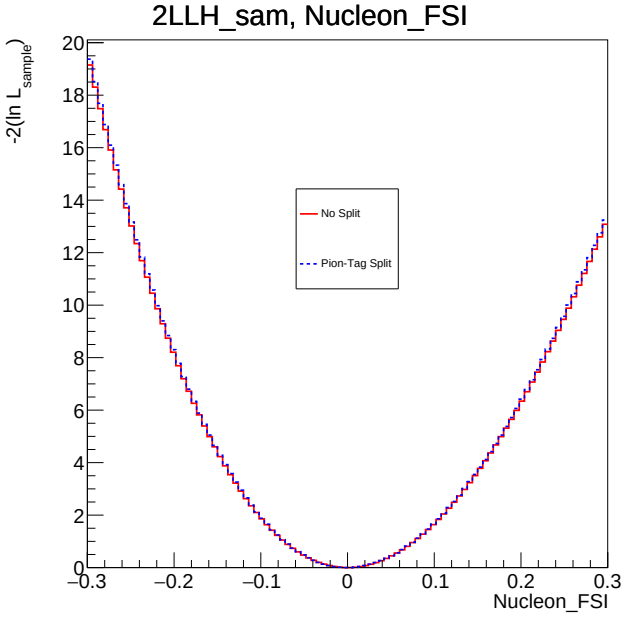
In Figure 7.14, a definite constraint can be seen above in the 4 FSI parameters: FEFQE (pion quasi-elastic scattering at low energy), FEFQEH (pion quasi-elastic scattering at high energy), FEFINEL (hadron production), FEFABS (pion absorption). There is a kink in the FSI Quasi-elastic (high energy) parameter LLH scan due to a weight cap, which is well known and studied [109]. It is expected that by splitting the CC1Pi histogram by pion tag, and increasing the

number of bins, events with pions are less washed out and the effect of varying FSI parameters are seen much clearer.



(a) FSI Charge Exchange (low energy) parameter

(b) FSI Charge Exchange (high energy) parameter



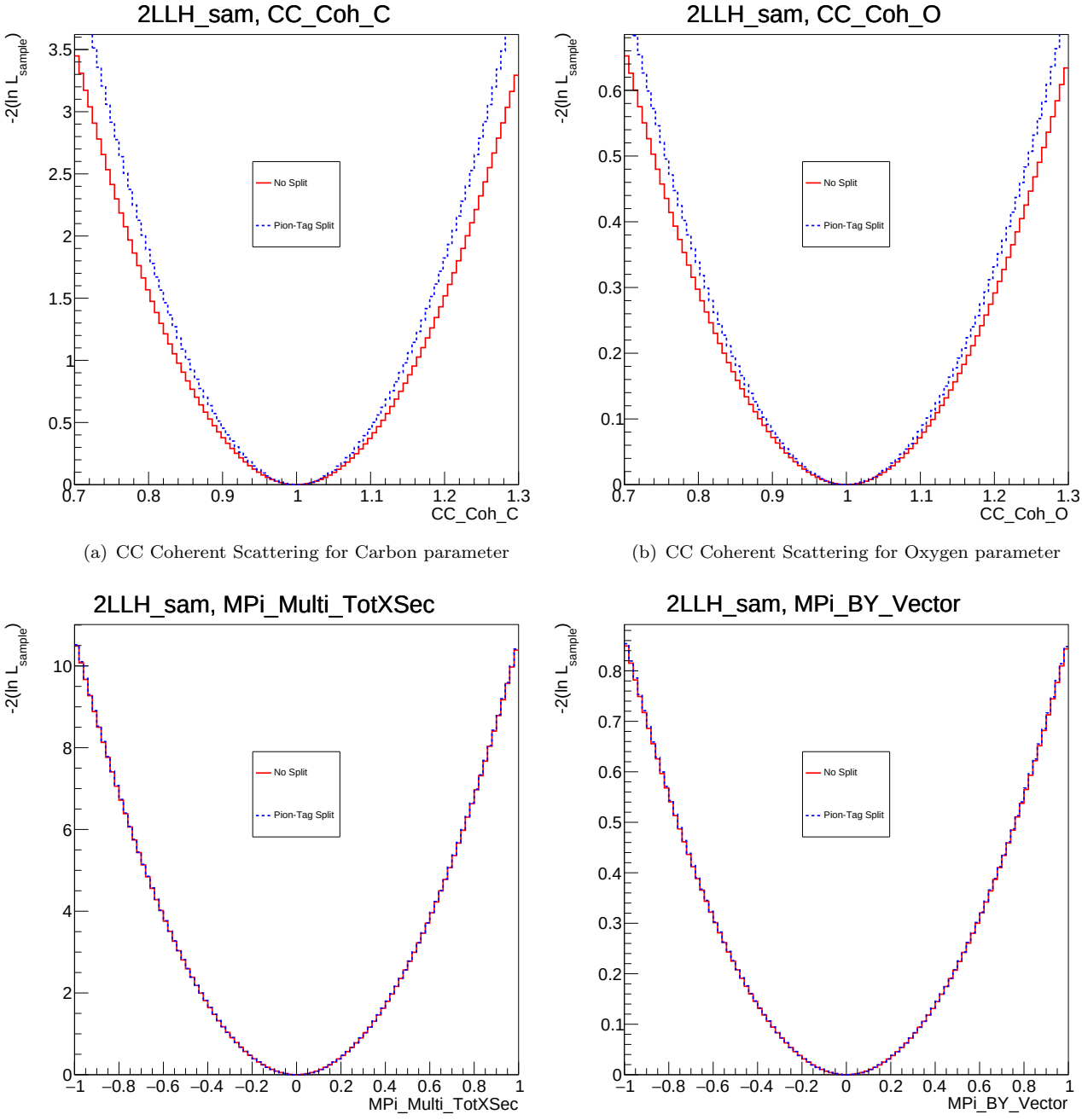
(c) Nucleon FSI parameter

FIGURE 7.15: FSI Parameters.

A constraint can be seen here in Figure 7.15 for the following 3 FSI parameters: FEFCX (single charge exchange at low energy), FEFCXH (single charge exchange at high energy), and Nucleon FSI. Significant constraints can be seen in all of the FSI parameters for the pion-tag split feature in Figure 7.14 and 7.15 above. This is due to the fact that separating by pion tag in lepton

kinematics increases the number of bins, separating between events and spreading them over the phase space. This means that the effect of the FSI re-weighting can be seen much clearer.

7.1.5 CCDIS, Multi-Pi, CC Coherent, Neutral Current, ν_e Parameters



(a) CC Coherent Scattering for Carbon parameter

(b) CC Coherent Scattering for Oxygen parameter

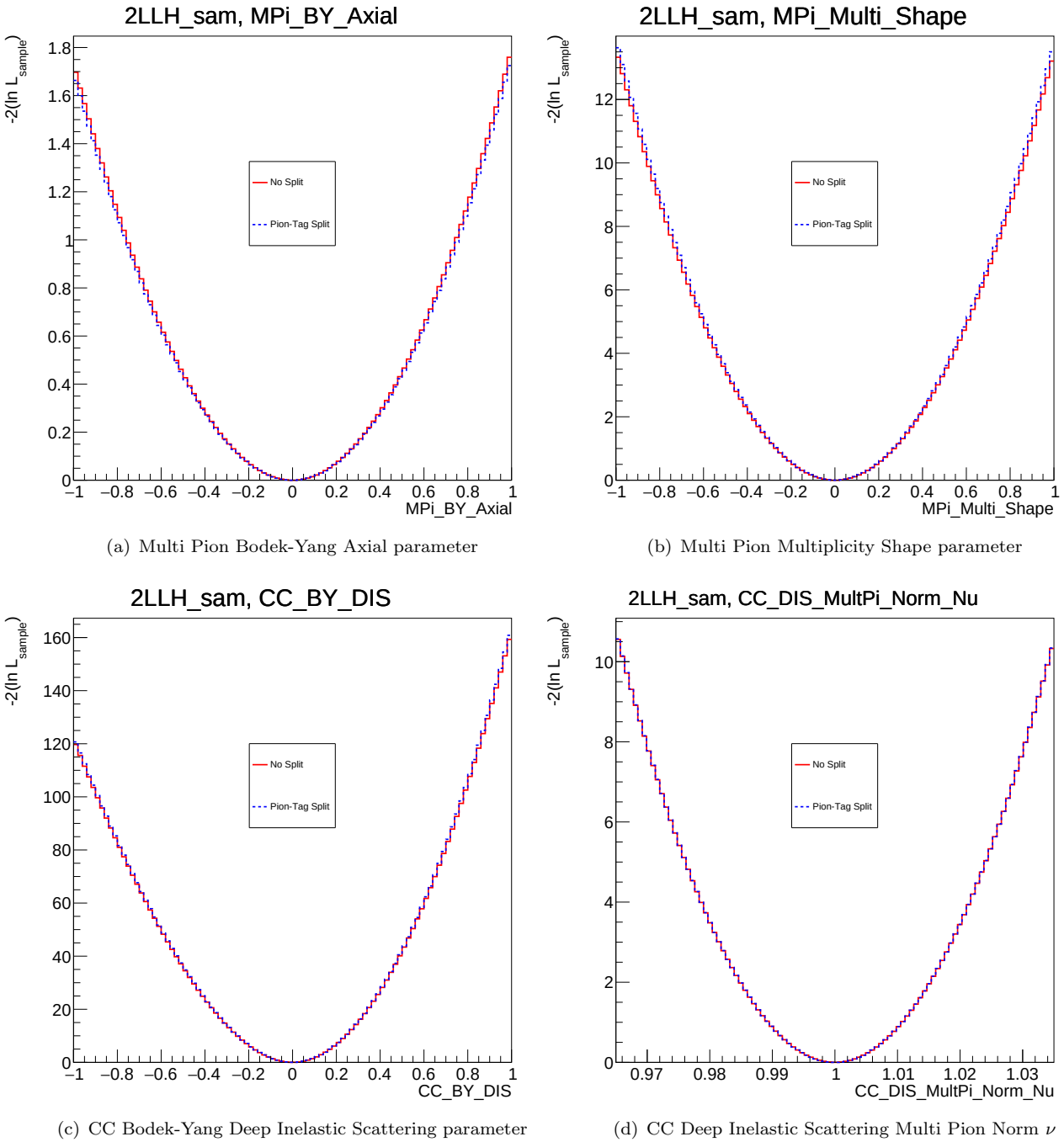
(c) Multi Pion Multiplicity Total XSec parameter

(d) Multi Pion Bodek-Yang Vector parameter

FIGURE 7.16: CC Coherent and CCDIS parameters.

In Figure 7.16 above, the coherent interaction is described by separate normalisation parameters for carbon and oxygen. Clear constraints are seen on the coherent parameters. This is due to the fact that separating by pion tag allows the TPC tag, with a significant coherent scattering contribution at T2K energies (see Chapter 6, Section 6.3.3, Figure 6.7), to become prominent as it does not get washed out by resonant pions at low energies from the other tags. From the

multi-pion parameters used to describe CCDIS interactions; (c) describes the total cross section for multiple particle production which models the phenomena using two different generators; one is the AKGY model and the second is the nominal NEUT model, Parameter (d) describes the Bodek-Yang model for DIS (the rest of the parameters shown below in Figure 7.17 below).



(a) Multi Pion Bodek-Yang Axial parameter

(b) Multi Pion Multiplicity Shape parameter

(c) CC Bodek-Yang Deep Inelastic Scattering parameter

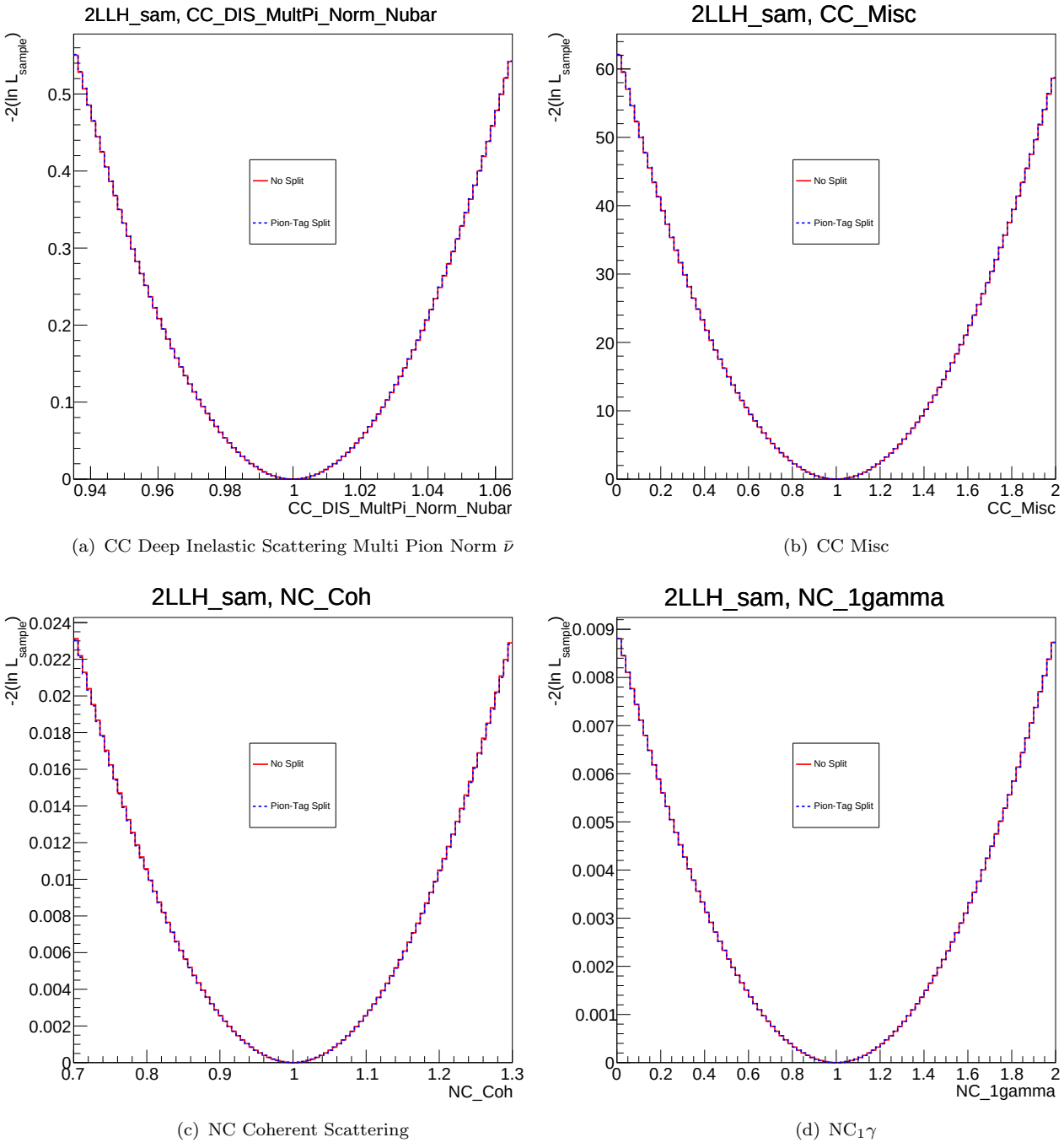
(d) CC Deep Inelastic Scattering Multi Pion Norm ν

FIGURE 7.17: CCDIS parameters.

In Figure 7.17 above, parameters (a) and (c) carry on describing the Bodek-Yang model for deep inelastic scattering. DIS interactions are modelled using Parton Distribution Functions (PDFs)

in NEUT with corrections from Bodek and Yang. Parameter (b) is an additional parameter which allows for the total cross section to remain fixed whilst cross section shape changes occur between the AKGY model and NEUT model.

In Figure 7.17 above, the final parameter (d) allows for a change in the amount of DIS interactions due to ν interactions; there is no difference in constraint between the split and the no-split feature. There are small changes seen in the scans, with parameter (a) showing a slightly better constraint with no tag splitting. Whereas parameters (b) and (c) show a slightly better constraint with the split feature. The constraints are the smallest seen with no large variations seen between the two features. This is expected as the majority, but not all, of CCDIS interactions end up in the CCOther sample where the binning between the two features are identical.



(a) CC Deep Inelastic Scattering Multi Pion Norm $\bar{\nu}$

(b) CC Misc

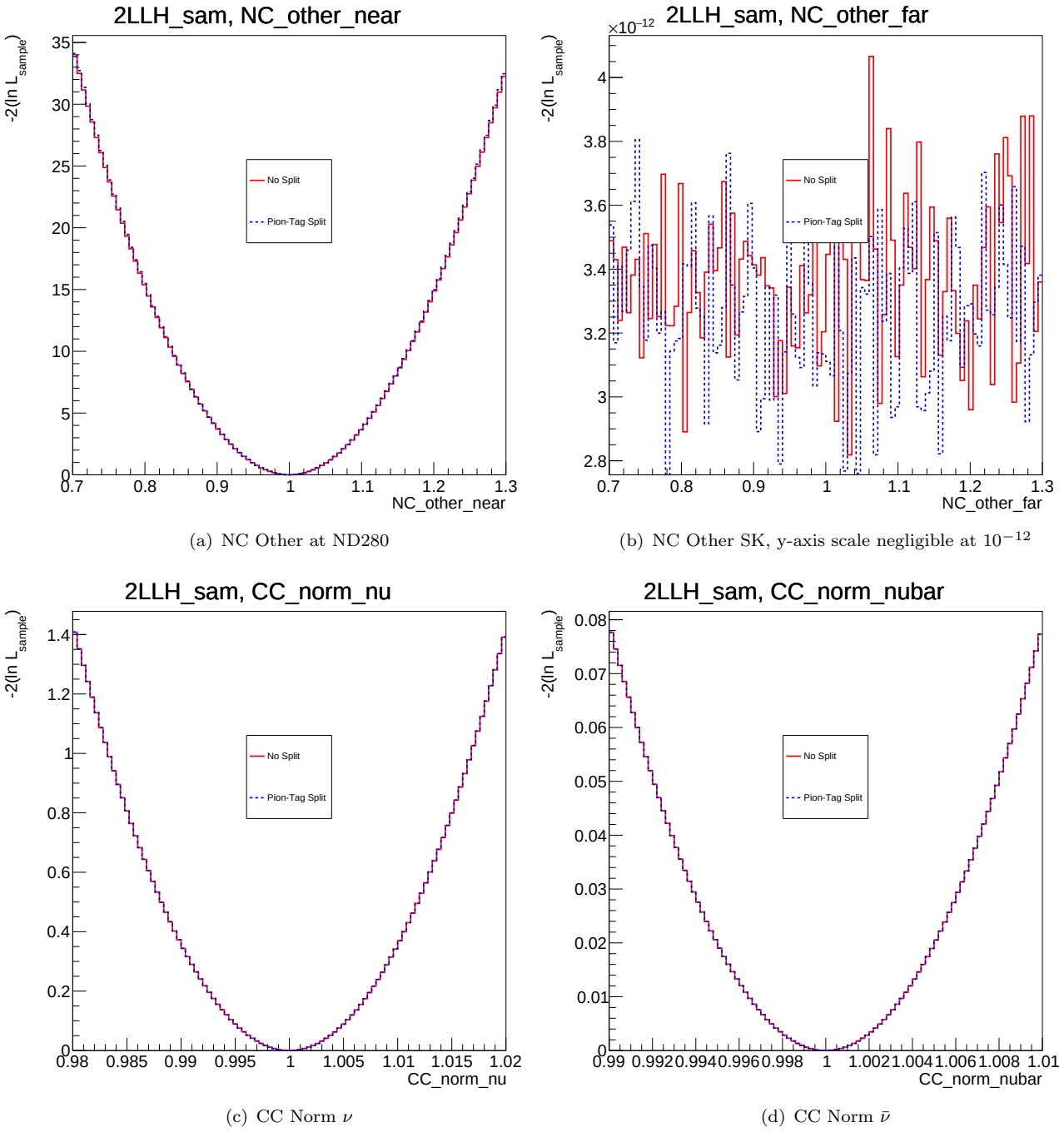
(c) NC Coherent Scattering

(d) $NC_1\gamma$

FIGURE 7.18: CCDIS, CC Misc, and NC Parameters.

In Figure 7.18 above parameter (a) allows for a change in the amount of DIS interactions due to $\bar{\nu}$ interactions. The CC Misc parameter (b) covers all of the interactions which produce η particles or Kaons. The NC parameters describe neutral current coherent interactions which has a normalisation parameter (c) and are modelled in the same way as charged current coherent interactions, however due to the small number of events, there are seperate parameters for oxygen or carbon. Parameter (d) is a normalisation parameter for the number of NC events which lead to a photon emission that are non-resonant. There are no additional constraints seen

in these parameters between the two features, which is expected as such events do not feature prominently within the CC1Pi No Photon sample.



(a) NC Other at ND280

(b) NC Other SK, y-axis scale negligible at 10^{-12}

(c) CC Norm ν

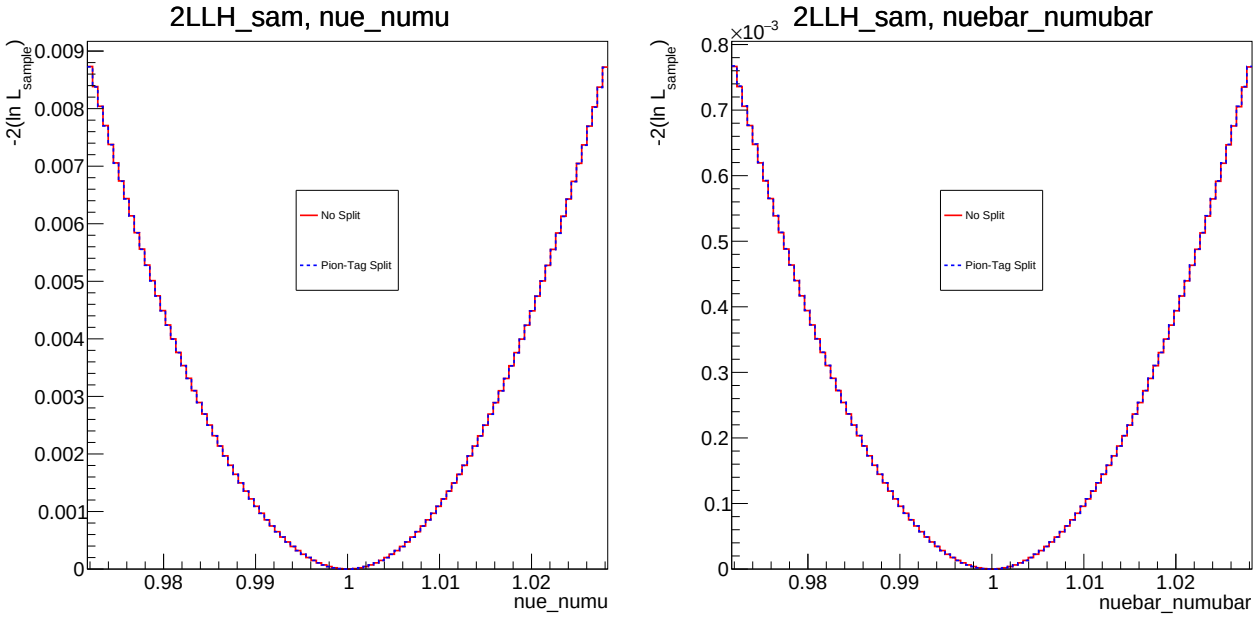
(d) CC Norm $\bar{\nu}$

FIGURE 7.19: NC and CC Parameters.

Parameter (a) and (b) in Figure 7.19 above displays the normalisation parameter for the NC Other interactions (DIS, Multi-Pi) at the near detector and far detector. There is a slight constraint seen at the NC Other parameter at ND280, which can be understood due to a

number of NC events within the CC1Pi No Photon sample. The NC Other parameter (b) is constrained at the far detector, as such only statistical fluctuations are seen in the LLH scan.

The final two parameters (c and d) in Figure 7.19 above are the CC Normalisation for the muon neutrino and muon antineutrino which adjust the lepton kinematics due to the Coulomb force of the nucleus. This potential affects electron flavoured neutrino and anti-neutrinos in a different way, hence there are separate parameters. There are no differences seen between the LLH scans of parameters (c) and (d).



(a) ν_e/ν_μ (b) $\bar{\nu}_e/\bar{\nu}_\mu$

FIGURE 7.20: CC Normalisation Parameters.

Both of the parameters above in Figure 7.20 are the CC Normalisation for the electron neutrino and electron antineutrino which adjust the lepton kinematics due to the Coulomb force of the nucleus. There are no differences seen between the LLH scans of parameters (c) and (d).

7.1.6 Parameter Constraints

Ultimately, many model parameters modelling CCRES, CCCoherent, and CCDIS interactions show improved constraints. Looking through the LLH scans, changes have been seen in the following parameters:

- **Rein-Seghal Delta Decay [C]**
- **RES $\mathbf{E_b}$ $\mathbf{^{12}C}$ ν_μ [C]**
- RES E_b ^{16}O ν_μ
- **FSI: QE [C], QEH [C], INEL [C], ABS [C], CX [C], CXH [C]**
- Nucleon FSI
- **CC Coherent Scattering for Carbon [C]**
- **CC Coherent Scattering for Oxygen [C]**
- Multi Pion Bodek-Yang Axial Parameter
- Multi Pion Multiplicity Shape Parameter
- CC Bodek-Yang Deep Inelastic Scattering
- NC Other at ND280

Out of 75 cross-section parameters, $\sim$20% of these parameters (16) show differences in constraints between the two features. Out of these 16, $\sim$94% (15) show better constraints with the CC1Pi split feature compared to the non-split feature. In order to verify the feature was implemented correctly, and to fully quantify these constraints in terms of parameter widths, an Asimov fit (see Section 7.2 below) is run.

The 1D distributions of the 10 parameters with the biggest differences in constraint between the two features will be shown (marked above with [C], in bold) in Section 7.3. These parameters are determined by looking at the 1 σ boundaries of the LLH scans and working out the percentage difference between the sample LLH value for both features at the boundary where the greatest difference is seen. A percentage difference of $>5.0\%$ is determined as significantly big and the parameter is marked with [C]. These marked parameters are from the CCRES, CC Coherent, and Final State Interaction parameters. These parameters are all pion-related as expected. The remaining 6 unmarked parameters are displayed in the Appendix (Chapter 9).

7.2 Asimov Fits Comparison

Asimov fits are where the MC is taken as the data, which in effect means that the fit is fitting the prediction to itself. This has many uses, the first of which is to test that the feature implementation has been done correctly and that the model responds appropriately. Secondly, any constraints in parameters can be seen by overlaying Asimov scans from both features, giving a view of improved sensitivity. The Asimov fit is an accurate description of the fit sensitivity, since it accounts for correlations between the parameters. In this thesis, the Asimov fit was run at the near detector only, with 2 million steps in total including a 200,000 step burn in.

The figures below show the prior distribution for each parameter (in shaded red) with a red dot indicating the central value, the posterior distribution for each parameter is represented by a line; with the solid black line representing the no-split feature in this thesis, and the blue dotted line representing the new split feature. The parameters are grouped by interaction type and all the distributions are overlaid, showing any constraints or shifts from the nominal value.

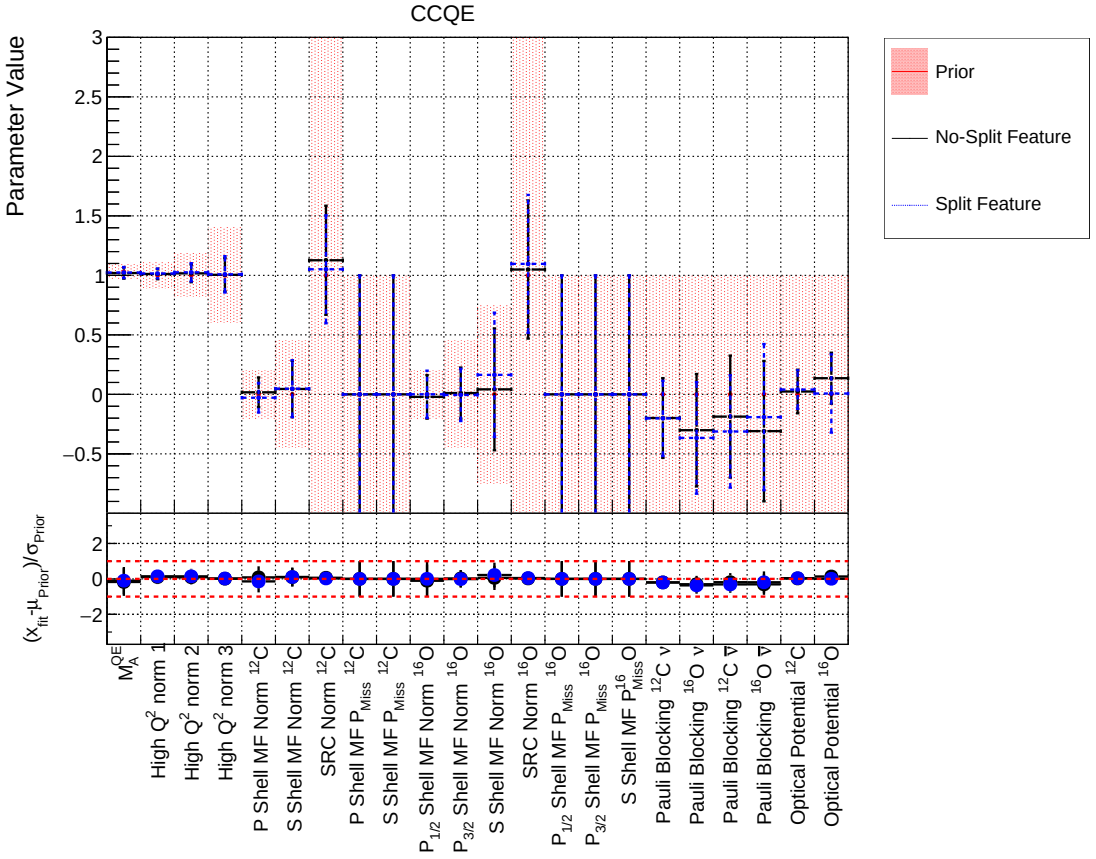


FIGURE 7.21: The two post-fit posteriors for the CCQE parameters overlaid (see Section 7.1.1).

Figure 7.21 shows the CCQE parameters post Asimov fit. Looking at the two features' posterior central values; most of the parameters are very similar in terms of their central value and width.

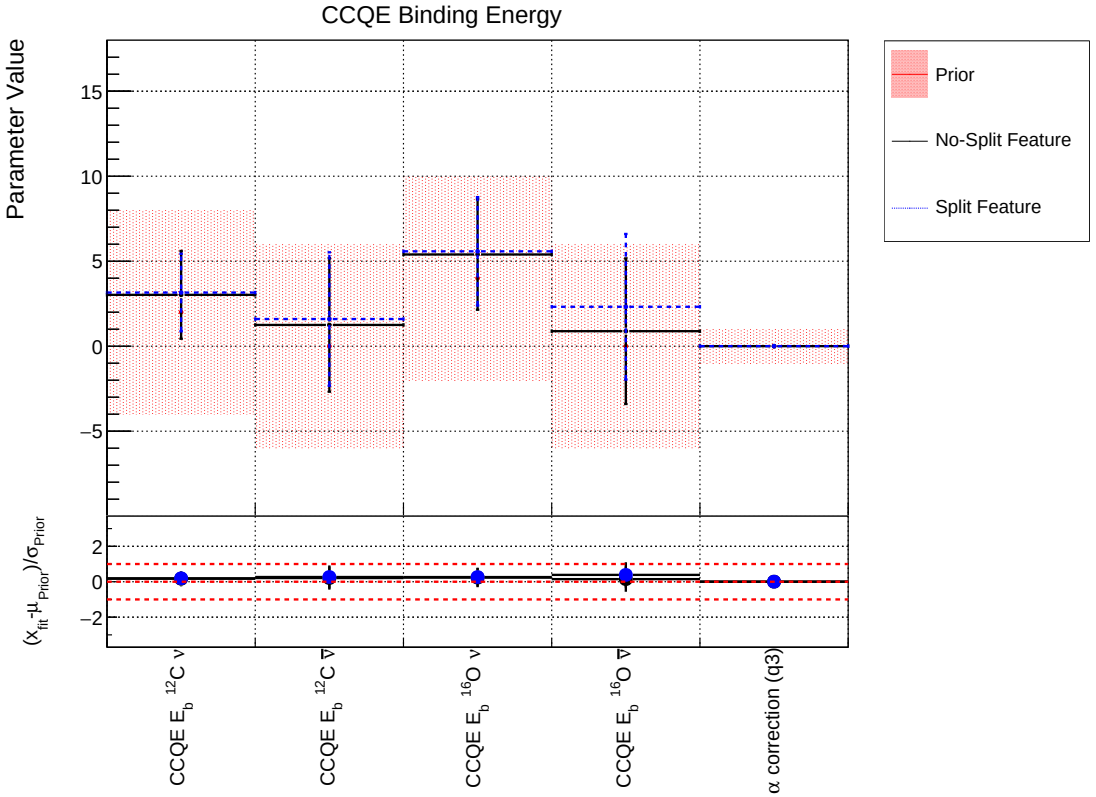


FIGURE 7.22: The two post-fit posteriors for the CCQE parameters overlaid (see Section 7.1.1).

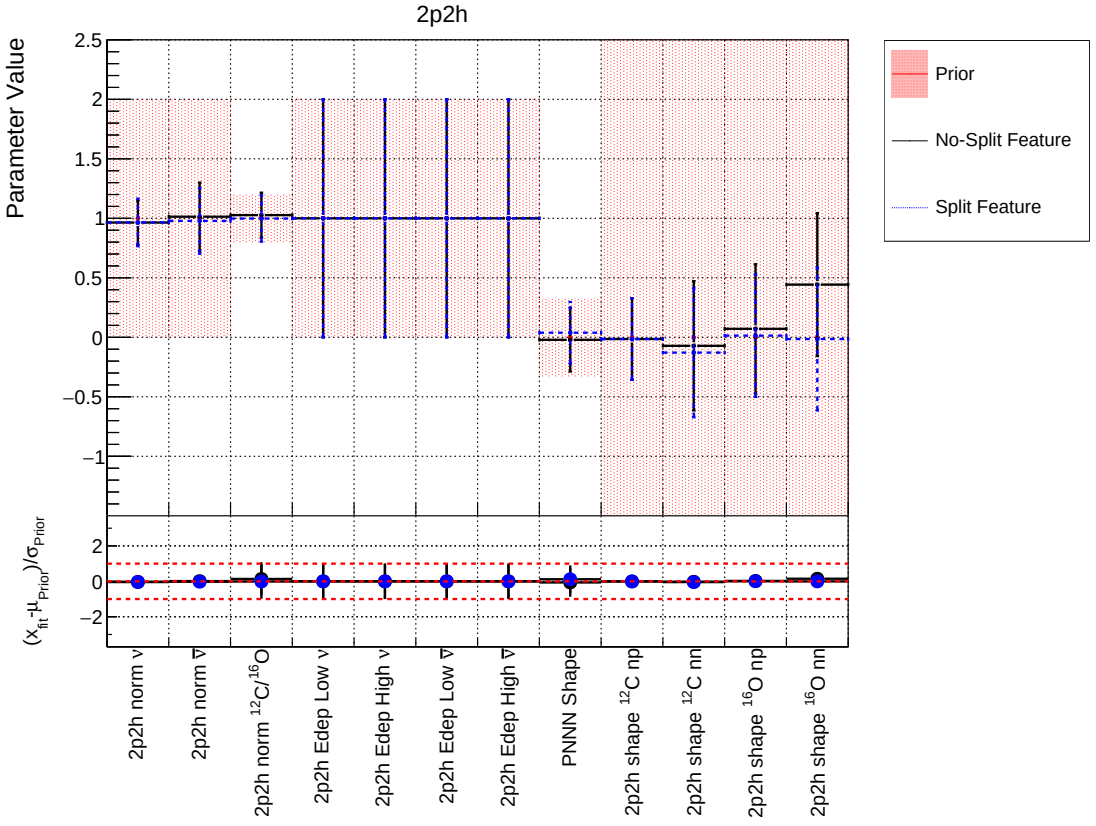


FIGURE 7.23: The two post-fit posteriors for the 2p2h parameters overlaid (see Section 7.1.2).

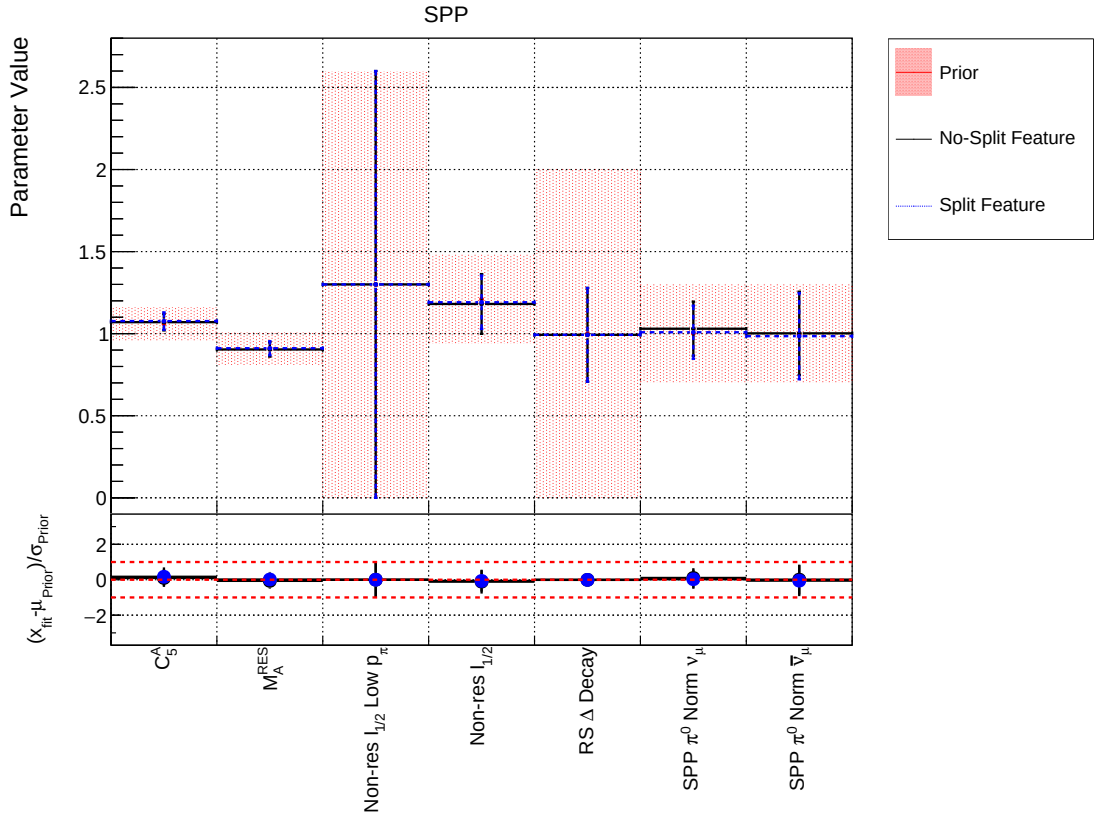


FIGURE 7.24: The two post-fit posteriors for the CCRES parameters overlaid (see Section 7.1.3).

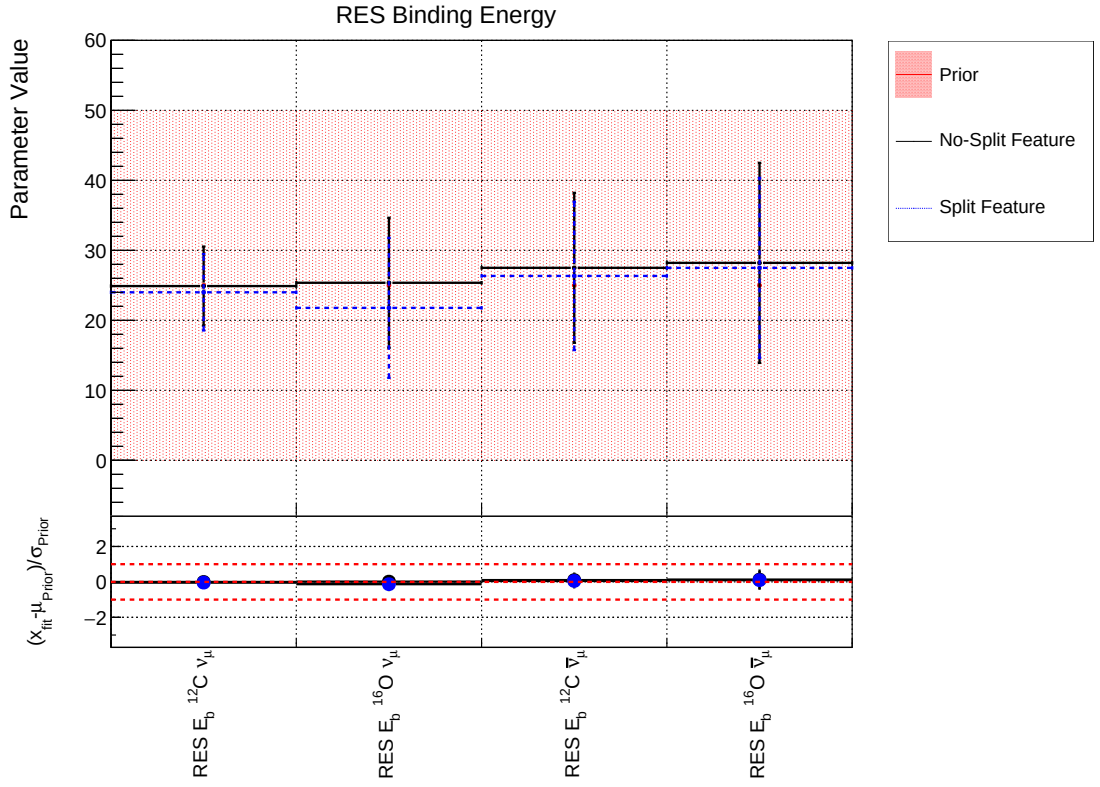


FIGURE 7.25: The two post-fit posteriors for the CCRES parameters overlaid (see Section 7.1.3).

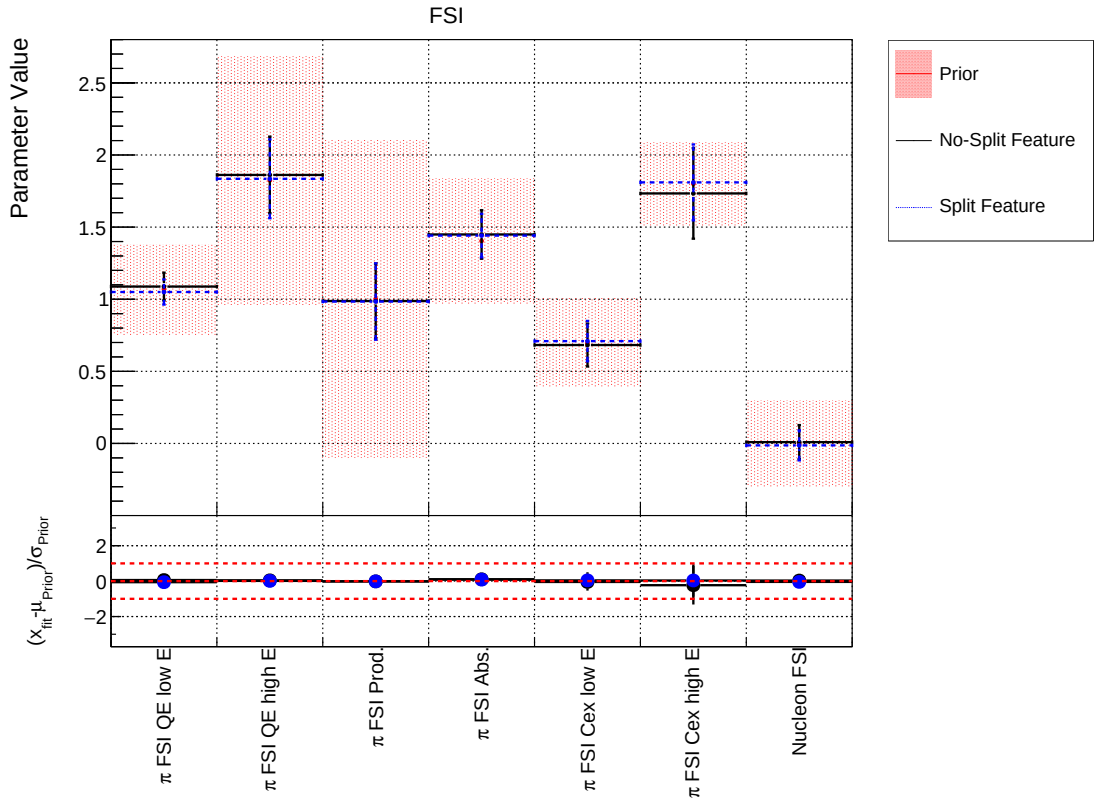


FIGURE 7.26: The two post-fit posteriors for the FSI parameters overlaid (see Section 7.1.4).

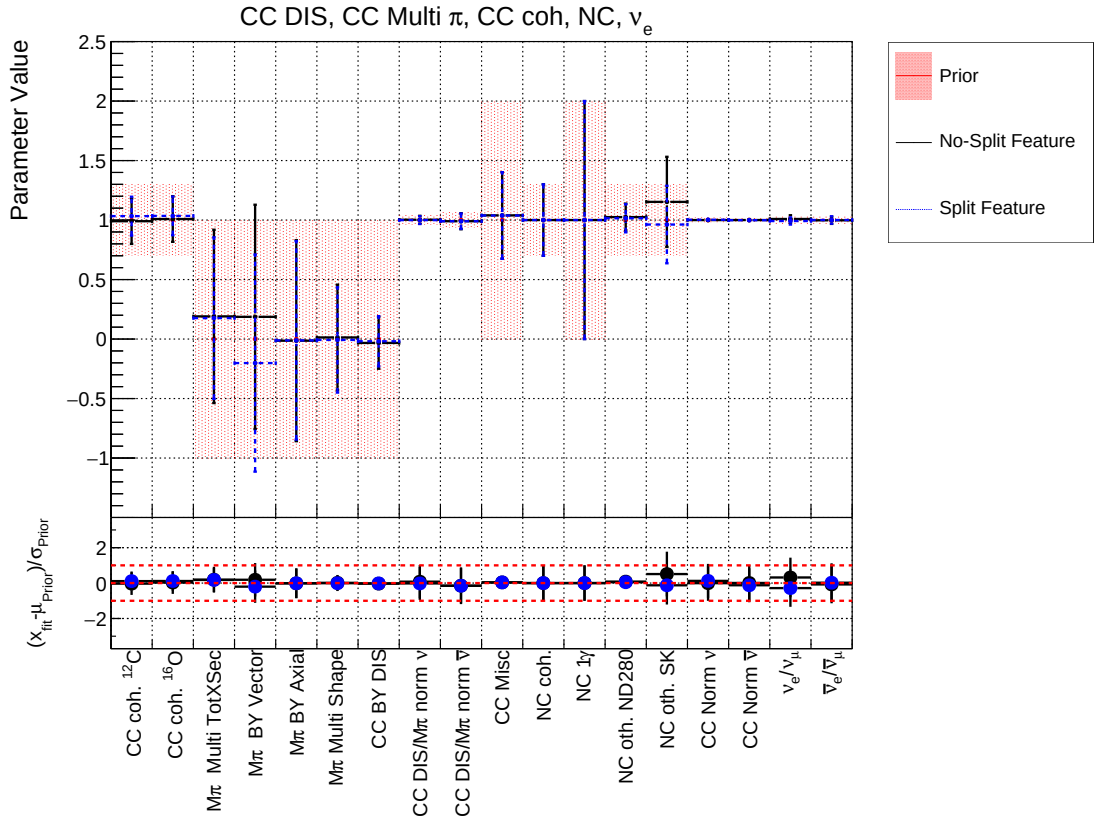


FIGURE 7.27: The two post-fit posteriors for the CCDIS parameters overlaid (see Section 7.1.5).

Figures 7.21 - 7.27 above illustrate that the parameters do not show any significant changes between the two features. Out of 75 parameters, ~ 2 parameters' central values differ by more than half a standard deviation - not exceeding 1 standard deviation of each other. This demonstrates that splitting the FHC CC1Pi sample by pion tag as a proxy to probe pion kinematics does not worsen parameter constraints, in fact it improves constraints in some cases. Ultimately, constraints are observed on the coherent scattering parameters and the FSI parameters by looking at the total width of the posterior distribution. The same constraints were also observed in the LLH scans prior (see Section 7.1).

It is also possible to determine the impact of this feature on the oscillation analysis. This can be inferred by looking at the far detector spectra which may yield more insights into what sort of sensitivity is being observed overall, bringing the far detector into the analysis (see Section 7.4).

7.3 1D Asimov Parameter Plots

To quantitatively understand the constraints seen in the LLH Scans for the above parameters in Section 7.1.6, the underlying distributions are observed. These distributions are presented below for the 10 aforementioned parameters. The plots shown below are taken from the same ND280 Asimov fit posterior mentioned in Section 7.2 previously. A total of 2 million steps are run, with a burn-in of 200,000. This posterior has been marginalised to a 1-dimensional distribution as described in Chapter 5, Section 5.4.2.

1-D Asimov fits clearly illustrate differences (constraints) in parameter width for certain cross-section parameters in comparison to a reference fit. Unlike in a data fit, the central value of the parameters are fixed in an Asimov fit, this makes it easier to observe constraints on different cross-section parameters. Based on prior knowledge and discussions; if the parameter widths of the two features are equivalent at 2 significant figures, then no significant constraint is seen. However, if there is a difference at the 2 significant figure level then this can be deemed as significant.

7.3.1 RS Delta Decay

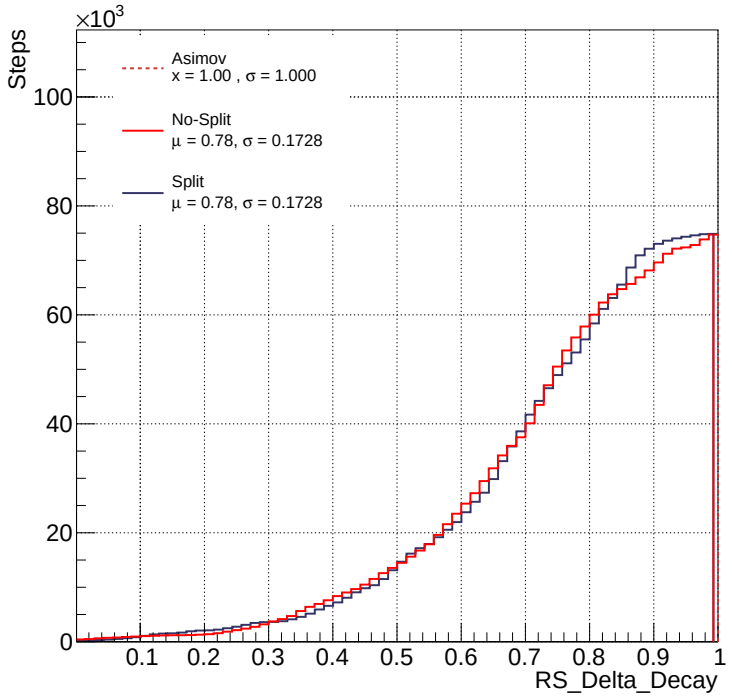


FIGURE 7.28: Rein-Sehgal Delta Decay parameter from the SPP parameters.

There is no additional constraint seen above in Figure 7.28 for this parameter, this seems to indicate a tension with the LLH scan of the parameter. However, the lack of constraint may be due to correlation terms (now included in the Asimov fit) that wash out any constraint due to the errors on the parameter itself (seen in the LLH scan in Section 7.1).

7.3.2 RES $E_b\,^{12}C\,\nu_\mu$

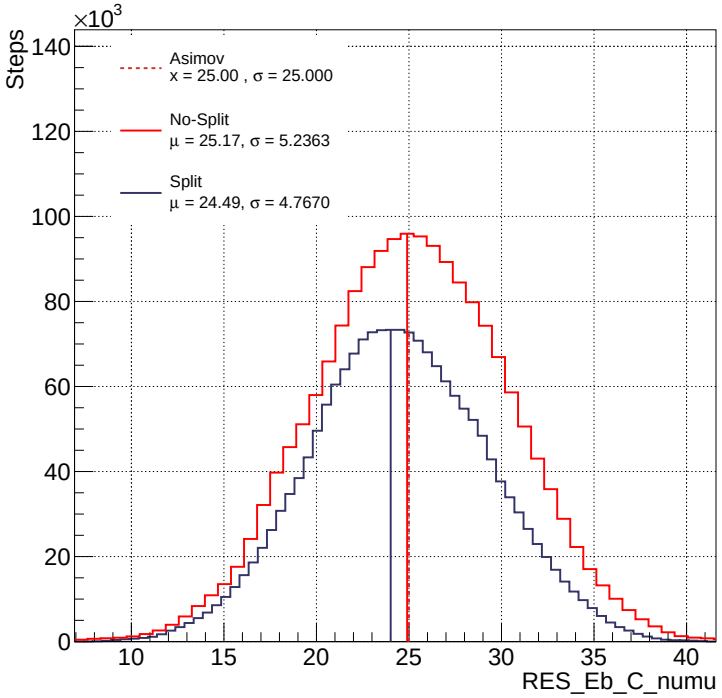


FIGURE 7.29: Resonant Binding Energy for ν_μ on Carbon from the SPP parameters.

A significant constraint is seen in Figure 7.29 for this parameter as mentioned in Section 7.1. A difference in bin width can be seen in the plot above between the two parameters. This is due to RES $E_b\,^{12}C\,\nu_\mu$ parameter gaining a tighter constraint (a narrower range) in the split case. As the number of bins are fixed, the resulting bin width is finer. Splitting by pion tag gives a tighter range and a better constraint for this CCRES parameter.

7.3.3 FSI QE

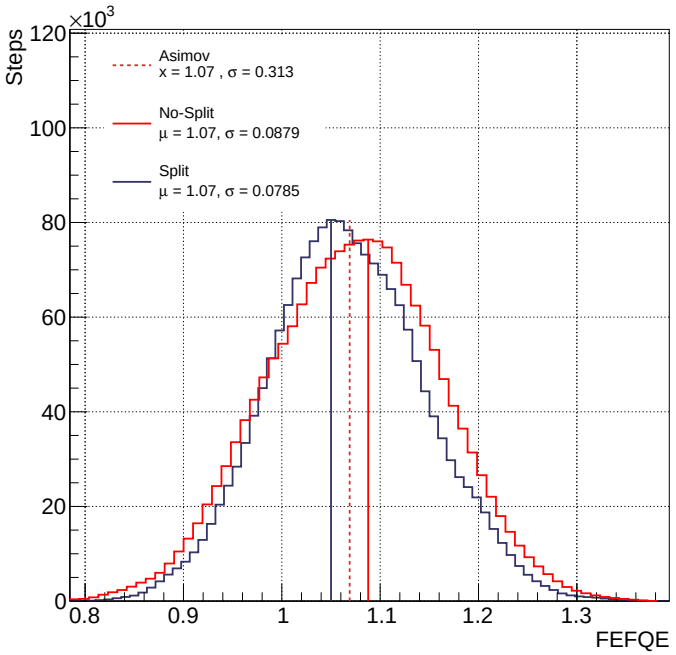


FIGURE 7.30: FSI Quasi-elastic scattering from the FSI parameters.

A significant constraint is seen in Figure 7.30 for this FSI parameter as mentioned in Section 7.1.

7.3.4 FSI QEH

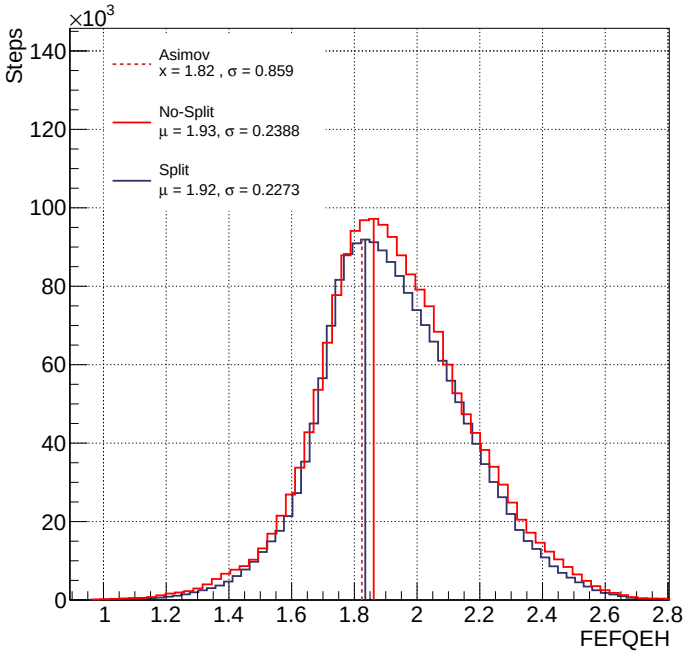


FIGURE 7.31: FSI Quasi-elastic scattering (at high energy) from the FSI parameters.

A significant constraint is seen above in Figure 7.31 for this FSI parameter as mentioned in Section 7.1.

7.3.5 FSI INEL

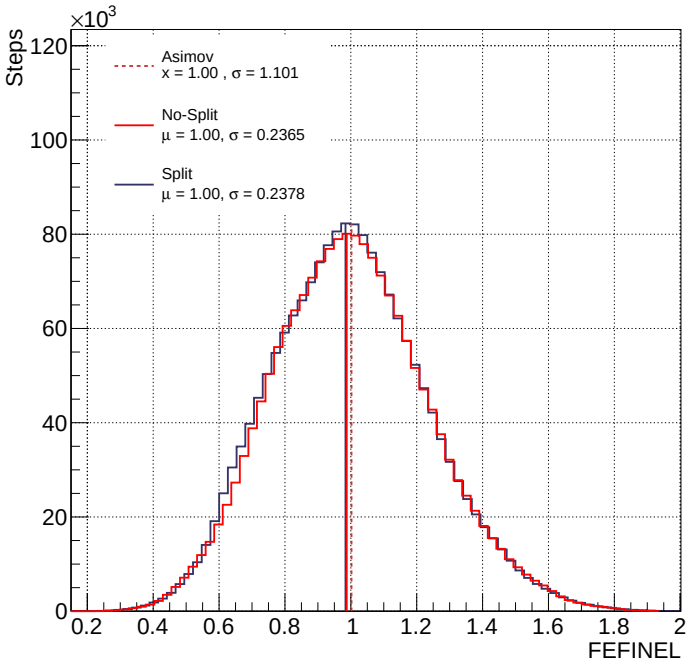


FIGURE 7.32: FSI inelastic scattering from the FSI parameters.

There is no significant constraint seen in Figure 7.32 for this FSI parameter for inelastic scattering. The lack of constraint may be due to correlation terms (now included in the Asimov fit) that wash out any constraint due to the errors on the parameter itself (seen in the LLH scan in Section 7.1)

7.3.6 FSI ABS

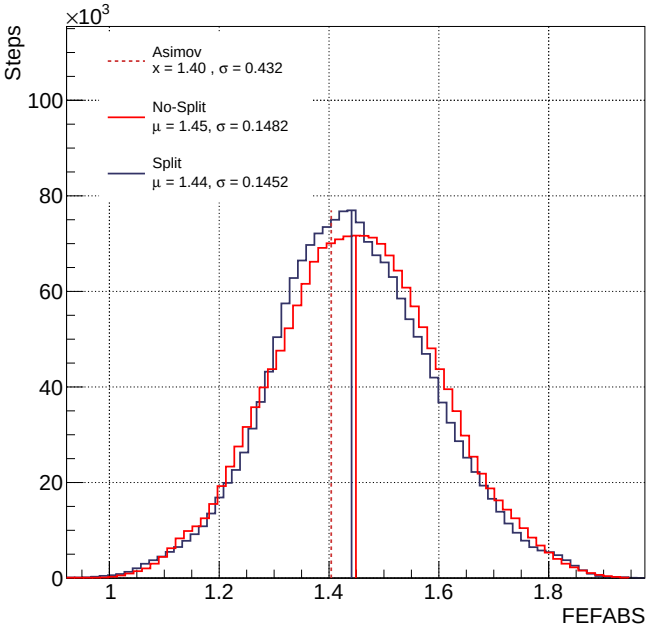


FIGURE 7.33: FSI absorption from the FSI parameters.

No significant constraint is seen in Figure 7.33 for this FSI parameter as mentioned in Section 7.1. Again, this lack of constraint may be due to correlation terms in the Asimov fit that wash out any constraint due to the errors on the parameter itself (seen in the LLH scan in Section 7.1).

7.3.7 FSI CX

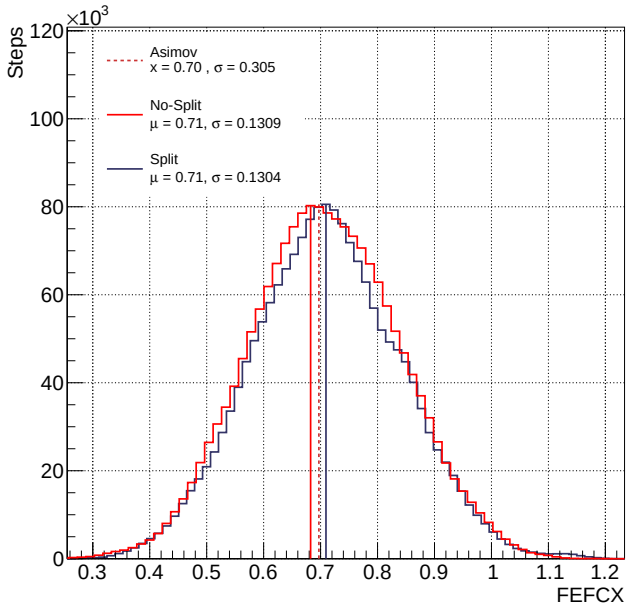


FIGURE 7.34: FSI charge exchange from the FSI parameters.

No significant constraint is seen in Figure 7.34 for this FSI parameter as mentioned in Section 7.1. This lack of constraint may be due to correlation terms (now included in the Asimov fit) that wash out any constraint due to the errors on the parameter itself (seen in the LLH scan in Section 7.1)

7.3.8 FSI CXH

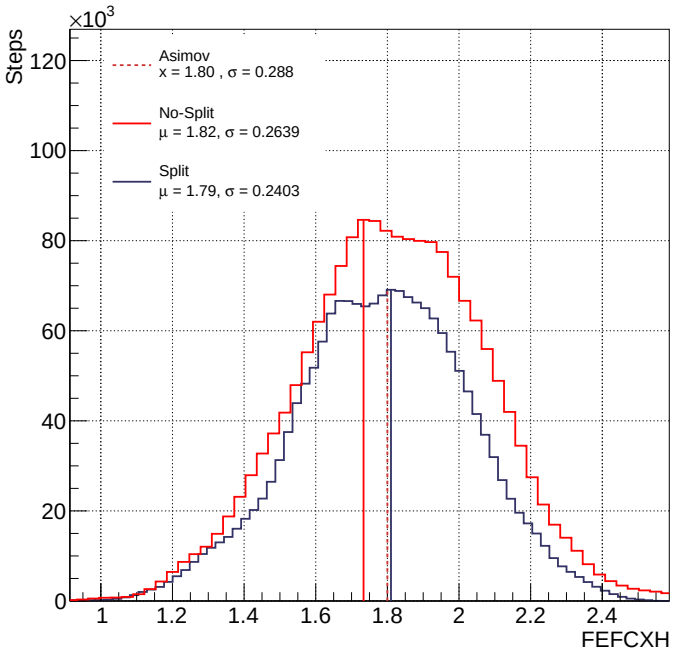


FIGURE 7.35: FSI charge exchange (at high energy) from the FSI parameters.

A significant constraint is seen in Figure 7.35 for this FSI parameter as mentioned in Section 7.1.

7.3.9 CC Coherent Scattering for Carbon

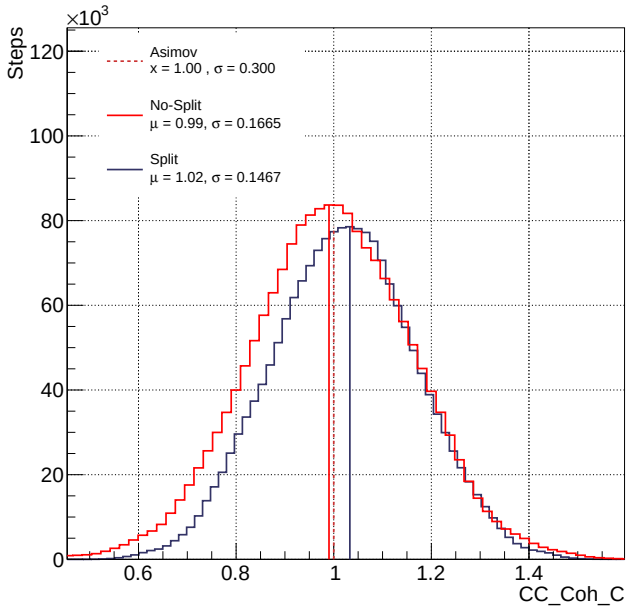


FIGURE 7.36: Coherent Scattering on Carbon from the CC Coherent parameters.

A significant constraint is seen in Figure 7.36 for this parameter as mentioned in Section 7.1.

7.3.10 CC Coherent Scattering for Oxygen

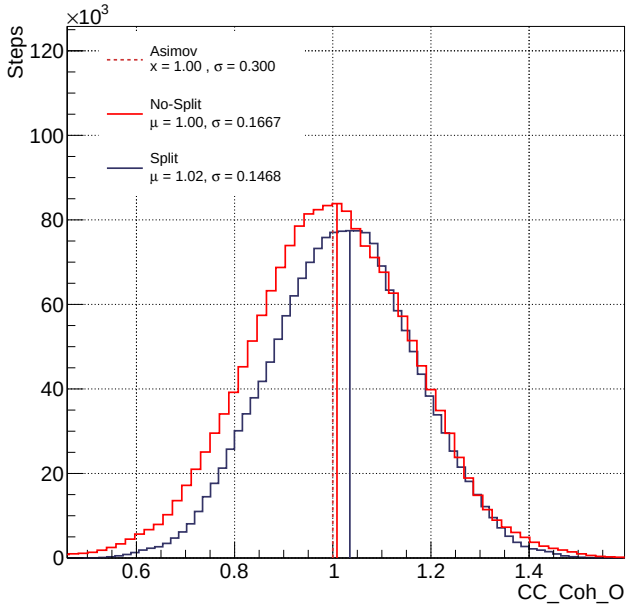


FIGURE 7.37: Coherent Scattering on Oxygen from the CC Coherent parameters.

A significant constraint is seen in Figure 7.37 for this parameter as mentioned in Section 7.1.

7.4 Super-K Event Rates (SK Posterior Predictives)

The far detector measures the number of events of a given sample per bin of reconstructed neutrino energy. A posterior predictive is run for the far detector which produces a best fit for the energy spectrum. The method is as follows:

1) A number of points are randomly selected from the posterior of the Asimov fit. In this thesis 3000 points are used. This is because finding the mode of the ~ 700 dimensional posterior is too computationally intensive in requirement.

2) Each point in parameter space carries a value for all the fit parameters; near detector, far detector, cross section, flux, and oscillation parameters. Points which are closer to highest posterior density will be selected with a greater likelihood.

3) For every bin in reconstructed neutrino energy, the 3000 points drawn from the posterior distribution are used to generate a total event number. This collection of 3000 numbers, each which represents a number of events, has a Gaussian fitted to the distribution peak. The central value of this distribution is used as the predicted value for the bin.

Given that the ND (near detector) parameter fits for both the features do not vary significantly, by comparing FD (far detector) spectra for two features any differences that are due to parameters of interest can be seen. Splitting the FHC CC1Pi sample at the near detector affects the far detector spectra for the 1 ring electron, 1 decay electron (1Re 1de) sample, as this involves single pion production (see Chapter 6, Section 6.6). The 1Re 1de sample is a ν_e appearance sample which is sensitive to δ_{CP}. Therefore, it is important to be able to model this sample accurately.

In this section, Asimov ND fits predicted to the far detector will be looked at, followed by Data ND fits propagated to the far detector for different tags.

7.4.1 Asimov: No-Split vs Split Feature

In this spectra, the near detector Asimov fit is taken and extrapolated to form a far detector prediction. As mentioned in Chapter 6, Section 6.6, the far detector sample examined is the 1Re 1de (1 ring electron, 1 decay electron) sample to see changes corresponding to splitting the CC1Pi sample at the near detector. The no-split feature is compared to the split feature in Figure 9.6 below.

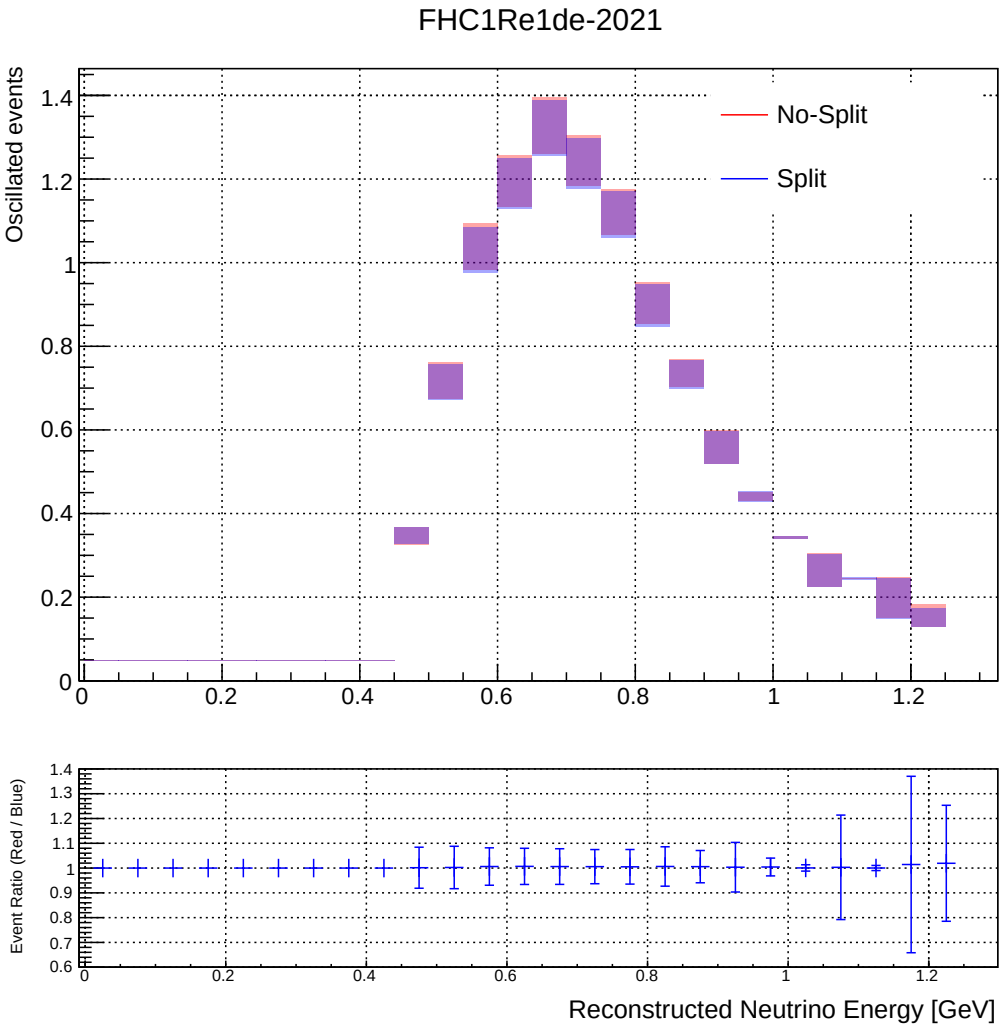


FIGURE 7.38: No-Split vs Split Asimov posterior predictive.

The far detector spectra agrees quite closely in width between the two features. This agreement allows us to see that running a full joint data-fit (ND + FD) would not show significant changes in far detector parameters constraints. Ultimately, what is shown is that the splitting lepton kinematics by pion tag doesn't change the far detector spectrum that much, therefore we expect no significant changes in neutrino oscillation sensitivities.

However, examining one or more of the pion tags individually begins to open up the possibility of modelling pion momenta better within certain regions of phase space.

7.4.2 Asimov: No-Split Feature vs Split [TPC Tag] Feature

The no-split feature is compared to the split feature with only the TPC tag events included in Figure 7.39 below.

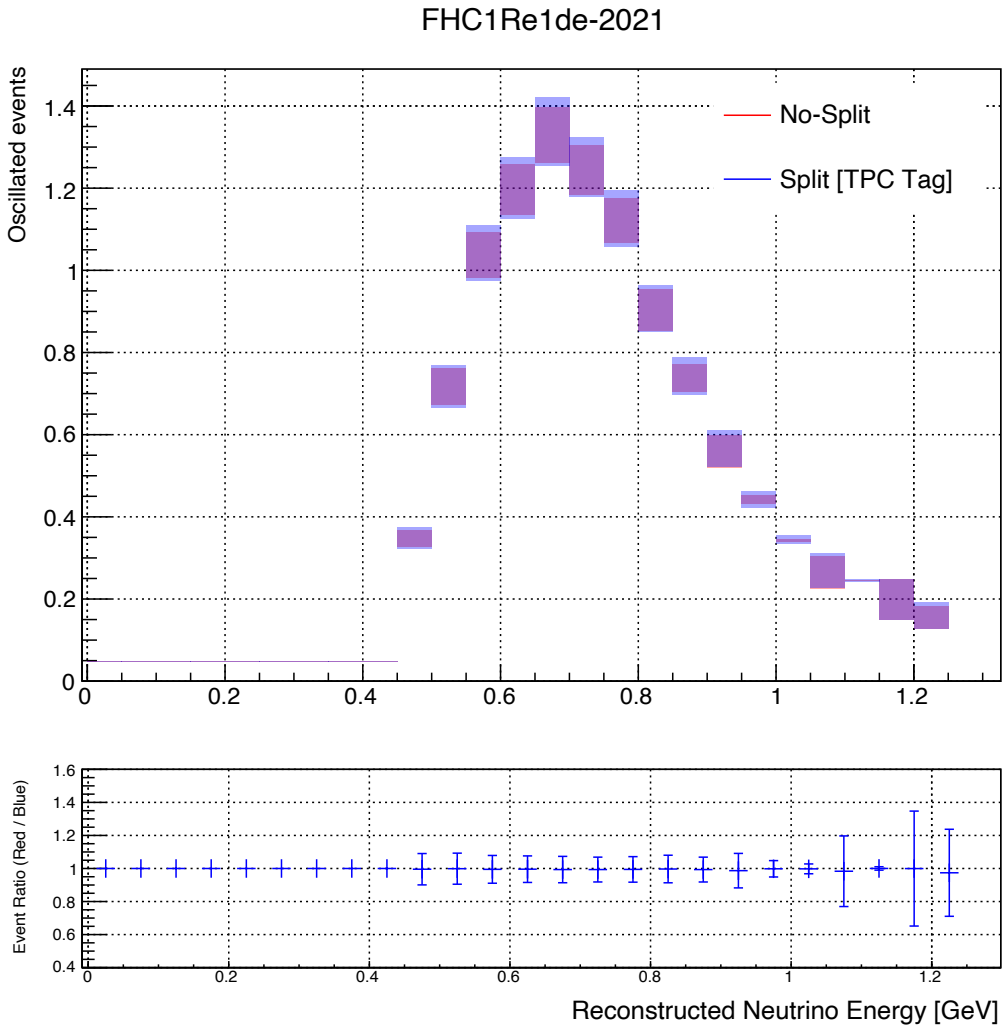


FIGURE 7.39: No-Split vs Split [TPC Tag] Asimov posterior predictive.

As seen in the spectra plot in Figure 7.39 above, the split feature is now compared to the split feature with only the TPC pion tagged leptons remaining which constitute 55% of the CC1Pi sample. There are minimal differences seen between both spectra, with the split TPC tag sample widths pulling up and having a greater spread as expected.

7.4.3 Asimov: No-Split Feature vs Split [Michel Tag] Feature

The no-split feature is compared to the split feature with only the Michel tag events included in Figure 7.40 below.

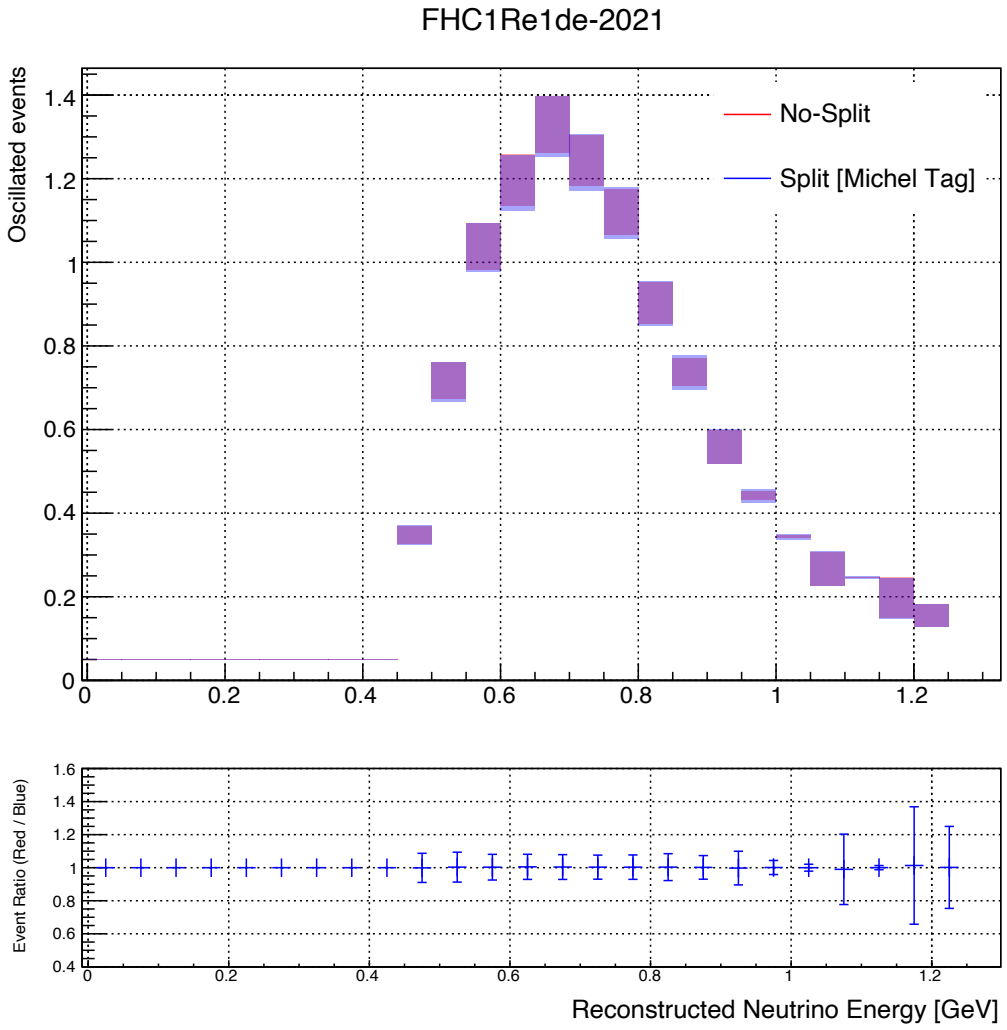


FIGURE 7.40: No-Split vs Split [Michel Tag] Asimov posterior predictive.

There is a smaller difference in Figure 7.40 between the far detector spectra when only including the Michel pion tag. For Michel tagged pion leptons, only 33% of events end up in the CC1Pi sample. However, in the above far detector spectra for the 1 ring electron, 1 decay electron sample, the Michel sample widths pull downwards, but changes in width between the two features are not strongly significant as expected.

7.4.4 Asimov: No-Split Feature vs Split [FGD Tag] Feature

The no-split feature is compared to the split feature with only the FGD tag in Figure 7.41 below.

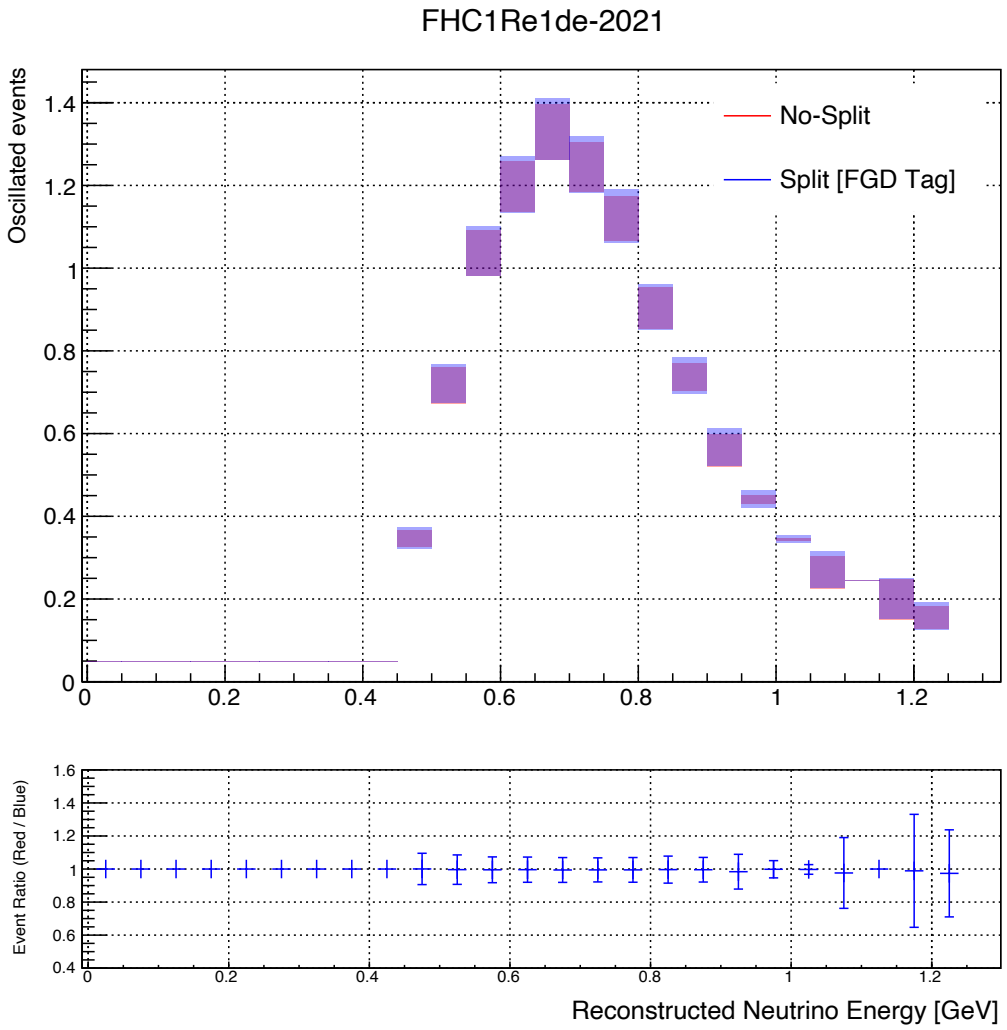


FIGURE 7.41: No-Split vs Split [FGD Tag] Asimov posterior predictive.

For FGD tagged pion leptons, only 12% of events end up in the CC1Pi sample. However, in the above far detector spectra for the 1 ring electron, 1 decay electron sample, the FGD sample widths pull upwards, but changes in width between the two features are not strongly significant as expected.

7.4.5 Data: No-Split vs Split Feature

In this spectra, the near detector data fit is taken and extrapolated to form a far detector prediction. As mentioned previously, the far detector sample examined is the 1Re 1de sample to see changes corresponding to splitting the CC1Pi sample at the near detector. The no-split feature is compared to the split feature in Figure 7.42 below.

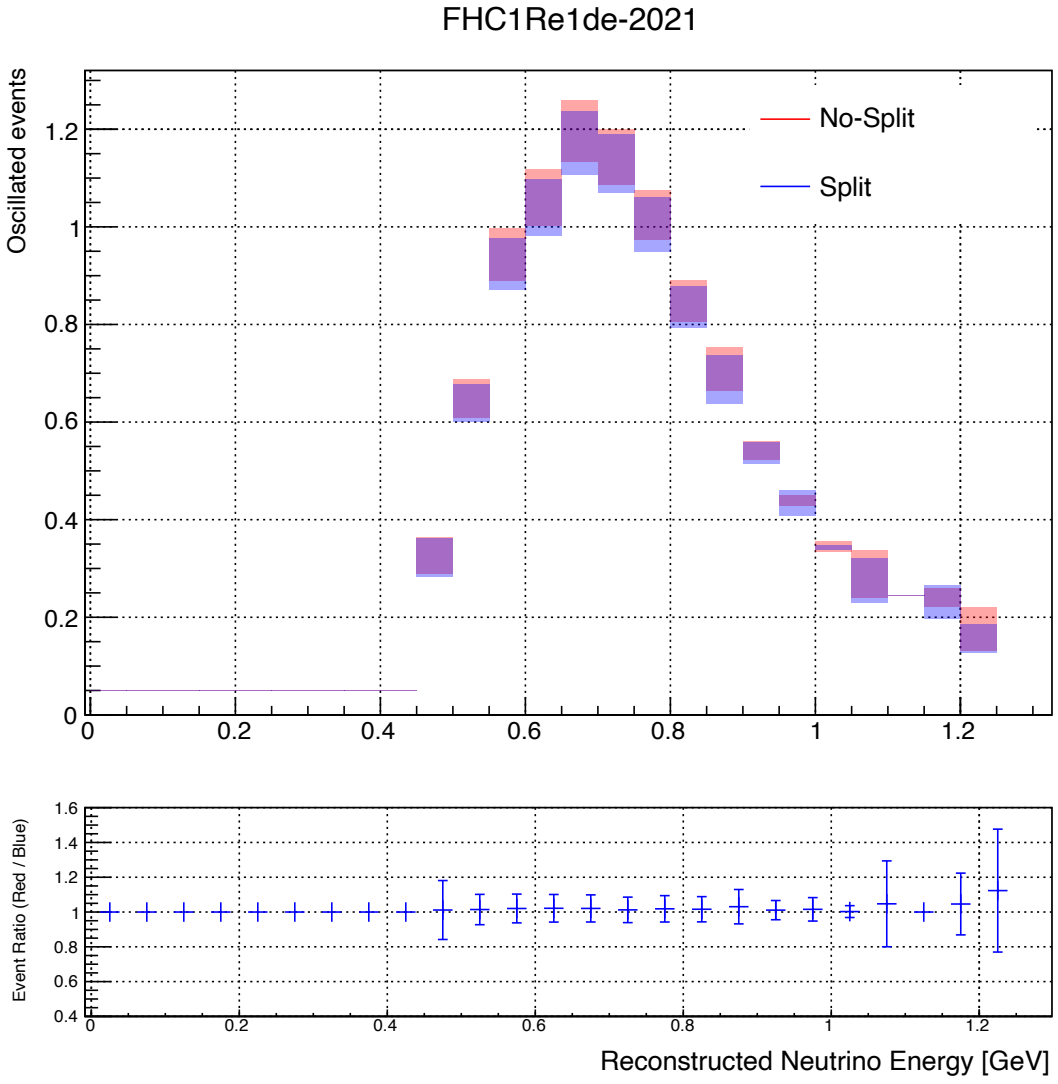


FIGURE 7.42: No-Split vs Split data posterior predictive.

As will be mentioned in Figures 7.43 - 7.45 forthcoming, the central values shift depending on the pion tag. However, when putting all the tags together, these changes mostly wash out giving the final spectra seen above in Figure 7.42.

7.4.6 Data: No-Split Feature vs Split [TPC Tag] Feature

The no-split feature is compared to the split feature with only the TPC tag in Figure 7.43 below.

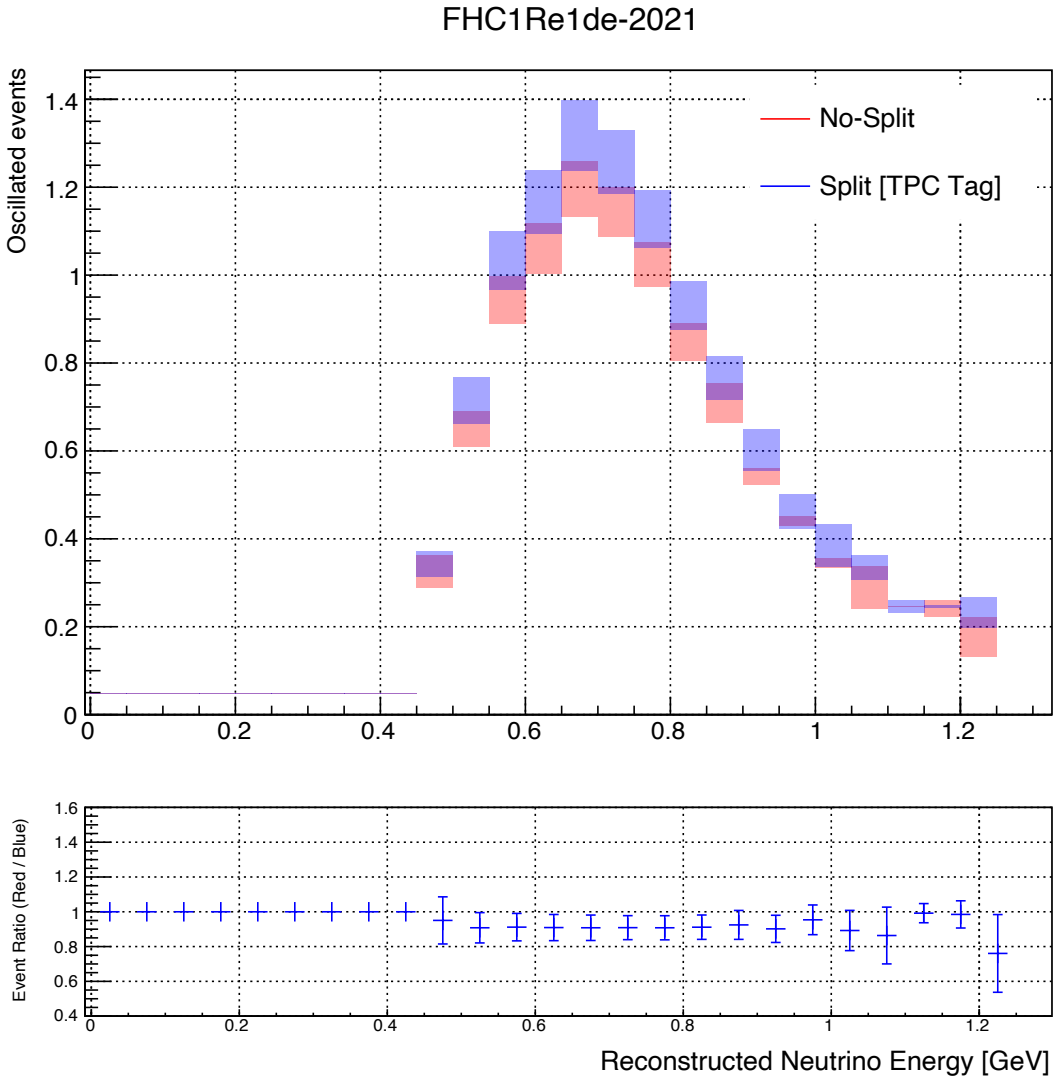


FIGURE 7.43: No-Split vs Split [TPC Tag] data posterior predictive.

For TPC tagged pion leptons, 55% of events end up in the CC1Pi sample. In the above far detector spectra for the 1 ring electron, 1 decay electron sample, the TPC sample pulls upwards.

7.4.7 Data: No-Split Feature vs Split [Michel Tag] Feature

The no-split feature is compared to the split feature with only the Michel tag in Figure 7.44 below.

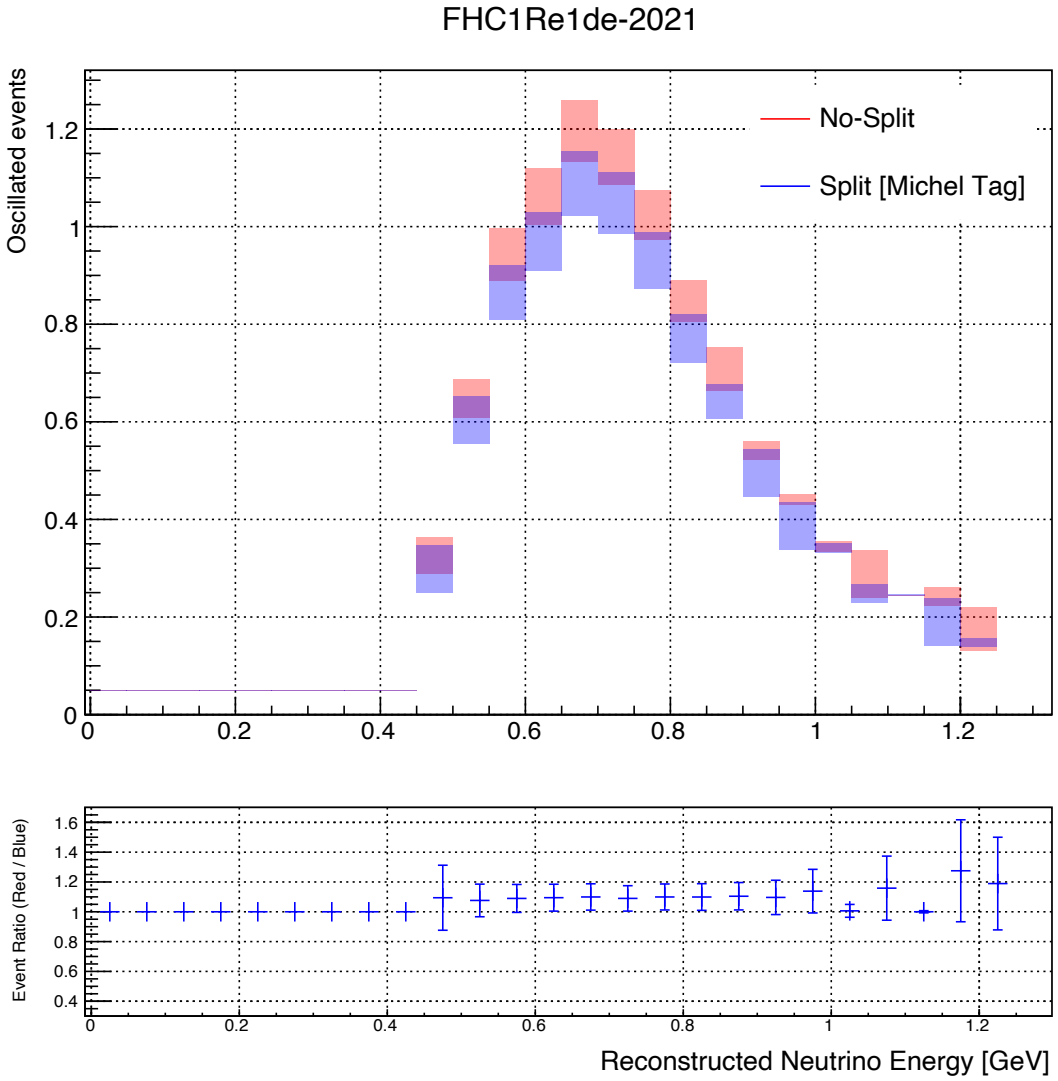


FIGURE 7.44: No-Split vs Split [Michel Tag] data posterior predictive.

For Michel tagged pion leptons, 33% of events end up in the CC1Pi sample. In the above far detector spectra for the 1 ring electron, 1 decay electron sample, the Michel sample is unique as it pulls downwards.

7.4.8 Data: No-Split Feature vs Split [FGD Tag] Feature

The no-split feature is compared to the split feature with only the FGD tag in Figure 7.45 below.

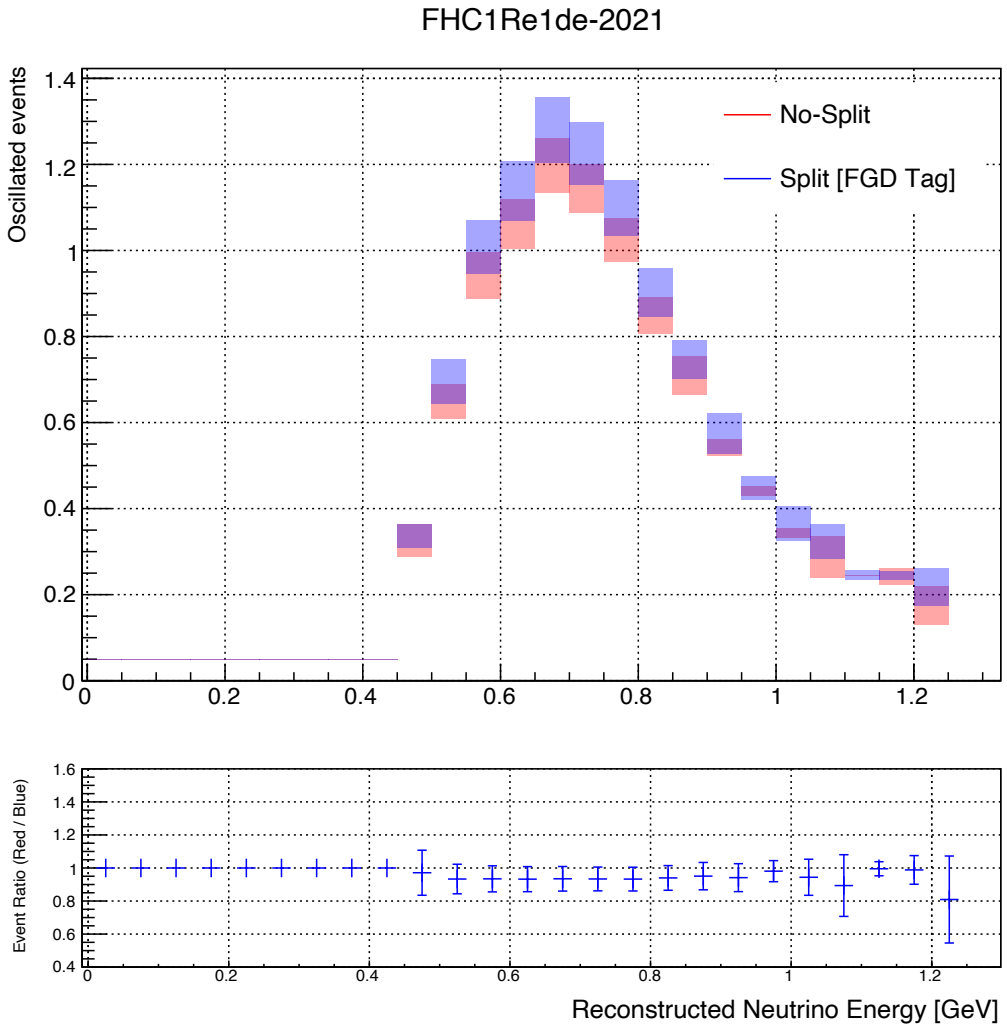


FIGURE 7.45: No-Split vs Split [FGD Tag] data posterior predictive.

For FGD tagged pion leptons, 12% of events end up in the CC1Pi sample. In the above far detector spectra for the 1 ring electron, 1 decay electron sample, the FGD sample pulls upwards.

The important conclusion from this study is that each individual tag pulls the spectra in its own way. The TPC and FGD tag pull the prediction up. Whereas, the Michel tag is unique in pulling the prediction down. It is interesting to note that when combining all the pion tagged events together the resultant sample is actually the original CC1Pi split sample and, as mentioned previously, in this sample, pulls from individual tags generally wash out as seen in the No-Split vs Split spectra (Figure 7.42).

Ultimately, using individual pion tagged events allows for more a pronounced far detector spectra. This is a clear indication that pions play a significant role in the 1Re 1de sample prediction. This underscores the importance of implementing pion kinematics proper at the ND280, along with cross-section model considerations for single pion production (SPP) in order to leverage this selection fully when fitting. It is hoped that this work has served as an intermediate step in bridging the gap to incorporate pion kinematics within the T2K oscillation analysis.

7.5 Joint Fit

A joint fit was performed to understand constraints in oscillation parameters between the two features. The near detector fits seem to indicate a small constraint from the split feature, whereas the far detector spectra (no-split vs split) between the two features seem to indicate strong agreement. In order to understand improvements to the oscillation parameters themselves, a joint fit is needed to truly verify any improvements.

For the split feature joint fit, 41.2 million steps were run at the near and far detector, with a burn-in of 60,000. These steps are sufficient enough to give good contours for the $1\,\sigma$ and $2\,\sigma$ credible regions. Drawing conclusive (smooth) contours at $2\,\sigma$ credible intervals requires around 20 - 30 million steps. For $3\,\sigma$, this requirement can exceed 100 million steps. The split feature is compared to the $\sim$250 million step no-split feature joint fit (used in the T2K 2022 oscillation analysis).

7.5.1 Joint Fit Comparison for No-Split vs Split Feature

Seven contours in total are presented below for the 3 oscillation parameter (2D) axes at T2K: $\sin^2\theta_{13}$ – δ_{CP}, $\sin^2\theta_{23}$ – δ_{CP}, and $\sin^2\theta_{23}$ – Δm^2_{32}. Additional graphs for these 2D contours are plotted assuming: a normal hierarchy, inverted hierarchy, or a combination (normal + inverted). These contours are without reactor constraint.

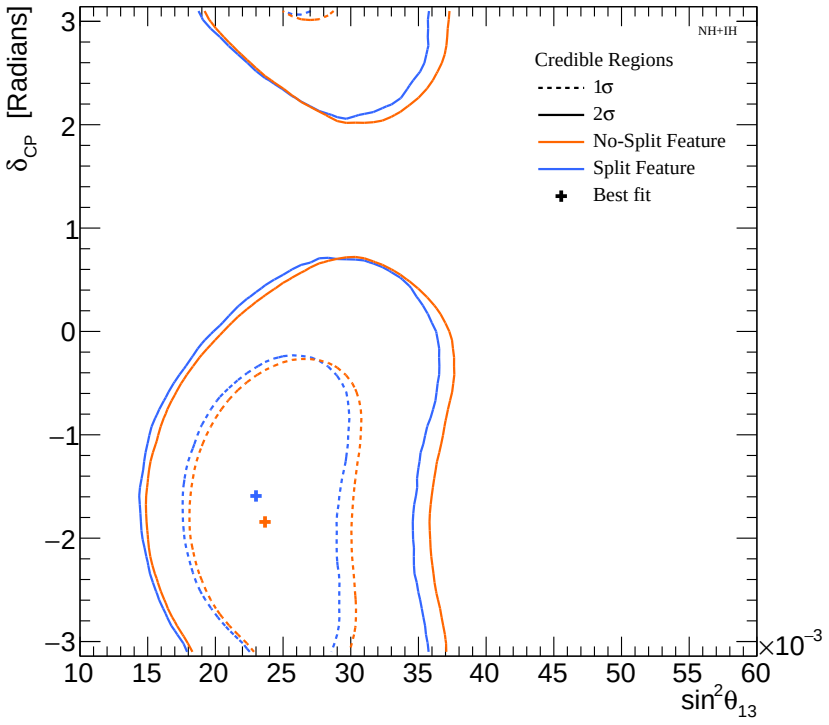


FIGURE 7.46: Contour plot in the $\sin^2\theta_{13}$–δ_{CP} space, normal + inverted hierarchy.

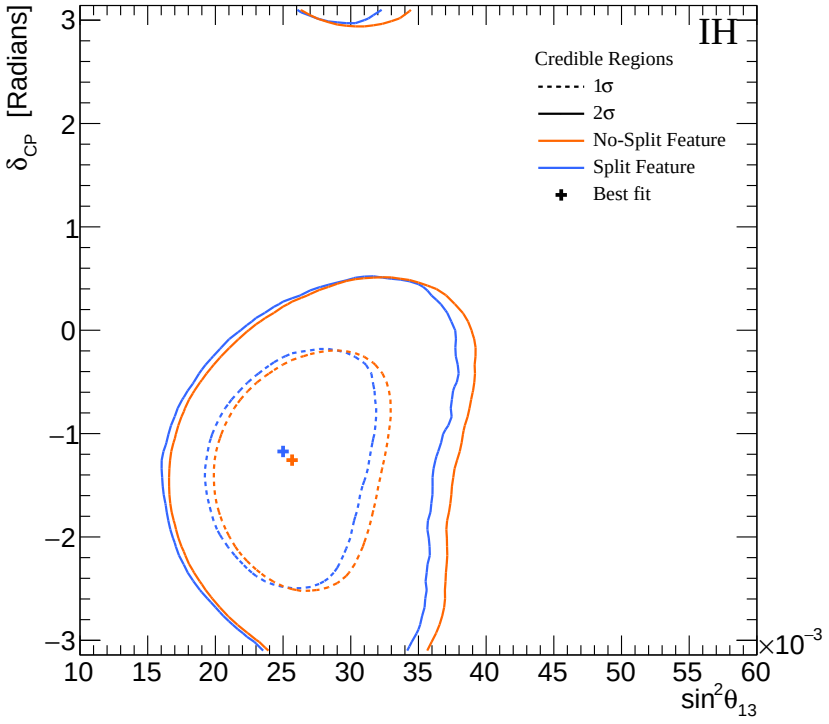


FIGURE 7.47: Contour plot in the $\sin^2\theta_{13}$–δ_{CP} space, inverted hierarchy.

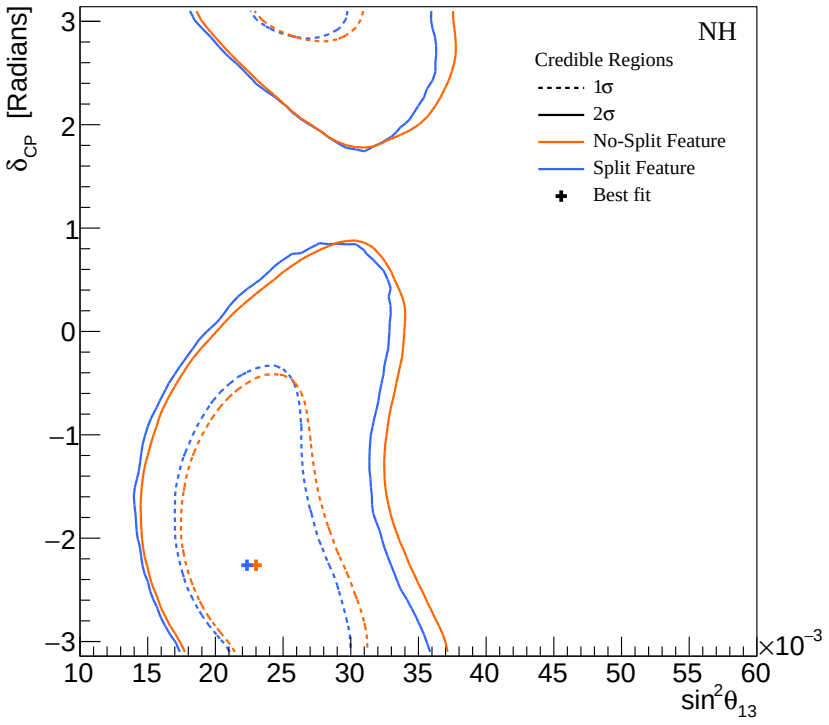


FIGURE 7.48: Contour plot in the $\sin^2\theta_{13}$–δ_{CP} space, normal hierarchy.

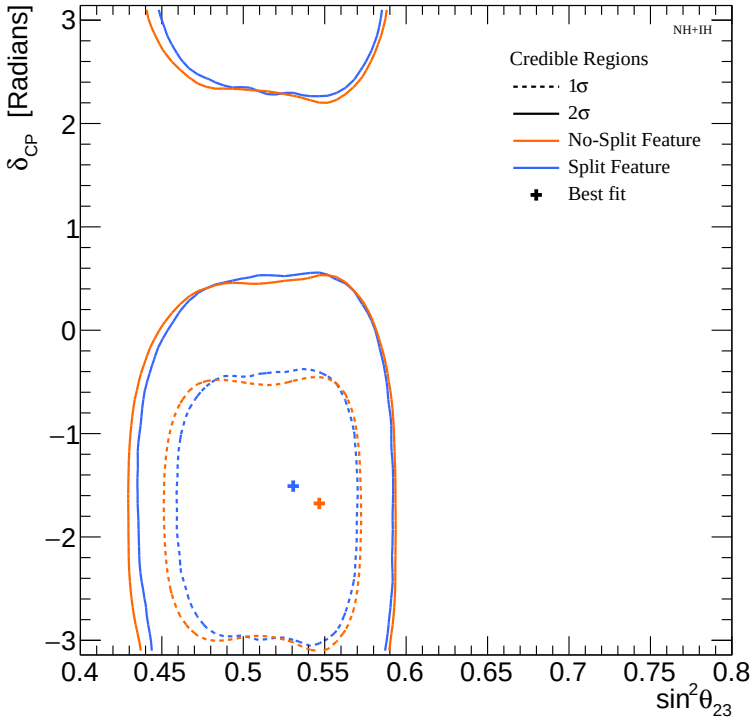


FIGURE 7.49: Contour plot in the $\sin^2\theta_{23}$–δ_{CP} space, normal + inverted hierarchy.

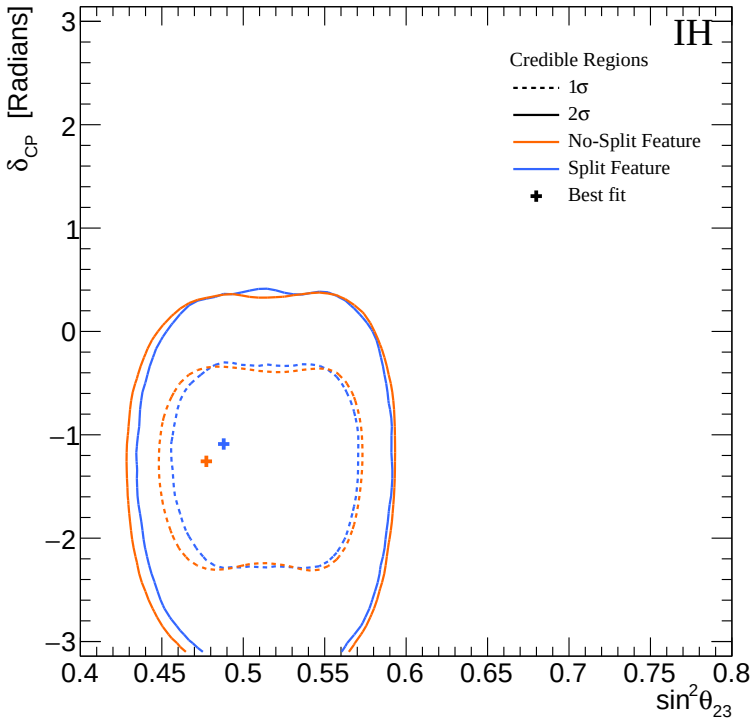


FIGURE 7.50: Contour plot in the $\sin^2\theta_{23}$–δ_{CP} space, inverted hierarchy.

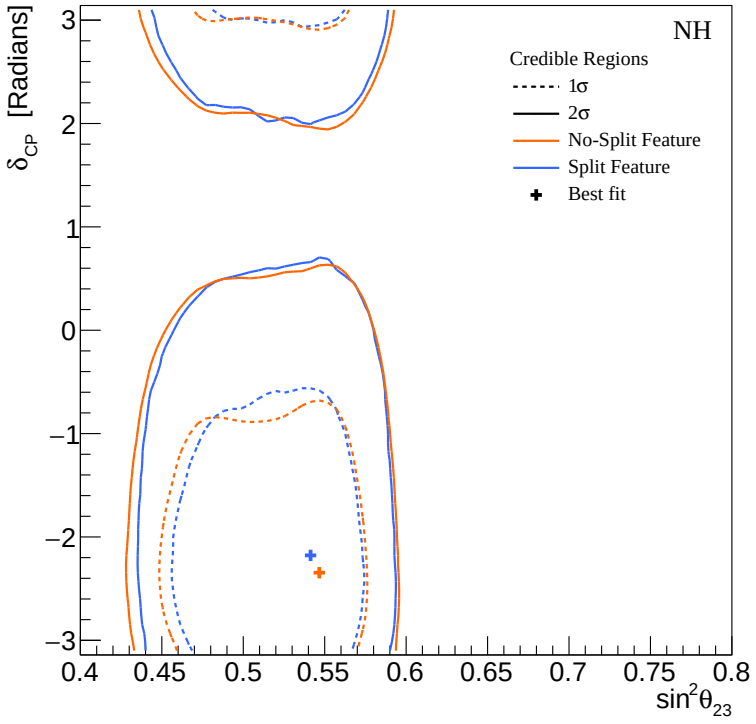


FIGURE 7.51: Contour plot in the $\sin^2\theta_{23}$–δ_{CP} space, normal hierarchy.

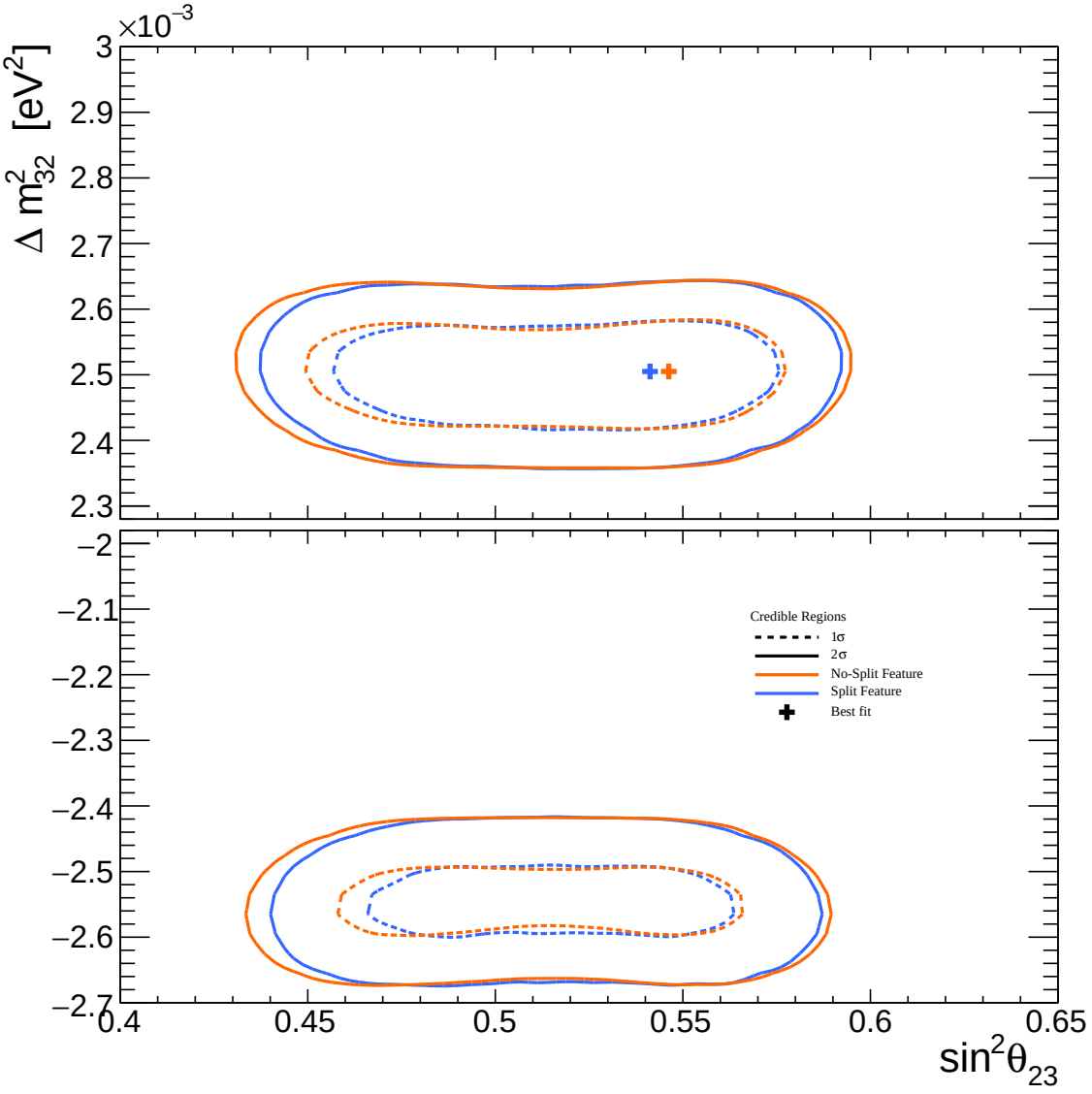


FIGURE 7.52: Contour plot in the $\sin^2\theta_{23} - \Delta m^2_{32}$ space. Normal (top) and inverted hierarchy (bottom) are both shown and compared.

Ultimately, a joint fit was run with a low number of steps in order to verify any changes in oscillation parameter sensitivity. Looking at Figures 7.46 - 7.52 above, the credible intervals for both features overlap at points in the 2D parameter space, as would be expected given the implementation and previous results. However, it is interesting that a small improvement in sensitivity for the new split-feature can be seen from the $2\,\sigma$ credible intervals in Figures 7.46 - 7.52 above. In order to quantify this fully, a higher number of steps need to be run.

Chapter 8

Conclusion and Summary

In this thesis, a new feature of the neutrino oscillation analysis was explored by splitting the FHC CC1Pi sample at the near detector by pion tag, making original contributions to the T2K experiment. The underlying pion distribution at the near detector was understood along with the distribution of pion tagged leptons. The splitting was implemented in the FHC CC1Pi sample using a novel stacked histogram method.

Detector systematic parameters were studied and developed, LLH scans were performed in order to investigate any improved sensitivities in cross-section parameters. Asimov and data fits were run at the near detector in order to quantitatively understand any cross-section parameter constraint shown in LLH Scans. An Asimov and data fit were performed, the result of this was extrapolated to the far detector, forming a Super-K posterior predictive in order to probe changes in far detector spectra and therefore any possible extra sensitivities in oscillation parameters. Additionally, a joint fit with the near and far detector was performed with a considerable number of steps in order to verify any changes in sensitivity.

The work done in this thesis is a necessary step in order to make an impact to the oscillation parameters in the future. Splitting the FHC CC1Pi sample by pion tag as a proxy to probe pion kinematics does not significantly worsen or shift parameter constraints; it actually improves the constraints in some cases. At ND280, better constraints were seen in the coherent scattering parameters, alongside many of the final state interaction (FSI) parameters. At the far detector, Super-K, the posterior predictive spectrum shows that splitting by pion tag does not significantly worsen the oscillation analysis. Therefore a net improvement is seen through implementing the novel pion tag split feature within lepton kinematics at the near detector.

The outcome of this work is to understand whether specific samples help us to better understand single pion production. Using individual pion tagged events makes a pronounced

difference in the far detector spectra. This is a clear indication that pions play a significant role in the 1Re 1de sample, highlighting the importance of implementing pion kinematics at the ND280. This thesis forms an important intermediate step which is necessary in being able to incorporate pion kinematics within the T2K oscillation analysis in the future.

Chapter 9

Appendix

9.1 1D Asimov parameter plots

The other 6 parameters mentioned in Chapter 7, Section 7.1.6 are listed below:

- RES E_b ^{16}O ν_μ
- Nucleon FSI
- Multi Pion Bodek-Yang Axial Parameter
- Multi Pion Multiplicity Shape Parameter
- CC Bodek-Yang Deep Inelastic Scattering
- NC Other at ND280

The parameters above all showed a constraint from the LLH Scans. However, as mentioned, the Asimov fit includes the correlation terms and therefore the 1D scans give a more accurate picture of the true constraints. The Asimov fits of these parameters are shown below:

9.1.1 RES E_b ^{16}O ν_μ

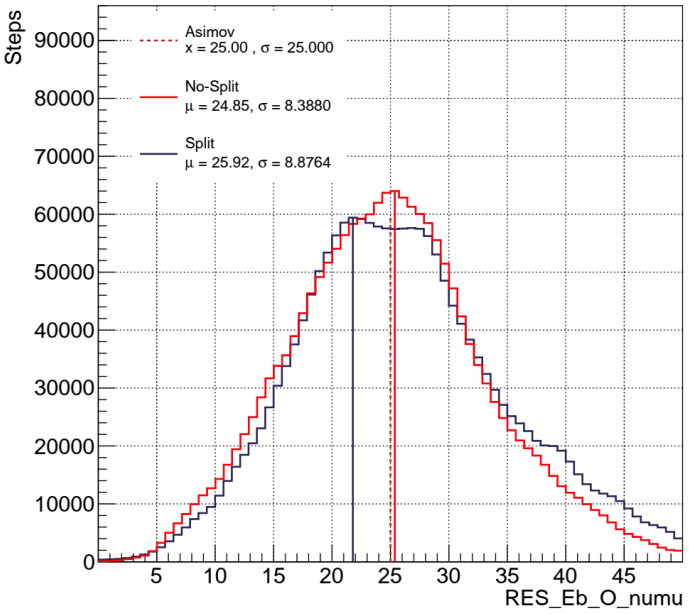


FIGURE 9.1: RES E_b ^{16}O ν_μ Parameter.

9.1.2 Nucleon FSI

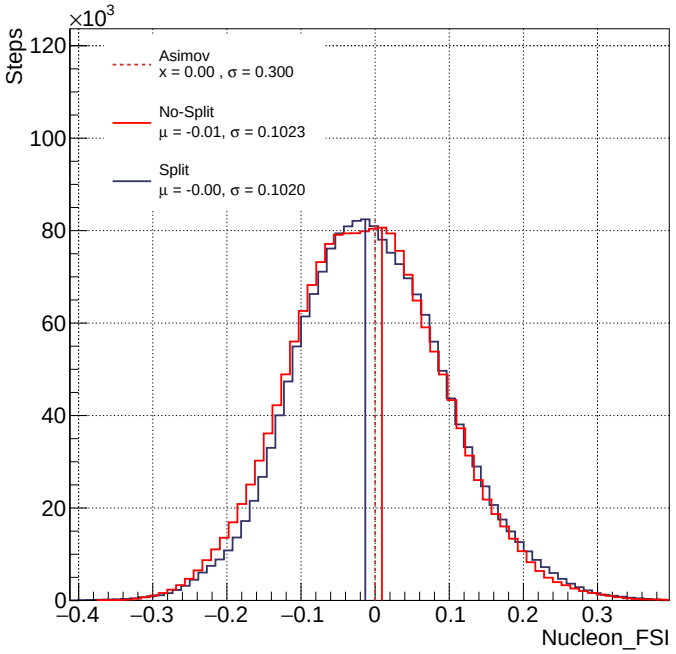


FIGURE 9.2: Nucleon FSI.

9.1.3 Multi Pion Bodek-Yang Axial Parameter

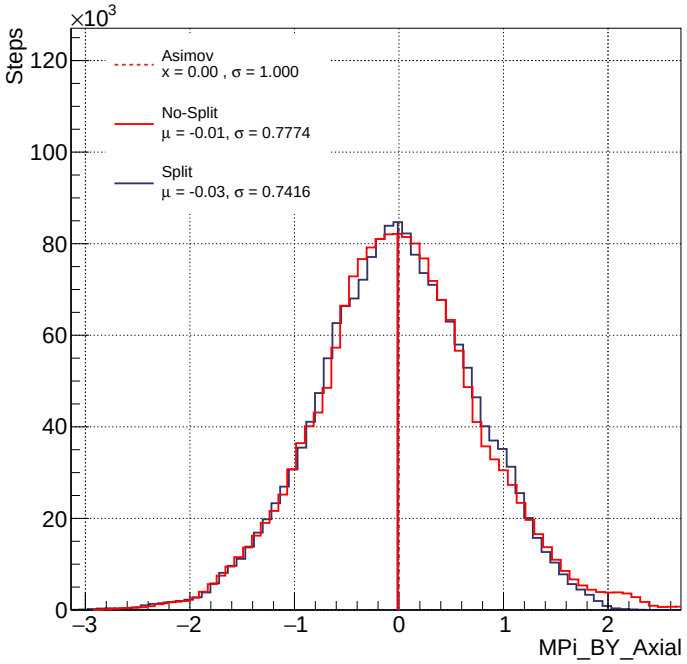


FIGURE 9.3: Multi Pion Bodek-Yang Axial Parameter.

9.1.4 Multi Pion Multiplicity Shape Parameter

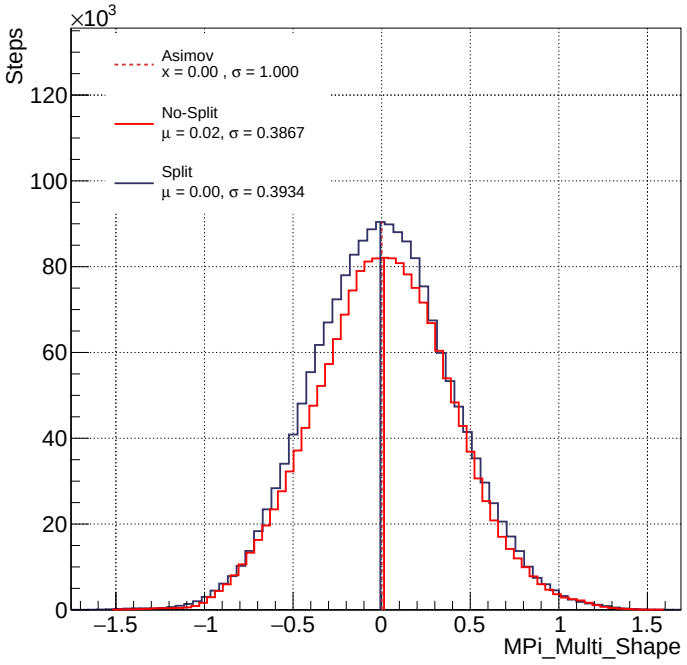


FIGURE 9.4: Multi Pion Multiplicity Shape Parameter.

9.1.5 CC Bodek-Yang Deep Inelastic Scattering

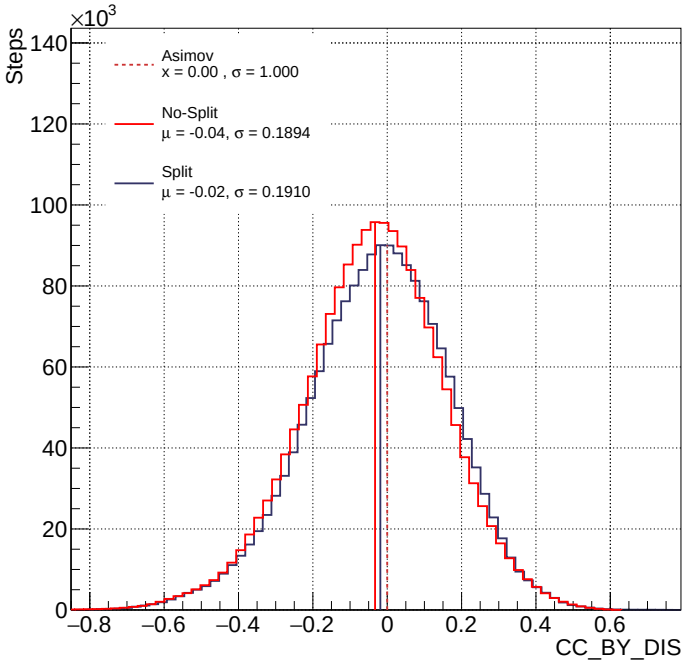


FIGURE 9.5: CC Bodek-Yang Deep Inelastic Scattering

9.1.6 NC Other at ND280

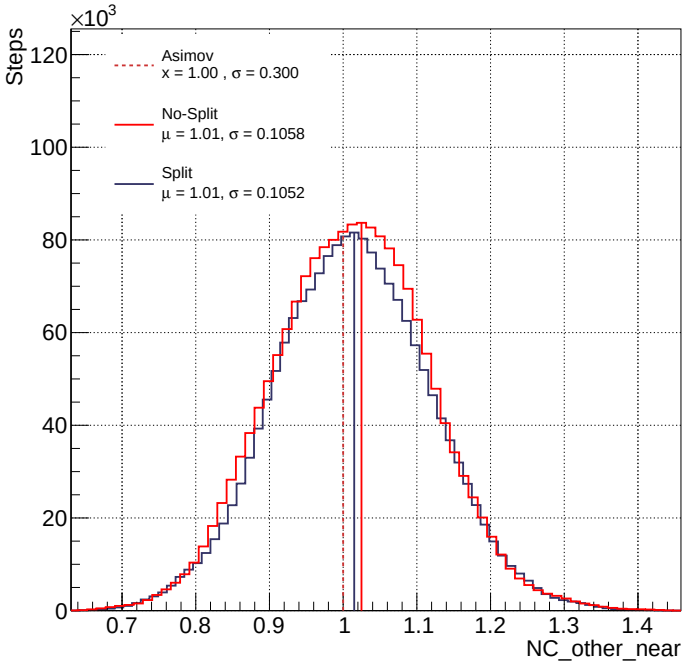


FIGURE 9.6: NC Other at ND280.

9.2 Near Detector (ND) Prior and Posterior Predictive

A table showing the uncertainties for different samples under each systematic group at ND280 is generated using prior and posterior predictives. The prior predictive values are found by using the detector covariance matrix to form sets of parameter variations and re-weighting the MC to those variations.

The posterior predictive values are used from the post-fit covariance matrix after an Asimov fit, this is used to form parameter variations. This is done for all systematic parameters by group (Flux, Cross section, and Detector). The table for the Split feature is shown below:

TABLE 9.1: Split feature ND prior and posterior prediction event rate uncertainty apportioned by ND sample and systematic type.

Sample	Flux		Cross Section		ND280		Total	
	Prior	Post	Prior	Post	Prior	Post	Prior	Post
FGD1 FHC CC0$\pi - 0p - 0\gamma$	5.2	2.9	11.9	3.0	1.8	1.22	12.80	0.7
FGD1 FHC CC0π-Np-0 γ	5.5	2.9	12.3	3.4	3.6	2.22	13.46	0.9
FGD1 FHC CC1 $\pi - 0\gamma$	5.3	2.9	9.3	3.0	3.1	1.46	11.16	0.9
FGD1 FHC CC-Other-0 γ	5.6	3.0	8.4	2.9	5.3	2.3	10.78	1.6
FGD1 FHC CC-Photon	5.6	2.9	8.8	3.0	3.0	1.8	10.45	0.8
FGD2 FHC CC0 $\pi - 0$p $- 0\gamma$	5.2	2.9	11.1	3.0	2.1	1.2	12.22	0.7
FGD2 FHC CC0 π-Np-0 γ	5.6	2.9	11.5	3.4	4.0	2.4	13.18	1.0
FGD2 FHC CC1 $\pi - 0\gamma$	5.3	2.9	9.2	3.0	3.7	1.6	10.93	0.9
FGD2 FHC CC-Other-0 γ	5.6	3.0	8.3	2.9	6.5	2.5	11.52	1.9
FGD2 FHC CC-Photon	5.5	2.9	8.7	3.0	2.6	1.6	10.31	0.8
FGD1 RHC CC0 π	5.0	3.2	11.3	3.2	2.0	1.2	12.13	0.9
FGD1 RHC CC1 π	4.7	3.1	10.5	3.5	4.3	2.5	11.79	2.0
FGD1 RHC CC-Other	4.5	2.9	9.2	3.1	3.6	1.9	10.41	1.5
FGD2 RHC CC0 π	5.0	3.2	10.0	3.2	2.1	1.2	10.93	0.9
FGD2 RHC CC1 π	4.7	3.0	10.3	3.3	4.0	2.1	11.28	1.7
FGD2 RHC CC-Other	4.6	2.9	9.7	3.2	3.0	1.7	10.62	1.5
FGD1 RHC BKG CC0 π	5.8	3.0	9.9	2.9	2.3	1.1	11.55	1.1
FGD1 RHC BKG CC1 π	5.6	2.9	8.4	2.8	3.4	1.6	10.67	1.2
FGD1 RHC BKG CC-Other	5.9	3.0	8.9	2.9	2.6	1.4	10.66	1.4
FGD2 RHC BKG CC0 π	5.7	3.0	9.5	2.9	2.3	1.1	12.05	1.1
FGD2 RHC BKG CC1 π	5.6	3.0	8.6	2.9	3.3	1.5	10.65	1.3
FGD2 RHC BKG CC-Other	5.9	3.0	9.0	2.9	2.6	1.3	10.70	1.4
Total	4.5	2.8	7.9	2.8	2.1	1.2	9.00	0.3

Comparing Table 9.1 to the no-split feature Table 9.2 below indicates there is no big disparity seen between the two. This is expected as the events within the histograms for the split feature are all the same, the only difference is that they are spread out in different tag regions over a

larger stacked histogram in CC1Pi. Therefore, seeing an agreement between these systematics serves as a validation that the feature was implemented correctly.

TABLE 9.2: No-Split feature ND prior and posterior prediction event rate uncertainty apportioned by ND sample and systematic type [123].

Sample	Flux		Cross Section		ND280		Total	
	Prior	Post	Prior	Post	Prior	Post	Prior	Post
FGD1 FHC CC0π – 0p – 0γ	5.0	2.7	11.8	2.8	1.8	1.2	12.8	0.6
FGD1 FHC CC0π-Np-0 γ	5.5	2.8	11.7	3.2	3.5	2.2	12.9	0.9
FGD1 FHC CC1 π – 0γ	5.2	2.7	9.1	2.7	3.0	1.4	10.6	1.0
FGD1 FHC CC-Other-0 γ	5.4	2.8	8.0	2.8	5.2	2.3	11.0	1.6
FGD1 FHC CC-Photon	5.5	2.8	8.5	2.8	2.8	1.8	10.5	0.8
FGD2 FHC CC0 π – 0p – 0γ	5.1	2.7	11.2	2.8	2.1	1.1	11.5	0.6
FGD2 FHC CC0 π-Np-0 γ	5.5	2.8	11.3	3.3	3.9	2.4	12.2	1.0
FGD2 FHC CC1 π – 0γ	5.2	2.7	9.0	2.7	3.6	1.6	10.5	1.0
FGD2 FHC CC-Other-0 γ	5.6	2.8	8.0	2.8	6.3	2.7	11.5	1.9
FGD2 FHC CC-Photon	5.4	2.8	8.3	2.8	2.5	1.6	10.4	0.8
FGD1 RHC CC0 π	4.9	3.2	11.3	3.2	1.9	1.2	12.2	0.9
FGD1 RHC CC1 π	4.6	3.1	10.3	3.0	4.2	2.6	11.4	1.9
FGD1 RHC CC-Other	4.5	2.9	9.3	3.0	3.5	2.0	10.5	1.5
FGD2 RHC CC0 π	4.8	3.2	10.4	3.0	2.1	1.2	13.8	0.9
FGD2 RHC CC1 π	4.6	3.0	9.9	3.2	3.9	2.3	10.9	1.9
FGD2 RHC CC-Other	4.6	2.9	9.7	3.1	2.9	1.8	11.3	1.4
FGD1 RHC BKG CC0 π	5.8	2.8	10.1	2.8	2.2	1.1	10.6	1.1
FGD1 RHC BKG CC1 π	5.6	2.8	8.0	2.5	3.3	1.6	11.2	1.3
FGD1 RHC BKG CC-Other	5.9	2.9	8.6	2.7	2.6	1.4	10.1	1.4
FGD2 RHC BKG CC0 π	5.8	2.8	9.5	2.8	2.2	1.1	10.4	1.1
FGD2 RHC BKG CC1 π	5.6	2.8	8.2	2.5	3.2	1.6	10.7	1.3
FGD2 RHC BKG CC-Other	5.9	2.8	8.6	2.7	2.5	1.4	10.6	1.4
Total	4.5	2.7	8.0	2.6	2.1	1.2	9.1	0.3

Bibliography

[1] K. Abe et al. Measurements of neutrino oscillation in appearance and disappearance channels by the T2K experiment with 6.6×10^{20} protons on target. *Phys. Rev. D*, 91: 072010, Apr 2015. doi: 10.1103/PhysRevD.91.072010. URL https://link.aps.org/doi/10.1103/PhysRevD.91.072010.

[2] Q. R. Ahmad et al. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the sudbury neutrino observatory. *Physical Review Letters*, 87 (7), 2001. doi: 10.1103/physrevlett.87.071301.

[3] Y. et al. Fukuda. Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998. doi: 10.1103/PhysRevLett.81.1562. URL https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[4] Joseph John Thomson. XL. Cathode rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897.

[5] Ernest Rutherford. LXXIX. The scattering of α and β particles by matter and the structure of the atom. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21(125):669–688, 1911.

[6] Vivian Ellesworth Pound. VI. The absorption of the different types of beta rays, together with a study of the secondary rays excited by the them. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 17(97):126–153, 1909.

[7] Allen Franklin. The Prehistory of the Neutrino. 2018. URL http://neutrinohistory2018.in2p3.fr/proceedings/franklin.pdf.

[8] James Chadwick. Possible existence of a neutron. *Nature*, 129(3252):312–312, 1932.

[9] Fred L Wilson. Fermi's theory of beta decay. *American Journal of Physics*, 36(12):1150–1160, 1968.

[10] Necia G Cooper. Los Alamos Science, Number 25–1997: Celebrating the neutrino. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 1997.

[11] Clyde L Cowan Jr, Frederick Reines, FB Harrison, HW Kruse, and AD McGuire. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956.

[12] F. Reines. NEUTRINOS TO 1960: PERSONAL RECOLLECTIONS. *J. Phys. Colloq.*, 43(C8):237–260, 1982. doi: 10.1051/jphyscol:1982815.

[13] Minkel, JR. Nobel Focus: Neutrino and X-ray Vision. *Physics*, 10:18, 2002.

[14] B Pontecorvo. Inverse–process. *Chalk River Laboratory Report PD-205*, 1946.

[15] Raymond Davis Jr. Nobel Lecture: A half-century with solar neutrinos. *Reviews of Modern Physics*, 75(3):985, 2003.

[16] Davis et al. Search for neutrinos from the sun. *Physical Review Letters*, 20(21):1205–1209, 1968. doi: 10.1103/physrevlett.20.1205.

[17] John N. Bahcall, Aldo M. Serenelli, and Sarbani Basu. New solar opacities, abundances, helioseismology, and neutrino fluxes. *The Astrophysical Journal*, 621(1):L85–L88, jan 2005. doi: 10.1086/428929. URL https://doi.org/10.1086%2F428929.

[18] Masayuki Nakahata. History of Solar Neutrino Observations, 2022. URL https://arxiv.org/abs/2202.12421.

[19] V Gribov and B Pontecorvo. Neutrino astronomy and lepton charge. *Phys. Lett., 28B: 493-6(Jan. 20, 1969).*, 1 1969. doi: 10.1016/0370-2693(69)90525-5. URL https://www.osti.gov/biblio/4835052.

[20] Arthur B. McDonald. Nobel lecture: The sudbury neutrino observatory: Observation of flavor change for solar neutrinos. *Rev. Mod. Phys.*, 88:030502, Jul 2016. doi: 10.1103/RevModPhys.88.030502. URL https://link.aps.org/doi/10.1103/RevModPhys.88.030502.

[21] P Anselmann, Wolfgang Hampel, G Heusser, J Kiko, T Kirsten, E Pernicka, R Plaga, U Rönn, M Sann, C Schlosser, et al. Solar neutrinos observed by GALLEX at Gran Sasso. *Physics Letters B*, 285(4):376–389, 1992.

[22] JN Abdurashitov, Sage Collaboration, 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. *arXiv preprint astro-ph/0204245*, 2002.

[23] B Pontecorvo. Conservation of lepton charge. *JETP*, 5:1297, 1957.

[24] Maki et al. Remarks on the unified model of elementary particles. *Progress of Theoretical Physics*, 28(5):870–880, 1962. doi: 10.1143/ptp.28.870.

[25] The SLD Electroweak, Heavy Flavour Groups, ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, et al. Precision electroweak measurements on the Z resonance. *Physics Reports*, 427(5-6):257–454, 2006.

[26] Gaillard Danby, Jean Maurice Gaillard, Konstantin Goulianos, Leon M Lederman, Nariman Mistry, Melvin Schwartz, and Jack Steinberger. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Physical Review Letters*, 9(1):36, 1962.

[27] K Kodama, N Ushida, C Andreopoulos, N Saoulidou, G Tzanakos, P Yager, B Baller, D Boehnlein, Walter Freeman, B Lundberg, et al. Observation of tau neutrino interactions. *Physics Letters B*, 504(3):218–224, 2001.

[28] Georges Aad, Tatevik Abajyan, B Abbott, J Abdallah, S Abdel Khalek, Ahmed Ali Abdelalim, R Aben, B Abi, M Abolins, OS AbouZeid, et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012.

[29] Donald Hill Perkins. Proton decay experiments. *Annu. Rev. Nucl. Part. Sci.*, 34(CERN-EP-84-07):1–52, 1984.

[30] Kohji S Hirata, Takaaki Kajita, T Kifune, K Kihara, Masayuki Nakahata, K Nakamura, S Ohara, Y Oyama, N Sato, M Takita, et al. Observation of b 8 solar neutrinos in the kamiokande-ii detector. *Physical Review Letters*, 63(1):16, 1989.

[31] GT Ewan, HC Evans, and HW Lee. Sudbury neutrino observatory proposal. Technical report, Queen's Univ., 1987.

[32] Jea et al. Boger. The Sudbury neutrino observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 449(1-2):172–207, 2000.

[33] Q Retal Ahmad, RC Allen, TC Andersen, JD Anglin, JC Barton, EW Beier, M Bercovitch, J Bigu, SD Biller, RA Black, et al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. *Physical review letters*, 89(1):011301, 2002.

[34] Hirata, KS and Kajita, T and Koshiba, M and Nakahata, M and Oyama, Y and Sato, N and Suzuki, A and Takita, M and Totsuka, Y and Kifune, T and others. Observation in the Kamiokande-II detector of the neutrino burst from supernova SN1987A. *Physical Review D*, 38(2):448, 1988.

[35] Y Totsuka. Kamiokande and Super-Kamiokande. Technical report, SCAN-9604117, 1996.

[36] K. Scholberg. Atmospheric Neutrinos at Super-Kamiokande, 1999. URL https://arxiv.org/abs/hep-ex/9905016.

[37] K Hidaka, M Honda, and S Midorikawa. Neutrino oscillations and the anomalous atmospheric neutrino flux. *Physical review letters*, 61(14):1537, 1988.

[38] Takaaki Kajita. Nobel lecture: Discovery of atmospheric neutrino oscillations. *Rev. Mod. Phys.*, 88:030501, Jul 2016. doi: 10.1103/RevModPhys.88.030501. URL https://link.aps.org/doi/10.1103/RevModPhys.88.030501.

[39] T. Kajita and Super-K Collaboration. Establishing atmospheric neutrino oscillations with Super-Kamiokande. *Nuclear Physics B*, 908:14–29, 2016. ISSN 0550-3213. doi: https://doi.org/10.1016/j.nuclphysb.2016.04.017. URL https://www.sciencedirect.com/science/article/pii/S0550321316300554. Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015.

[40] Andrea Taroni. Nobel prize 2015: Kajita and McDonald. *Nature Physics*, 11(11):891–891, 2015.

[41] P Adamson, C Ader, M Andrews, N Anfimov, I Anghel, K Arms, E Arrieta-Diaz, A Aurisano, DS Ayres, C Backhouse, et al. First measurement of electron neutrino appearance in NOvA. *Physical review letters*, 116(15):151806, 2016.

[42] Babak Abi, Roberto Acciarri, Mario A Acero, Giorge Adamov, David Adams, Marco Adinolfi, Zubayer Ahmad, Jhanzeb Ahmed, Tyler Alion, S Alonso Monsalve, et al. Deep underground neutrino experiment (dune), far detector technical design report, volume ii: Dune physics. *arXiv preprint arXiv:2002.03005*, 2020.

[43] Masashi Yokoyama. The hyper-kamiokande experiment. *arXiv preprint arXiv:1705.00306*, 2017.

[44] Giganti et al. Neutrino oscillations: the rise of the pmns paradigm. *Progress in Particle and Nuclear Physics*, 98:5, 2018.

[45] F. P. An et al. Improved measurement of electron antineutrino disappearance at Daya Bay. *Chinese Physics C*, 37(1):011001, jan 2013. doi: 10.1088/1674-1137/37/1/011001. URL https://doi.org/10.1088%2F1674-1137%2F37%2F1%2F011001.

[46] Kayser and Boris. Neutrino physics, Jun 2005. URL https://arxiv.org/abs/hep-ph/0506165.

[47] N. Abgrall et al. Time projection chambers for the t2k near detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 637(1):25–46, 2011. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2011.02.036. URL https://www.sciencedirect.com/science/article/pii/S0168900211003421.

[48] Atmospheric Neutrinos and Neutrino Oscillations, Date of Access: Sept 2019. URL https://www.ipmu.jp/sites/default/files/imce/E02_Feature.pdf.

[49] Ivan Esteban, M.C. Gonzalez-Garcia, Michele Maltoni, Thomas Schwetz, and Albert Zhou. The fate of hints: updated global analysis of three-flavor neutrino oscillations. *Journal of High Energy Physics*, 2020(9), sep 2020. doi: 10.1007/jhep09(2020)178. URL https://doi.org/10.1007%2Fjhep09%282020%29178.

[50] Ivan Esteban, M. C. Gonzalez-Garcia, Alvaro Hernandez-Cabezudo, Michele Maltoni, and Thomas Schwetz. Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of 23, CP, and the mass ordering. *Journal of High Energy Physics*, 2019(1), jan 2019. doi: 10.1007/jhep01(2019)106. URL https://doi.org/10.1007%2Fjhep01%282019%29106.

[51] Kirsty Duffy. *Measurement of the neutrino oscillation parameters* $\sin^2\theta_{23}, \Delta^2_{32}, \sin^2\theta_{13}$, *and* δ_{CP} *in neutrino and anti–neutrino oscillation at T2K*. PhD thesis, University of Oxford, 2016.

[52] J. A. Formaggio and G. P. Zeller. From ev to eev: Neutrino cross sections across energy scales. *Rev. Mod. Phys.*, 84:1307–1341, Sep 2012. doi: 10.1103/RevModPhys.84.1307. URL https://link.aps.org/doi/10.1103/RevModPhys.84.1307.

[53] Sara Bolognesi Luke Pickering, Yoshinari Hayato and Kevin McFarland. TN344. NIWG model and uncertainties for 2019-2020 oscillation analysis. *Internal T2K Technote*, October 2019.

[54] L. Alvarez-Ruso et al. Nustec1 1neutrino scattering theory experiment collaboration http://nustec.fnal.gov. white paper: Status and challenges of neutrino–nucleus scattering. *Progress in Particle and Nuclear Physics*, 100:1–68, 2018. ISSN 0146-6410. doi: https://doi.org/10.1016/j.ppnp.2018.01.006. URL https://www.sciencedirect.com/science/article/pii/S0146641018300061.

[55] M Glück, E Reya, and A Vogt. Dynamical parton distributions revisited. *The European Physical Journal C-Particles and Fields*, 5(3):461–470, 1998.

[56] Arie Bodek and UK Yang. Higher twist, ξw scaling, and effective LO PDFs for lepton scattering in the few GeV region. *Journal of Physics G: Nuclear and Particle Physics*, 29 (8):1899, 2003.

[57] Kevin McFarland Sara Bolognesi, Callum Wilkinson and Yoshinari Hayato. TN315. Updated recommendation of the 2017 NIWG parameters. *Internal T2K Technote*, April 2017.

[58] Teppei Katori. Meson Exchange Current (MEC) Models in Neutrino Interaction Generators, 2013. URL https://arxiv.org/abs/1304.6014.

[59] Callum Wilkinson. NEUT development for T2K and relevance of updated 2p2h models. In *NuFact 2014, XVIth International Workshop on Neutrino Factories and Future Neutrino Beam Facilities*, 2014.

[60] Julia Tena Vidal. Comparison of Validation Methods for Final State Interactions in Hadron Production Experiments. University Lecture, March 2021.

[61] ES Pinzon Guerra, C Wilkinson, S Bhadra, S Bolognesi, J Calcutt, P De Perio, S Dolan, T Feusels, GA Fiorentini, Y Hayato, et al. Using world π±-nucleus scattering data to constrain an intranuclear cascade model. *Physical Review D*, 99(5):052007, 2019.

[62] Max Aker, Konrad Altenmüller, Marius Arenz, Woo-Jeong Baek, John Barrett, Armen Beglarian, Jan Behrens, Anatoly Berlev, Uwe Besserer, Klaus Blaum, et al. First operation of the katrin experiment with tritium. *The European Physical Journal C*, 80(3):1–18, 2020.

[63] Aharon Davidson and Kameshwar C Wali. Universal seesaw mechanism? *Physical Review Letters*, 59(4):393, 1987.

[64] Josephine Paton. Neutrinoless Double Beta Decay in the SNO+ Experiment, 2019. URL https://arxiv.org/abs/1904.01418.

[65] ST Petcov and M Piai. The lma msw solution of the solar neutrino problem, inverted neutrino mass hierarchy and reactor neutrino experiments. *Physics Letters B*, 533(1-2): 94–106, 2002.

[66] Simpson et al. Strong bayesian evidence for the normal neutrino hierarchy. *Journal of Cosmology and Astroparticle Physics*, 2017(06):029, 2017.

[67] RN Cahn, DA Dwyer, SJ Freedman, WC Haxton, RW Kadel, Yu G Kolomensky, KB Luk, P McDonald, GD Gann, and AWP Poon. White paper: measuring the neutrino mass hierarchy. *arXiv preprint arXiv:1307.5487*, 2013.

[68] Carlos Blanco, Dan Hooper, and Pedro Machado. Constraining sterile neutrino interpretations of the LSND and MiniBooNE anomalies with coherent neutrino scattering experiments. *Physical Review D*, 101(7), April 2020. doi: 10.1103/physrevd.101.075051. URL https://doi.org/10.1103%2Fphysrevd.101.075051.

[69] R Acciarri, C Adams, R An, C Andreopoulos, AM Ankowski, M Antonello, J Asaadi, W Badgett, L Bagby, B Baibussinov, et al. A proposal for a three detector short-baseline neutrino oscillation program in the Fermilab booster neutrino beam. *arXiv preprint arXiv:1503.01520*, 2015.

[70] K Nakamura, ST Petcov, et al. Neutrino masses, mixing, and oscillations. 2018.

[71] K. Abe, N. Abgrall, H. Aihara, Y. Ajima, J.B. Albert, D. Allan, P.-A. Amaudruz, C. Andreopoulos, B. Andrieu, M.D. Anerella, and et al. The T2K experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 659(1):106–135, Dec 2011. ISSN 0168-9002. doi: 10.1016/j.nima.2011.06.067. URL http://dx.doi.org/10.1016/j.nima.2011.06.067.

[72] K. Abe et al. The T2K experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 659(1):106 – 135, 2011. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2011.06.067. URL http://www.sciencedirect.com/science/article/pii/S0168900211011910.

[73] K Abe. T2K neutrino flux prediction. *Phys. Rev. D*, 87:012001, Jan 2013. doi: 10.1103/PhysRevD.87.012001. URL https://link.aps.org/doi/10.1103/PhysRevD.87.012001.

[74] J-PARC website, Date of Access: Jun 2020. URL https://j-parc.jp/Neutrino/en/nu-facility.html.

[75] Takahiro Hiraki. First measurement of muon anti-neutrino disappearance by the T2K experiment. *Proceedings of Flavor Physics CP Violation 2015 — PoS(FPCP2015)*, 2016. doi: 10.22323/1.248.0068.

[76] Updated results from T2K experiment with 3.13×10^{21} POT, Date of Access: Sept 2019. URL http://www.icrr.u-tokyo.ac.jp/icrr_seminar/seminars18/20190110_Okumura.pdf.

[77] K Suzuki et al. Measurement of the muon beam direction and muon flux for the T2K neutrino experiment. *Progress of Theoretical and Experimental Physics*, 2015(5), 05 2015. ISSN 2050-3911. doi: 10.1093/ptep/ptv054. URL https://doi.org/10.1093/ptep/ptv054. 053C01.

[78] K. Abe et al. The T2K experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 659(1):106 – 135, 2011. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2011.06.067. URL http://www.sciencedirect.com/science/article/pii/S0168900211011910.

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

[80] S. Aoki et al. The T2K Side Muon Range Detector (SMRD). *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 698:135 – 146, 2013. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2012.10.001. URL http://www.sciencedirect.com/science/article/pii/S0168900212011242.

[81] P.-A. Amaudruz et al. The T2K fine-grained detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 696:1 – 31, 2012. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2012.08.020. URL http://www.sciencedirect.com/science/article/pii/S0168900212008789.

[82] Kei Ieki. MPPC for T2K fine-grained detector. In *International Workshop on New Photon Detectors*, volume 90, page 023. SISSA Medialab, 2010.

[83] N. Abgrall et al. Time projection chambers for the t2k near detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 637(1):25 – 46, 2011. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2011.02.036. URL http://www.sciencedirect.com/science/article/pii/S0168900211003421.

[84] S. Fukuda et al. The Super-Kamiokande detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 501(2):418 – 462, 2003. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(03)00425-X. URL http://www.sciencedirect.com/science/article/pii/S016890020300425X.

[85] Super-K Gallery, 1993-1996. URL https://www-sk.icrr.u-tokyo.ac.jp/en/sk/experience/gallery/.

[86] N Abgrall, O Andreeva, A Aduszkiewicz, Yasir Ali, T Anticic, N Antoniou, B Baatar, F Bay, A Blondel, J Blumer, et al. Na61/shine facility at the cern sps: beams and detector system. *Journal of Instrumentation*, 9(06):P06005, 2014.

[87] Alfredo Ferrari, Paola R Sala, Alberto Fasso, Johannes Ranft, U Siegen, et al. FLUKA: a multi-particle transport code. Technical report, Stanford Linear Accelerator Center (SLAC), 2005.

[88] René Brun, F Bruyant, Federico Carminati, Simone Giani, M Maire, A McPherson, G Patrick, and L Urban. *GEANT: Detector Description and Simulation Tool; Oct 1994*. CERN Program Library. CERN, Geneva, 1993. doi: 10.17181/CERN.MUHF.DMJ1. URL https://cds.cern.ch/record/1082634. Long Writeup W5013.

[89] C Zeitnitz and TA Gabriel. GCALOR. In *Proc. of International Conference on Calorimetry in High Energy Physics*, 1993.

[90] Nicolas Abgrall. NA61-SHINE: Hadron Production Measurements for Cosmic Ray and Neutrino Experiments, 2010.

[91] Yoshinari Hayato. A NEUTRINO INTERACTION SIMULATION PROGRAM LIBRARY NEUT. *Acta Physica Polonica B*, 40(9), 2009.

[92] S Berkman, P de Perio, A Konaka, et al. fitqun: A new reconstruction algorithm for super-k. Technical report, T2K-TN-146, 2013.

[93] Andrew D. Missert. Improving the T2K Oscillation Analysis With fiTQun: A New Maximum-Likelihood Event Reconstruction for Super-Kamiokande. *Journal of Physics: Conference Series*, 888:012066, sep 2017. doi: 10.1088/1742-6596/888/1/012066. URL https://doi.org/10.1088/1742-6596/888/1/012066.

[94] Juan Estrada. MCMII and the TriP chip. 12 2003. doi: 10.2172/820406.

[95] P. A. S. Ritt and K. Olchanski. MIDAS (Maximum Integration Data Acquisition). 2001.

[96] Y. Giomataris, Ph. Rebourgeard, J.P. Robert, and G. Charpak. Micromegas: a high-granularity position-sensitive gaseous detector for high particle-flux environments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 376(1):29 – 35, 1996. ISSN 0168-9002. doi: https://doi.org/10.1016/0168-9002(96)00175-1. URL http://www.sciencedirect.com/science/article/pii/0168900296001751.

[97] R. Asfandiyarov et al. Proposal for SPS beam time for the baby MIND and TASD neutrino detector prototypes, 2014.

[98] 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 Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 656(1):69 – 83, 2011.

ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2011.07.022. URL http://www.sciencedirect.com/science/article/pii/S0168900211014513.

[99] M. Thorpe, C. Angelsen, G. Barr, C. Metelko, T. Nicholls, G. Pearce, and N. West. The T2K Near Detector Data Acquisition Systems. *IEEE Transactions on Nuclear Science*, 58(4):1800–1806, 2011.

[100] Antonin Vacheret, Sarah Greenwood, Matt Noy, Mark Raymond, and Alfons Weber. The front end readout system for the T2K-ND280 detectors. In *2007 IEEE Nuclear Science Symposium Conference Record*, volume 3, pages 1984–1991. IEEE, 2007.

[101] D. Orme, N. Nagai, A. Minamino, T. Nakaya, M. Yokoyama, T. Nakadaira, T. Murakami, M. Tanaka, F. Retiere, A. Vacheret, and Yu. Kudenko. Development of multi-pixel photon counters for the t2k long baseline neutrino experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 623(1):321–323, 2010. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2010.02.235. URL https://www.sciencedirect.com/science/article/pii/S0168900210005206. 1st International Conference on Technology and Instrumentation in Particle Physics.

[102] T2K Collaboration. Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580(7803):339–344, apr 2020. doi: 10.1038/s41586-020-2177-0. URL https://doi.org/10.1038%2Fs41586-020-2177-0.

[103] Fred James et al. Minuit. *CERN Program Library Long Writeup D*, 506:1993, 1994.

[104] Charles J Geyer. Practical markov chain monte carlo. *Statistical science*, pages 473–483, 1992.

[105] Roger Barlow and Christine Beeston. Fitting using finite monte carlo samples. *Computer Physics Communications*, 77(2):219–228, 1993.

[106] J. S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra. *PHYSTAT 2011, CERN*.

[107] W Keith Hastings. Monte Carlo sampling methods using Markov chains and their applications. 1970.

[108] J. E. Gubernatis. Marshall rosenbluth and the metropolis algorithm, January 2005. URL https://doi.org/10.1063/1.1887186.

[109] T. Doyle et al. TN423. Constraining the flux and cross section models for the 2021 oscillation analysis using ND280 data. *Internal T2K Technote*, May 2022.

[110] T. Doyle et al. TN421. ND280 FHC samples with Photon and Proton information for the 2021 Oscillation Analysis. *Internal T2K Technote*, May 2022.

[111] Justyna agoda. Probing kaon-originated neutrinos with the muons produced outside of the T2K near detector. In *Journal of Physics: Conference Series*, volume 888, page 012217. IOP Publishing, 2017.

[112] T. Doyle et al. TN421. ND280 FHC samples with Photon and Proton information for the 2021 Oscillation Analysis. Executive Summary. *Internal T2K Technote*, May 2022.

[113] S. J. Jenkins. TN417. Updated measurement of numuCC1+ cross-section on hydrocarbon and water. *Internal T2K Technote*, Oct 2021.

[114] D. Barrow et al. TN399. Super-Kamiokande Data Quality, MC, and Systematics in Run10. *Internal T2K Technote*, June 2020.

[115] C. Vilela R. Wendell S. M . Lakshmi, Y. Nakajima. TN447. Super-Kamiokande ν_μ multi-ring samples for the 2021 Oscillation Analysis. *Internal T2K Technote*, May 2022.

[116] G Battistoni, PR Sala, M Lantz, Alfredo Ferrari, and G Smirnov. Nuclear effects in neutrino interactions and their implementation in the fluka monte carlo code. *Ric. Sci. Educ. Perm., Suppl.*, 126:497–506, 2006.

[117] L. Berns et al. TN401. Flux Prediction and Uncertainty with NA61/SHINE 2010 Replica-Target Data. *Internal T2K Technote*, March 2022.

[118] S. King et al. TN407. FHC muon neutrino charged current multi-pion samples in the ND20 tracker for summer 2020 0A inputs. P6T NEUT D. Run 2+3+4+8. *Internal T2K Technote*, June 2020.

[119] Yoshinari Hayato and Luke Pickering. The NEUT neutrino interaction simulation program library. *The European Physical Journal Special Topics*, 230(24):4469–4481, oct 2021. doi: 10.1140/epjs/s11734-021-00287-7. URL https://doi.org/10.1140%2Fepjs%2Fs11734-021-00287-7.

[120] J Wolcott, L Aliaga, O Altinok, L Bellantoni, A Bercellie, M Betancourt, A Bodek, A Bravar, H Budd, T Cai, et al. Measurement of Electron Neutrino Quasielastic and Quasielasticlike Scattering on Hydrocarbon at¡ E ν¿= 3.6 GeV. *Physical review letters*, 116(8):081802, 2016.

[121] E. Atkin et al. TN429. Measuring Oscillation Parameters in a joint ND280+SK analysis using Markov Chain Monte Carlo Methods. *Internal T2K Technote*, Sep 2022.

[122] Dieter Rein and Lalit M Sehgal. Neutrino-excitation of baryon resonances and single pion production. *Annals of Physics*, 133(1):79–153, 1981. ISSN 0003-4916. doi: https://doi.org/10.1016/0003-4916(81)90242-6. URL https://www.sciencedirect.com/science/article/pii/0003491681902426.

[123] R. Gran, J. Nieves, F. Sanchez, and M. J. Vicente Vacas. Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 gev. *Phys. Rev. D*, 88:113007, Dec 2013. doi: 10.1103/PhysRevD.88.113007. URL https://link.aps.org/doi/10.1103/PhysRevD.88.113007.

[124] M. Martini, M. Ericson, G. Chanfray, and J. Marteau. Unified approach for nucleon knock-out and coherent and incoherent pion production in neutrino interactions with nuclei. *Physical Review C*, 80(6), dec 2009. doi: 10.1103/physrevc.80.065501. URL https://doi.org/10.1103%2Fphysrevc.80.065501.

[125] Ch. Berger and L. M. Sehgal. Partially conserved axial vector current and coherent pion production by low energy neutrinos. *Physical Review D*, 79(5), mar 2009. doi: 10.1103/physrevd.79.053003. URL https://doi.org/10.1103%2Fphysrevd.79.053003.

[126] Arie Bodek and Un-ki Yang. A Unified Model for inelastic e- N and v- N cross sections at all Q 2. In *AIP Conference Proceedings*, volume 792, pages 257–260. American Institute of Physics, 2005.

[127] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1–19, 2011.

[128] C. Bronner et al. TN414. NIWG Model Recommendations for OA 2021. *Internal T2K Technote*, April 2021.