Measurement of the Muon Neutrino Charged Current Single Positive Pion Production Cross Section on Water in the ND280 Near Detector

The University Of Sheffield.

Sam Jenkins

Supervisor: Dr. Susan Cartwright

Department of Physics and Astronomy,
The University of Sheffield

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

March 31, 2022

Abstract

T2K is a long-baseline neutrino oscillation experiment, situated in Japan, which measures the oscillation probability of muon neutrinos produced at the J-PARC accelerator complex and detected at the water Cherenkov far detector, Super-Kamiokande. Essential to measuring neutrino oscillation probabilities is the understanding of the underlying neutrino-nucleus interaction cross section, which is currently a leading systematic uncertainty in oscillation measurements. It is therefore crucial to reduce this uncertainty, particularly as neutrino physics moves towards next generation experiments, where measurements will no longer be statistically limited.

This thesis presents a measurement of the muon neutrino charged current single positive pion cross section on both water and hydrocarbon targets, made using data taken by the T2K near detector ND280, with a simultaneous extraction using log-likelihood minimisation. Although such measurements exist, this is the first of its kind to be performed in the form of a double differential cross section as a function of the outgoing pion kinematics, which are sensitive to both the underlying nuclear ground state model and the neutrino interaction model. The initial unblinding to data shows a reasonable level of tension with the nominal Monte Carlo model used, with a χ^2/Ndf value of 4.31. In addition to this measurement, an attempt is made to measure the 4-dimensional differential cross section, in terms of both muon and pion kinematics simultaneously. Whilst successful, this measurement is seen to be statistically limited. Both measurements also feature a novel method for reconstructing the kinematics of pions detected through their decay to muons and subsequently Michel electrons, providing access to a region of pion kinematics previously unmeasured by T2K.

It's a dangerous business, Frodo,
going out your door.
You step onto the road, and if you don't keep your feet,
there's no knowing where you might be swept off to.

– J.R.R. Tolkien, The Lord of the Rings

Acknowledgements

My first thanks go to my supervisor, Susan Cartwright, both for giving me this opportunity to pursue a PhD and supporting me through it. While never overbearing, you have always been available to provide insight on my work, along with encouraging me to take up additional opportunities throughout. Your meticulous proof-reading of this thesis has also been incredibly helpful. I would also like to thank Lee Thompson, whose door has always been open to me when I've been struggling. Your advice has always been useful, whether it be about neutrino physics or the best ramen spots in Tokyo.

Four years in the department would have been far less entertaining if it weren't for the various occupants of D36a during my time. Ample cake, communal office jigsaws and a beverage bell should be mandatory in all workplace environments. I extend my particular thanks to Liam O'Sullivan, for his guidance on both cross-section measurements and *onsen* etiquette, as well as Celeste Pidcott for her continued support during the PhD. Your belief that I am a capable researcher has definitely prevented me from quitting at least twice, and I can't thank you enough for suffering through highland with me.

While there are too many people on T2K, Super-K and Hyper-K to thank individually, I would like to express my thanks to all I have worked with and who have made collaboration meetings so enjoyable. Particular thanks go to Andrew Cudd, for helping me understand and debug the fitter software; Luke Pickering, for making my brief foray into T2KReWeight slightly more bearable; Xianguo Lu for his suggestions and guidance in the Michel electron reconstruction work; and Sophie King, for both her help understanding and debugging systematics, as well as letting me rant to her about many, many things. Special thanks also go to Stephen Dolan, for putting up with my many questions as I got to grips with how the fitter works, and for his advice over the last few months as I attempted to consolidate this work into a cross-section measurement. I must also thank Helen O'Keeffe, who has been like an additional supervisor to me over these years; I'm sure you didn't realise what you were getting yourself into when you agreed to be my master's supervisor. Thank you for your insight and knowledge on the timeslip systematics, general support over the years, and of course for indoctrinating me into neutrino physics in the first place.

My time living in Tōkai wouldn't have been nearly as fun if it weren't for the friends I was with: Jordan, Tristan, Joe, Charlie, Yue and Gabriel, known collectively as the 'ECAL RMM3 Appreciation Society'. Frequent early mornings and long nights dealing with detector issues were balanced by the many Mario Kart tournaments, meals at *Gyū-Kaku*, karaoke sessions and trips to Joyful Honda. The Human League will forever remind me of late-night trips to 7-Eleven in the DAQ-mobile.

Thank you to Kathryn, Heather, Ellie H., Ellie P., Alex and Ashley for your continued support and encouragement. Additional thanks to Harry, Joe and Drew, both for the aforementioned, and for the weekly games of D&D that kept me going whilst the world was falling apart around us. I'd also like to thank the friends that I have lived with during my time in Sheffield: Sarah, for your support and understanding on how hard a PhD can be,

and in particular, Ollie, Rob and Jordan. If it weren't for the large amount of board games we played, along with frequent D&D and rockband sessions, being stuck in the house for 18 months would have driven me insane. Instead, it was actually kind of fun.

Sophie, thank you for your continued love and support, and of course for putting up with my constant stress, which has been particularly bad over the last few months. Luckily this thesis isn't made from fondant icing, else I'd be very close to throwing it straight in the bin. Finally I'd like to thank my parents, for their encouragement throughout both this PhD and anything else I've done. You've always been supportive of me to do anything I wanted in life, and I can't thank you enough for that.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Neutrinos are neutral leptons, which interact only through the weak nuclear force and gravity. The name itself means 'little neutral one', originating from the Italian diminutive of 'neutral', and goes some way to describing the important properties of the neutrino.

1.1 A Brief History of Neutrinos

In 1930, Austrian physicist Wolfgang Pauli did something that he himself later described as "terrible". In his letter of the 4th of December, which he addressed to the "Radioactive Ladies and Gentlemen" at a meeting in Tübingen he could not attend, he proposed a new particle as a "desperate remedy" to the apparent issues that plagued studies of beta decay at the time [1]. From its first measurement in 1911, all the way through to the early 1930s, the observed continuous energy spectrum of the measured outgoing electron from beta decay was a source of contention; in what was assumed to be a two body decay, the electron should have a narrow distribution of kinetic energy, associated with the energy difference between the initial and final nuclear states. Such a persistent observation had several notable physicists at the time even questioning conservation of energy, including Niels Bohr who suggested it may only be conserved on a statistical basis, and thus could be violated in a single decay. Additionally, investigation into the spin of ^{14}N atoms revealed an integer value, which was not possible under the current model, implying non-conservation of angular momentum.

In the letter, Pauli proposed that an additional particle existed within the nucleus, which he named the 'neutron'. He hypothesised that these 'neutrons', which were required to be electrically neutral, spin $\frac{1}{2}$, and obey his exclusion principle, would have a mass on the order of magnitude of that of the electron. Assuming then that this 'neutron' were also ejected from the nucleus in beta decay, it would carry sufficient energy that the sum of energies of the electron and 'neutron' would be constant. Despite the potential for this explaining the continuous beta decay spectrum, Pauli was deeply uncomfortable with postulating "a particle that cannot be detected" [2].

In 1933, Enrico Fermi proposed his theory of beta decay, which included Pauli's neutrino – now renamed to avoid confusion with James Chadwick's neutron discovered the year before. He supposed that beta decay was a four-way interaction, in which a neutron within

the nucleus would decay to a proton, producing an electron and neutrino with it. It would later be discovered that in order to conserve lepton number, this neutrino must specifically be an electron antineutrino, such that beta decay proceeds as:

$$n \longrightarrow p + e^- + \bar{\nu}_e. \tag{1.1}$$

This was a major step forward for the application of quantum mechanics to matter particles, hypothesising that the electron and neutrino were created in the interaction, rather than contained within the nucleus as had previously been thought. Fermi's paper on the theory of beta decay was published a year later[1]. At the same time, experimental evidence was presented [4] showing a hard upper limit on the energy spectra of electrons in beta decay, going against Bohr's idea that energy was only conserved on a statistical basis.

Several early attempts at finding experimental evidence of the neutrino, such as that of Crane and Halpern [5] and Nahmias [6], focused on indirect detection methods, where the recoil energy of the beta-ray emitting nucleus was compared with that of the electron. But it wasn't until 1956 that the neutrino was directly detected. Following his work on the Manhattan project, Fred Reines joined with Los Alamos colleague Clyde Cowan Jr. to set up "Project Poltergeist" [7] in 1951, which was dedicated to detecting these ghostly particles. The principle of detection was simple. From his work with atomic weapons, Reines knew that a bomb would be an ideal neutrino source; not only would a large number of neutrinos be produced from the decay of short-lived fission products, the blast would take place over such a short timescale that the effect of backgrounds on the detector would be minimal. The detection method relied on the inverse of the beta decay process predicted by Fermi. An antineutrino, on interacting with a proton, can transform the proton to a neutron, with a positron also emitted:

$$\bar{\nu}_e + p \longrightarrow n + e^+. \tag{1.2}$$

To provide the target protons for the inverse beta decay interactions, a detector filled with several tons of liquid scintillator would be used. The scintillation properties of some organic liquids had recently been discovered, and provided the opportunity to measure the energy of the detected particles from the intensity of light produced. The remaining issue came from protecting the sensitive detector from the force of a kiloton-scale nuclear bomb, and resulted in the initial experimental scheme shown in Figure 1.1. A tunnel would be dug into the ground near the bomb test site, and the detector suspended halfway down, before filling the upper half of the tunnel to protect from radiation. At the moment of detonation, the suspension line would be cut, and the detector would fall freely for several seconds through an evacuated chamber whilst the shock wave passed, landing in a pit filled with feathers and foam rubber to protect it from the impact. Despite being approved by the

[1]Fermi initially submitted the paper to the journal Nature, which rejected it on the grounds that "it contained speculations too remote from reality to be of interest to the reader" [3]. Despite revised versions being accepted by German and Italian publications, the initial rejection and general lack of early enthusiasm for the theory ultimately caused Fermi to move from theoretical to experimental physics for some time. Nature would later go on to describe this as its greatest blunder.

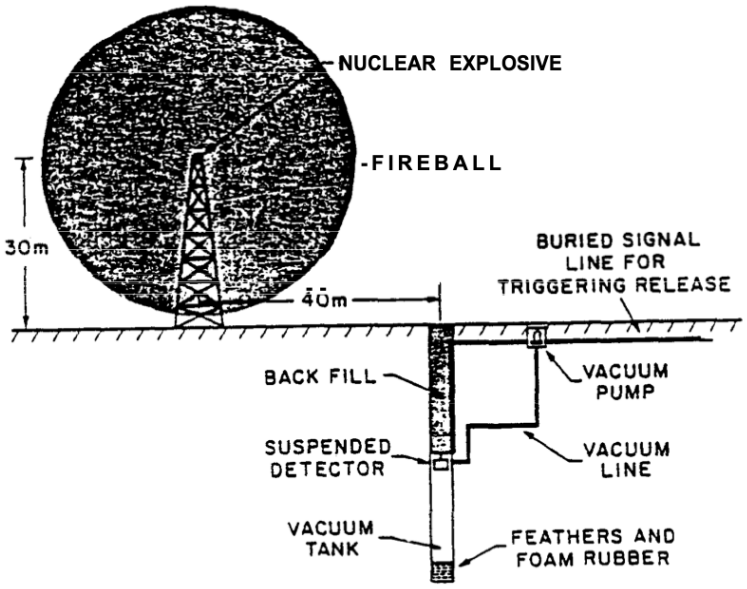


Figure 1.1: Schematic showing the first proposed experimental setup for Project Poltergeist, where antineutrinos from a nuclear explosion would be detected by a liquid scintillation detector in free fall [7].

Los Alamos Laboratory director and construction being started on both the detector and pit, Reines and Cowan ultimately decided to instead use a nuclear reactor as the neutrino source, starting with a prototype detector at the Hanford reactor in 1953. Despite the lower flux than from a nuclear bomb, using a reactor as a neutrino source was deemed preferable as it could be more controlled, and would enable a much longer period of data taking with repeat measurements. In addition, the interaction of neutrinos in the liquid scintillator tank could be further confirmed by the delayed detection of produced neutrons. This was done by dissolving $CdCl_2$ in the scintillator, due to cadmium having a high neutron capture cross section. The neutron produced from the interaction, after losing energy in the tank, would be captured by the cadmium:

$$n + {}^{108}\mathrm{Cd} \longrightarrow {}^{109\mathrm{m}}\mathrm{Cd} \longrightarrow {}^{109}\mathrm{Cd} + \gamma, \tag{1.3}$$

which on dropping down to a lower energy state released a delayed gamma ray. This was then detected in coincidence a few microseconds later [8]. The results from the Hanford prototype experiment were published in Reines and Cowan's 1953 paper [9], where they suggested detection had been "probable", but further work was required. They later moved to the Savannah River plant, where the detector design was overhauled considerably, and the underground placement of the detector reduced background sources. Instead of the single cylindrical tank used in the Hanford experiment, they moved to having two flat water tanks, which were the target for the inverse beta decay process. Cadmium chloride was dissolved in the water, in order to obtain a greater neutron capture cross section, and the water target panels were layered between three large tanks of liquid scintillator, used to detect the resulting gamma rays from positron annihilation and cadmium de-excitation. The new

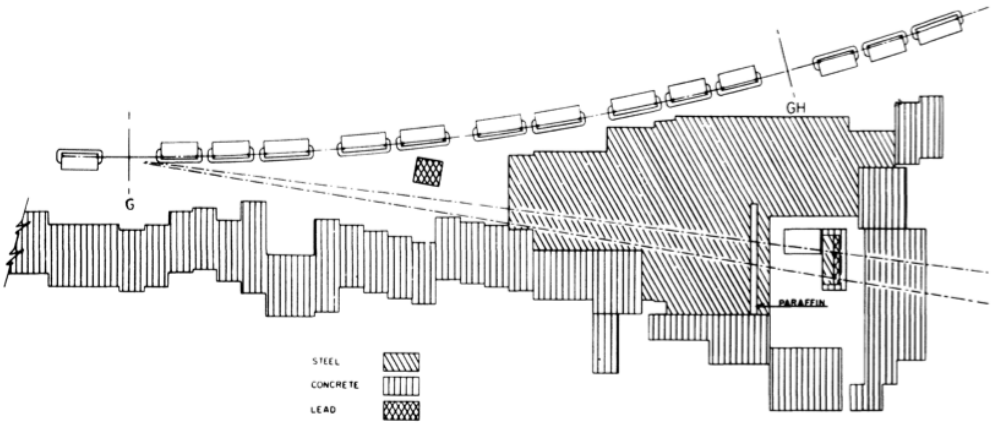


Figure 1.2: Diagram of the experimental setup at AGS. Pions produced from proton collisions on a beryllium target at G decayed in flight to produce muons and muon neutrinos, the latter of which penetrated the iron shielding to reach the spark chamber on the opposite side [12].

configuration, where an interaction in a target panel generally only produced signal in the two neighbouring scintillator tanks, was found to be much better at separating signal from background caused by cosmic rays. The detection of neutrinos was confirmed and published in 1956 [10], 26 years after Pauli first suggested their existence, and earned Reines the 1995 Nobel Prize in Physics [11].

Only seven years after the confirmation of the electron (anti)neutrino, the muon neutrino was discovered. Led by Leon Lederman, Melvin Schwartz and Jack Steinberger [12], a team at the Brookhaven National Laboratory used the Alternating Gradient Synchrotron (AGS) to create a beam of 15 GeV protons, which on striking a beryllium target produced a beam of pions, moving in the same general direction as the proton beam. These pions decayed in flight as

$$\pi^{\pm} \longrightarrow \mu^{\pm} + \nu/\bar{\nu}, \tag{1.4}$$

producing a beam of both muons and neutrinos, where the majority of the muons were stopped by a 13.5 m thick wall of iron shielding. A diagram of the experimental setup at AGS is presented in Figure 1.2. After passing through the shielding, the produced neutrinos reached a 10 ton aluminium spark chamber. Neutrino interactions in the aluminium layers of the spark chamber caused production of muons, which left sparks that could be captured by surrounding cameras. Along with identifying the presence of muons from neutrino interactions, Lederman, Schwartz and Steinberger additionally compared the number of identified muons in the chamber to that of electrons. If the neutrinos produced from the pion decay were the same as the electron neutrinos previously discovered, muons and electrons should have been produced in the spark chamber at a roughly equal rate. The experiment identified 34 single muon events, of which five were considered to be from cosmic ray backgrounds, but only six shower events consistent with electron behaviour. From this

they concluded that the muon neutrino[2] ν_μ produced from pion decay was indeed different to the previously identified ν_e. For the discovery of the muon neutrino, and the observation that there were two types of neutrino which coupled to electrons and muons separately, Lederman, Schwartz and Steinberger received the 1988 Nobel Prize in Physics [13]. It was also the first use of an accelerator to create a beam of neutrinos, a technique now employed by many modern neutrino experiments.

Following the discovery of the tau lepton at SPEAR in 1975 [14], the existence of a third type of neutrino was expected. Several experiments over the next 10 years were able to identify the presence of the ν_τ via measurements of τ decay and lifetime [15, 16, 17], but it wasn't until 2000 when the direct observation of the ν_τ through interactions was discovered by the DONUT collaboration (Direct Observation of the Nu Tau) [18]. This involved colliding 800 GeV protons from the TeVatron proton accelerator at Fermilab with a metre-long tungsten beam dump, producing a wide variety of mesons, including the strange-charmed meson D_s, which decays as

$$\mathrm{D}_s^- \longrightarrow \tau^- + \bar{\nu}_\tau, \tag{1.5}$$

with a branching fraction of $(5.48 \pm 0.23)\%$ [19]. The subsequent decay of the tau lepton produces further ν_τ. 36 m downstream from the beam dump, the neutrinos would interact in a nuclear emulsion tracker, where $\sim$5% of the neutrino interactions were predicted to originate from ν_τ. The remaining neutrino interactions were expected to be caused by neutrinos originating from the decay of other hadrons produced in the collisions with the beam dump. Charged current ν_τ interactions produce τ leptons which, due to their short-lived nature, were identified by searching for 'kinks' in the emulsion tracks of outgoing particles, where the τ had decayed to a single charged daughter[3]. Out of the total 203 recorded neutrino interactions, four passed all criteria to be evidence of ν_τ interactions. This was consistent with the standard model, and represented the first direct detection of tau neutrinos.

In the unification of the electromagnetic and weak nuclear forces into electroweak theory by Glashow, Salam and Weinberg, three massive gauge bosons were predicted as mediators[4]. These mediators, the $W^\pm$ and Z^0 bosons were predicted to have masses on the order of 80–90 GeV, and thus it wasn't until the emergence of high energy proton-antiproton colliders that experiments could start to investigate them. Following the observations of weak neutral current interactions by Gargamelle in 1973 [20], and the subsequent detection of the $W^\pm$ and Z^0 bosons by the UA1 and UA2 collaborations in 1983 [21], work began to use the existence of the neutral Z^0 boson to experimentally confirm the number of light neutrinos that should exist. Around the Z^0 pole, the $e^+e^- \longrightarrow Z^0 \longrightarrow f\bar{f}$ annihilation cross section

[2] The muon neutrino was initially named the 'neutretto' for distinction from the previously discovered neutrino, but the naming convention never stuck.

[3] 86% branching fraction.

[4] Along with one massless mediator, the photon.

is given by

$$\sigma_f = \frac{12\pi}{M_Z^2} \frac{s\Gamma_{ee}\Gamma_{f\bar{f}}}{(s - M_Z^2)^2 + s^2 \frac{\Gamma_Z^2}{M_Z^2}}, \tag{1.6}$$

which around the Z^0 mass resonance peak ($\sqrt{s} = M_Z$) reduces to

$$\sigma_f = \frac{12\pi}{M_Z^2} \frac{\Gamma_{ee}\Gamma_{f\bar{f}}}{\Gamma_Z^2}, \tag{1.7}$$

where Γ_Z is the total resonance width, $\Gamma_Z = \Gamma_{\text{had}} + \Gamma_{ee} + \Gamma_{\mu\mu} + \Gamma_{\tau\tau} + N_\nu \Gamma_{\nu\nu}$. It is clear then that the cross section is sensitive to the number of light neutrinos, N_ν [22]. Given that the cross section had been experimentally measured, and found to be in good agreement with the standard model prediction [23], this can be used to calculate the number of light neutrinos allowed by Z^0 decays. Employing lepton universality ($\Gamma_{ee} = \Gamma_{\mu\mu} = \Gamma_{\tau\tau} = \Gamma_{\ell\ell}$), and defining the invisible width due to neutrinos as $\Gamma_{\text{inv}} = N_\nu \Gamma_{\nu\nu}$, the experimentally measured ratio of invisible to leptonic widths around the pole,

$$R_{\text{inv}}^0 \equiv \frac{\Gamma_{\text{inv}}}{\Gamma_{\ell\ell}} = N_\nu \left(\frac{\Gamma_{\nu\nu}}{\Gamma_{\ell\ell}} \right)_{\text{SM}}, \tag{1.8}$$

allows for calculation of the number of light neutrinos. The first measurement of this by the ALEPH collaboration was reported in 1989, yielding a value of

$$N_\nu = 3.27 \pm 0.30, \tag{1.9}$$

ruling out a fourth type of light neutrino at the 98% confidence limit [24]. Further combined measurements from the experiments at LEP found a best fit value of

$$N_\nu = 2.9840 \pm 0.0082, \tag{1.10}$$

comprehensively ruling out any number of light neutrinos other than 3 [23]. Figure 1.3 shows the experimental data from these experiments, along with the standard model predictions of the cross section for varying types of light neutrinos, making it clear that only three exist.

1.2 The Solar Neutrino Problem

After the initial discovery of the electron neutrino, work began separately to start studying the properties of what were at the time very mysterious particles. In particular, in 1964 Ray Davis and John Bahcall proposed the Homestake experiment, which was designed to test theoretical calculations of the neutrino flux originating from the core of the Sun. The main[5] nuclear fusion process by which energy is generated within the Sun is referred to as the 'solar p–p chain'. Whilst the process can proceed along several different branches of the chain, the first steps are always the same. First, two protons are fused into deuterium,

[5] Approximately 99% of the output energy comes from the solar p–p chains, whilst the remaining 1% comes from the CNO cycle.

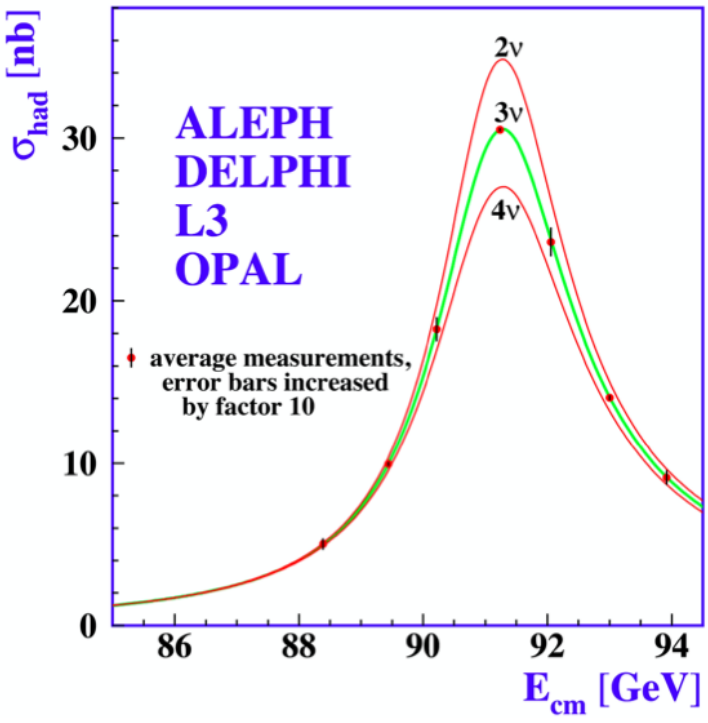


Figure 1.3: $e^+e^- \longrightarrow Z^0 \longrightarrow$ hadrons cross section as a function of centre of mass energy, measured by multiple experiments at LEP. Curves show the standard model predictions of the cross section, for two, three and four sets of light neutrinos. Averaged experimental measurements show a clear preference for only three light neutrino species [22].

which involves the emission of a positron and neutrino:

$$p + p \longrightarrow d + e^+ + \nu_e. \tag{1.11}$$

The deuterium subsequently fuses with another proton creating the ${}^3_2\mathrm{He}$ isotope,

$$d + p \longrightarrow {}^3_2\mathrm{He} + \gamma, \tag{1.12}$$

where the fusion of two of these helium isotopes enables the creation of ${}^4_2\mathrm{He}$. This whole chain involves the creation of two neutrinos, which are required to be ν_e for lepton number conservation. Regardless of which branch of the p–p chain by which the reaction continues, exactly two neutrinos will always be emitted. Due to the fact neutrinos are very weakly interacting, they are then able to escape the core of the Sun with relative ease, carrying a small amount of the energy from the reactions with them.

In order to confirm the understanding of the solar p–p chain model, Bahcall used it to calculate the flux of neutrinos that should arrive at Earth from the Sun, and the Homestake experiment was used to detect them. Homestake used a radiochemical detection method originally put forward by Bruno Pontecorvo [25], whereby a neutrino interacting with an

isotope of chlorine would induce an inverse beta decay reaction, producing an argon atom:

$$\nu_e + {}^{37}\mathrm{Cl} \longrightarrow {}^{37}\mathrm{Ar} + e^-. \tag{1.13}$$

The experiment consisted of a large tank, containing 390 m^3 of tetrachloroethylene (C_2Cl_4), a standard cleaning fluid. As naturally occurring chlorine is composed of roughly 25% ^{37}Cl, this provided sufficient material for between two and seven captures per day. For several months at a time, the tank was left exposed to the neutrino flux from the Sun, allowing a number of ^{37}Ar atoms to accumulate, before it was flushed through with helium in order to remove the argon atoms. At the same time, a known quantity of the stable ^{36}Ar atoms was added, so that the efficiency of the argon removal by helium could be monitored. Following removal, the argon sample was purified and counted, with Ray Davis confident that he could count several dozen ^{37}Ar atoms out of the total $\sim 10^{30}$ in the tank [26]. In fact, Davis was already considerably practiced at this technique, using the same method with reactor antineutrinos in the 1950's to show that neutrinos and antineutrinos were not the same [27]. It is important to note that the reaction in Equation (1.13) requires a neutrino with energy of at least 0.814 MeV to proceed. This in fact means that Homestake was not sensitive to the neutrinos produced in the beginning of the p–p chain, and instead only to those higher energy ones which come from the p–p II and p–p III branches, which involve electron capture on ^{7}Be and beta decay of ^{8}B, respectively. A full description of the branches of the solar p–p chain, the energy of the neutrinos they produce and the capture cross section on ^{37}Cl can be found in [28].

Despite setting out to experimentally confirm the nuclear fusion model of the Sun, Homestake instead detected approximately one third of the number of electron neutrinos that were expected [29]. Despite efforts of theorists to check and refine the solar models, and Homestake's sensitivity improved, the discrepancy persisted, and was later independently verified by additional neutrino experiments. Two of these were the GALLEX and SAGE experiments, which both used gallium as the detector medium, rather than chlorine. The principle of detection still relied on inverse beta decay,

$$\nu_e + {}^{71}\mathrm{Ga} \longrightarrow {}^{71}\mathrm{Ge} + e^-, \tag{1.14}$$

but this interaction has a much lower threshold energy, at only 0.233 MeV [30], meaning it is also sensitive to neutrinos from the first branch of the p–p chain in Equation (1.11).

GALLEX (Gallium Experiment) operated between 1991 and 1997 at the Laboratori Nazionali del Gran Sasso, using a target in the form of a $GaCl_3$ solution, equating to 30.3 tons of gallium. At the end of each solar run (3–4 weeks), the tank was flushed with nitrogen to remove the germanium atoms, which were then counted in a similar manner to that of Homestake. After completing 53 solar runs, GALLEX reported a result of $69.7 \pm 6.7(\mathrm{stat.})^{+3.9}_{-4.5}(\mathrm{syst.})$ SNU, where 1 SNU (solar neutrino unit) is defined as 10^{-36} capture/atom/second [31]. This was found to be approximately 55% of the value predicted by the standard solar model.

SAGE (Soviet-American Gallium Experiment) began operation in 1990, and still con-

tinues to perform monthly extractions to this day. Instead of a gallium solution, the target material used by SAGE is liquid gallium metal, kept just above its melting point of 29.8°C by heaters provided to the chemical reactors. Similarly to GALLEX and Homestake, following extraction by a mix of chemicals the produced germanium atoms were counted, observing a capture rate of $73^{+18}_{-16}(\text{stat.})^{+5}_{-7}(\text{syst.})$ SNU, equating to only 55–60% of the expected value [32].

It was believed by many that the solar neutrino deficit observed by Homestake was likely a mistake, but the emergence of these results, along with additional ones such as that from Kamiokande-II [33], showed similar discrepancies with systematically different approaches. With the experimental results in agreement about a solar neutrino deficit, and additional verification of the standard solar models not finding any errors, the only remaining explanation was *'nu'* physics.

The eventual solution to the solar neutrino problem didn't come until over 30 years after its initial discovery, with the emergence of results from the Super-Kamiokande and SNO collaborations. Super-Kamiokande (described in detail in Section 3.3) measured the ratio of the observed to expected solar neutrino rate to be $0.473^{+0.010}_{-0.009}(\text{stat.})^{+0.017}_{-0.014}(\text{syst.})$ [34]. At the same time, analysis of atmospheric neutrinos (those created from cosmic ray interactions in the upper atmosphere) found a high dependency on zenith angle; the number of neutrinos detected from below the detector and therefore coming through the earth was roughly half of those coming from above [35]. The final evidence for neutrino oscillations came with the results from the Sudbury Neutrino Observatory (SNO), which through its use of heavy water was capable of detecting neutral current neutrino reactions, and was thus sensitive to all flavours of light neutrino. While the number of detected electron neutrinos through charged current interactions showed the now-expected deficit, the number of interactions from all light neutrinos detected via neutral current interactions matched the standard solar model predictions [36]. These results showed that both the neutrino experiments and solar models had been correct, and required the missing neutrinos to have changed flavour during flight, earning both Takaaki Kajita and Arthur McDonald the 2015 Nobel Prize in Physics [37].

1.3 Neutrino Oscillations

The concept of neutrino oscillation was first proposed by Bruno Pontecorvo as early as 1957, who suggested that there could be oscillations between neutrinos and antineutrinos in the same way as those proposed to take place between neutral kaons by Gell-Mann and Pais [38]. Despite this never being observed, Pontecorvo presented the idea that the neutrino and antineutrino were composed of two different mass states [39], which paved the way for the theory of neutrino flavour oscillations by Maki, Nakagawa and Sakata in 1962 [40]. The theory of neutrino oscillations states that the neutrino flavour states (ν_e, ν_μ and ν_τ) which are observed in weak interactions are in fact superpositions of the three individual mass states (ν_1, ν_2 and ν_3) which propagate through time and space. The relation between the neutrino flavour eigenstates $|\nu_\alpha\rangle$ ($\alpha \in \{e, \mu, \tau\}$) and mass eigenstates $|\nu_k\rangle$ ($k \in \{1, 2, 3\}$) can

be expressed as

$$|\nu_\alpha\rangle = \sum_k U^*_{\alpha k} |\nu_k\rangle \,, \tag{1.15}$$

where $U_{\alpha k}$ is a 3×3 unitary mixing matrix that describes the superposition of states, and is known as the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) matrix. The unitarity condition is imposed on the matrix to insure that the probability of a neutrino being in any state is always one, and an individual mass eigenstate can be trivially expressed as a superposition of flavour eigenstates by applying $U_{\alpha k}$. The full form of the PMNS matrix can be expressed as

$$\mathbf{U} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \tag{1.16}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha_1/2} & 0 \\ 0 & 0 & e^{i\alpha_2/2} \end{pmatrix} \tag{1.17}$$

where s_{ij} and c_{ij} represent the sine and cosine of θ_{ij}, the mixing angle of the i^{th} and j^{th} mass states ($i \neq j$) within a given flavour state, and δ represents the Dirac CP (charge-parity) violation phase [41]. The additional Majorana CP violation phases α_1 and α_2 are non-zero only if neutrinos are their own antiparticle. In the case where neutrinos are not Majorana fermions, the final matrix reduces simply to the identity matrix.

The neutrino mass eigenstates are eigenstates of the Hamiltonian,

$$\hat{H} |\nu_k\rangle = E_k |\nu_k\rangle \,, \tag{1.18}$$

where the energy of the eigenstate is given by the eigenvalues, $E_k = \sqrt{p^2 + m_k^2}$. In order to determine the time-evolution of the neutrino mass state, the time-dependent Schrödinger equation is used:

$$|\nu_k(t)\rangle = e^{-iE_k t} |\nu_k\rangle \,, \tag{1.19}$$

and consequently the time-evolution of the neutrino flavour state can be obtained using Equation (1.15), giving

$$|\nu_\alpha(t)\rangle = \sum_k U^*_{\alpha k} e^{-iE_k t} |\nu_k\rangle \,. \tag{1.20}$$

Using the inverse of Equation (1.15), it is then simple to see that at any time after $t = 0$, a

neutrino of flavour α is in fact a superposition of all three flavour states:

$$|\nu_\alpha(t)\rangle = \sum_\beta \sum_k U^*_{\alpha k} e^{-iE_k t} U_{\beta k} |\nu_\beta\rangle . \tag{1.21}$$

In the ultra-relativistic limit ($p \gg m_k$), and treating the neutrino as a single plane wave, the energy of the eigenstate E_k can be approximated to

$$E_k = \sqrt{p^2 + m_k^2} \approx E + \frac{m_k^2}{2E}, \tag{1.22}$$

where $E \approx p$ is the energy of the neutrino in a massless approximation, where in natural units $c = \hbar = 1$. Using this, along with the additional approximation $L \approx t$, the probability that the neutrino of initial flavour α from Equation (1.21) will later be observed as flavour β can be evaluated as

$$P_{\nu_\alpha \to \nu_\beta}(t) = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 \tag{1.23}$$

$$= \left| \sum_k U^*_{\alpha k} U_{\beta k} e^{-iE_k t} \right|^2 \tag{1.24}$$

$$= \sum_k \sum_l U^*_{\alpha k} U_{\beta k} U_{\alpha l} U^*_{\beta l} \exp\left(-i \frac{\Delta m^2_{kl} L}{2E} \right), \tag{1.25}$$

where $\Delta m^2_{kl} = m^2_k - m^2_l$ is the squared mass difference between the mass eigenstates, and propagation is assumed to be occurring within a vacuum. The oscillation probability can further be rewritten as

$$P_{\nu_\alpha \to \nu_\beta}(t) = \delta_{\alpha\beta} - 4 \sum_{k<l} \mathfrak{Re}(U_{\alpha k} U_{\beta l} U^*_{\alpha l} U^*_{\beta l}) \sin^2 \left(\frac{\Delta m^2_{kl} L}{2E} \right) \tag{1.26}$$

$$\pm 2 \sum_{k<l} \mathfrak{Im}(U_{\alpha k} U_{\beta l} U^*_{\alpha l} U^*_{\beta l}) \sin^2 \left(\frac{\Delta m^2_{kl} L}{2E} \right), \tag{1.27}$$

which splits the probability into two separate terms, where the real part is CP conserving and the imaginary part is CP violating [42]. The sign of the imaginary term is positive (negative) for neutrinos (antineutrinos), representing an enhancement (diminution) of the oscillation probability. It is clear that in order for the phenomenon of neutrino oscillations to be present, neutrinos are required to have non-zero mass.

In the simplified 2-flavour neutrino regime, where the mixing matrix is given simply by

$$\mathbf{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{1.28}$$

with only a single mixing angle θ, the oscillation probability in Equation (1.24), which will describe a the probability neutrino of initial flavour ν_μ to be later observed as ν_e, reduces

to:

$$P(\nu_\mu \to \nu_e) = \left| U^*_{\mu 1} U_{e1} + U^*_{\mu 2} U_{e2} \exp\left(-i\frac{\Delta m^2_{21} L}{2E}\right)\right|^2 \tag{1.29}$$

$$= \left|\cos\theta \sin\theta - \sin\theta \cos\theta \exp\left(-i\frac{\Delta m^2_{21} L}{2E}\right)\right|^2 \tag{1.30}$$

$$= \sin^2 2\theta \sin^2\left(\frac{\Delta m^2_{21} L}{4E}\right). \tag{1.31}$$

From this, it can be seen that there is a fixed value of $\frac{L}{E}$ for which the neutrino oscillation probability will be maximal. Long baseline neutrino oscillation experiments often fix their neutrino energy and baseline, so that the detector lies within the oscillation maximum, providing the greatest sensitivity to oscillation effects.

1.4 Matter Oscillations and the MSW Effect

The formalism of neutrino oscillations described in Section 1.3 is for those occurring within a vacuum. As neutrinos propagate through real matter however, the interactions that they can undergo in flight must be taken into account. Whilst all flavours of neutrino can participate in neutral current elastic forward scattering $\nu_\ell + X \longrightarrow \nu_\ell + X$, for electron neutrinos scattering off electrons, $\nu_e + e^- \longrightarrow \nu_e + e^-$, there is an additional W exchange process available. As electrons are naturally abundant in regular matter, whereas muons and taus are not, the addition of the charged current forward scattering interactions causes a difference in the potentials, defined as

$$V = V_e - V_a = \sqrt{2} G_F n_e, \tag{1.32}$$

where V_e and V_a are the potentials for electron and non-electron flavour neutrinos respectively, G_F is the Fermi coupling constant, and n_e is the number density of electrons in the matter being traversed. This leads to an additional contribution to the interaction Hamiltonian, which is modified from $\hat{H}_0$ as

$$\hat{H}_0 \to \hat{H} = \hat{H}_0 + \hat{V}, \tag{1.33}$$

and thus also modifies the mass eigenstates and mixing angle, which now become dependent on the electron density in the medium, so that the latter can be expressed as a function of the vacuum mixing angle θ as

$$\tan 2\theta_m = \frac{\sin 2\theta}{\cos 2\theta - \sqrt{2} G_F n_e \frac{E}{\Delta m^2}}. \tag{1.34}$$

It is clear from Equation (1.34) that the effective mixing parameter will exhibit a resonance quality at the point that $\cos 2\theta = \sqrt{2} G_F n_e \frac{E}{\Delta m^2}$, corresponding to an angle of $\theta_m = \frac{\pi}{4}$, where the depth of oscillations becomes maximal [43]. It can also be seen that matter effects on neutrino oscillations introduce sensitivity to the *sign* of the mass splitting, Δm^2,

where previously only the magnitude mattered. It is through matter effects that future long baseline experiments hope to be able to determine the ordering of the three neutrino mass states, otherwise known as the mass hierarchy, which still remains unknown.

It is also interesting to consider what happens when neutrinos, such as those produced in the p–p chain reactions within the Sun, traverse a medium with changing density ($n_e = n_e(t)$). At the very high density in the centre of the Sun where the neutrinos are produced, n_e becomes large and thus $\tan 2\theta_m \to 0 \implies \theta_m \to \frac{\pi}{2}$. In this limit the neutrino mass states, which are represented as

$$|\nu_{1m}\rangle = \cos\theta_m |\nu_e\rangle + \sin\theta_m |\nu_a\rangle , \tag{1.35}$$

$$|\nu_{2m}\rangle = -\sin\theta_m |\nu_e\rangle + \cos\theta_m |\nu_a\rangle , \tag{1.36}$$

reduce simply to

$$|\nu_{1m}\rangle \approx |\nu_a\rangle , \tag{1.37}$$

$$|\nu_{2m}\rangle \approx -|\nu_e\rangle , \tag{1.38}$$

and so electron neutrinos are comprised almost solely of $|\nu_{2m}\rangle$. If the density of the medium changes slowly enough that it may be considered adiabatic, as is the case in the Sun, then subsequent transitions between $\nu_{1m} \leftrightarrow \nu_{2m}$ can be neglected. As the $|\nu_{2m}\rangle$ state propagates, the mixing angle $\theta_m(t)$ changes, reaching a point at the resonance where the ν_e and ν_a components of $|\nu_{2m}\rangle$ are equal. Once the neutrinos escape the Sun and the density is low, the matter eigenstates tend back to the vacuum eigenstates. It follows then that the survival probability of an electron neutrino produced in the Sun is

$$P(\nu_e \to \nu_e) = |\langle \nu_e | \nu_2 \rangle|^2 \approx \sin^2\theta, \tag{1.39}$$

which is a result of non-oscillatory transition [44]. This is known as the Mikheyev-Smirnov-Wolfenstein [45, 46] effect, or MSW effect. The value of Equation (1.39) for θ_{12} evaluates to only 0.03 away from the ratio of charged current to neutral current neutrino interactions observed by SNO,

$$\frac{\Phi_{\mathrm{CC}}}{\Phi_{\mathrm{NC}}} = 0.340 \pm 0.023^{+0.029}_{-0.031}, \tag{1.40}$$

where the difference is due to small oscillation effects in the Sun and regeneration of ν_e caused by matter effects in the Earth [44].

1.5 Mass Hierarchy

Throughout the preceding sections, many references are made to the three neutrino mass states, along with the squared mass splittings between them, Δm^2. Both the value and sign of the splitting between the first and second mass states, Δm^2_{12}, have been known for some time, with a current best fit value of $\Delta m^2_{21} = 7.42^{+0.21}_{-0.20} \times 10^{-5}$ eV^2 [47]. However, the mass

of the third mass eigenstate in relation to the first two still remains unclear. The ordering of the mass states is known as the mass hierarchy, where there are two possible scenarios. The first, known as 'Normal Hierarchy' (NH), is where the mass of the eigenstates increases with number – $\nu_1 < \nu_2 < \nu_3$. The other scenario is known as the 'Inverted Hierarchy' (IH), where the third mass state is the lowest in mass – $\nu_3 < \nu_1 < \nu_2$. This is shown diagrammatically in Figure 1.4. In order to gain sensitivity to the mass hierarchy, future

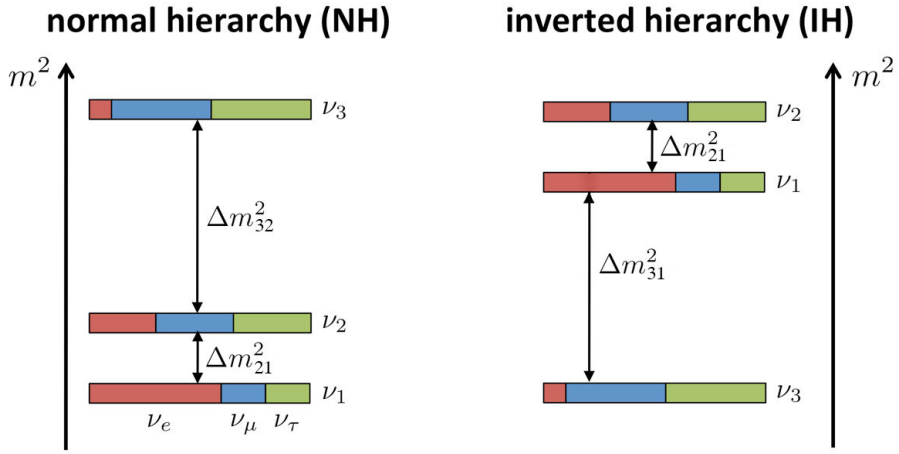


Figure 1.4: Diagram showing the two possible neutrino mass hierarchies: normal ($\nu_1 < \nu_2 < \nu_3$) and inverted ($\nu_3 < \nu_1 < \nu_2$). Figure adapted from [48].

long baseline experiments aim to probe matter effects, which as shown in Equation (1.34) have some dependence on the sign of the mass splitting. In addition to determining the neutrino mass hierarchy for the sake of building a full understanding of neutrino properties, other experiments rely on this knowledge. Measurements of δ_{CP}, the Dirac CP violation phase which describes if, and to what extent, neutrinos and antineutrinos oscillate differently, rely on deviations from expected values. Therefore, the effect of the mass hierarchy and δ_{CP} in long baseline experiments must be disentangled in order to gain full sensitivity. The mass hierarchy also has important implications on measurements of the absolute neutrino mass, along with determining whether neutrinos are Majorana fermions. A comprehensive overview of the methods for gaining sensitivity to the hierarchy, along with current and future experiments that employ these, can be found in [49].

1.6 State of the ν-tion

In the nearly two decades since the confirmation of neutrino oscillations by Super-Kamiokande and SNO, knowledge of neutrino oscillation parameters has progressed rapidly, with many experiments dedicated to their understanding. The currently running NOνA [50] and T2K [51] experiments are long baseline accelerator oscillation experiments, which study the disappearance of muon neutrinos and appearance of electron neutrinos in a muon neutrino beam. This gives them sensitivity to the two mixing angles θ_{23} and θ_{13}, along with the

mass splitting Δm^2_{31} and CP violation phase, δ_{CP}. To a lesser extent the older experiments MINOS [52] and K2K [53] also provided some sensitivity to the atmospheric parameters, and are still used in global data fits [54]. In 2020, T2K became the first experiment to find evidence of a non-conserving value of δ_{CP}, with a 3σ confidence interval of $[-3.41, -0.03]$ for NH and $[-2.54, -0.32]$ for IH [55]. Although tension between the NOνA and T2K results was presented at the NEUTRINO 2020 conference [50], fits to updated results have found a reduction in the tension, which can be seen in Figure 1.5. Looking towards future

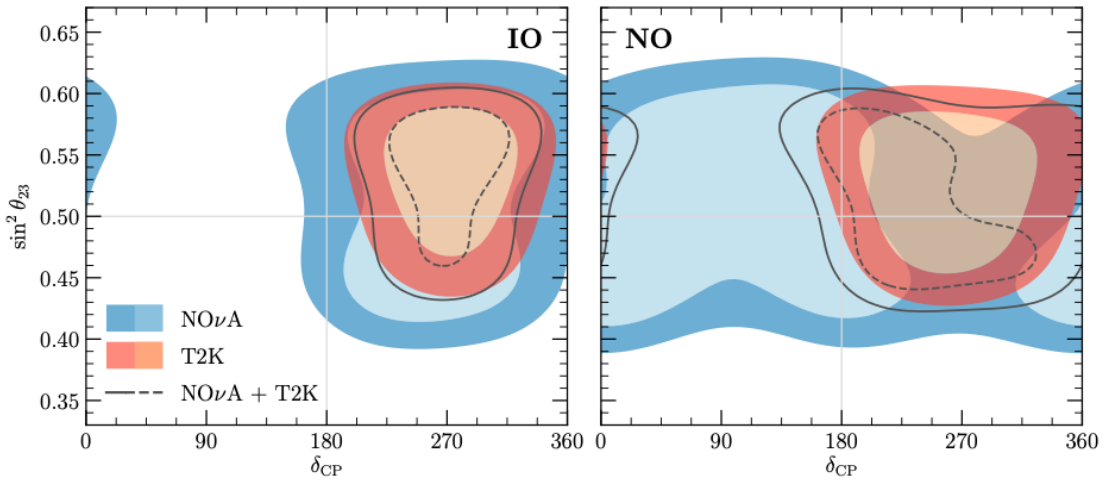


Figure 1.5: 1σ and 2σ allowed regions for T2K (red) and NOνA (blue) results, whilst black curves show the combination of results. Results and figure from NuFIT 5.0 [47]. Contour regions are shown for inverted hierarchy (left) and normal hierarchy (right).

accelerator experiments, Hyper-Kamiokande [56] and DUNE [57] aim to measure the value of δ_{CP} by greatly reducing current statistical uncertainties, whilst the significant matter effects that DUNE will see over its 1300 km baseline will also make it particularly sensitive to the mass hierarchy.

The reactor sector of neutrino experiments can be split up into two sections: short baseline ($\sim$1 km), of which the main experiments are Daya Bay [58], RENO [59] and Double Chooz [60]; and the long baseline reactor experiment KamLAND [61]. The short baseline reactor experiments are particularly sensitive to θ_{13} and Δm^2_{31}, where the former was the last of the three mixing angles to be experimentally measured – latest results from Daya Bay and RENO put the value of $\sin^2 2\theta_{13}$ at 0.0856 ± 0.0029 [62] and 0.0892 ± 0.0063 [63], respectively. This is a particularly important result, as the generally used formalism of the PMNS matrix (Equation (1.17)) couples $\sin\theta_{13}$ with δ_{CP}, requiring a non-zero mixing angle to be able to determine the CP violation phase in this case. While the short baseline reactor experiments have also measured values of Δm^2_{31} (Daya Bay made an independent measurement yielding a result of $(2.42 \pm 0.11) \times 10^{-3}$ eV^2 [64]), future experiments such as JUNO [65] and RENO-50 [66] hope to be able to determine the mass hierarchy with high-precision measurements [67].

The KamLAND experiment, situated in the old KamiokaNDE chamber in the Mozumi drift mine, observes 53 commercial nuclear reactors, for an averaged baseline of ∼180 km. Despite being a reactor experiment, KamLAND's long baseline gives it the most sensitivity to the solar parameters, and in particular the mass splitting Δm^2_{21}, measured to be $7.50^{+0.19}_{-0.20} \times 10^{-5}$ eV^2 [61]. However, a recent global assessment of neutrino oscillation parameters [54] finds this in some contention with combined results from the solar neutrino experiments GALLEX/GNO [68], SAGE [69], Homestake [70], Borexino [71], Super-Kamiokande [72] and SNO [73], which have a best fit value of $\Delta m^2_{21} = 4.8 \times 10^{-5}$ eV^2 (although there is overlap in the 90% confidence limits). However these experiments provide a much tighter constraint on the mixing angle $\sin^2\theta_{12}$, with a current best fit value of $0.304^{+0.013}_{-0.012}$ [47].

The final set of oscillation experiments are those which, alongside the long baseline accelerator experiments, focus on measuring the atmospheric parameters $\left|\Delta m^2_{3\ell}\right|$ and θ_{23}. The major atmospheric neutrino experiments are Super-Kamiokande, IceCube [74] and ANTARES [75]. Comparison of recent results between Super-Kamiokande and IceCube finds some contention, with Super-Kamiokande slightly preferring a best-fit angle in the lower octant of θ_{23} ($\sin^2\theta_{23} < 0.5$) [72], and IceCube preferring a best-fit value in the upper octant ($\sin^2\theta_{23} > 0.5$) [76]. The obtained values of $\left|\Delta m^2_{3\ell}\right|$, however, are mostly in agreement. The future atmospheric detectors, such as Hyper-Kamiokande [56], PINGU [77] and KM3NeT-ORCA [78] seek to determine the neutrino mass hierarchy, which should be possible with enough exposure regardless of δ_{CP}, although will be significantly aided by combination with future accelerator measurements.

A comprehensive overview of the global best-fit oscillation parameters, from the 2020 NuFIT 5.0 analysis [47], is presented in Table 1.I.

1.7 Thesis Outline

This chapter has introduced the long, exciting history of neutrino physics, along with presenting an overview of the mechanism by which neutrinos oscillate and covering the current state of the field itself. The structure of the rest of this thesis is as follows:

- Chapter 2 covers the theory of neutrino-nucleus interactions, the difficulty in measuring them, and presents some recent results in the specific field of neutrino interactions.
- Chapter 3 describes the T2K experiment itself, from the beam production to the far detector, with particular emphasis placed on the ND280 near detector used in this thesis.
- Chapter 4 details the work performed to understand the source of errors in time of flight information between ND280 subdetectors and how often this has an effect on recorded data, before evaluating a systematic uncertainty to cover these errors. The initial treatment developed by collaboration members is described, whilst the new treatment presented is the work of the author, following discussion with collaborators.
- Chapter 5 gives an overview of the cross-section analysis strategy that will be used in the measurement which is the main topic of this thesis.

		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 2.7$)	
		bfp $\pm1\sigma$	3σ range	bfp $\pm1\sigma$	3σ range
without SK atmospheric data	$\sin^2\theta_{12}$	$0.304^{+0.013}_{-0.012}$	$0.269 \to 0.343$	$0.304^{+0.013}_{-0.012}$	$0.269 \to 0.343$
	$\theta_{12}/^\circ$	$33.44^{+0.78}_{-0.75}$	$31.27 \to 35.86$	$33.45^{+0.78}_{-0.75}$	$31.27 \to 35.87$
	$\sin^2\theta_{23}$	$0.570^{+0.018}_{-0.024}$	$0.407 \to 0.618$	$0.575^{+0.017}_{-0.021}$	$0.411 \to 0.621$
	$\theta_{23}/^\circ$	$49.0^{+1.1}_{-1.4}$	$39.6 \to 51.8$	$49.3^{+1.0}_{-1.2}$	$39.9 \to 52.0$
	$\sin^2\theta_{13}$	$0.02221^{+0.00068}_{-0.00062}$	$0.02034 \to 0.02430$	$0.02240^{+0.00062}_{-0.00062}$	$0.02053 \to 0.02436$
	$\theta_{13}/^\circ$	$8.57^{+0.13}_{-0.12}$	$8.20 \to 8.97$	$8.61^{+0.12}_{-0.12}$	$8.24 \to 8.98$
	$\delta_{\rm CP}/^\circ$	195^{+51}_{-25}	$107 \to 403$	286^{+27}_{-32}	$192 \to 360$
	$\frac{\Delta m^2_{21}}{10^{-5}\ \text{eV}^2}$	$7.42^{+0.21}_{-0.20}$	$6.82 \to 8.04$	$7.42^{+0.21}_{-0.20}$	$6.82 \to 8.04$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ \text{eV}^2}$	$+2.514^{+0.028}_{-0.027}$	$+2.431 \to +2.598$	$-2.497^{+0.028}_{-0.028}$	$-2.583 \to -2.412$
		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 7.1$)	
		bfp $\pm1\sigma$	3σ range	bfp $\pm1\sigma$	3σ range
with SK atmospheric data	$\sin^2\theta_{12}$	$0.304^{+0.012}_{-0.012}$	$0.269 \to 0.343$	$0.304^{+0.013}_{-0.012}$	$0.269 \to 0.343$
	$\theta_{12}/^\circ$	$33.44^{+0.77}_{-0.74}$	$31.27 \to 35.86$	$33.45^{+0.78}_{-0.75}$	$31.27 \to 35.87$
	$\sin^2\theta_{23}$	$0.573^{+0.016}_{-0.020}$	$0.415 \to 0.616$	$0.575^{+0.016}_{-0.019}$	$0.419 \to 0.617$
	$\theta_{23}/^\circ$	$49.2^{+0.9}_{-1.2}$	$40.1 \to 51.7$	$49.3^{+0.9}_{-1.1}$	$40.3 \to 51.8$
	$\sin^2\theta_{13}$	$0.02219^{+0.00062}_{-0.00063}$	$0.02032 \to 0.02410$	$0.02238^{+0.00063}_{-0.00062}$	$0.02052 \to 0.02428$
	$\theta_{13}/^\circ$	$8.57^{+0.12}_{-0.12}$	$8.20 \to 8.93$	$8.60^{+0.12}_{-0.12}$	$8.24 \to 8.96$
	$\delta_{\rm CP}/^\circ$	197^{+27}_{-24}	$120 \to 369$	282^{+26}_{-30}	$193 \to 352$
	$\frac{\Delta m^2_{21}}{10^{-5}\ \text{eV}^2}$	$7.42^{+0.21}_{-0.20}$	$6.82 \to 8.04$	$7.42^{+0.21}_{-0.20}$	$6.82 \to 8.04$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ \text{eV}^2}$	$+2.517^{+0.026}_{-0.028}$	$+2.435 \to +2.598$	$-2.498^{+0.028}_{-0.028}$	$-2.581 \to -2.414$

Table 1.I: Recent neutrino parameter global best-fit values, from NuFIT 5.0 [47].

- Chapter 6 details the development of the CC1π^+ selection for signal and background samples, with a particular focus on work performed to measure pion kinematics in previously inaccessible regions of interest. Section 6.1 describes the previously developed selection, whilst the following sections describing the improved selection and the implementation of reconstruction for pions detected from Michel electrons is the work of the author.

- Chapter 7 presents an overview of the systematic uncertainties associated with the selection, along with efficiency studies and the binning schemes used for the input signal samples and final unfolded cross-section result. The systematic treatments described are the work of many collaborators over many years, and the author's work was to evaluate them for the analysis presented. The efficiency and binning studies are all the work of the author.

- Chapter 8 gives an overview of validation tests performed on the cross-section fitter, in order to check its stability before proceeding to unblind to data. While the fitting

framework was pre-existing, this was heavily modified by the author for this analysis, and all studies presented are their own work.

- Chapter 9 presents the final measured cross-section results, with some concluding remarks.

Chapter 2

Neutrino Interactions

Chapter 1 provides an overview of the current main questions surrounding the field of neutrinos, specifically relating to determination of the oscillation parameters. Whilst oscillation results are, by definition, the main focus of oscillation experiments, there is large amount of further understanding required to reach these results. In a general oscillation experiment, the rate of interactions observed after oscillation can be expressed as

$$R(\vec{x}) = \Phi(E_\nu) \times \sigma(E_\nu, \vec{x}) \times \epsilon(\vec{x}) \times P(\nu_\alpha \rightarrow \nu_\beta), \tag{2.1}$$

where $\Phi(E_\nu)$ is the initial neutrino flux as a function of neutrino energy E_ν, $\sigma(E_\nu, \vec{x})$ is the neutrino interaction cross section for ν_β, which is a function of both the neutrino energy and the final state reconstructed kinematics $\vec{x}$, ϵ is the detector efficiency and $P(\nu_\alpha \rightarrow \nu_\beta)$ is the oscillation probability. It is clear from this that precise determination of the oscillation probability relies on good knowledge of the other underlying variables that determine the rate.

Figure 2.1 shows the total charged current neutrino interaction cross section as a function of energy for the intermediate neutrino energy regime, which is relevant for long baseline accelerator experiments. The black curve gives the total predicted cross section, whilst the coloured ones give the predicted cross section for three of the underlying channels: quasi-elastic scattering (red), resonant interactions (blue) and deep inelastic scattering (green). It can be seen that across this energy range the total cross section is formed of multiple processes, each of which has to be well understood to predict the overall cross section. The neutrino beam used by the T2K experiment for oscillation analyses peaks in flux at an energy of 600 MeV; although at this value the cross section is dominated by Charged Current Quasi-Elastic interactions, there is a large contribution from resonant interactions, and deep inelastic scattering events are also beginning to turn on. This chapter gives an overview of how different neutrino-nucleon interaction channels such as these are modelled, and why making the jump to neutrino-nucleus interactions is non-trivial.

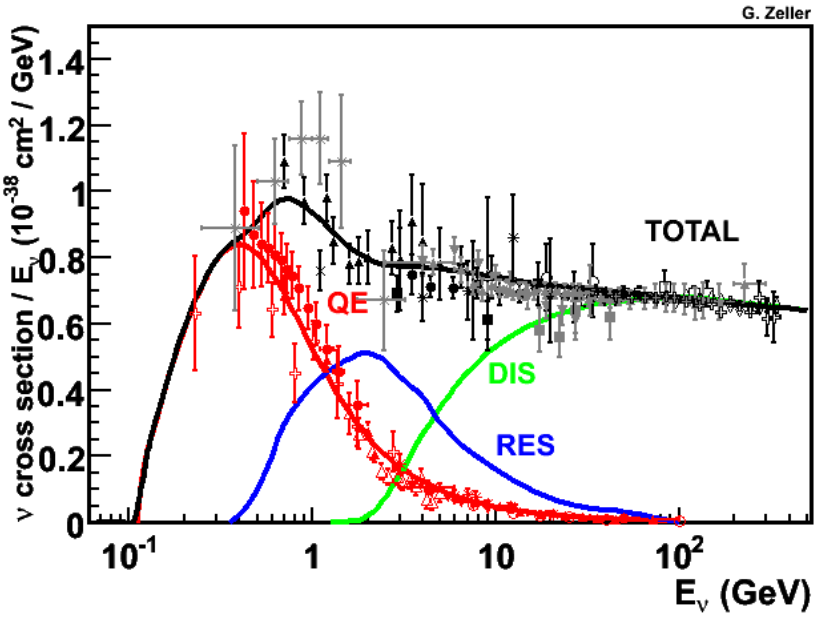


Figure 2.1: Neutrino charged current interaction cross section as a function of energy, with theoretical model predictions for underlying processes [79].

2.1 Charged Current Quasi-Elastic Scattering

At the lower range of energies spanned by accelerator neutrino beams, the interaction of a neutrino and a nucleon results in the neutrino scattering off the nucleon, which remains intact. The charged current version of this process proceeds via the exchange of a $W^{\pm}$ boson, and thus requires the neutrino to have enough energy that its respective leptonic partner can be produced:

$$\nu_\ell + n \longrightarrow \ell^- + p, \tag{2.2}$$

$$\bar{\nu}_\ell + p \longrightarrow \ell^+ + n. \tag{2.3}$$

This interaction channel is known as Charged Current *Quasi*-Elastic scattering (CCQE), as the change from the initial baryon to its isospin doublet partner means the outgoing state is not exactly the same as the initial. CCQE interactions are particularly important in neutrino physics, as the two-body interaction allows the kinematics of the event to be fully reconstructed, and thus the initial neutrino energy to be calculated. The Feynman diagram for such an interaction is given in Figure 2.2.

To calculate the CCQE neutrino-nucleon scattering differential cross section, the regular approach of writing down the scattering amplitudes following the Feynman rules cannot be used, due to the fact the the nucleon is an extended object. Instead, the Llewellyn-Smith formalism is used [80]. This is given by

$$\frac{d\sigma}{dQ^2} = \frac{M_N^2 G_F^2 \cos^2\theta_c}{8\pi E_\nu^2}\left[A(Q^2) \pm \frac{B(Q^2)(s-u)}{M_N^2} + \frac{C(Q^2)(s-u)^2}{M_N^4}\right], \tag{2.4}$$

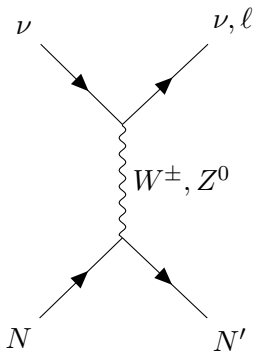


Figure 2.2: Feynman diagram for elastic scattering of a neutrino from a nucleon.

where M_N is the nucleon mass, G_F is the Fermi coupling constant, θ_c is the Cabbibo angle and E_ν is the neutrino energy. $(-)+$ represents the differential cross section for (anti)neutrinos, and the difference between the Mandelstam variables s and u equates to $s - u = 4M_N E_\nu - Q^2 - m_l^2$, where m_l is the charged lepton mass and Q^2 the squared momentum transfer ($Q^2 = -q^2$). The factors A, B and C are given by the relations

$$A(Q^2) = \frac{(m_\ell^2 + Q^2)}{M_N^2}\Bigg\{ \left[(1+\tau)F_A^2 - (1-\tau)F_1^2 + \tau(1-\tau)F_2^2 + 4\tau F_1 F_2\right] - \frac{m_\ell^2}{4M_N^2}\left[(F_1+F_2)^2 + (F_A+2F_P)^2 - 4(1+\tau)F_P^2\right]\Bigg\}, \tag{2.5}$$

$$B(Q^2) = \frac{Q^2}{M_N^2} F_A(F_1+F_2), \tag{2.6}$$

$$C(Q^2) = \frac{1}{4}\left[F_A^2 + F_1^2 + \tau F_2^2\right], \tag{2.7}$$

where $\tau = Q^2/4M_N^2$. In the above three equations, the four form factors F_1, F_2, F_P and F_A are used to parametrise the spatial distribution of the nucleus. F_1 and F_2 are the two vector form factors of the weak interaction, which can be reliably measured in electron scattering experiments by employing the Conserved Vector Current (CVC) hypothesis [79]. This is also done in the modelling of the form factors, which the BBBA05 [81] model expresses in terms of the electric and magnetic form factors of the proton and neutron.

Whilst often neglected in earlier neutrino scattering experiments due to appearing in Equation (2.5) multiplied by $\frac{m_\ell^2}{M_N^2}$ [82], the pseudoscalar form factor can be obtained by invoking the Partial Conservation of Axial Current (PCAC) hypothesis. This results in the Goldberger-Treiman relation [83], which links the pseudoscalar form factor F_P to the axial-vector form factor by

$$F_P(Q^2) = \frac{2M_N}{Q^2 + m_\pi^2} F_A(Q^2), \tag{2.8}$$

where m_π is the mass of the pion.

The axial-vector form factor F_A is assumed to take a dipole form,

$$F_A(Q^2) = \frac{g_A}{\left(1 + \frac{Q^2}{M_A^2}\right)^2}, \tag{2.9}$$

where $g_A = F_A(Q^2 = 0)$ can be accurately determined from β decay [84], and M_A is known as the axial mass. M_A can be determined in two possible ways. The first is from measurements of (quasi)elastic scattering of neutrinos and antineutrinos off protons or nuclei, whilst the second involves fitting for the Q^2 dependence of charged pion electroproduction off protons [85]. These measurements were a large focus between the 1960s and 1990s, resulting in world averages [85] of

$$M_A = (1.026 \pm 0.021)\ \text{GeV (neutrino scattering)}, \tag{2.10}$$

$$M_A = (1.069 \pm 0.016)\ \text{GeV (electroproduction)}. \tag{2.11}$$

Despite the large number of experiments in this period finding a value of M_A around 1 GeV, more recent neutrino scattering measurements have proven contentious. Low-energy ν_μ scattering from nucleons bound in carbon (MiniBooNE) and oxygen (K2K, T2K) find a CCQE cross-section between 20–30% larger than is predicted [86]. This is thought to be due to additional nuclear effects (see Section 2.5), which can be parameterised by a larger effective axial mass, on the order of 1.3 GeV (MiniBooNE found $M_A^{\text{eff}} = 1.35 \pm 0.17$ GeV [87]). Figure 2.3 shows cross-section measurements obtained by MiniBooNE, along with predictions from theoretical models using differing values of M_A. This disagreement in values of M_A at low and high values of Q^2 is referred to as the MiniBooNE large axial mass anomaly, and also led theorists to consider non-dipole forms of the form factor, such as the z-expansion parametrisation [89].

2.2 Neutral Current Elastic Scattering

Neutral current (NC) interactions of a neutrino with a nucleon proceed in much the same way as the charged current interaction, but with the interaction mediated by the neutral Z^0 boson. At low energies, these interactions can be considered truly elastic, as the initial and final states are identical:

$$\nu_\ell + N \longrightarrow \nu_\ell + N. \tag{2.12}$$

To model the cross section of NC elastic scattering interactions, the Llewellyn-Smith formalism is again used, albeit with some modification. The nuclear form factors present in Equations (2.5) to (2.7) are altered in order to introduce additional coupling factors and a contribution from strange quarks. In particular, the axial-vector form factor changes to

$$F_A(Q^2) = \frac{g_A \tau_3}{2\left(1 + \frac{Q^2}{M_A^2}\right)^2} - \frac{F_A^s(Q^2)}{2}. \tag{2.13}$$

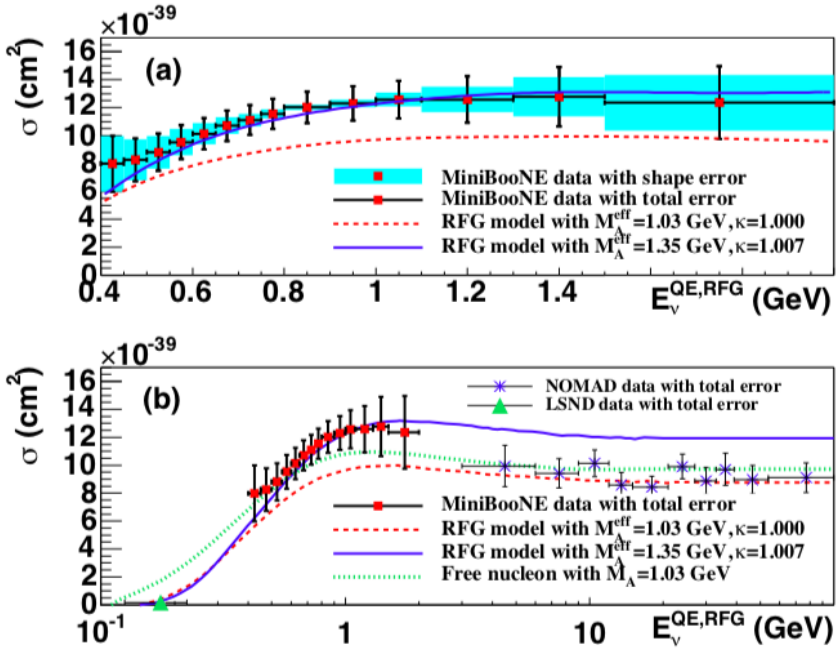


Figure 2.3: MiniBooNE ν_μ CCQE cross section per neutron as a function of energy. (a) shows the MiniBooNE measurements along with predictions from a `NUANCE` simulation using different parameter values. (b) extends the energy range to include data from NOMAD, as well as predictions for scattering from free nucleons [88].

where $\tau_3 = +1(-1)$ for proton (neutron) scattering, and the strange axial-vector form factor is given as

$$F_A^s(Q^2) = \frac{\Delta s}{\left(1 + \frac{Q^2}{M_A^2}\right)^2}, \tag{2.14}$$

where Δs is the contribution of the strange quark to nuclear spin [79].

2.3 Resonant Pion Production

As the energy imparted to the struck nucleon by the incoming neutrino rises, it becomes possible that the nucleon will be excited into a resonant state. The first of these resonances which is accessible is the $\Delta(1232)$ resonant state, but in the 1–2 GeV region where resonant pion production is particularly important there are 18 such resonances which are generally taken into account [90]. After the resonant state is produced it swiftly decays again, still within the nuclear medium, to produce a pion and nucleon. For CC interactions, the most

common decay chains available are

$$\nu_\mu + p \longrightarrow \mu^- + p + \pi^+ \qquad \bar{\nu}_\mu + n \longrightarrow \mu^+ + n + \pi^-$$
$$\nu_\mu + n \longrightarrow \mu^- + p + \pi^0 \qquad \bar{\nu}_\mu + p \longrightarrow \mu^+ + n + \pi^0$$
$$\nu_\mu + n \longrightarrow \mu^- + n + \pi^+ \qquad \bar{\nu}_\mu + p \longrightarrow \mu^+ + p + \pi^-$$

where the left column is given for muon neutrino interactions, and the right for muon antineutrinos. An example Feynman diagram for one of these typical resonant pion production interactions is given in Figure 2.4.

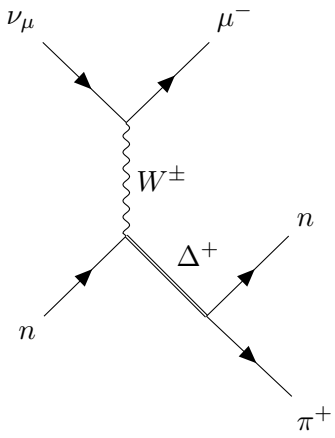


Figure 2.4: Feynman diagram of one possible resonant π^+ production channel through the $\Delta(1232)$ resonance.

The majority of generators model resonant pion production using the Rein-Sehgal model [91], which describes resonant interactions in the region up to $W = 2$ GeV using the Feynman-Kislinger-Ravndal relativistic quark model [92]. The Rein-Sehgal model expresses the cross section for a resonance of mass W as

$$\frac{d\sigma}{dq^2 dv} = \frac{G_F^2}{4\pi^2}\left(\frac{-q^2}{Q^2}\right)\frac{W^2 - M_N^2}{2M_N}\left(u^2\sigma_+ + v^2\sigma_- + 2uv\sigma_0\right), \tag{2.15}$$

where σ are the partial cross sections for left-handed $(+)$, right-handed $(-)$ or longitudinal (0) propagator polarisations, and

$$u = \frac{E + E' + Q}{2E} \qquad v = \frac{E + E' - Q}{2E}. \tag{2.16}$$

This model also introduces nuclear form factors in order to parametrise the spatial distribution of the nucleus, which are required to extend the cross-section calculation to nuclear resonances below 1.75 GeV. Similarly to in the Llewellyn-Smith formalism, the vector form factors can be measured with relative ease. Of the six axial-vector form factors, the non-vanishing ones can be expressed as a function of the final remaining Graczyk-Sobczyk form-factor, known as C_5^A [93]. This must be determined through fits to neutrino experimental data, along with the resonant axial mass normalisation parameter M_A^{RES}, and the normalisation constant $I_{\frac{1}{2}}$ used to describe the non-resonant background.

Though used in many neutrino interaction generators, the Rein-Sehgal model has several drawbacks. As standard, the lepton is assumed to be massless (although some generators such as NEUT [94] implement lepton mass using the later published Berger-Sehgal model [95]), and the helicity amplitudes are not a function of pion angle. Multiple generator implementations also only include the Δ resonance, as opposed to all 18 that should be included. An updated treatment is the MK model [90], which along with non-zero lepton mass effects, incorporates the non-resonant interactions contributing to single pion production and the related interference terms, and predicts the cross section as a function of outgoing pion angle. When implemented into the NEUT generator, the MK model shows a better fit to data from bubble chambers than the Rein-Sehgal model does.

2.4 Deep Inelastic Scattering

All of the interactions described up until this point involve the neutrino in question interacting with a nucleon as a whole. At higher energies however, the weak probe has enough energy for it to resolve individual quarks, breaking up the nucleon as it does so. This is referred to as Deep Inelastic Scattering (DIS). While it can be seen from Figure 2.1 that the neutrino cross section is not dominated by DIS interactions until neutrino energies reach the order of 10 GeV, the turn on for DIS events occurs at approximately 1 GeV, which is of particular relevance for accelerator neutrino experiments. While the higher energy DIS region is fairly well understood, the low-energy transition region between resonant interactions and DIS (often referred to as Shallow Inelastic Scattering or non-resonant multi pion production), which is particularly relevant for the T2K energy range, is not [96].

Due to the depth and complexity of this topic, full details are not presented here. Instead, a thorough description can be found in the review of Conrad, Shaevitz and Bolton [97]. The predicted cross section is expressed in terms of nuclear structure functions, which in turn are broken down into parton distribution functions in order to describe the interaction of a nucleon with a single quark.

2.5 Neutrino-Nucleus Interactions

The aforementioned interaction channels have been discussed with the assumption that the neutrino interaction takes place with an individual nucleon (or quarks in the case of DIS). For the early scattering experiments which were performed on hydrogen or deuterium, this assumption is valid, as the nucleons are effectively free. In recent years however, neutrino experiments have moved to using heavier targets such as carbon, oxygen and argon, in order to increase the rate of interactions. In doing so, we move from modelling interactions of neutrinos with simple, quasi-free nucleons to nuclei, which are a bound state of multiple nucleons.

2.5.1 Modelling the Nucleus

In an attempt to deal with the extension of cross-section predictions on free nucleons to entire nuclei, many neutrino interaction generators make use of the Impulse Approximation [98]. The Impulse Approximation makes two assumptions. The first is that the incoming probe imparts high enough momentum $|\vec{q}|$ that it can resolve the individual nucleons within the nucleus, which can be treated as independent. The second is that the particles that are produced at the interaction vertex subsequently evolve independently of one another, with the $(A-1)$-nuclear system effectively treated as a spectator. This neglects any final state interactions or statistical corrections. The cross section can then be expressed as a sum of the individual processes for each nucleon.

Whilst this approximation drastically simplifies things, the Pauli exclusion principle makes it clear that nucleons within a nucleus cannot exist in identical states, and it therefore makes no sense to treat each individual nucleon as at rest. Instead, models known as spectral functions are used, which describe the probability that a nucleon will have a certain initial momentum and associated nuclear binding energy required to liberate it from the nucleus.

One of the most simplistic and widely used spectral functions in neutrino event generators is the Relativistic Fermi Gas (RFG) model, developed by Smith and Moniz [99]. The RFG model describes the nucleus as an interaction-free system of multiple plane wave nucleons, where all possible momentum states are filled up to the Fermi momentum p_F, beyond which no states exist. The probability for a state to exist with momentum $\vec{p}$ and energy E is defined as

$$P_{\mathrm{RFG}}(|\vec{p}|, E) = \frac{6\pi^2 A}{p_F^2}\Theta(p_F - |\vec{p}|), \tag{2.17}$$

where A is the atomic number of the nucleus, and Θ is the Heaviside step function. As all states up to the Fermi momentum are occupied, outgoing nucleons must have momentum greater than p_F, otherwise they are subject to Pauli blocking. Due to this hard cut-off, the effect of Pauli blocking by the model is often over-exaggerated.

A relatively simple extension of this model is the Local Fermi Gas (LFG) model, which instead makes the Fermi momentum radially dependent, by expressing it as

$$p_F(r) = \left[3\pi^2\rho(r)\right]^{1/3}, \tag{2.18}$$

where $\rho(r)$ is the local nucleon density [100]. It can be seen from comparison of the green (RFG) and red curves (LFG) in Figure 2.5 that the LFG model provides a much smoother probability distribution, as opposed to the sharp 'Fermi cliff' of RFG.

A more sophisticated spectral function employed in some recent versions of neutrino interaction generators is the Benhar-Fantoni (BF) model [102]. This is a two-dimensional function which combines a mean field potential description of single nucleons within the nucleus with a short-range correlation term between nucleon pairs. The probability distribution for the BF spectral function for oxygen is given in Figure 2.6. The mean field structure, at low momentum and removal energy values, is seen to display a band-like structure, cor-

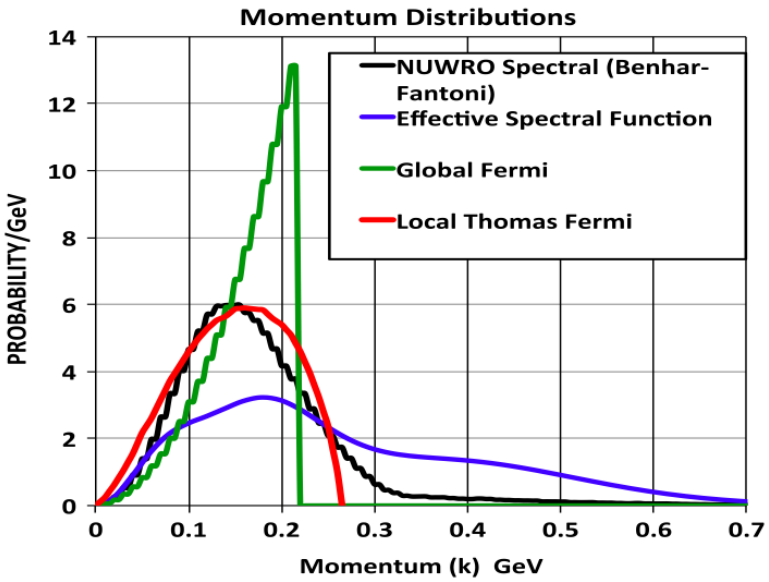


Figure 2.5: Comparison of probability distribution at different momentum values for four spectral functions: Benhar-Fantoni (black), Effective Spectral Function (blue), RFG (green) and LFG (red). Figure from [101].

responding to the different nuclear shells of oxygen. This accounts for approximately 80% of the cross section. The remaining ∼20% comes from the addition of short-range correlations, which gives rise to a large tail in high momentum values, extending out beyond the previous cut off value of the Fermi momentum p_F, as can be seen in Figure 2.5. Although the BF spectral function includes correlations between bound nucleons, these correlations only affect the momentum and binding energy distribution of the nucleons, with the Impulse Approximation still being used to model the interaction as taking place between a neutrino and a single nucleon. As such, only one nucleon is ejected from the nuclear system.

2.5.2 Modelling the Interaction

Whilst valid at higher values of Q^2, the Impulse Approximation described and used throughout Section 2.5.1 breaks down at lower momentum transfers, as the de Broglie wavelength of the neutrino is large enough that it cannot resolve multiple bound nucleons in the nucleus when it interacts. This can significantly modify the cross section, and can result in multiple nucleons being ejected from the nucleus. This can be a particular issue for the energy range at which accelerator neutrino experiments operate.

One of the possibilities suggested for the higher than expected cross-section measurements seen in heavy target experiments such as MiniBooNE is the effect of Meson Exchange Currents (MEC) between nucleons. This is when the nucleons interact with each other via exchange of colourless mesons. In these interactions, scattering involving these bound nucleons can remove multiple nucleons, a process often referred to as N-particle N-hole (Np-Nh), with the most common form being 2p-2h. The MEC effect is expected to increase with nuclear mass as the probability of having bound nucleons increases, making it a good ex-

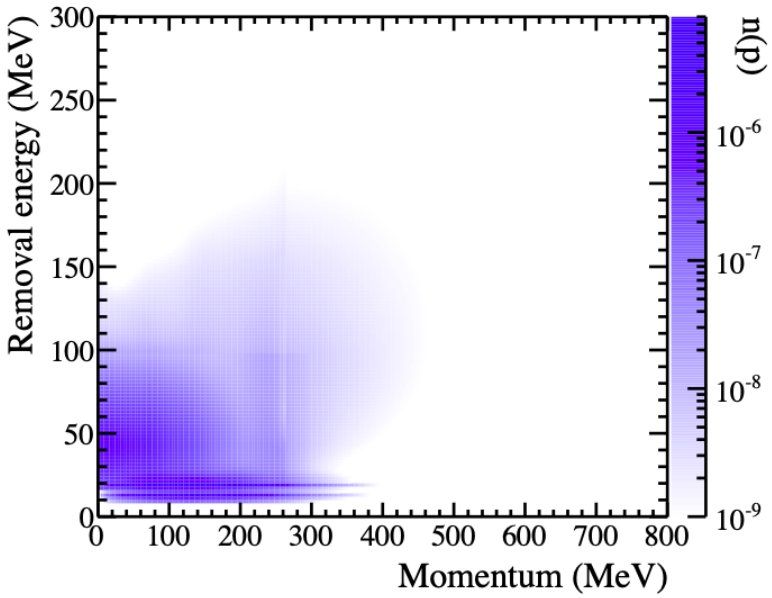


Figure 2.6: Probability distribution for initial state protons within an oxygen nucleus, as predicted by the Benhar-Fantoni spectral function. Figure from [103].

planation for the observed difference between the older bubble chamber data performed with light targets and the more recent, heavy target experiments [104]. For a more detailed discussion of neutrino-nucleus interaction models, see [105].

2.5.3 Final State Interactions

Once the interaction between the neutrino and nucleon(s) has taken place, the particles produced from the interaction still have the non-trivial job of escaping from the nuclear medium. As these particles travel, they have a non-zero probability of re-interacting with either the nuclear medium or other products of the interaction. Produced pions can undergo charge exchange, scatter both elastically and inelastically, or be reabsorbed by a pair of nucleons. These secondary interactions, referred to as Final State Interactions (FSI), can drastically change the distribution of nucleons observed outside of the nucleus, making it impossible to determine the reaction that happened within the nucleus based on the final state particles. A somewhat exaggerated example of how the observed final state can differ from the original due to FSI effects is given in Figure 2.7.

The majority of interaction generators model FSI using a cascade model, where each particle leaving the interaction vertex is independently propagated through the nucleus in small time increments. At each point the probability of a re-interaction is calculated, and if one is said to occur, one of the interaction channels described above is randomly chosen based on their relative cross section. The resultant particles are then added to the cascade, and the process will continue until all particles have either left the nucleus or been absorbed.

The effect of final state interactions on the observed final state in the detector makes measuring neutrino interactions extremely difficult. For this reason, experiments have made

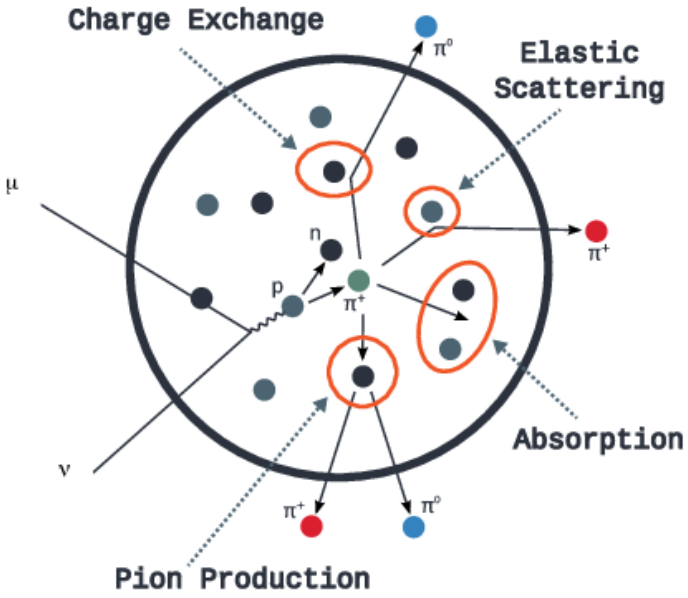


Figure 2.7: Example diagram of Final State Interactions (FSI), showing how the observed final state can look very different to particles produced at the interaction vertex. Figure from [106].

the move to classifying interactions by their final state topology, such as CC0π (CCQE-*like*) or CC1π. This allows cross-section measurements to be reported based solely on the observed final state, instead of making any assumptions of the underlying processes. In doing so, a large amount of dependence on general models can be removed from the measurement.

2.6 Recent Measurements

Over the last decade or so, there has been a wealth of new data on neutrino cross sections, motivated both by their importance to oscillation experiments and for nuclear modelling. In contrast to the early bubble chamber experiments, which generally used light targets of hydrogen or deuterium, more modern experiments use heavier targets, such as water, hydrocarbon or argon. As mentioned in Section 2.1, the move to heavier targets found contention in CCQE cross-section results, which is parametrised by an inflation of the axial mass parameter. Despite becoming known as the 'MiniBooNE large axial mass anomaly', the first experiment to observe such an increase in the cross section of CCQE events was K2K, which performed a measurement of M_A using data from the SciFi (scintillating fibre) [107] near detector. This analysis found a value of $M_A = 1.20 \pm 0.12$ GeV, with a χ^2/Ndf of $\frac{261}{235}$ [108]. Following this, and the results from MiniBooNE, additional experiments measuring CCQE interactions also found increased cross sections that were parameterised by inflated axial mass values [109, 110, 111].

After the suggestion that the anomaly could be caused by correlation between bound nucleons, the Transverse Enhancement Model (TEM) was published by Bodek, Budd and Christy [86]. This modifies the transverse component of the CCQE cross section, with the enhancement attributed to Meson Exchange Currents. CCQE cross-section results on

hydrocarbon, reported by the MINERνA [112] collaboration at a higher mean energy of E_ν $\sim$ 3.5 GeV, found the observed data were best described by an RFG model with $M_A = 0.99$ GeV, with the addition of TEM [113]. Additional reanalyses of MiniBooNE data presented in [114, 115] supported this, showing that the inclusion of multi-nucleon processes gave agreement with data, without having to inflate the axial mass parameter. From these results, it can be seen that the interaction process being measured by MiniBooNE *et al.* was not purely CCQE, as it involved additional effects such as 2p-2h interactions, and highlights the importance of reporting cross-section measurements in terms of the observed final state topology, as described in Section 2.5.3.

Many experiments have since started doing this, with a large number of measurements for the CC0π/CCQE-*like* cross section being reported by MINERνA [116, 117] and T2K [118, 119, 120]. These results are particularly useful, as they are reported in the form of double-differential cross sections in terms of the final state kinematics. This allows testing of the underlying models in a two-dimensional space, and provides more opportunities to resolve model differences. Similar leaps forward have been made in the single pion final state (CC1π) regime, with MINERνA [121, 122, 123], MiniBooNE [124, 125, 126] and T2K [127, 128] performing differential measurements, some of which are made in pion final state kinematics. Measurements of the outgoing pion can provide greater sensitivity to FSI effects, and so will be an important test for interaction model separation going forward.

A particular region of interest in the study of underlying nuclear effects is the measurement of transverse kinematic variables, proposed in [129]. In the proposed scheme, a plane is formed transverse to the incoming neutrino direction, which is shown in Figure 2.8 . The outgoing final state lepton and hadron can then be projected onto this plane to obtain the components transverse to the neutrino direction: $\vec{p}_T^{\,\ell'}$ and $\vec{p}_T^{\,\mathrm{N}'}$, respectively. If the struck initial-state nucleon were free and at rest, then the final-state momenta of the lepton and hadron are required to be equal and opposite. Calculating the difference between these,

$$\delta\vec{p}_T \equiv \vec{p}_T^{\,\ell'} + \vec{p}_T^{\,\mathrm{N}'}, \tag{2.19}$$

provides sensitivity to the initial nucleon state. Additionally, the transverse boosting angle $\delta\alpha_T$ can be defined as:

$$\delta\alpha_T \equiv \arccos\frac{-\vec{p}_T^{\,\ell'}\cdot\delta\vec{p}_T}{p_T^{\ell'}\delta p_T}. \tag{2.20}$$

In the case of no FSI, to a first approximation, $\delta\vec{p}_T$ would be independent of neutrino energy, and the distribution of $\delta\alpha_T$ would be flat. This therefore makes these variables an important probe for the effects of FSI and for separation between models, with recent differential measurements made by MINERνA [130] and T2K [131]. In pion production modes it is also possible to assess the double-transverse momentum imbalance δp_{TT}, which is transverse to the lepton scattering plane, where asymmetry between the outgoing pion and proton is a clear indicator of nuclear effects. Publications from both MINERνA [132] and T2K [133] show interesting model separation in the tails of the cross-section distribution, differential in δp_{TT}. It is clear that as we enter an era of precision oscillation measurements,

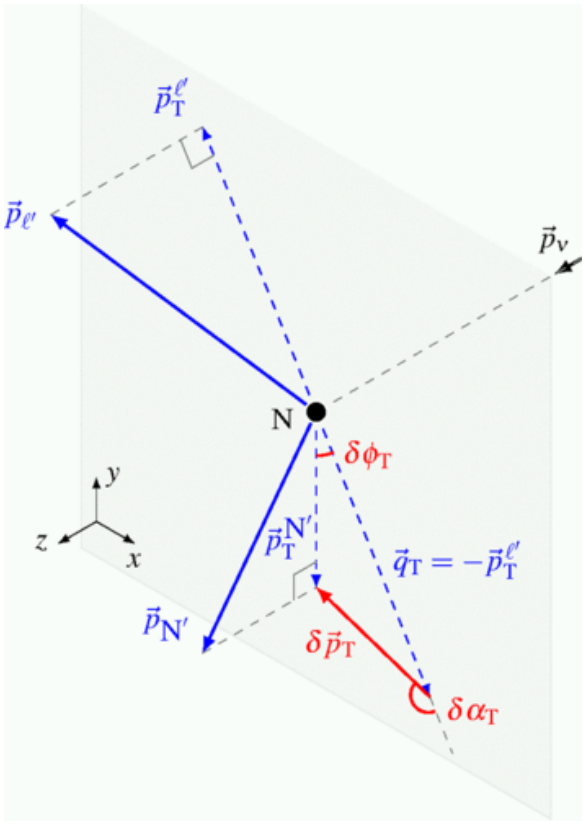


Figure 2.8: Geometric view of the plane formed transverse to the neutrino direction, used to make measurements of neutrino cross sections in the transverse kinematic variables. Transverse kinematic variables in which measurements are made are shown in red. Figure from [129].

such novel variables will be crucial to driving down cross-section systematic uncertainties.

While there has been an influx of new cross-section results in recent years, made on differing nuclear targets, it is noted that none of the published measurements described report double differential measurements in both pion momentum and angle, which should be particularly sensitive to FSI effects. This work aims to address that, along with exploring the possibility of making higher dimensional differential measurements, in terms of both the outgoing lepton and meson.

Chapter 3

The T2K Experiment

The T2K (Tokai to Kamioka) experiment is a long baseline accelerator neutrino oscillation experiment, situated in Japan. The goal of T2K is to make precision measurements of neutrino oscillation parameters, through both appearance and disappearance signals, using multiple detectors and a high intensity muon neutrino beam.

A schematic diagram of the T2K experiment is presented in Figure 3.1. A high purity beam of muon neutrinos is produced at the Japan Proton Accelerator Research Complex (J-PARC) in Tōkai-mura (東海村), Ibaraki-ken, on the eastern coast of Japan. The neutrino beam is generated from proton beam collisions on a stationary target. The near detector complex sits 280 m downstream from the beam production point, where measurements of the unoscillated beam are performed. 295 km further downstream, near Hida-shi, Gifu-ken, the Super-Kamiokande (Super-K) detector is used to measure the flavour composition of the beam after oscillation has occurred. T2K makes use of a 'narrow-band' neutrino beam, where the beam centre is directed 2.5° off-axis from Super-K, yielding a lower, yet tighter, neutrino energy spectrum than the on-axis alternative. In order to maximise neutrino oscillation probability for the 295 km baseline, the off-axis angle of the beam is tuned to produce a neutrino energy spectrum that peaks at 600 MeV. This chapter gives an overview of all sections of the T2K experimental setup: the T2K neutrino beam is discussed in Section 3.1, and details of the T2K near detector complex are given in Section 3.2. For completeness, Section 3.3 gives a brief description of the T2K far detector, Super-Kamiokande, although this was not used in the analysis presented in this thesis.

3.1 T2K Neutrino Beam

An overview of the T2K beamline components at J-PARC is given in Figure 3.2. Section 3.1.1 describes the multi-stage production of the 30 GeV proton beam, whilst Section 3.1.2 gives details of the target station and hadron decay volume. The off-axis configuration of T2K, briefly described above, is discussed in more detail in Section 3.1.3.

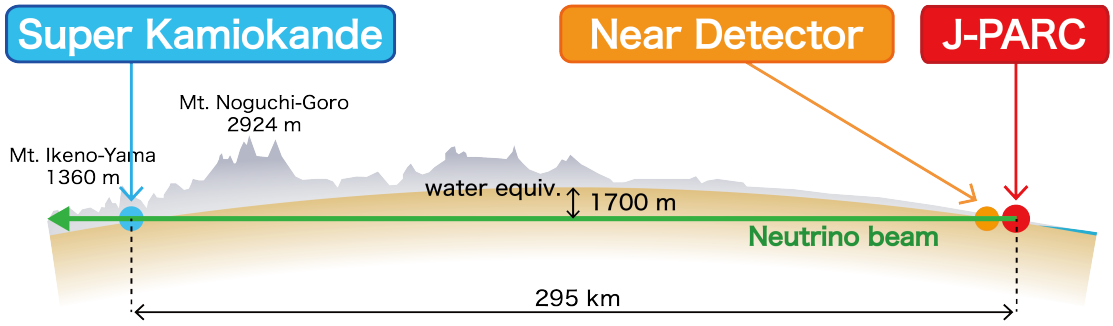


Figure 3.1: Diagram of the T2K experimental layout, taken from [134].

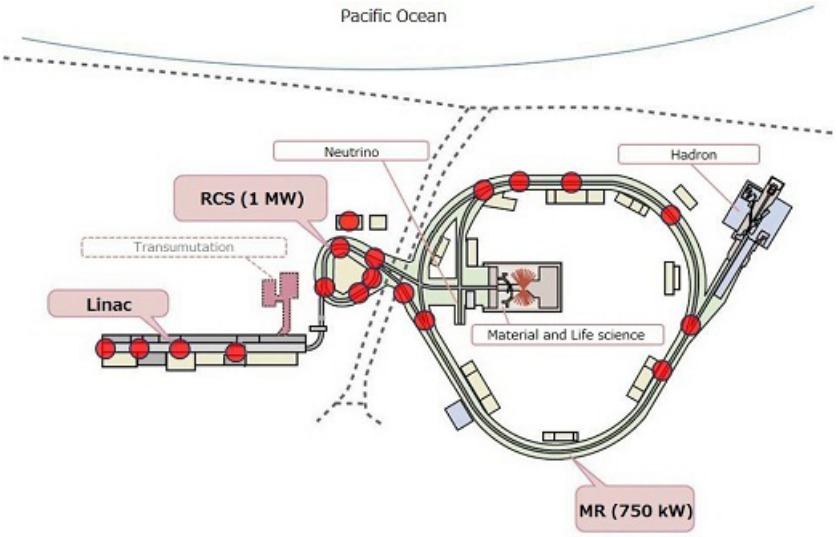


Figure 3.2: J-PARC beamline diagram, showing the multiple accelerators used in production of the T2K neutrino beam, taken from [135].

3.1.1 J-PARC Proton Accelerator

As is the case with many large-scale particle accelerators around the world, proton acceleration at J-PARC is a multiple stage process. In the first stage of the acceleration process, a 249 m linear accelerator (LINAC) is used to accelerate negative H^- ions up to 400 MeV [136]. Upon injection into the Rapid Cycling Synchrotron (RCS), charge stripping foils are used to convert the H^- ions to a proton beam, which is then accelerated up to 3 GeV at a 25 Hz cycle, with two bunches in each. Only ∼5% of these bunches are injected into the 1567 m circumference main ring (MR) synchrotron; the majority are instead supplied to other beamlines in the Material and Life Science Facility [137]. After being injected into the MR, the proton beam is accelerated up to 30 GeV, with eight bunches in each 0.5 Hz cycle. Along the MR, two extraction points exist, which use different extraction modes: bunches for the hadron facility use a slow extraction, whereas extraction for the neutrino facility is

done in fast extraction mode. The latter extracts all eight proton bunches from the main ring in a single turn, using kicker magnets at the extraction point.

3.1.2 T2K Beamline

After extraction from the MR, the proton beam is handled by the T2K beamline at the J-PARC neutrino facility. This consists of two main sections, the primary and secondary beamlines, and is shown in Figure 3.3.

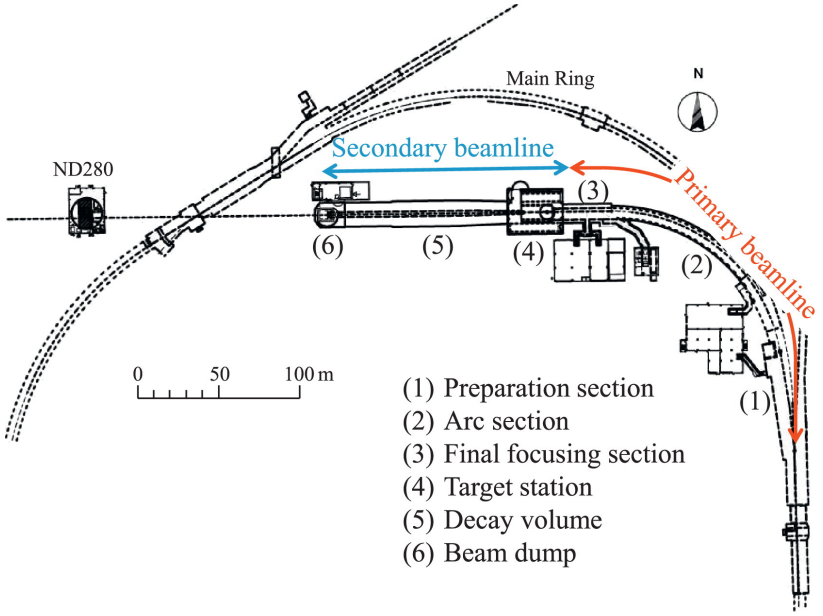


Figure 3.3: Schematic diagram of the T2K beamline, showing the various subsections of the primary and secondary beamline. Figure taken from [137].

The primary beamline itself consists of three sub-sections. The first, the 54 m long preparation section, receives the extracted proton beam and tunes it so that it can be accepted by the arc section. This is done using a series of regular conducting magnets. The arc section then uses 14 doublets of superconducting magnets in order to bend the proton beam by 80.7° into the direction of the Super-K detector. Finally, the focusing section directs the beam 3.657° downwards from the horizontal, angling it towards the target station. Along the primary beamline are a series of beam loss monitors, used to check the intensity, position and loss along the beamline.

The secondary beamline is shown in more detail in Figure 3.4. The first part of the secondary beamline that the proton beam reaches is the target station. This consists of a helium vessel evacuated down to 50 Pa, containing the baffle, optical transition radiation monitor (OTR), target and magnetic horns. The baffle collimates the beam in order to protect the magnetic horns from radiation damage, whilst the OTR can be used to monitor the beam profile before the target, by viewing the optical transition radiation produced as the beam passes through a thin foil made of titanium alloy. The core of the target itself is a 914 mm long (1.9 interaction lengths), 26 mm diameter graphite rod. Graphite is used as a target due to its relatively low density (1.8 g/cm^3); this is important for dissipation of the

heat generated from the proton interactions. The target and its surrounding graphite tube are encased in a titanium casing, with the whole assembly being supported within the first magnetic horn and cooled by helium gas flow.

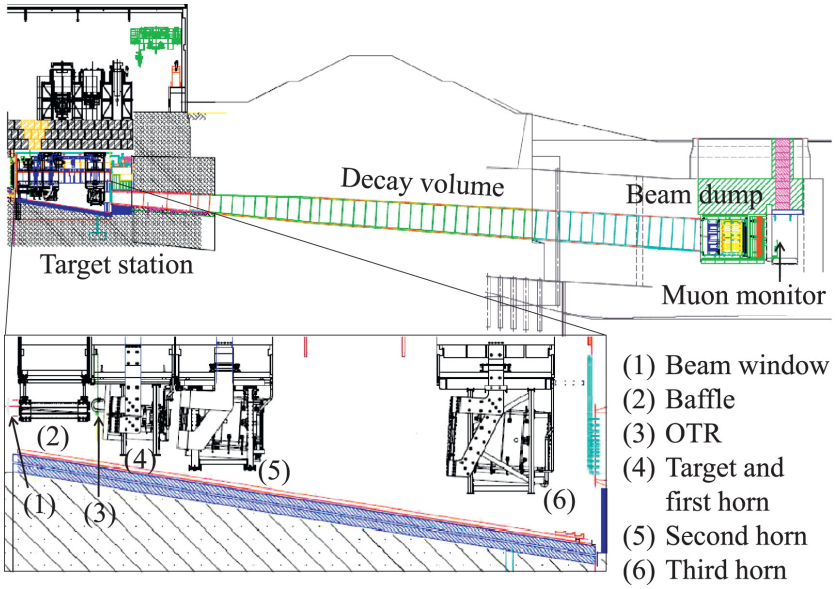


Figure 3.4: Side-on diagram of the T2K secondary beamline, sections 4 - 6 of Figure 3.3. Beam travels from left to right of the diagram. Figure taken from [137].

Upon striking the graphite target, the high energy proton interactions cause a slew of different hadrons to be produced. The three magnetic horns are pulsed at 320 kA in time with the arriving beam spills, equating to a maximum field strength of 2.1 T, in order to select out positively (negatively) charged pions, which will decay in flight to produce muon (anti)neutrinos and positively (negatively) charged muons, with a branching ratio of 99.98770% [19]. The magnetic horns can be run in two different modes, by reversing the polarity on the magnets. These are Forward Horn Current (FHC) and Reverse Horn Current (RHC), which are used to select out π^+ and π^- respectively. The use of these magnetic focusing horns also has the advantage of defocusing oppositely charged hadrons, reducing wrong sign contamination in the eventual neutrino beam.

Pions selected and focused by the magnetic horns are then directed towards the decay volume. This is a 96 m long, ~1500 m^3 steel tunnel, surrounded by concrete shielding. The volume itself is filled with helium gas, held at 1 atm, in order to reduce pion absorption during flight, as well as suppressing tritium and NO_x production from the beam [138]. At the end of the decay volume, still contained within the helium vessel, is the beam dump. With a total weight of 75 tons, this 3.174 m deep block of graphite is used to absorb the majority of the muons from pion decay, along with hadrons remaining in the beam. To aid in this, the graphite is followed by 17 iron plates, totalling 2.4 m thickness. This combination of materials generally stops muons below ~5 GeV from reaching the muon monitor pit.

Further downstream, behind the beam dump, sits the muon monitor, which by measuring the produced muon distribution profile is able to provide information on the neutrino beam intensity and direction on a bunch-by-bunch basis. This is important as the neutrino beam

direction is then defined to be in the direction between the carbon target and the centre of the muon profile. The muon monitor is designed to measure the muon profile centre to a precision of 3 cm, corresponding to a precision of 0.25 mrad (0.014°) on the neutrino beam direction. In addition to this, a nuclear emulsion tracker is placed downstream of the monitor, which through two separate detector configurations is able to make measurements of the absolute flux and momentum distribution of the muons [139].

3.1.3 Off-axis configuration

As mentioned previously, a key feature of T2K is that the produced neutrino beam is not directed at the far detector target. Instead, the beam is angled away from Super-K by 2.5°. Since decay of the pion is a two-body interaction, the energy of the resultant neutrino can easily be expressed as

$$E_\nu = \frac{m_\pi^2 - m_\mu^2}{2(E_\pi - p_\pi \cos\theta_{\mathrm{OA}})}, \tag{3.1}$$

where m_π, E_π and p_π are the mass, energy and momentum of the pion respectively, m_μ is the mass of the muon, the mass of the neutrino is negligible, and θ_{OA} is the off-axis angle of the beam [140]. It can be seen from Equation (3.1) that an increase in the off-axis angle causes a decrease in the neutrino energy as a function of the initial pion momentum, and a reduction in the width of the neutrino energy spectrum. Figure 3.5 shows the resultant neutrino energy as a function of the initial pion momentum, for varying off-axis angles. Naturally this off-

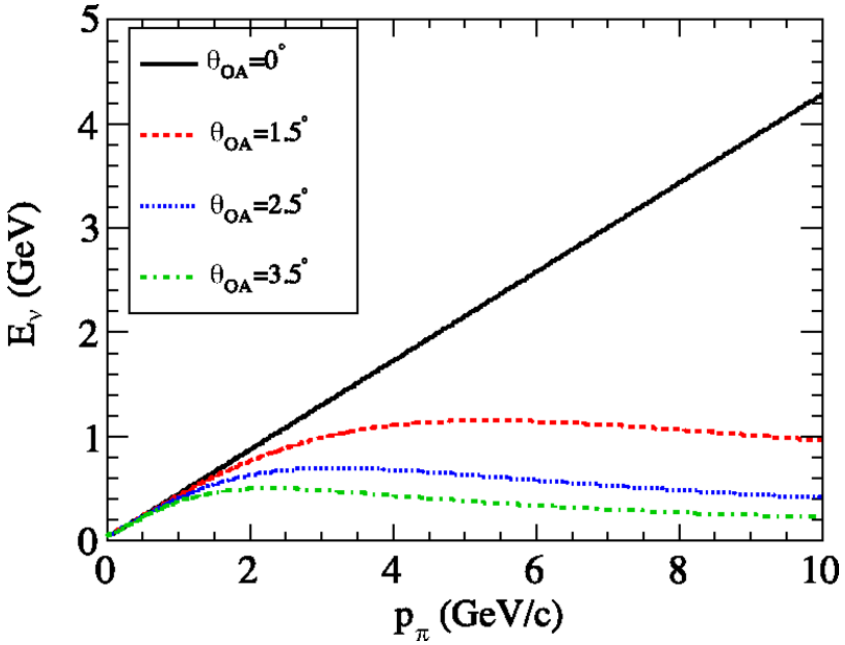


Figure 3.5: Neutrino energy as a function of primary pion momentum, for different values of the off-axis angle θ_{OA}. Figure taken from [140].

axis approach also reduces the flux of neutrinos directed at the far detector in comparison to on-axis, but the narrower neutrino energy distribution is seen as advantageous since

it maximises the oscillation probability, and therefore sensitivity to the mixing angle θ_{13}, which the experiment initially aimed to measure. Figure 3.6 shows the predicted neutrino

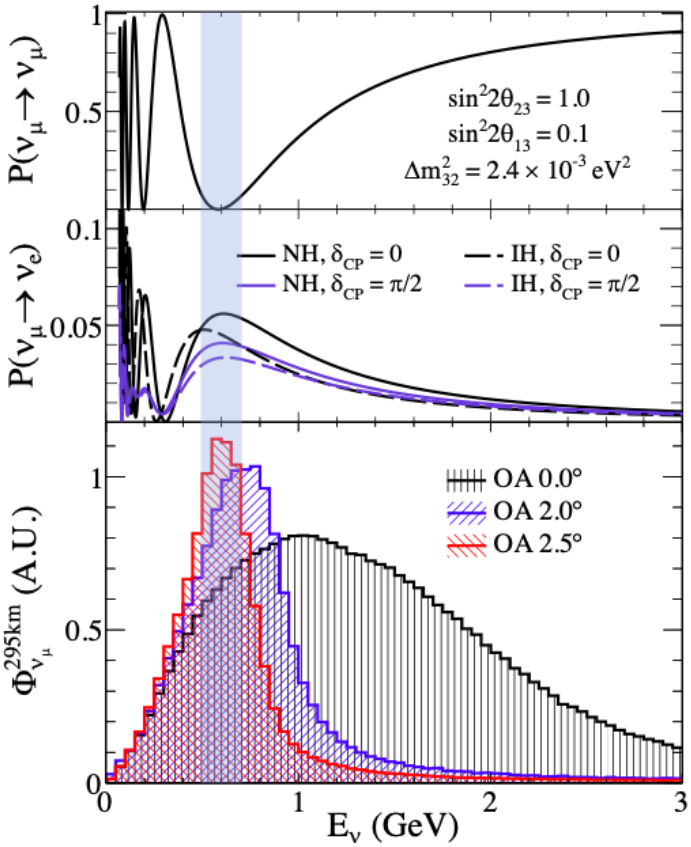


Figure 3.6: Scaled muon neutrino flux distributions (bottom) at Super-K, for an on-axis configuration, and two possible off-axis angles. The off-axis configuration gives a 'narrow band' beam, which has a much tighter neutrino energy distribution compared to on-axis. The energy of the beam is tuned to coincide with minimal survival probability (top) and maximal oscillation probability (middle) at the far detector. Using the 2.5° off-axis configuration, the T2K neutrino beam peaks at an energy of 600 MeV. Figure taken from [141].

flux at Super-K for an on-axis, 2.0° off-axis and 2.5° off-axis configuration, with the scaled fluxes given in arbitrary units. Whereas the on-axis configuration has a large spread in possible energy values, the 2.5° off-axis configuration used offers a lower, yet much narrower, energy spectrum, peaking at 600 MeV. This off-axis angle was tuned to give a neutrino energy that peaks at the maximal oscillation probability with the 295 km baseline to the far detector. This also shows the importance of precise knowledge of the beam direction, as small deviations in angle can cause large changes in the energy spectrum of neutrinos at the far detector.

3.1.4 Flux Simulation

Despite the near detector complex being used to constrain the flux of the neutrino beam, it is still important to be able to provide a good prediction of the flux, due to differences in the near and far detector target materials and detection efficiencies [142]. Additionally, an accurate flux prediction is necessary for neutrino cross-section measurements, such as the one presented in this thesis.

The prediction of the neutrino beam is done in several separate steps, utilising a combination of external data sets and Monte Carlo simulations. Firstly, measurements of the primary beam profile are used to specify the input trajectories of the protons entering the target and baffle, where the interactions are simulated using `FLUKA` [143, 144]. Magnetic horn measurements are also used as an input, with the horn-focusing done using `JNUBEAM`, which is a `GEANT3` [145] based MC simulation. The out-of-target interactions up to the decay neutrinos are then simulated using `GCALOR` [146].

Once simulations are done, the interaction chain is run through again, and has weights applied in order to tune the generator output to external hadron production data, mostly from NA61/SHINE (SPS Heavy Ion and Neutrino Experiment) at CERN, which provides hadron production measurements for various long baseline neutrino experiments [147].

3.2 Near Detector Complex

280 m downstream from the beam production point is the T2K near detector complex, composed of multiple detectors. The complex is housed inside a pit of 37 m depth, which unlike the far detector has no overburden [137]. A diagram of the near detector complex, which spans three floors, is presented in Figure 3.7. The on-axis INGRID detector is described in Section 3.2.1, whilst full details of the off-axis detector relevant to this analysis, ND280, are given in Section 3.2.2. The more recent addition to the near detector complex, WAGASCI, is described in Section 3.2.3.

3.2.1 INGRID

The Interactive Neutrino GRID (INGRID) detector is designed to sample the neutrino beam prior to oscillation, to provide information on the neutrino flux and energy spectrum. Although this is also done by the muon monitor downstream of the beam dump, only the very high energy muons ($\gtrsim$ 5 GeV) that can penetrate the beam dump itself will be detected there, meaning that the muon monitor covers a very specific phase space that is not indicative of the general T2K energy spectrum [148]. Instead, the on-axis INGRID detector is used to measure the beam direction and intensity by detecting neutrino interactions, covering a phase space much closer to that of the off-axis detectors.

Figure 3.8 shows the structure of the INGRID detector, which is formed of 16 identical detector modules aligned around the neutrino beam centre. Each module is a layered structure, formed from nine iron target plates between 11 scintillator tracking planes. The tracking planes themselves are each formed of one horizontal and one vertical layer, made from 24 scintillator bars. Neutrino interactions in the target cause scintillation light to be

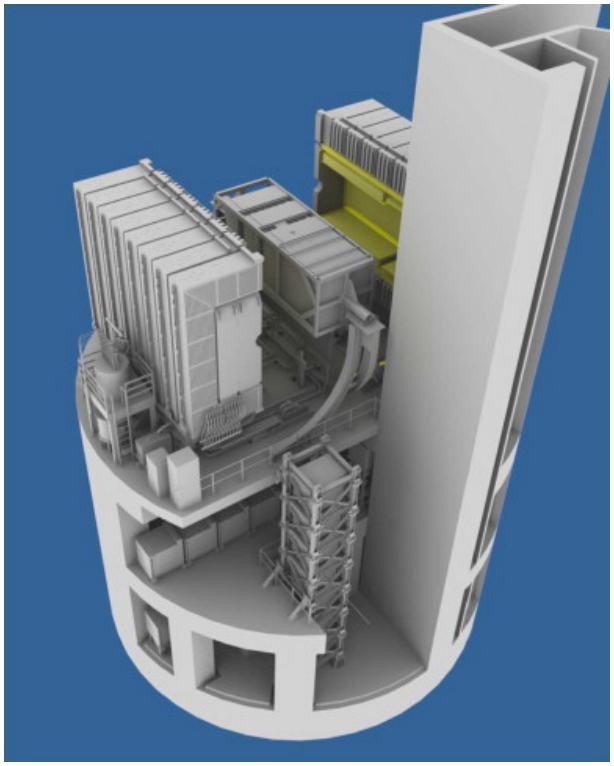


Figure 3.7: Diagram of the near detector complex. Top floor houses the ND280 near detector and magnet, centre floor houses the horizontal INGRID modules. The INGRID vertical modules span the centre and lower floor, whilst the lower floor also houses the WAGASCI-BabyMIND detector. Neutrino beam enters from the bottom right of the diagram. Figure taken from [137].

generated in the bars, and the light is collected and transported using wavelength shifting fibres attached to Multi-Pixel Photon Counters (MPPCs) [149]. The MPPCs then record the charge collected, corresponding to the interaction signal [150].

The 16 modules are arranged in a cross shape, with seven modules aligned vertically, in front of another seven horizontally. The beam detector is aligned so that the beam centre passes through the centre of the cross, where two modules overlap. The remaining two modules are placed above the horizontal arm, in order to measure the axial symmetry of the beam. Each arm of the cross extends roughly 5 m out from the centre of the detector, which corresponds to the spatial width (1σ) of the beam [148]. This enables INGRID to measure neutrino interactions from the beam at off-axis angles ranging between 0° and 1.1° [140].

An additional detector component was added to INGRID several years after running first started. The INGRID proton module was designed to detect both the muons and protons from neutrino interactions; this allowed identification of quasi-elastic interactions, in order to compare reaction channel rates to Monte Carlo simulation. To provide better resolution tracking compared to the standard INGRID detector modules, the proton module uses scintillator bars of a narrower width, and removes the iron target planes, making it a

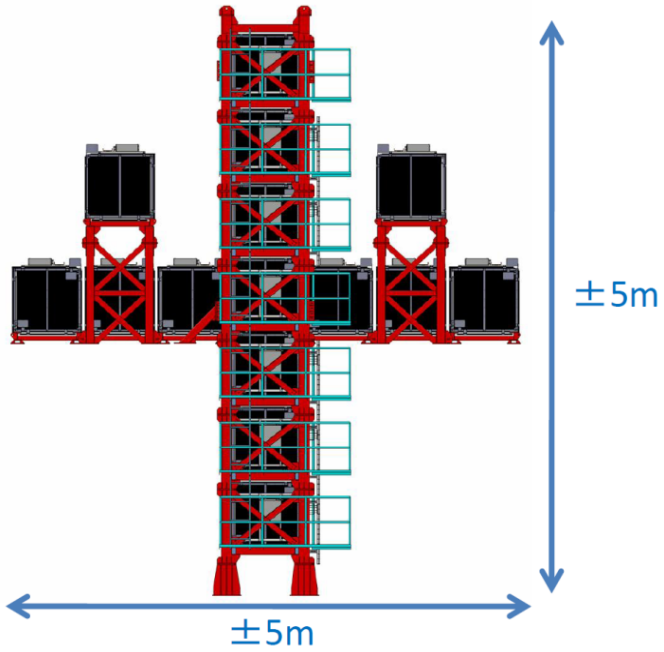


Figure 3.8: INGRID near detector diagram. INGRID is formed of 16 identical modules, made of layered iron target plates and scintillator tracking planes, positioned in a stack of seven vertical, seven horizontal and two off axis modules. The T2K beam centre goes through the centre of INGRID, where two modules overlap. Figure taken from [148].

fully active detector. The tracking planes are then surrounded by six veto planes [151]. The proton module was installed between the vertical and horizontal arms of INGRID, so as to be able to use the downstream INGRID module as another tracking detector. The proton module was installed in October 2010, and started taking data in the November, but in 2017 was removed, and at time of writing now forms part of the WAGASCI detector setup, detailed in Section 3.2.3.

3.2.2 ND280

Also 280 m downstream from the carbon target, but at the same 2.5° off-axis angle as Super-K, sits the ND280 near detector [137]. The goal of ND280 is to minimise T2K oscillation measurement systematic errors by constraining relevant beam and neutrino interaction parameters. This includes [152]:

- measuring the muon neutrino energy spectrum,
- determining the rate of charged-current (CC) interactions,
- measuring the ratio of quasi-elastic (QE) and non-QE cross sections, along with studying non-QE processes that produce pions below the Super-K Cherenkov threshold, and
- measuring neutral-current (NC) π^0 production rates.

To achieve these goals, ND280 is composed of two main sections, where the upstream section is used to detect neutral particles, and the downstream section is primarily used for the detection and tracking of particles from charged-current interactions. An exploded view of ND280, showing the various subdetectors it comprises, is given in Figure 3.9. Both sections of the detector are housed within the refurbished UA1 dipole magnet, which provides a magnetic field of 0.2 T [153], for the purposes of charge discrimination and momentum measurements. Full details of the individual subdetectors are given in the following subsections.

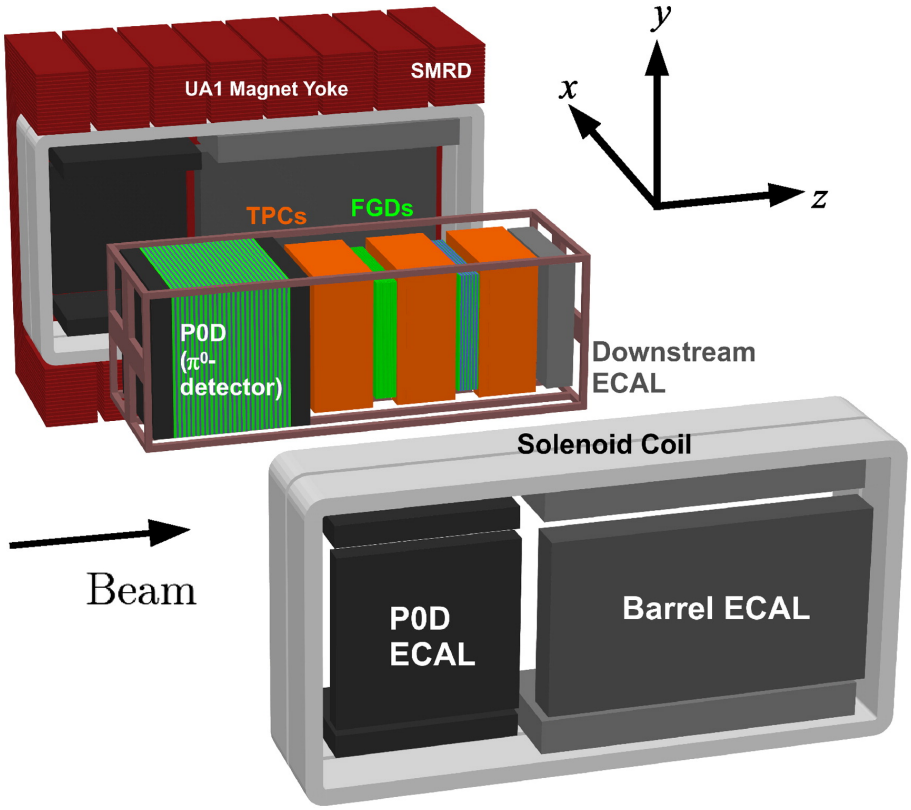


Figure 3.9: Exploded view of the ND280 off-axis near detector, showing the individual subdetectors comprising it. Neutrino beam enters from left of figure. Figure taken from [137].

3.2.2.1 Pi-zero Detector

The pi-zero detector, otherwise known as the PØD, is the most upstream subdetector in ND280 [154]. In order to precisely measure the rate of ν_e appearance at the far detector, which proceeds via the process $\nu_e + n \longrightarrow e^- + p$, it is essential that the neutral current production rate of π^0s ($\nu_\mu + N \longrightarrow \nu_\mu + \pi^0 + X$) is understood. Second to the irreducible ν_e contamination in the ν_μ beam, π^0 production at Super-K is the dominant background to the ν_e appearance analysis [155], due to the fact that electrons and photons are indistinguishable in water Cherenkov detectors. Neutral pions produced from these neutral current interactions will rapidly decay to photon pairs (where in many cases only one photon will be detected), or a singular photon with a electron-positron pair, mimicking the signal of

charged current quasi-elastic ν_e interactions. The PØD is designed to measure the rate of these neutral current π^0 interactions, using a water target for similarity with the far detector.

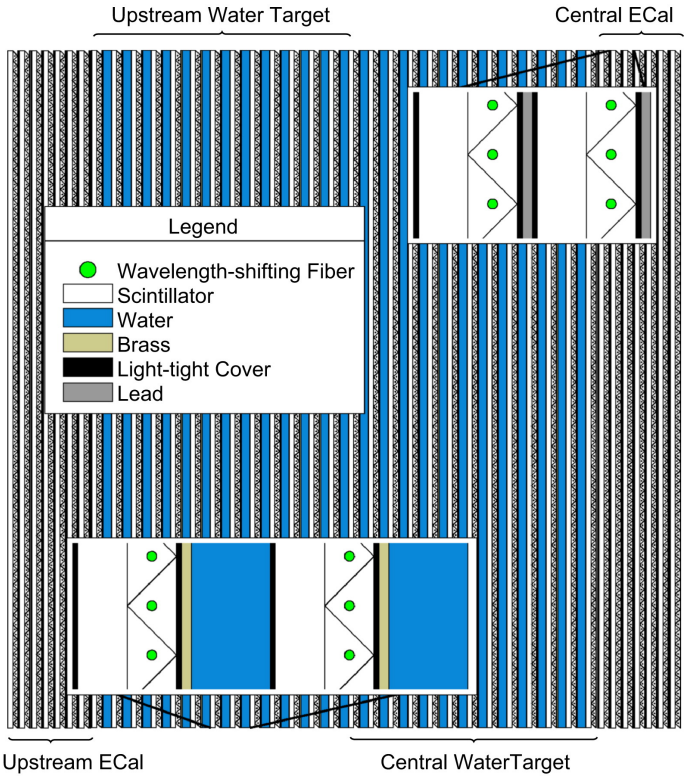
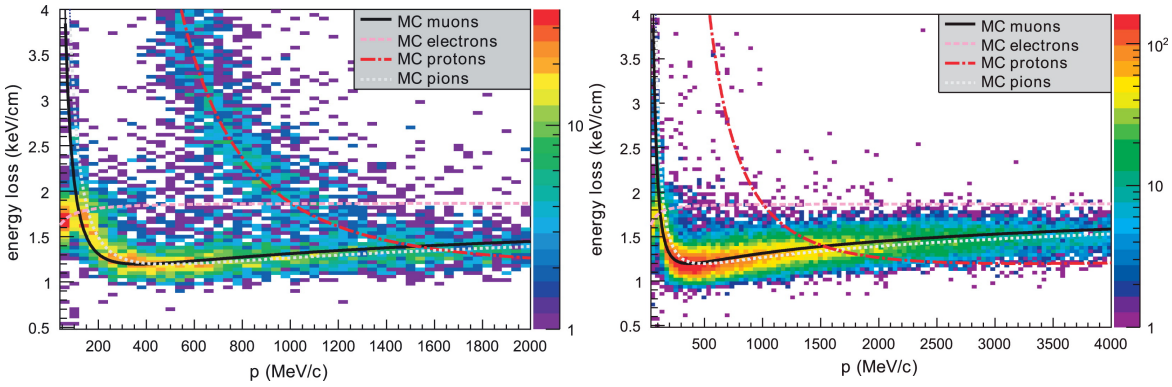


Figure 3.10: Side-on diagram of the active region of the PØD, which is composed of four SuperPØDules: the upstream ECal, upstream water target, central water target and central ECal. Close-up views of the scintillator bar structure in both the ECals and water targets are shown. Neutrino beam enters from the left of the diagram. Figure taken from [154].

In order to achieve these measurements of the interaction rate on water, a large water target mass is required. The design of the active section of the PØD, which is made up of four main parts (or Super-PØDules), is shown in Figure 3.10. The central region is formed of alternating scintillator planes, water bags, and brass sheets, where each scintillator plane (or PØDule) is itself made of two perpendicular arrays of triangular scintillator bars, threaded with wavelength shifting fibres. The upstream (central) water Super-PØDules are formed from 13 PØDules, alternating with 13 (12) water target bags, and 13 (12) brass sheets, where the brass sheets act as radiators to induce electromagnetic showers. To determine the rate of interactions on water, statistical subtraction is performed using data taken from separate runs with and without the water bags being filled. The remaining two Super-PØDules, which sit either side of the water target Super-PØDules, are Electromagnetic calorimeters (ECals), formed from alternating scintillator planes and lead sheets. These provide containment and measurement of electromagnetic showers, and act as a veto region for interactions entering from outside the PØD.

3.2.2.2 Time Projection Chambers

Downstream from the PØD is the main charged particle tracking section of ND280. This consists of three time projection chambers (TPCs) for precision tracking, alternating with two fine-grained detectors (FGDs) providing target mass. The TPCs have three main uses. Firstly, their ability to reconstruct events in 3D with high precision makes them useful for determining the multiplicity of particles resulting from different types of interactions. Secondly, the magnetic field provided by the UA1 magnet causes curvature of charged particles passing through the TPCs. By measuring the curvature of the tracks, the momentum of the traversing particles can be calculated. This is particularly useful in the case of quasi-elastic interactions, where it allows the energy of the incoming neutrino to be calculated. Curvature of the tracks due to the magnetic field also allows the sign of the particle charge to be determined, which is essential to event selection. Finally, using a combination of the measured momentum and the amount of ionisation left by different charged particles, particle identification can be performed. Along with allowing different event topologies to be selected with a high purity, this also contributes to measuring the contamination of electron neutrinos in the beam before oscillation [137]. Obtained data distributions for momentum against energy loss compared to MC expectation for selected positive and negative particles are shown in Figures 3.11a and 3.11b respectively, showing that the majority of particles detected in the TPCs have a momentum such that they are classed as minimally ionising. A minimum ionising particle (MIP) is defined as one whose mean energy loss through matter is close to the minimum of the curves seen in Figures 3.11a and 3.11b. The exception to this is the proton, which can be seen in the TPC data to exhibit a wider range of energy loss values. The exact method for using these variables to perform particle identification is discussed in Section 6.1.6.



(a) Positive tracks

(b) Negative tracks

Figure 3.11: TPC momentum measurements against energy loss from the T2K first physics run, for (a) positive and (b) negative tracks. Expected distributions from MC simulation are also plotted for comparison. Figures taken from [156].

The TPCs are constructed using a double box design (shown in Figure 3.12), where the inner box wall forms the electric field cage, and the outer walls are at ground potential. CO_2 is used as an insulator between the two, and the inner box is filled with a gas mixture

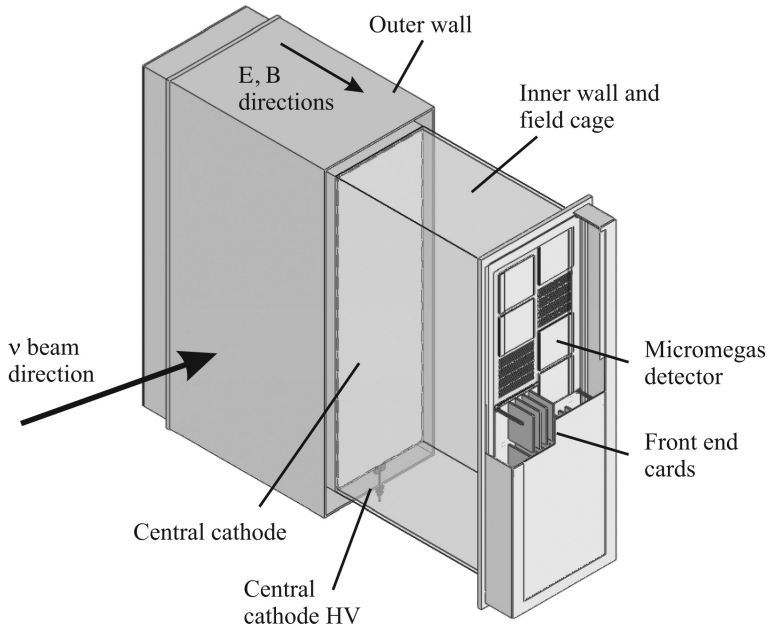


Figure 3.12: Simplified conceptual diagram of one of the ND280 TPCs. Charged particles passing through a TPC ionise the gas, leaving electrons which drift away from the central cathode to one of the outer readout planes, where they are detected by the Micromegas. Figure taken from [156].

composed of 95% argon, 3% tetrafluoromethane (CF_4) and 2% isobutane (iC_4H_{10}) [156]. Charged particles passing through the TPC gas mixture cause ionisation, and the resultant ionisation electrons drift in the electric field away from the central cathode, towards one of the readout planes. Here they are detected with bulk Micromegas detectors [157], where the pattern of signals and the arrival time are combined to create high-quality 3D images of the particle trajectories. Each of the six readout planes in a single TPC contains 12 Micromegas modules, which are arranged in two vertical columns that are slightly offset, so that the inactive regions between modules do not align. The three ND280 TPCs are shown in orange in Figure 3.9.

3.2.2.3 Fine Grained Detectors

Interleaved between the three ND280 TPCs are two fine-grained detectors (FGDs), which form the main target mass of the ND280 tracker. In the design of the detector, it was required that the tracker region should contain ∼1 ton of target mass for neutrino interactions, in order to achieve a statistically significant sample of events [158]. In particular, part of the target mass is also required to be made of water. As Super-K is a water Cherenkov detector, it is important to be able to make cross section and relevant event rate measurements on the same target medium; nuclear effects mean that interaction cross sections are dependent upon the target nucleus, and these effects cannot be corrected for without relying heavily on model assumptions. The FGDs aim to provide precise particle tracking and vertex reconstruction, but are made to be thin enough that the majority of particles should escape into a TPC. Due to the differing structures of FGD1 and FGD2, they are described separately

here.

FGD1

FGD1 is a fully active scintillation detector, comprising 5760 individual extruded polystyrene scintillator bars; each bar has dimensions of 9.61 mm × 9.61 mm × 1864.3 mm. The bars feature a reflective TiO_2-containing coating, which helps to limit light propagation between bars, and the centre of each is instrumented with a wavelength shifting fibre, which collects the scintillation light and transports it to one end, where an MPPC digitises the signal. The opposite end of each bar is mirrored by vacuum deposition of aluminium, to ensure as much scintillation light as possible is registered [158]. In order to provide the best possible tracking capabilities, 192 scintillator bars are aligned in the x direction, followed by 192 aligned in the y direction, forming what is known as an XY module. This ensures that all bars are perpendicular to the neutrino beam direction, and thus should maximise the number of bars that a forward-going charged particle passes through. FGD1 is composed of 15 sequential XY modules, totalling 30 individual layers. An LED-based light injection system is included in order to perform calibration of the photosensor response in situ, and the light yield is measured on a yearly basis to monitor degradation due to scintillator ageing. A recent study over 10 years of data taking found an annual light yield reduction in the FGD scintillator of 1.2 ± 0.2% [159]. The modules are glued together, with 0.25 mm thick G10 fibreglass sheets attached to both the upstream and downstream faces of each in order to provide additional mechanical rigidity. Each module is held by five stainless steel straps, which support them inside the aluminium dark box used to prevent light from outside causing false signals. Finally, the front-end electronics are housed within 24 minicrates, surrounding the FGD on all four long edges. An image of the FGD design can be seen in Figure 3.13.

FGD2

FGD2 differs from FGD1 in that along with plastic scintillator, it also includes water layers as part of the target mass, which was one of the initial design requirements. FGD2 comprises seven XY scintillator layers, identical to those in FGD1, interleaved with six water target modules, corresponding to a total water target depth of 150 mm. The water modules are made of thin corrugated polycarbonate of 25 mm thickness, hollowed out to be filled with water. The ends are sealed with a polyurethane sealant, and the water within the targets is held at a pressure below atmosphere by a vacuum pump. Should a leak in the water bags occur, this ensures that air is sucked into the bags, rather than water leaking out into the surrounding FGD electronics. Comparing event rates between FGD1 and FGD2 allows the rate of interactions on water to be determined. This is the technique used in this analysis, and is described in more detail in Section 5.2.3. Both FGDs have the same outer dimensions, and other than the difference in target layers were built with the same geometry and readout for ease of operation [158].

Over the years of operation, extended running has unfortunately taken its toll on elements of FGD2. In 2015, an accident during routine draining of the target water layers caused the most upstream one to rupture. Although no harm to the surrounding electronics was caused, due to the large amount of work that would be required to replace the water

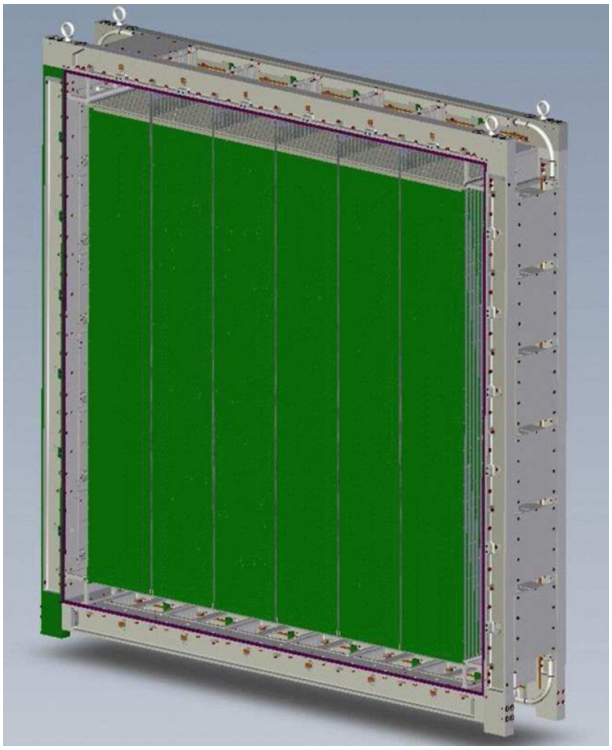


Figure 3.13: Representation of FGD1 with the front cover removed. Scintillator bars are shown in green, with the five sets of stainless steel strips visible between the scintillator plane and the surrounding aluminium dark box frame. On the top of the frame, six uncovered minicrates can be seen, whilst the six down the right edge are covered by the cooling lids. Figure taken from [137].

panel and the possible additional risks associated, it was decided that the panel would not be replaced. This means that for the FHC data-taking runs used in this analysis, run 8 only has five water targets, and thus the statistics on water for that run are reduced by one sixth. This is accounted for in the Monte Carlo simulation used for this analysis. Similarly a leak in water panel 5 occurred in 2020, which required the panel to be drained; the FGD continued to be run with four out of six water panels for run 10. However this has no effect on this analysis, as data from run 10 is not used due to additional issues with the magnet.

3.2.2.4 Electromagnetic Calorimeters

The two separate sections of ND280, the PØD and the tracker region, are both surrounded by electromagnetic calorimeters (ECals). These can be grouped into three main parts: the PØD ECal surrounding the PØD, the barrel ECal (BrECal) surrounding the tracker region (FGDs and TPCs), and the downstream ECal (DsECal), which sits just downstream of TPC3. The primary aim of the ECal modules is to provide measurements of neutral particles[6] and electron-positron showers, as well as energy reconstruction and particle identification. It is also possible to use the lead radiator to make neutrino cross-section measurements [160].

[6]In the case of the PØD ECal, it assists where showers are not fully contained within the PØD.

Each ECal module is composed of layers of polystyrene scintillator bars with a cross section of 40 mm × 10 mm, glued to 1.75 mm thick (4 mm in the PØD ECal) sheets of lead, which act as an absorber to incite electromagnetic showers. The width of both the scintillator and lead layers, along with the number of each, were chosen to optimise the π^0 reconstruction efficiency and required number of radiation lengths, X_0, although the design was ultimately limited by space: the PØD ECal and BrECal modules are contained within the space between the inner detector frame (known as the basket) and the magnet, to which the modules are attached. The DsECal is also held within the basket. In total, the ECal consists of 13 independent modules. The BrECal has six modules, with two on the top and bottom where the magnet yoke opens for access, as can be seen in Figure 3.9. Each BrECal module is formed of 31 lead-scintillator layers, which corresponds to a total of 9.7 X_0. The single DsECal module has 34 lead-scintillator layers, equating to 10.6 X_0, where each layer is formed of 50 bars, which run in alternating directions between layers. The final six modules, making up the PØD ECal, have a slightly different design, with six scintillator layers separated by five 4 mm lead layers, giving 3.6 X_0. The number of radiation lengths required is much fewer here, as the PØD is already a dedicated π^0 detector. An example of the ECal module design is given in Figure 3.14.

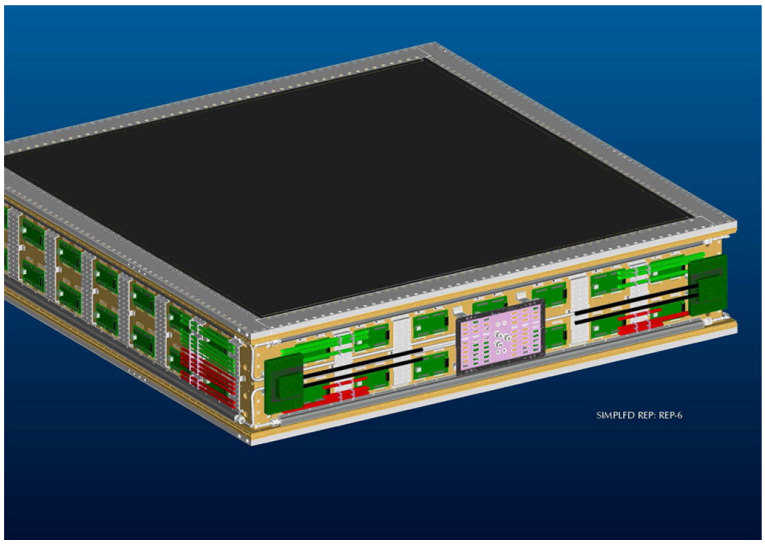


Figure 3.14: ECal module design, showing scintillator bars running horizontally. The aluminium plates which frame the ECal module are used to hold the readout electronics, whilst the grey on the top shows the carbon fibre layer used to provide support. Figure taken from [137].

Similar to the FGD design, each bar of polystyrene scintillator has a TiO_2 coating, and is threaded with a wavelength shifting fibre to carry the signal to the ends, where they are detected by MPPCs. The MPPC signals are then read out by Trip-T [161] front-end electronics boards (TFBs), which are held on the outside of the module. A carbon fibre layer is glued to the lead-scintillator on the inner side of each module, in order to provide mechanical stability without putting too much material in the way of the ECal target [137].

3.2.2.5 Side Muon Range Detector

The side muon range detector (SMRD) is another scintillation light detector, which is embedded in the gaps between the separate sections of the UA1 magnet yoke. The SMRD has several aims. Firstly, placing the SMRD components in the air gaps within the magnet ensures that muons escaping the tracker region at high angles do not get missed by ND280. Secondly, the fact that the neutrino beam is not point-like means that interactions are able to happen outside of the FGD targets. The SMRDs can help to identify beam interactions that have occurred outside the cavity walls, or in the high density iron of the magnet, which can help to veto magnet events that then enter the tracker region. Finally, they are also used to monitor external events, such as from cosmic rays which penetrate ND280 [162]. The layered scintillator used is much the same in design as that of the FGDs in Section 3.2.2.3 and ECals in Section 3.2.2.4, albeit with a slab-like structure depending on the size of the magnet yoke gaps, and thus is not described here.

3.2.2.6 Trip-T Electronics Configuration

Due to the similar structure of INGRID and the ND280 ECal, PØD and SMRD subdetectors, all four systems make use of the Trip-T electronics read-out system [160], the structure of which is described here. Each of the MPPCs which read out the signals from wavelength shifting fibres are connected to Trip-T application-specific integrated circuits (ASICs), which can have at most 16 MPPCs connected to each. Each Trip-T ASIC is then read out to a Trip-T front end board (TFB), where one TFB is capable of reading out four Trip-T ASICs simultaneously. The back-end section of the electronics readout starts with the Readout Merger Modules (RMMs), each of which is capable of merging the data received from up to 48 TFBs. Along with combining and reading out the signals received from TFBs, each RMM is also capable of controlling and issuing triggers to the TFBs it is linked to.

Trigger control is distributed initially from the single ND280 Main Clock Module (MCM), which can receive triggers from various sources. The main one is the trigger received when a signal is issued from the beamline GPS clock system, indicating that a beam spill has been sent. Beam spill triggers have priority over all other triggers. Whilst outside of a beam spill, the MCM will cycle through other triggers, including pedestal readout, cosmic ray triggers and light injection [163]. When a trigger is registered by the MCM, it gets fanned out to the various Secondary Clock Modules (SCMs). Each of the Trip-T detectors has one SCM assigned to it. The SCMs then pass the trigger to the RMMs, which control the TFB hierarchy are described above. It is also possible for the individual SCMs to be taken out of the global configuration and run separately, which is useful both for calibration runs and debugging exercises. In this case, the SCM of the subdetector acts like the MCM.

3.2.2.7 Data Taking and Event Simulation

The T2K experiment first began taking data in January of 2010 [164], and has been running regularly ever since with only a small number of unscheduled interruptions[7]. Data-taking is quantified by the number of protons on target (POT), which at time of writing has reached 2.17×10^{21} in FHC mode and 1.65×10^{21} in RHC mode [165]. Figure 3.15 shows the J-PARC MR beam power and the associated accumulated POT for T2K's 11 years of running, with each run period shown by red bands.

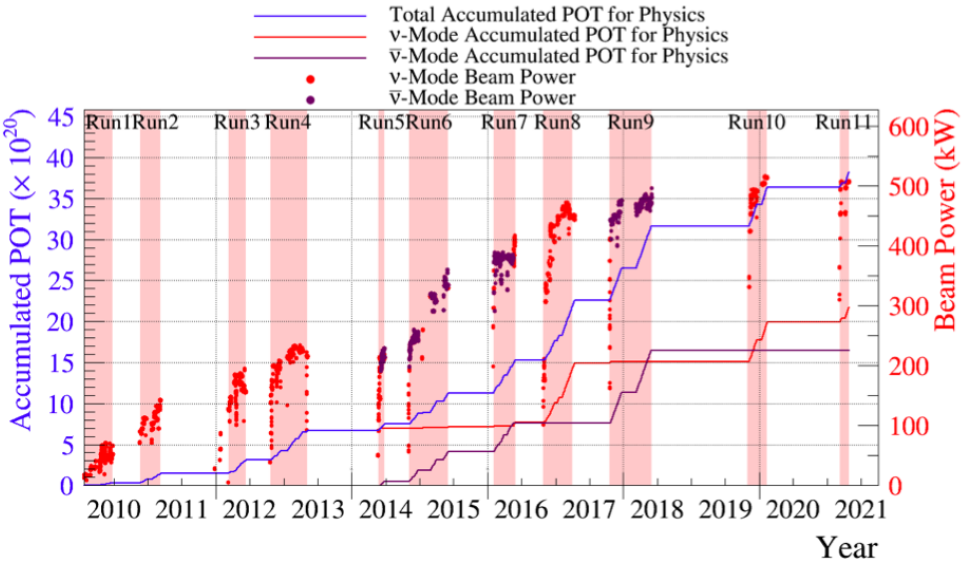


Figure 3.15: Plot of J-PARC MR beam power, and the obtained POT from the beam, as a function of year, from the start of T2K operation. The total accumulated POT is shown along with the individual ν-mode (FHC) and $\bar{\nu}$-mode (RHC). Red bands show the separate T2K runs.

As this thesis focuses on a measurement of a *neutrino* cross section, the relevant data runs are those where the magnetic horns were run in FHC mode: runs 2, 3, 4, 8, 10 and 11. However, run 10 is excluded from this analysis due to cooling problems with the UA1 magnet, which resulted in it not being switched on for that entire data-taking period; the magnetic field it usually provides is crucial for the charge discrimination required to identify negative muons and positive pions in the analysis presented. Additionally, run 11 occurred during the height of the COVID-19 global pandemic, when the majority of international travel was suspended. The inability of detector experts to get to Japan to carry out maintenance on and operation of ND280 resulted in it not being used in run 11, and as such there is no relevant data available. Therefore, the T2K run periods used are runs 2, 3, 4 and 8. Individual POT for each run is discussed further in Section 5.5.

In order to properly develop selections and assess the efficiency without being biased towards underlying data distributions, it is important to have a sophisticated Monte Carlo

[7] T2K run 2 was stopped short by the 2011 Tōhoku earthquake, often referred to in Japan as the 'Great earthquake disaster of East Japan' (東日本大震災).

simulation of the detector and the neutrino interactions which can occur within. This is achieved through a multiple stage process. First, flux files are used to propagate neutrino vectors through a flux plane in front of ND280 and across the detector. For each vector, the probability of interaction is calculated, which depends upon the target material that the vector is propagated through, along with the exact off-axis angle of the neutrino in question. Looping back over the vectors, the one which maximises the interaction probability is chosen, and running the neutrino interaction simulation program `NEUT` [94] provides the list of outgoing particles from the interaction at a particular set of coordinates, where final state interactions (FSI) are simulated using a cascade model [166].

Following the simulation of the specific neutrino interactions, the output particles are propagated through the detector simulation using `Geant4` [167], with the associated energy loss also being simulated. This is then passed to `elecSim`, the ND280 detector response simulation, where the response of the detector to each event is simulated, with the effect of dead electronics channels included. The output of this simulation is therefore functionally identical to the output of the ND280 response to real data, albeit with the true information also included for reference.

In addition to neutrino interactions taking place within the ND280 detector, they can also occur in the sand surrounding the pit that the near detector complex lies in. These interactions create so-called 'sand muons', which on entering the detector can be reconstructed and possibly confused with muons from in-detector events. To assess the affect of the beam-related background, separate sand muon simulations are produced and processed in the manner described above (see also Sections 7.1.1.10 and 7.1.1.13).

3.2.3 WAGASCI-BabyMIND

Despite the success of the near detectors at reducing the systematic error on predicted event rates for different neutrino oscillation modes, some uncertainties still remain. The largest of these is a non-cancelling cross-section model uncertainty, which is caused by the difference in target material between the near and far detectors, along with the limited coverage of ND280, whereas Super-K has full angular coverage [168]. The new WAter Grid And SCIntillator (WAGASCI) detector has been added to the near detector complex, with the long-term goal of using measurements of the cross-section ratio between water and scintillator targets to reduce the T2K systematic error.

The WAGASCI detector itself has two main elements. The central part of the detector, which is the target for neutrino interactions, is composed of plane and grid scintillators, where the gaps in the grid scintillator are filled with water. This is shown diagrammatically in Figure 3.16, and leads the target mass to be a roughly 80:20 ratio of water to hydrocarbon. Similarly to the scintillator subdetectors used in ND280, the plastic scintillator bars are instrumented with wavelength shifting fibres which carry the light to MPPCs. However, the MPPCs used for WAGASCI are a new generation version, featuring a lower dark noise rate and low rate of afterpulses [168]. In order to identify and measure the momentum of muons produced in interactions, the main target mass of WAGASCI is flanked on both the left and right sides by muon range detectors, referred to as WallMRDs. The WallMRDs are

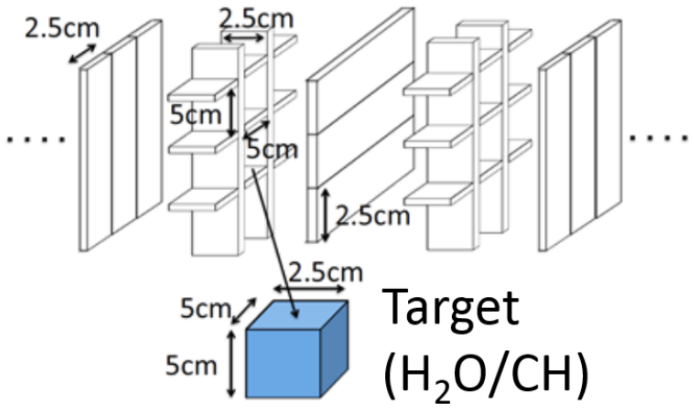


Figure 3.16: Diagram showing the main target mass of WAGASCI, composed of scintillator layers, and scintillator grids which are filled with water. Neutrino beam axis follows the dotted lines. Figure taken from [169].

made of layered iron and scintillator, and are able to estimate the direction of reconstructed tracks using time of flight information.

Since it began running, WAGASCI has been through several different configuration changes. Whilst it originally started with an on-axis placement, the detector now sits at 1.5° off-axis, putting it closer to the neutrino flux seen by ND280 and Super-K. The peak of the neutrino energy spectrum at this off-axis angle is ∼700 MeV. In addition, several other detector modules have been added to the WAGASCI configuration. First is the prototype Magnetised IroN Detector (BabyMIND), which is made of alternating layers of plastic scintillator and iron plates [170]. The iron plates in BabyMIND are magnetised using regular conducting aluminium coils on the surface of each plate, which allows the charge of traversing particles to be identified via their trajectories.

Secondly, the proton module, originally used as part of INGRID and as such described in Section 3.2.1, now sits sandwiched between two WAGASCI target modules [171]. As a completely hydrocarbon target with high resolution tracking, the proton module is used to aid in subtraction of background events to neutrino interactions on the water targets of WAGASCI. The first physics run of the newest WAGASCI-BabyMIND setup has now been completed, and cross-section results using the data are beginning to emerge, such as the one presented in [172].

3.3 Far Detector

The far detector for the T2K experiment is the Super-Kamiokande detector [173], which is situated 295 km to the west of the neutrino beam production point, at the first neutrino oscillation maximum for the ∼600 MeV beam. Super-K is a 50 kton cylindrical water Cherenkov detector, built 1 km underneath Mt. Ikeno (池ノ山), within the Mozumi Mine of the Kamioka Mining and Smelting Co. A diagram of the detector and the access tunnels

used to reach it is presented in Figure 3.17. Super-K consists of two sections, known as the Inner Detector and Outer Detector (ID and OD respectively), which are separated by a cylindrical scaffold covered in plastic sheets, in order to keep the two regions optically separate [174]. Both the ID and OD are instrumented with a total of roughly 13,000 photo-

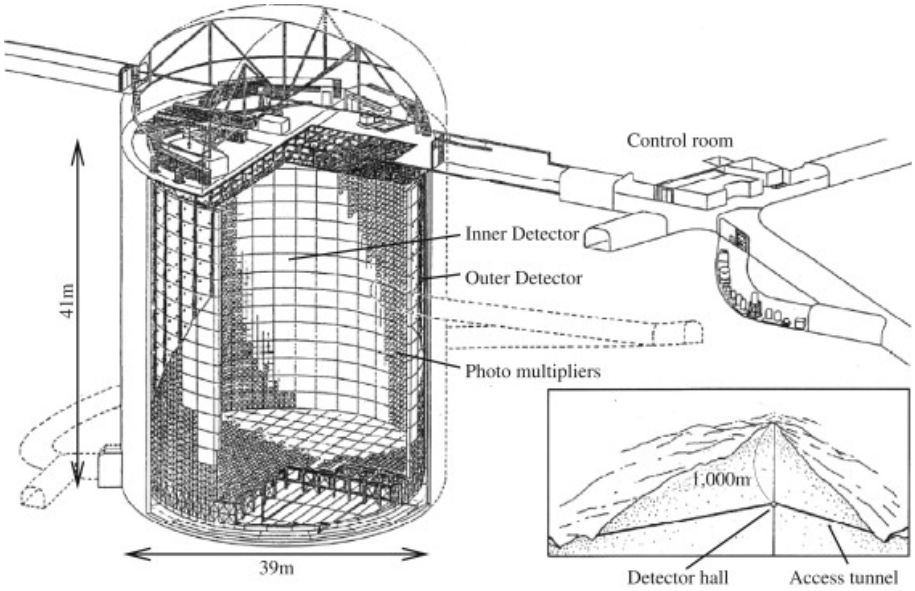


Figure 3.17: Diagram of the Super-K structure and associated access tunnels.

multiplier tubes (PMTs), which are used to image neutrino interactions taking place within the target medium. In the ID, 11,129 50 cm diameter Hamamatsu R3600 PMTs [175] are used, which have a combined quantum and collection efficiency of roughly 20%. This number of PMTs equates to roughly 40% photocathode coverage, which provides high enough spatial resolution to infer lepton flavour type. Neutrinos interacting in the water can produce charged particles, which depending on the incoming neutrino energy may travel faster than the speed of light in water. If they do, the resulting Cherenkov radiation forms ring-like patterns on the walls of the detector, which are observed using the PMTs. PMT images of the ring-like structures are used to separate electron- and muon-induced rings, where muon rings are observed to be relatively clean, as opposed to the 'fuzzy' rings from electrons due to the multiple scatterings they undergo as they traverse the detector medium. This can be observed in Figure 3.18. Temporal and charge information from the ring pattern can also be used to extract information about the event vertex position and the momentum of product particles [137].

The Super-K OD, a 2 m space surrounding the ID on all sides, is instrumented with 1,885 20 cm PMTs, facing outwards. To compensate for the much lower PMT coverage, the walls of the OD are lined with reflective Tyvek sheeting, so that photons striking the walls will have a chance of reflecting and reaching an OD PMT. The sparsity of the OD PMTs is due to the OD's initial purpose of acting as a veto region for cosmic ray muons and other backgrounds entering the detector. However, selecting events occurring in coincidence with the T2K beam allows beam events to be separated from background. The OD information is then used in the T2K data reduction. Neutrino interaction events are labelled as 'fully-

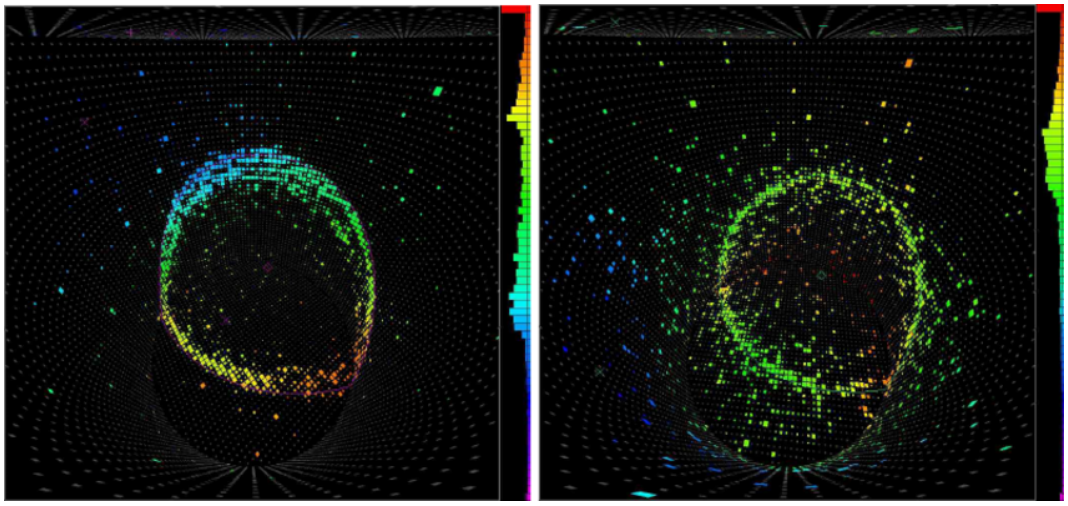


Figure 3.18: Comparison of simulated Cherenkov rings in Super-K. Left: a muon leaves a ring with a 'clean' outer edge. Right: an electron leaves a 'fuzzy' ring, due to the multiple scatterings and electromagnetic showers it can undergo. Colouring corresponds to time of signals. Figure taken from [176].

contained' (FC) when they stop within the ID; for this to be achieved, FC events are required to have no more than 15 hits in the largest OD cluster. Events which fail this cut are instead classed as 'outer-detector' events.

As this thesis focuses on a measurement using the T2K near detector ND280, further details on the Super-K MC simulation and reconstruction softwares are omitted here. A comprehensive overview of these, and their uses in the T2K oscillation analysis, can be found in [137, 177, 178].

Chapter 4

ND280 Timeslip Systematics

This chapter describes the assessment of systematic uncertainties associated with timeslips between ND280 subdetectors, which occur randomly during operation. The effect of these timeslips on time of flight information between subdetectors when simulated in Monte Carlo is examined, and an updated systematic treatment applicable to the current error calculation is developed. Section 4.1 describes the use of time of flight in ND280 analyses, along with an overview of the observed data–MC differences and how they are currently dealt with. Section 4.2 gives a brief overview of the ND280 electronics and how timeslips can occur in data, whilst Section 4.3 estimates how often this happens. Finally, Section 4.4 looks at how to take into account the effect of timeslips within analyses.

4.1 Time of Flight Measurements in T2K

In reconstructing neutrino interactions in a detector, many different variables and quantities can play important roles. Whilst tracks and showers are generally reconstructed from the charge deposition they leave, without timing information the direction of tracks is hard to estimate; a line between two subdetectors could conceivably go in either direction, until we know which detector the hits occurred in first. Similarly, curvature of tracks in the surrounding magnetic field is used to discern the polarity of the associated particle, but this again relies on timing information, as a positive track travelling forwards in the detector will exhibit the same curvature as a negative particle travelling backwards. Time of flight measurements can be additionally useful in reducing background and out of fiducial volume contributions to selected signals, by vetoing objects whose time of flight between subdetectors is far from an expected value.

The ND280 detector is mostly suited to forward-going measurements, due to its limited efficiency in high-angle and backward-going regions. However, as the Super-K far detector has a full angular acceptance of 4π, it is important to try and work with the high-angle and backward-going regions as much as practically possible, as these regions may contain interesting physics which is currently less accessible. A good example of this can be seen in a recent CC1π^+Np cross-section measurement by T2K, of transverse kinematic imbalance [133]. As discussed in Section 2.6, transverse kinematic variables can be a very useful probe for examining final state interactions in cross-section measurements. The result reported,

however, found less sensitivity to the transverse boosting angle $\delta\alpha_T$, due to angular phase space constraints ($\theta < 70°$) applied to the outgoing muon, pion and proton. Access to these high-angle regions should give more discriminating power between various generator models. Upcoming cross-section analyses aim to perform measurements with full angular acceptance, for which precise knowledge of the detector timing and its associated uncertainties is required.

4.1.1 Time of Flight Calculation

In order to reconstruct the sense of a track, the raw timing information from ND280 has to be corrected for the following two effects:

1. Light propagation time inside the detector and fibres.

2. Timing offset of each of sub-detector slave clock modules (SCMs).

The first correction is calculated simply using the position of hits within the detector, giving the distance that the light has to travel. To calculate the SCM offsets, sand muons crossing the PØD, all TPCs, FGDs and the DsECal are used. The expected time of flight (ToF) assuming a speed of c is compared with the measured time of flight after the first correction, in order to find the offset. Once the initial corrections have been applied, the time and position of each hit within a subdetector are fitted with a linear function. The values of this function at the first and last hit position are taken as the start and end times, T_1 and T_2 respectively. The average of these values,

$$T = \frac{T_1 + T_2}{2}, \tag{4.1}$$

is taken as the hit time within the sub-detector, allowing the time of flight between detectors X and Y to be be defined as

$$\mathrm{ToF}_{XY} = T_Y - T_X. \tag{4.2}$$

Initially all tracks are forced to be forward going (travelling from lower z position to higher z position) by the reconstruction. In this case, true backward going tracks will exhibit a negative ToF. The sign of the ToF value, and thus the direction the track travels in, is often also referred to as *sense*, which will be used henceforth.

Figure 4.1 shows a schematic with two possible event geometries in ND280 for a muon travelling between FGD1 and the BrECal: forward-going or backward-going. A hit in the FGD is recorded at time t_1, whilst a hit in the BrECal is recorded at t_2. For an event such as this where the muon *truly* goes forward from the FGD1 and interacts downstream in the BrECal, the ToF value can be defined as

$$\mathrm{ToF} = t(z_{\mathrm{end}}) - t(z_{\mathrm{start}}) \tag{4.3}$$

$$\mathrm{ToF}_{\mathrm{fwd}} = t_2 - t_1, \tag{4.4}$$

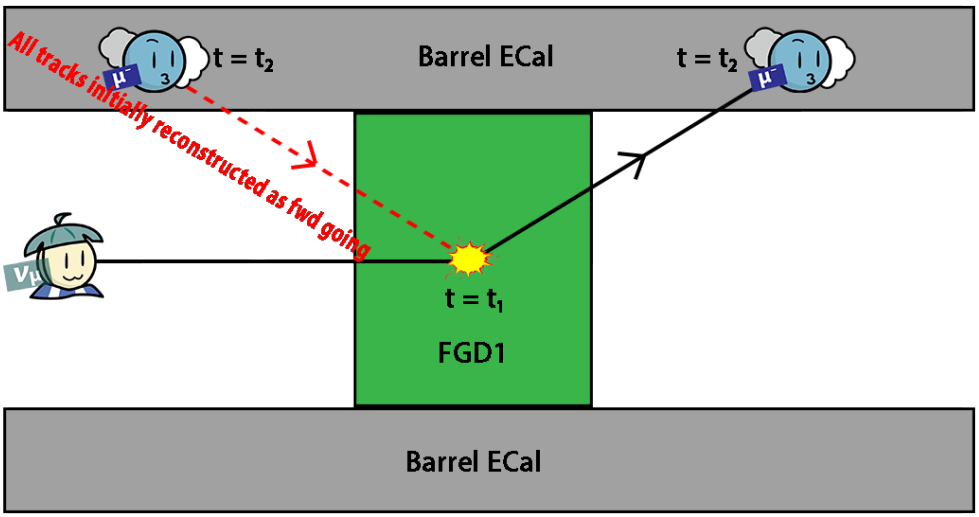


Figure 4.1: Schematic showing two possible event geometries in ND280. The solid black line shows a muon travelling forward from FGD1 to the BrECal, whereas the red dashed line shows a muon travelling backward from FGD1 to BrECal, but is initially reconstructed as going from the BrECal to FGD1 in the forward direction.

where $t_2 - t_1$ must be greater than 0 ns. However looking instead at the *truly* backwards going track, from FGD1 to a point further upstream in the BrECal, the ToF here is defined using Equation (4.3) as

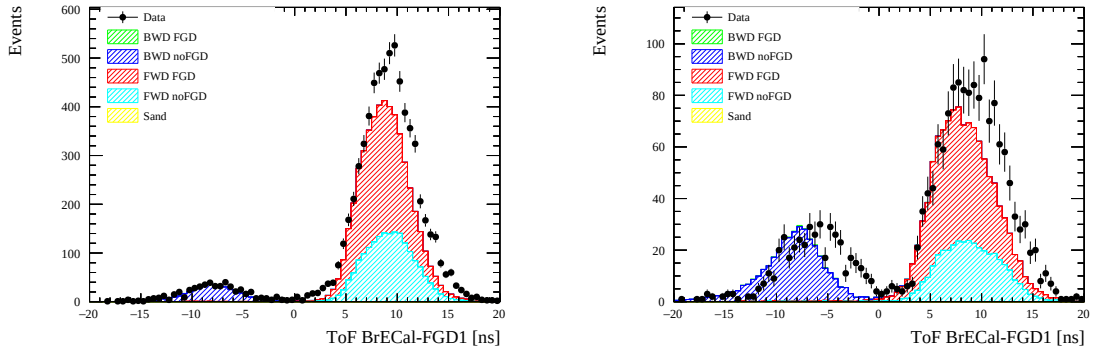
$$\mathrm{ToF}_{\mathrm{bwd}} = t_1 - t_2, \tag{4.5}$$

where $t_1 - t_2$ must be less than 0 ns, as the neutrino interaction truly occurs in the FGD before the muon travels backwards. This negative ToF value tells us that the track is actually backward-going, and therefore should have its sense corrected by being flipped from its initially forward-going state. If instead a beam neutrino were to interact in the BrECal, and the resultant muon travel forwards towards FGD1, the ToF value would again be defined by $t_1 - t_2$ as for Equation (4.5), but in this case would exhibit a positive ToF.

4.1.2 Data–MC Discrepancies

As described previously, the ToF value within ND280 is useful for event topology information, along with reducing external background. However, event reconstruction features (such as differing hit patterns from showers, particle curvature in the magnetic field, and intrinsic detector timing resolution) within each subdetector can cause differences in the calculated ToF value, meaning that timing distributions can vary a lot depending on the features of the events used. Comparing the ToF distributions from Monte Carlo prediction and real data shows this effect. Although this is somewhat expected, it is also found that there are often significant differences between data and MC ToF distributions, even when

track features are similar between both. An example of this is shown in Figure 4.2 for track and shower-like samples in Run 3, specifically for forward going tracks between FGD1 and the Barrel ECal (BrECal). The full set of data–MC distributions is provided in Appendix A.



(a) Track-like (b) Shower-like

Figure 4.2: ToF distribution for forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like. Run 3 MC is plotted and normalised to run 3 data POT, showing the data-MC differences.

4.1.3 Current Approach

The current method for dealing with the data–MC discrepancy is simply to bias and smear the ToF distribution in MC until it resembles the data distribution. To do this, both data and MC distributions are fitted with a sum of two Gaussians, representing the peaks for forward and backward-going beam muon tracks. A correction following a Gaussian distribution is calculated as

$$P(T = t) = \frac{1}{\sigma_{\text{corr}}\sqrt{2\pi}} \exp\left(\frac{(t - \mu_{\text{corr}})^2}{2\sigma^2_{\text{corr}}}\right), \tag{4.6}$$

where μ_{corr} and σ^2_{corr} are given by the differences between the parameter values from the respective data and MC fits:

$$\mu_{\text{corr}} = \mu_{\text{data}} - \mu_{\text{MC}}, \tag{4.7}$$

$$\sigma^2_{\text{corr}} = \sigma^2_{\text{data}} - \sigma^2_{\text{MC}}. \tag{4.8}$$

The calculated correction is then added to the nominal ToF value in MC, giving the corrected value in MC:

$$\text{ToF}_{\text{corrected}} = \text{ToF} + T. \tag{4.9}$$

To calculate these values, a control sample of beam muons is used. This is selected by applying the following cuts:

- **Event Quality Cut** – This cut first checks that the data are usable. ND280 subruns

are assigned data quality flags based on the operational status of sub-detectors at the time of data taking. Data quality flags are required to be good to pass this stage. The events are also required to be in time with the beam trigger.

- **Track Multiplicity Cut** – Events are required to have at least one reconstructed track in one of the FGDs.

- **Track Candidate Available Cut** – All events with tracks which pass the previous cuts have the general TPC track quality cut applied. A good quality TPC track is defined as having over 18 vertical hit clusters in the TPC, in order to reject short tracks where the reconstruction is less reliable. For all tracks in an event that pass this, the detector volume for the start and end positions is checked, allowing the tracks to be sorted into forward and backward going tracks, for FGD1 or FGD2. For tracks that fail the track quality cut, these are again sorted according to their start and end positions for those that have valid ECal tracks. These are the high-angle (HA) forward and backward going samples. Each of the samples has its tracks sorted according to momentum, and the main track in the event is set to the highest momentum track in the first available sample, in the order forward, backward, HA-forward, HA-backward. If the main track exists, then the event passes this cut. A summary of the requirements for each sample is described in Table 4.I.

Sample	Position	Track Quality Requirement
Fwd	Start in FGD1(2)	>18 TPC Hits
Bwd	Stop in FGD1(2)	>18 TPC Hits
HAFwd	Start in FGD1(2) – Stop in ECal	≤18 TPC Hits
HABwd	Stop in FGD1(2) – Start in ECal	≤18 TPC Hits

Table 4.I: Overview of track start/end position and quality requirements for the different control samples identified in the initial selection.

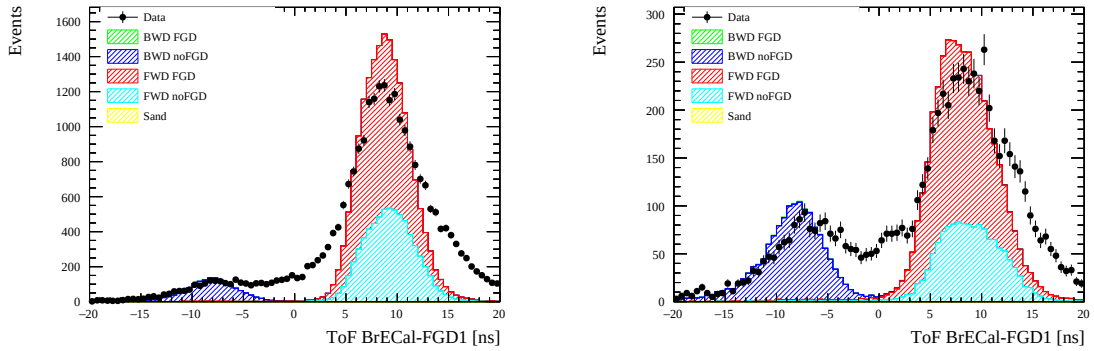
Using the information on the number of ECal segments in each of the samples previously described, a full set of control samples is formed, outlined in Table 4.II. For each of these control samples, the correction necessary to smear MC to data is calculated. This is then applied as a correction at the analysis level, and a systematic uncertainty related to these correction values is calculated.

Although the method described has so far been sufficient in correcting for the data–MC discrepancies observed, issues were observed in Run 8 data–MC comparisons, showing much larger discrepancies than had previously been observed. Investigations into this concluded that the increase in disagreement was large enough compared to previous runs that run-dependent corrections would be required [179]. The ToF distributions with comparisons between data and MC for all samples in Table 4.II are given in Figures 4.3 to 4.7, Figures 4.8 to 4.11, Figures 4.12 and 4.13 and Figures 4.14 and 4.15, for Fwd, Bwd, HAFwd and HABwd

Sample	ToF Sample	
	FGD1	FGD2
Fwd	FGD1–FGD2	–
	FGD1–BrECal track-like	FGD2–BrECal track-like
	FGD1–BrECal shower-like	FGD2–BrECal shower-like
	FGD1–DsECal track-like	FGD2–DsECal track-like
	FGD1–DsECal shower-like	FGD2–DsECal shower-like
Bwd	FGD1–PØD track-like	–
	FGD1–PØD shower-like	FGD1–FGD2
	FGD1–BrECal track-like	FGD2–BrECal track-like
	FGD1–BrECal shower-like	FGD2–BrECal shower-like
HAFwd	FGD1–BrECal track-like	FGD2–BrECal track-like
	FGD1–BrECal shower-like	FGD2–BrECal shower-like
HABwd	FGD1–BrECal track-like	FGD2–BrECal track-like
	FGD1–BrECal shower-like	FGD2–BrECal shower-like

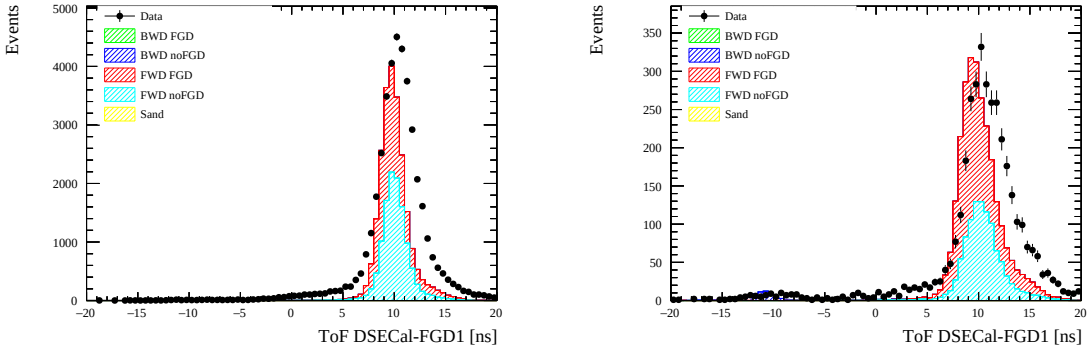
Table 4.II: Full list of samples for which data and MC ToF values are compared.

samples respectively. Each figure is split by whether the main track identified causes a track or shower, for all detector topologies aside from FGD1–FGD2.



(a) Track-like (b) Shower-like

Figure 4.3: Run 8 ToF distribution in data and MC for forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure 4.4: Run 8 ToF distribution in data and MC for forward-going events between FGD1 and DsECal. Samples are split into (a) track-like and (b) shower-like.

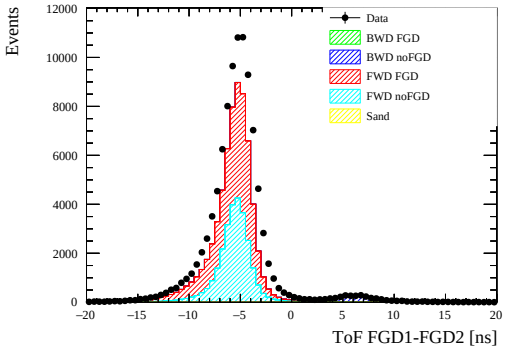
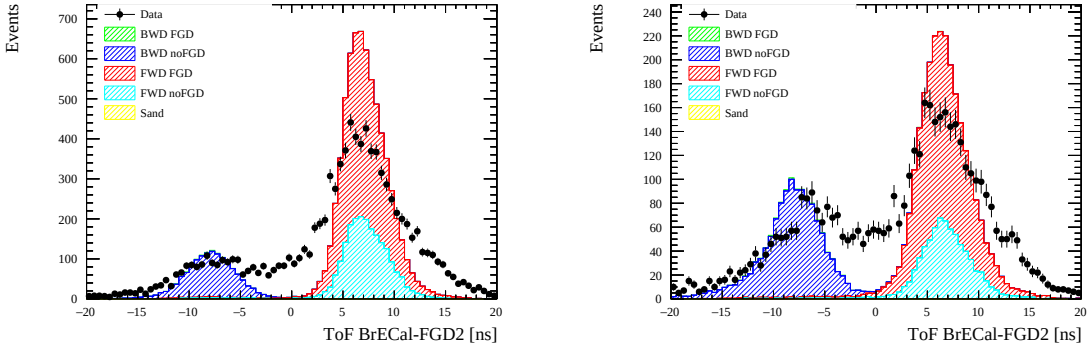
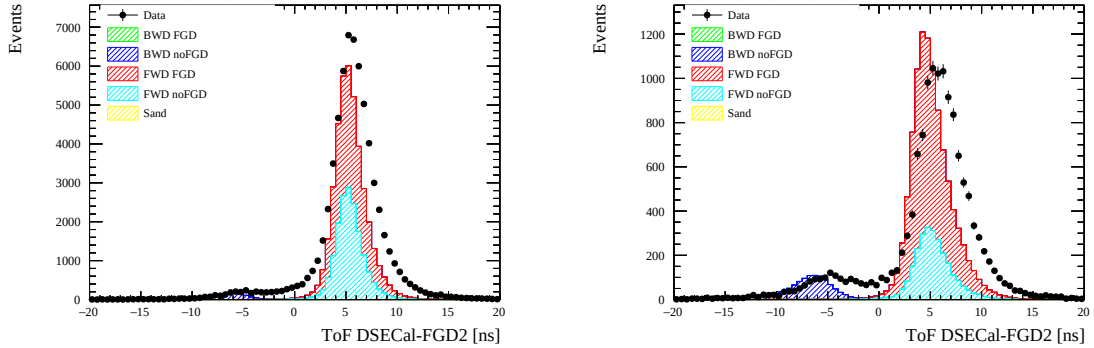


Figure 4.5: Run 8 ToF distribution in data and MC for forward-going events between FGD1 and FGD2.



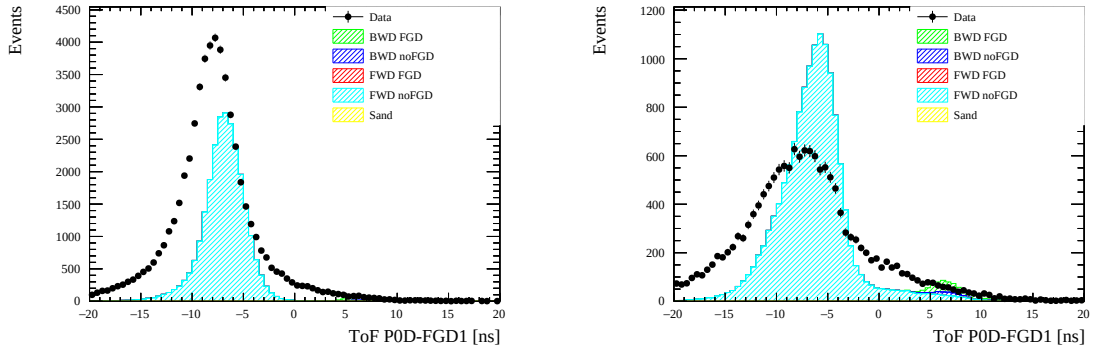
(a) Track-like (b) Shower-like

Figure 4.6: Run 8 ToF distribution in data and MC for forward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.



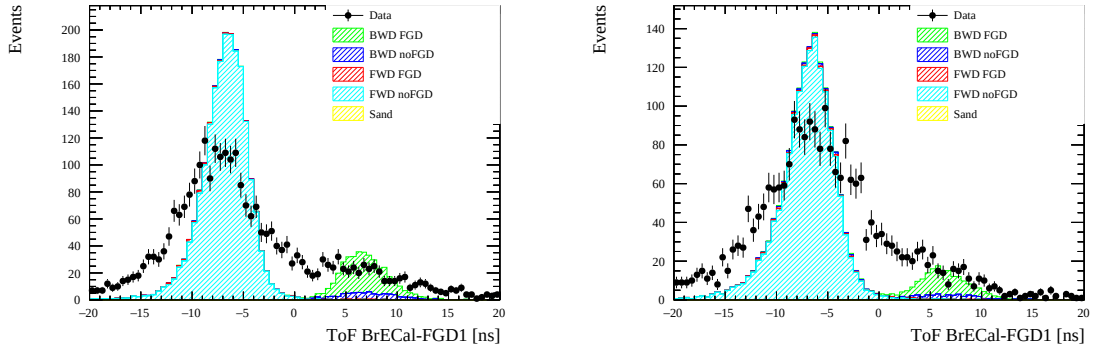
(a) Track-like (b) Shower-like

Figure 4.7: Run 8 ToF distribution in data and MC for forward-going events between FGD2 and DsECal. Samples are split into (a) track-like and (b) shower-like.



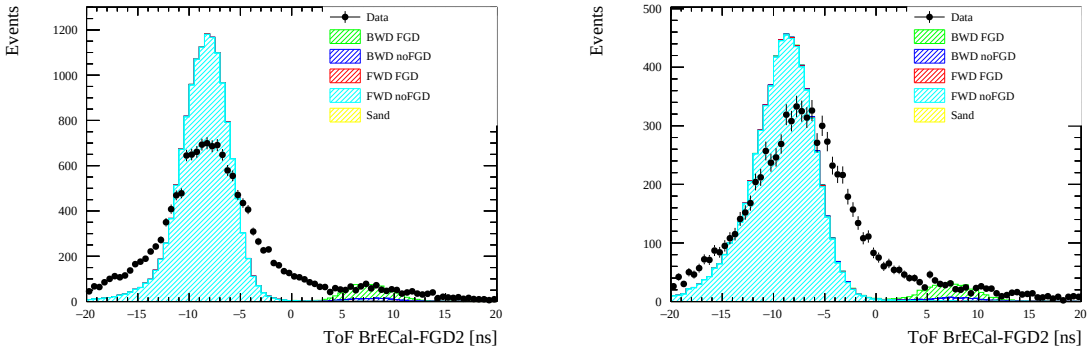
(a) Track-like (b) Shower-like

Figure 4.8: Run 8 ToF distribution in data and MC for backward-going events between FGD1 and PØD. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure 4.9: Run 8 ToF distribution in data and MC for backward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure 4.10: Run 8 ToF distribution in data and MC for backward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.

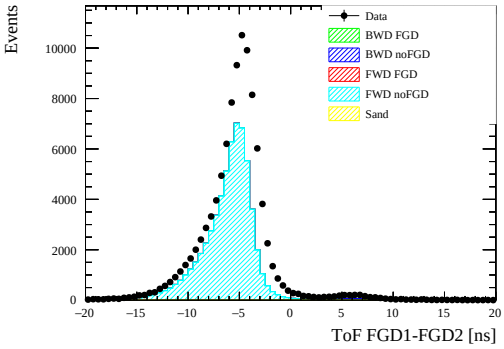
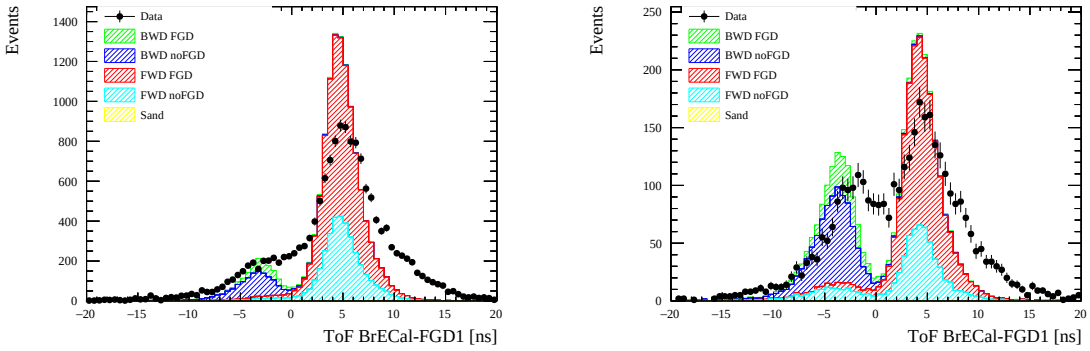
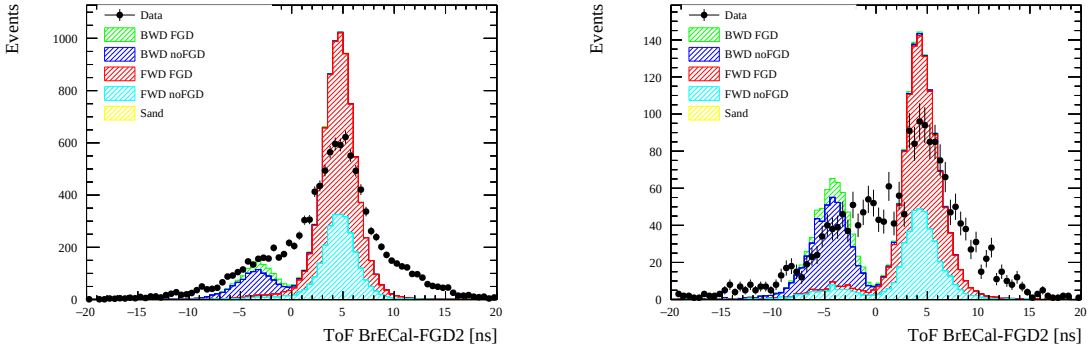


Figure 4.11: Run 8 ToF distribution in data and MC for backward-going events between FGD1 and FGD2.



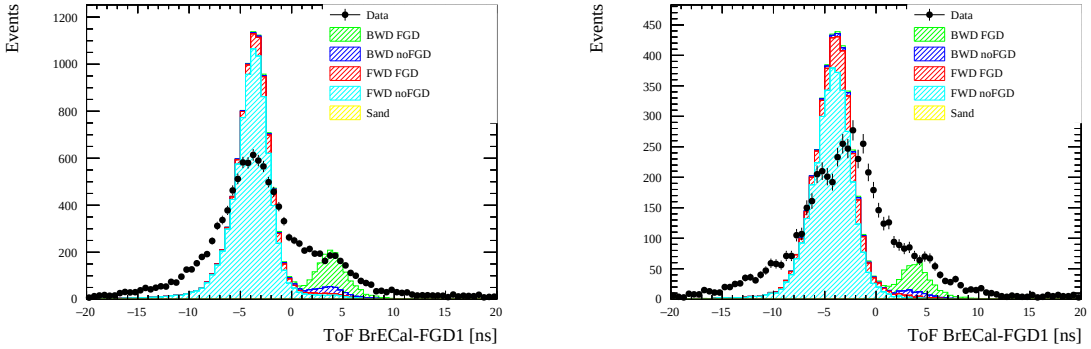
(a) Track-like (b) Shower-like

Figure 4.12: Run 8 ToF distribution in data and MC for high-angle forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



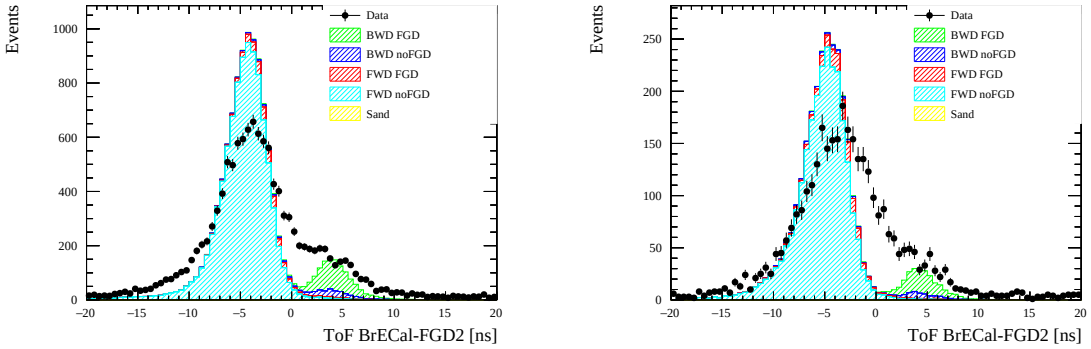
(a) Track-like (b) Shower-like

Figure 4.13: Run 8 ToF distribution in data and MC for high-angle forward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure 4.14: Run 8 ToF distribution in data and MC for high-angle backward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure 4.15: Run 8 ToF distribution in data and MC for high-angle backward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.

4.2 Electronics Setup and Timeslip Appearance

The ND280 ECal, PØD and SMRD subdetectors all make use of the Trip-T electronics read-out system [160], for which the structure is described fully in Section 3.2.2.6. The front-end section of the readout consists of detector signals being collected by Trip-T front end board (TFBs). The back-end section of the electronics readout starts with the Readout Merger Modules (RMMs), which can each collect and merge incoming data from up to 48 TFBs. Along with combining and reading out the signals received from TFBs, each RMM is also capable of controlling and issuing triggers to the TFBs it is linked to. Trigger control is distributed initially from the single ND280 Main Clock Module (MCM) to each ND280 subdetector, via the Secondary Clock Modules (SCMs). Each ND280 detector has one SCM assigned to it. In the case of the Trip-T detectors, the SCMs pass the trigger to the RMMs, which control the TFB hierarchy as described previously.

In order to account for the timing offset caused by differences in cable lengths between SCMs and RMMs, and RMMs and TFBs, calibration was performed using cosmic ray data. Whilst the study in [163] finds the TFB calibration constants to be generally consistent to within ± 2.5 ns, for the RMM calibration constants seemingly random jumps on the order of ± 10 ns are observed. The origin of the timing instability between RMMs is understood to be due to the nature of the connection between them and their respective SCMs, which takes place through Xilinx RocketIO™ transceivers [180]. The connection has a phase-lock loop with a rate of 100 MHz, 4 times slower than the rate at which the TFB clocks operate. Every time the connection between an RMM and SCM is established, the phase lock between the boards also has to be established. When this happens, it is possible for the timing between boards to be shifted by 10 ns by the phase lock, resulting in the appearance of a timing offset. Although this generally happens during power-cycling of the boards, it is also possible that it will occur randomly due to link loss across the optical fibres when the link is restablished. This can result in timing offsets appearing between boards during regular detector operation. The same connections are also used to link the individual SCMs to the MCM, and the MCM to the CTM (cosmic trigger module), which also allows the possibility of an entire subdetector to slip relative to another subdetector. While the analysis in [163] made an attempt at correcting for these timeslips, the developed algorithm often misses slips that happen in quick succession, as opposed to one slip occurring and then multiple ND280 sub-runs featuring the same offset, which is easier to identify. In particular for run 8, where the link loss issue was more prevalent, time slips became so frequent that they were often missed, hence the worse resolution observed there.

4.3 Analysis Procedure

As described in Section 4.2, electronics link losses between the individual RMMS, SCMs and MCM can cause them to 'slip' in time with respect to one another, putting them out by 10 ns. It is thought that these timeslips between boards may be responsible for the data–MC disagreement in the inter-detector time of flight (ToF) values. This particularly came to light after the collection of run 8 data, where significant timing issues were observed during

data taking, and an increase in data–MC disagreement was also observed.

Taking the example in Section 4.1.1 a step further, we can then imagine the effect that timeslips between individual detector SCMs would have on the inter-detector timing. If the SCM on the FGD slips forward by 10 ns, the hit in the FGD is registered 10 ns later, and so the resultant ToF value is reduced by 10 ns. The same effect is observed if the ECal slips backward by 10 ns. This effect can be simulated by taking the nominal MC distribution, and reducing the ToF value by 10 ns. Conversely, the FGD slipping backward by 10 ns or the ECal slipping forward by 10 ns will result in an increase in the ToF by 10 ns. It is also possible to consider the cases where both SCMs slip at the same time. However this is not included in the full analysis due to initial investigations finding relatively small contribution from this ($< 1\%$), and the fact that half of the cases would result in the SCMs slipping in the same direction and no relative time difference being observed.

Due to the fact that these timeslips are not included within the standard MC production, to test the contribution to the observed data, the MC must be shifted and included manually. In order to do this, the MC distribution is simply copied and adjusted to change the ToF. Examples of the nominal MC shifted up and down are given in Figure 4.16. In

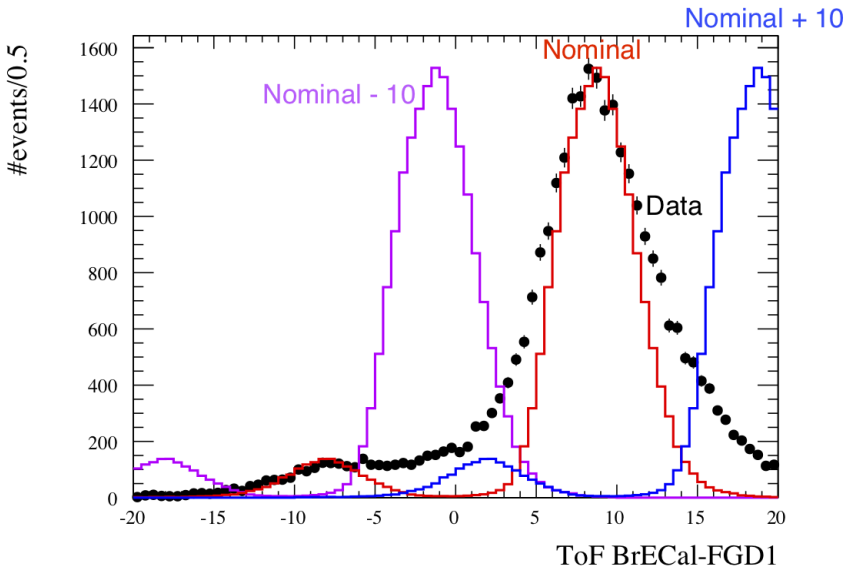


Figure 4.16: ToF distribution in Monte Carlo for the nominal (red), and shifted (purple and blue) values, prior to fitting to data. Data is given in black points. Distribution is shown for the forward-going FGD1–BrECal

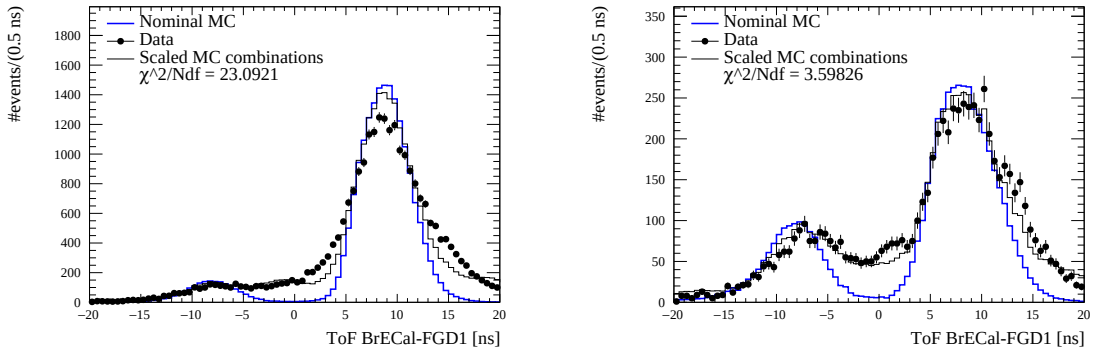
order to find the contribution of the two shifted histograms to the data distribution, the `TFractionFitter` [181] package is used. This allows us to take the three MC samples (nominal, nominal + 10 and nominal − 10), and fit them to the data to find the relative contributions of each. Because timeslips are not observed between FGDs, the FGD1–FGD2 samples are not included in this analysis. The results of these fits for run8 data and MC are given in Figures 4.17 to 4.27. Run 8 was used for initial testing and development due to the more prevalent timing issues observed there. Table 4.III provides the relative contributions

of the nominal and shifted MC to the data distribution for each sample. The described figures show the effect that inclusion of the timeslipped MC has, and in all cases agreement between MC and data is improved by including the scaled contributions from the MC ± 10 ns. Continuing to use the Fwd FGD1–BrECal sample as an example, Figures 4.17a and 4.17b show the distributions for track and shower samples in this detector combination, respectively. Although the timeslip corrected MC does not fit fully to the data, mainly in the forward-going peak where there is still some disagreement, it can be seen that it is in overall closer agreement with the data. This is particularly apparent in the region around 0 ns where the corrected MC is mostly in agreement with data, and is arguably the most important region to understand as the sense is defined by the sign of the ToF. In the shower sample however, the corrected MC is almost completely in agreement with the data points, even at the peaks of both the forward and backward-going distributions. This is a trend that is continued throughout the different samples; the corrected MC generally fits the data slightly better in the shower sample than for the track, although improved agreement is observed in both. Several potential explanations exist for this.

Firstly, the shower samples overall have lower statistics in comparison to track samples, as the muon is a minimum-ionising particle, and is less likely to leave shower-like signals in the ECals. This means simply that the statistical error on data points is higher, and therefore agreement within error is easier to achieve. However, it can be seen from Figure 4.17b that even if the error bars are ignored, the fit to the shower-like sample is still better. Secondly, ECal showers are messy, with hits registering in many parts of the detector due to the scattering of electrons and low-energy pions, which can make the ToF calculation less accurate. The remaining disagreement between corrected MC and data can also potentially be explained by the effect of the TFBs. As well as the 10 ns timeslips between SCMs, it is possible for individual TFBs to slip by 2.5 ns with respect to one another. Due to the number of TFBs associated with each RMM, properly modelling this is not possible. It could potentially be corrected for by applying a suitable Gaussian smearing to each bin, but since the majority of the effect of timeslips seems to come from the SCM slips, TFB timeslips are discounted from the analysis for now.

Sample		MC /%		
Direction	Detector Topology	Nominal	Nominal + 10	Nominal −10
Fwd	FGD1–BrECal track-like	83.9 ± 0.9	8.8 ± 0.3	7.4 ± 0.2
	FGD1–BrECal shower-like	79.2 ± 1.6	11.4 ± 0.7	9.4 ± 0.7
	FGD1–DsECal track-like	97.4 ± 0.8	1.2 ± 0.1	1.4 ± 0.1
	FGD1–DsECal shower-like	96.6 ± 2.5	1.7 ± 0.3	1.7 ± 0.4
	FGD2–BrECal track-like	83.2 ± 1.4	7.2 ± 0.4	9.6 ± 0.5
	FGD2–BrECal shower-like	78.1 ± 2.1	9.0 ± 0.8	12.9 ± 1.0
	FGD2–DsECal track-like	97.8 ± 0.6	1.0 ± 0.1	1.2 ± 0.1
	FGD2–DsECal shower-like	96.4 ± 1.2	2.0 ± 0.2	1.6 ± 0.4
Bwd	FGD1–PØD track-like	92.5 ± 0.7	2.2 ± 0.1	5.2 ± 0.1
	FGD1–PØD shower-like	79.6 ± 1.1	5.9 ± 0.4	14.5 ± 0.4
	FGD1–BrECal track-like	79.7 ± 2.6	10.9 ± 1.0	9.3 ± 0.9
	FGD1–BrECal shower-like	79.9 ± 2.9	12.0 ± 1.3	8.1 ± 0.9
	FGD2–BrECal track-like	83.4 ± 1.1	9.4 ± 0.3	7.2 ± 0.3
	FGD2–BrECal shower-like	79.4 ± 1.5	15.2 ± 0.7	5.4 ± 0.4
HAFwd	FGD1–BrECal track-like	91.5 ± 1.2	5.9 ± 0.3	2.6 ± 0.4
	FGD1–BrECal shower-like	93.5 ± 2.4	6.5 ± 0.7	0.0 ± 0.3
	FGD2–BrECal track-like	90.6 ± 1.4	5.2 ± 0.3	4.2 ± 0.4
	FGD2–BrECal shower-like	93.3 ± 3.3	4.5 ± 0.8	2.2 ± 1.1
HABwd	FGD1–BrECal track-like	89.3 ± 1.3	5.9 ± 0.4	4.8 ± 0.3
	FGD1–BrECal shower-like	88.3 ± 2.0	6.6 ± 0.6	5.0 ± 0.4
	FGD2–BrECal track-like	89.2 ± 1.3	7.1 ± 0.4	3.7 ± 0.3
	FGD2–BrECal shower-like	90.2 ± 2.5	7.5 ± 1.0	2.3 ± 0.5

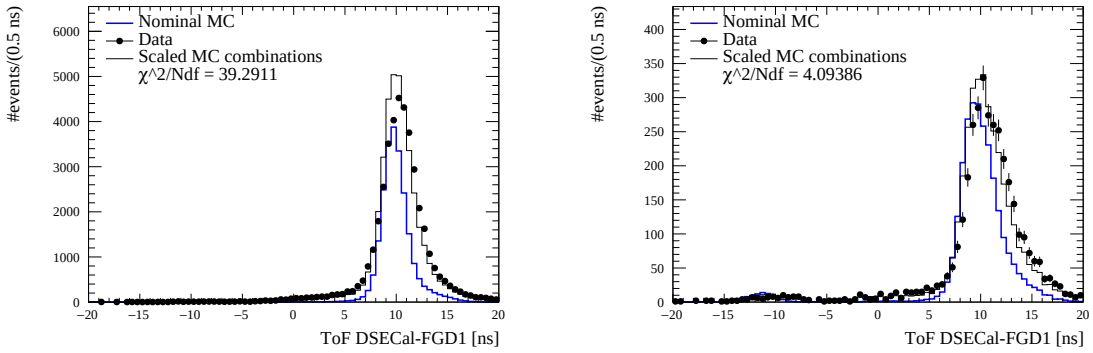
Table 4.III: Relative contributions of the nominal and shifted MC to the data distribution for each detector and track topology sample, for comparisons between run 8 data and MC.



(a) Track-like

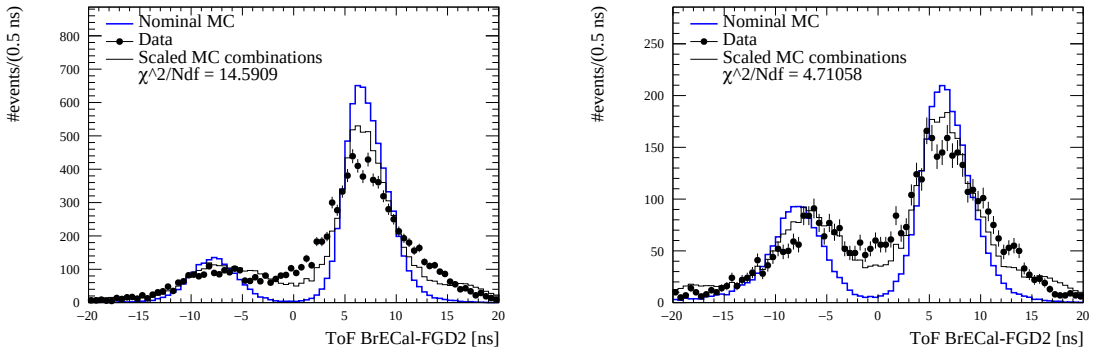
(b) Shower-like

Figure 4.17: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



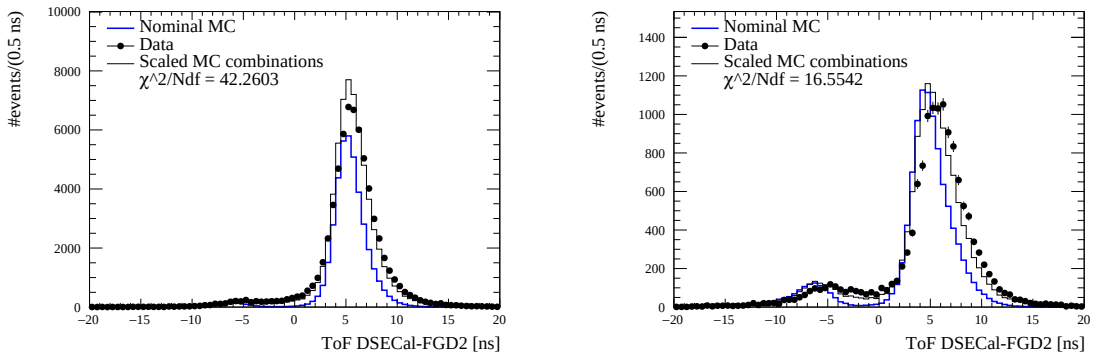
(a) Track-like (b) Shower-like

Figure 4.18: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for forward-going events between FGD1 and DsECal. Samples are split into (a) track-like and (b) shower-like.



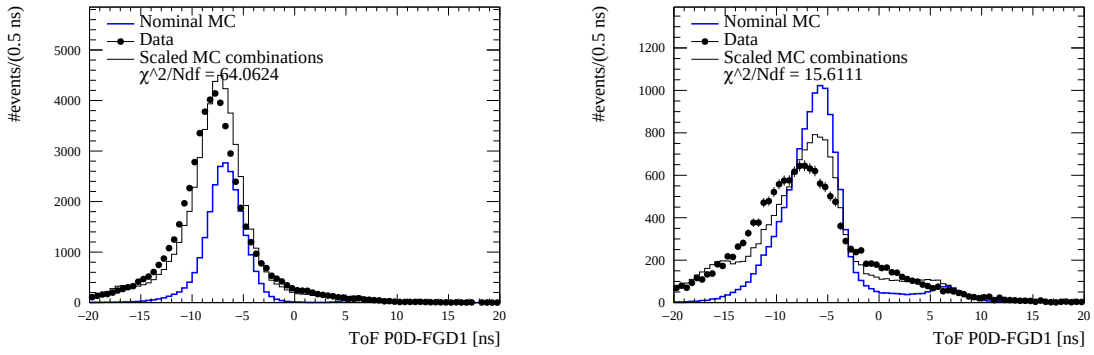
(a) Track-like (b) Shower-like

Figure 4.19: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for forward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.



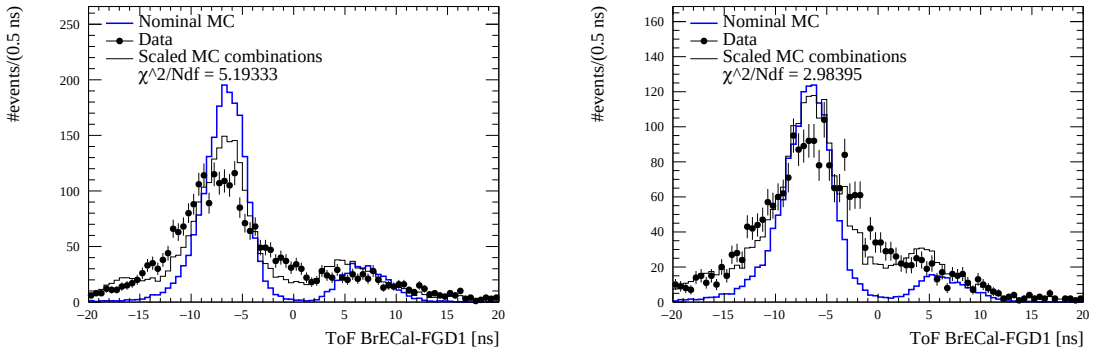
(a) Track-like (b) Shower-like

Figure 4.20: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for forward-going events between FGD2 and DsECal. Samples are split into (a) track-like and (b) shower-like.



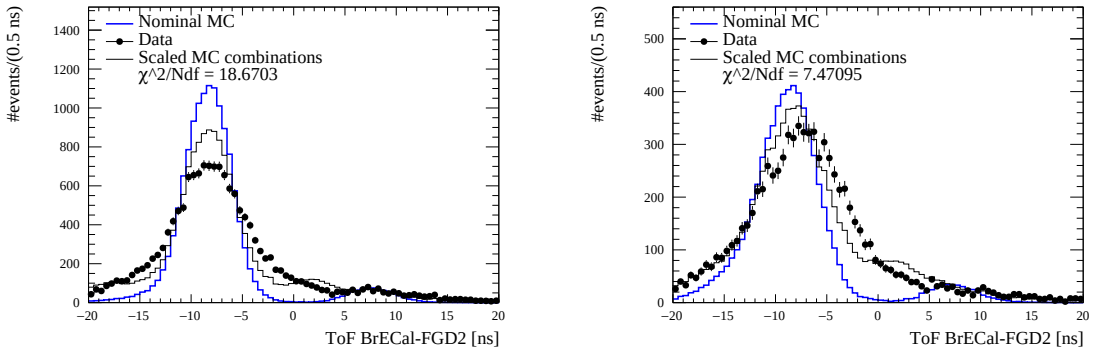
(a) Track-like (b) Shower-like

Figure 4.21: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution in data and MC for backward-going events between FGD1 and PØD. Samples are split into (a) track-like and (b) shower-like.



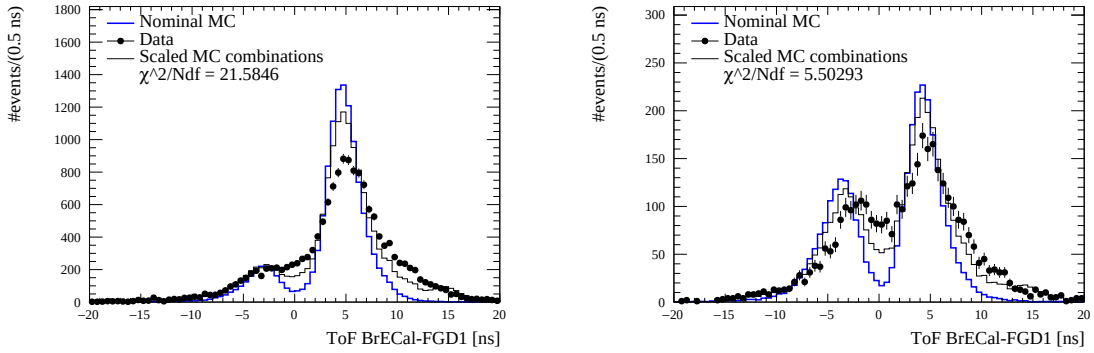
(a) Track-like (b) Shower-like

Figure 4.22: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for backward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



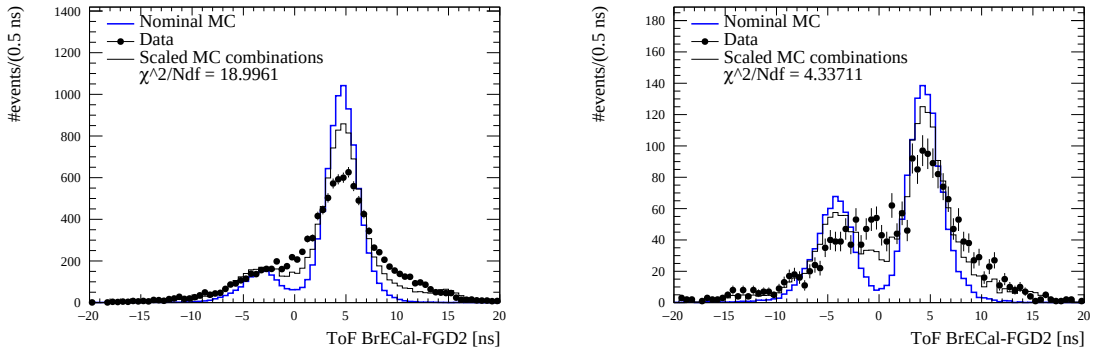
(a) Track-like (b) Shower-like

Figure 4.23: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for backward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.



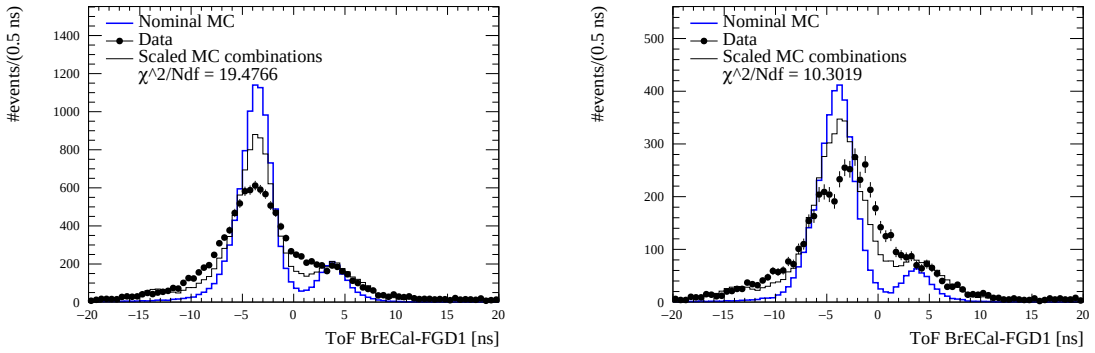
(a) Track-like (b) Shower-like

Figure 4.24: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for high-angle forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



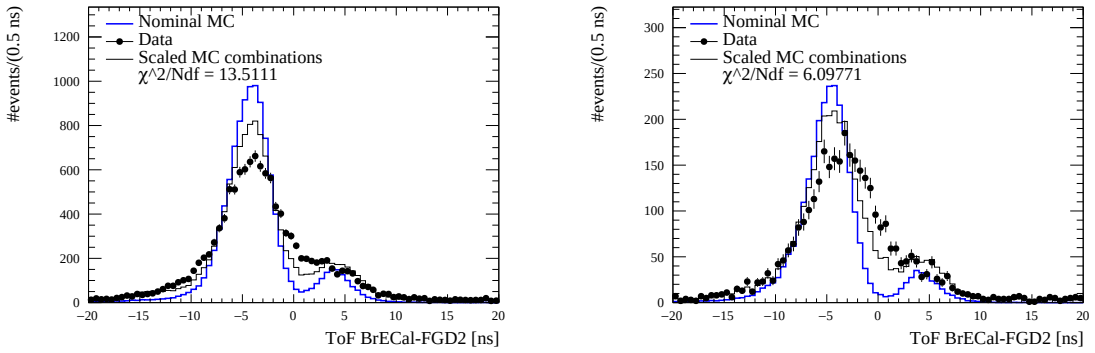
(a) Track-like (b) Shower-like

Figure 4.25: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for high-angle forward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure 4.26: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for high-angle backward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure 4.27: Nominal MC (blue line), scaled contribution from nominal plus shifted MC (black line) and data (black points) for run 8 ToF distribution for high-angle backward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.

4.4 Systematics Evaluation

With the understanding of how timeslips can cause changes in the ToF and how often they are likely to happen, it is important to recalculate the systematic uncertainty on the sense of a track based on these contributions. In order to do this, we have to manually apply timeslips to the nominal Monte Carlo production, and count how often a timeslip causes a change in sense. This is done in two separate steps, the first in order to calculate the percentage of tracks with incorrect sense due to timeslips, the second translating this into a form that can be used by the `psyche` analysis framework to propagate the uncertainty.

4.4.1 Incorrect Track Sense Estimation

As mentioned in Section 4.3, nominal MC production does not simulate the effect of timeslips between SCMs. Therefore in order to calculate how often timeslips cause track sense to be reconstructed incorrectly, they again have to be applied manually. This is achieved using the percentage contributions of shifted MC to data found previously in Table 4.III, which multiplied by the nominal number of MC events gives the number of events which should have their ToF values shifted up or down in time. The method used to properly simulate the timeslips and their effect on track sense is as follows:

- Retrieve the nominal MC distribution and form the associated cumulative event distribution. The `TRandom3` [182] package using the Mersenne Twister algorithm is used to generate a pseudo-random event number, with the bin number and associated ToF value found using the cumulative event distribution. The event number is generated randomly in order to reflect the effect of timeslips, which should occur randomly in the electronics, and not be biased towards any specific events.

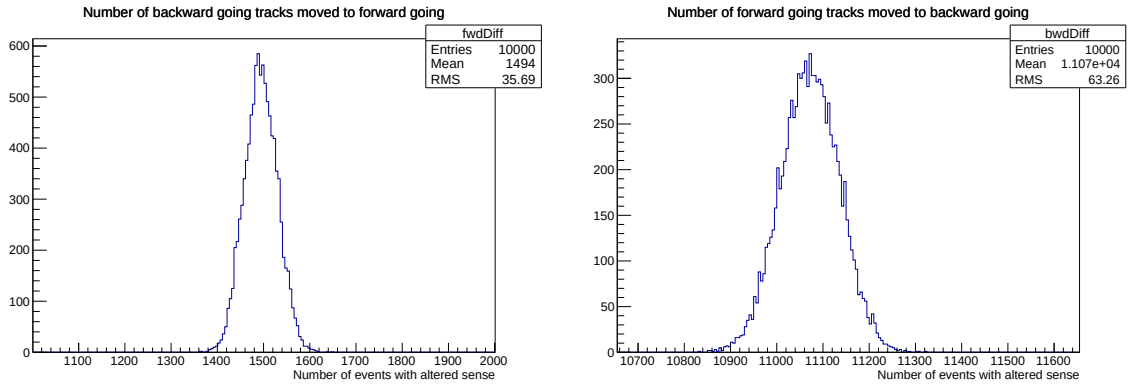
- For the randomly chosen event, the ToF is increased by 10 ns. If the timeslip causes a change of sense in the event, this is recorded. The affected event is then removed from the list of available events. This ensures that timeslips are applied on an event-by-event basis, making it impossible for an event to have a simulated timeslip applied multiple times. This process is repeated for the necessary number of events that should have their ToF increased.

- The same process is applied to the remaining events that have not been affected, this time reducing the ToF of the required number of events by 10 ns. If a change in sense in the opposite direction occurs, this is recorded.

The above process gives an estimation of the number of tracks that change sense due to timeslips, separated into tracks which move from positive to negative ToF and vice versa. The process is repeated a total of 10,000 times, allowing distributions for both sense change totals to be built. Taking the mean of the distributions then gives an accurate estimate of the number of tracks that have incorrect sense due to timeslips.

Returning to the forward-going FGD1–BrECal track-like sample, the method in Section 4.3 finds a contribution of 8.8% from MC shifted up by 10 ns, and 7.4% from MC

shifted down by 10 ns (see Table 4.III). The remainder comes from the nominal MC. Using these fractions and the number of events in the nominal MC distribution ($N_{\text{tot}} = 242471$) we find the number of events required to be shifted up by 10 ns as 21337, and down by 10 ns as 17943. These values are rounded to the nearest integer number. After applying the timeslips on an event-by-event basis, the distributions describing the number of tracks which have their sense changed from negative to positive and vice versa are built; these are given in Figure 4.28.



(a) $- \rightarrow +$ (b) $+ \rightarrow -$

Figure 4.28: Histograms of the number of tracks which have their sense incorrectly changed from (a) negative to positive and (b) positive to negative by applying timeslips to events. Each simulation is thrown 10000 times in order to obtain an accurate mean value.

This gives an average of 1494 events crossing from negative to positive ToF, and 11071 moving from positive to negative. Normalising these by the nominal number of tracks of each sense separately (22124 and 220347, respectively), the percentages of tracks which have their sense changed from negative to positive or positive to negative are found to be 6.75% and 5.02% respectively. The random nature of the timeslips occurring within the electronics indicates that there should be no preference for the direction in which the SCMs slip, but the different characteristics of the forward- and backward-going track samples means that this does not imply that the fraction of tracks which change sense should be the same in both. However, to match the current systematic propagation (described more in Section 4.4.2), we define the percentage of tracks for which a timeslip causes a change of sense, regardless of which direction that change happens in:

$$\%_{\text{change tot.}} = \frac{\mu_{-\rightarrow+} + \mu_{+\rightarrow-}}{N_{\text{tot}}}. \tag{4.10}$$

For the forward-going FGD1–BrECal track-like sample, this gives a result of 5.18% of the total number of tracks having an incorrect sense due to timeslips. The percentages for all samples, both split by sense sign and total, are given in Table 4.IV.

Sample		Tracks change sense /%		
Direction	Detector Topology	$- \to +$	$+ \to -$	Total
Fwd	FGD1–BrECal track-like	6.75	5.02	5.18
	FGD1–BrECal shower-like	7.94	6.47	6.85
	FGD1–DsECal track-like	0.38	0.75	0.75
	FGD1–DsECal shower-like	0.40	0.89	0.87
	FGD2–BrECal track-like	5.85	8.57	8.04
	FGD2–BrECal shower-like	6.71	11.39	9.80
	FGD2–DsECal track-like	0.94	1.20	1.20
	FGD2–DsECal shower-like	1.73	1.52	1.54
Bwd	FGD1–PØD track-like	2.07	4.90	2.14
	FGD1–PØD shower-like	5.15	13.20	5.78
	FGD1–BrECal track-like	9.93	8.30	9.67
	FGD1–BrECal shower-like	10.40	7.22	10.03
	FGD2–BrECal track-like	7.04	6.10	6.98
	FGD2–BrECal shower-like	9.59	4.62	9.26
HAFwd	FGD1–BrECal track-like	5.65	2.57	3.05
	FGD1–BrECal shower-like	6.23	0.00	2.46
	FGD2–BrECal track-like	5.02	4.09	4.24
	FGD2–BrECal shower-like	4.40	2.14	2.95
HABwd	FGD1–BrECal track-like	5.78	4.73	5.62
	FGD1–BrECal shower-like	6.50	4.87	6.30
	FGD2–BrECal track-like	6.86	3.54	6.45
	FGD2–BrECal shower-like	7.16	2.22	6.63

Table 4.IV: Percentage of tracks for which manually simulating timeslips causes a change to the reconstruction sense of the track.

4.4.2 Gaussian Width Approximation

Because the previous form of the correction applied a Gaussian smear to the nominal Monte Carlo in order to make it resemble the data, the systematic propagation within `psyche` is set up to take a Gaussian width as the argument. Although Section 4.4.1 gives an estimation of how often timeslips cause incorrect sense reconstruction as a percentage, it is preferable to translate this into a Gaussian width so as to avoid changing the current framework. In order to do this, we wish to apply a Gaussian smear to the nominal MC in order to cause tracks to change sense, and find the width of the Gaussian distribution that causes the correct number of tracks to change to an incorrect sense. To do this, the following procedure is followed:

- For a certain value w, form a Gaussian distribution centred at 0 with width w in ns.
- For each event in the nominal MC, add a shift to the ToF, where the shift value is randomly sampled from the formed Gaussian distribution. This is again achieved using the `TRandom3` package to sample a random correction.
- Adding the shift to the ToF, check whether the shift causes a change in sense, whether

negative to positive or vice versa. In total this provides a percentage of tracks that have their sense changed due the the application of the Gaussian smear.

This is done for each integer value of w between 0 and 10 ns, with each being checked 10 times in order to find an average percentage of tracks which have incorrect sense. This coarse search finds an integer range within which the optimal value of w lies, which achieves the percentage previously found in Section 4.4.1. The process is repeated again between the optimal set of integers found, in steps of 0.1 ns, with each value of w checked 100 times to achieve a more accurate average.

Plotting the set of w values against the obtained percentages of incorrect track sense, the resultant plot can then be fitted. This is done using a fifth order polynomial, which has no physical meaning but simply represents the response function between the two variables. The function can then be solved for the desired percentage of incorrect sense tracks, in order to find the optimal width w to smear the nominal MC by to achieve this.

For the forward-going FGD1–BrECal track-like sample, the percentage of total tracks which have a change in sense from timeslips is found to be 5.18%. Integer values are tested for w, finding an optimal value between 4 and 5 ns. The analysis then steps back through 4.0 to 4.9 in steps of 0.1, and all widths along with their respective shifted fractions of tracks are plotted. This is shown in Figure 4.29, with the response function obtained from fitting a fifth order polynomial represented by the red curve. The six parameters obtained for this fit are presented in Table 4.V. Setting a fifth order polynomial of this form equal to the required fraction of tracks moving sense then yields the width of the Gaussian distribution required to do this, which in this case is 4.87 ns. Tabulated width values for all samples are presented in Table 4.VI.

Parameter	Value
p_0	-5.19×10^{-3}
p_1	1.38×10^{-2}
p_2	-1.20×10^{-2}
p_3	4.43×10^{-3}
p_4	5.17×10^{-4}
p_5	2.02×10^{-5}

Table 4.V: Fit parameters for a fifth order polynomial, fitted to the simulated data points in Figure 4.29.

4.5 Conclusions

The analysis presented provides a method for which to parametrise the effect of timeslips between ND280 hardware on the data taken by the detector. While a systematic treatment existed for this already, it consisted of simply smearing the nominal MC to match the data, without attempting to understand why the data appears the way it does. Along with showing that the cause of the data–MC agreement is the effect of these observed timeslips, we

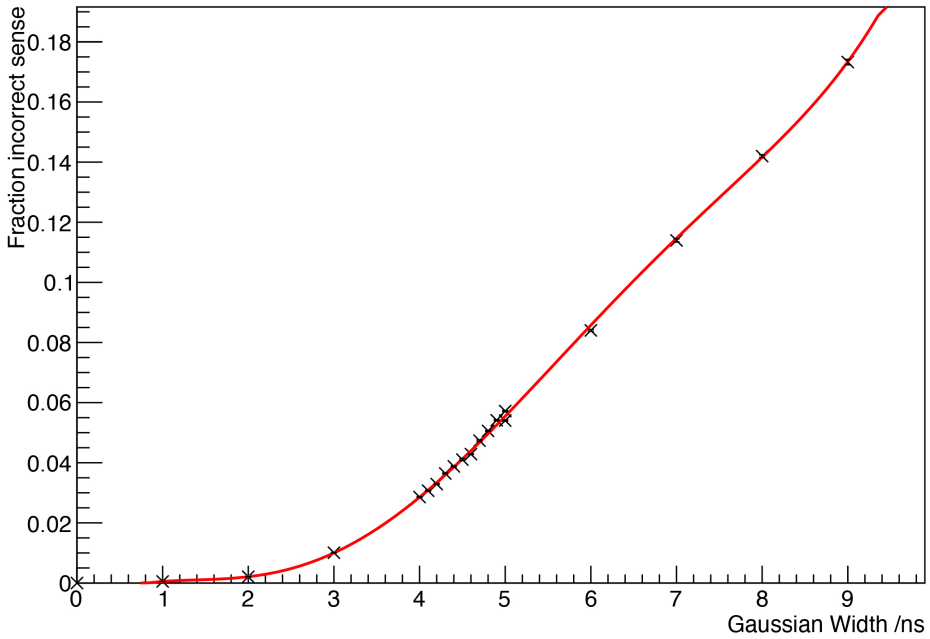


Figure 4.29: Relationship between the width of the Gaussian distribution used to smear events in the forward-going FGD1–BrECal track ToF distribution, and the percentage change in sense that the smearing causes. Black points show simulated response in MC, whilst the red curve is the result of fit with a fifth order polynomial to obtain a response function. Error bars are neglected as the fit is simply used for interpolation.

are able to quantify how often the data are affected, before translating this into a systematic uncertainty that can be propagated in the same manner as the previous treatment.

This analysis is performed on run 8 data, as this was known to suffer most from timing problems. Following application of the analysis, a significant increase in agreement between data and the corrected MC is seen, which suggests that it should be easily applicable to the other ND280 runs which were less severely affected. As timeslips are run-dependent, the described method will have to be applied to all 10 available ND280 runs separately, to calculate the uncertainties for all possible track topologies.

Sample		Gaussian width w /ns
Direction	Detector Topology	
Fwd	FGD1–BrECal track-like	4.87
	FGD1–BrECal shower-like	4.92
	FGD1–DsECal track-like	3.80
	FGD1–DsECal shower-like	3.70
	FGD2–BrECal track-like	4.81
	FGD2–BrECal shower-like	4.88
	FGD2–DsECal track-like	1.91
	FGD2–DsECal shower-like	1.71
Bwd	FGD1–PØD track-like	2.81
	FGD1–PØD shower-like	3.31
	FGD1–BrECal track-like	4.80
	FGD1–BrECal shower-like	4.80
	FGD2–BrECal track-like	5.23
	FGD2–BrECal shower-like	6.06
HAFwd	FGD1–BrECal track-like	1.75
	FGD1–BrECal shower-like	1.38
	FGD2–BrECal track-like	1.95
	FGD2–BrECal shower-like	1.67
HABwd	FGD1–BrECal track-like	1.80
	FGD1–BrECal shower-like	2.08
	FGD2–BrECal track-like	2.20
	FGD2–BrECal shower-like	2.67

Table 4.VI: Optimal values for the width of a Gaussian distribution, which when used to smear uncorrected ToF values gives the percentage of tracks with incorrect sense calculated in Table 4.IV.

Chapter 5

Analysis Strategy

5.1 Signal Definition

The signal required for this analysis is ν_μ charged current (CC) interactions, where the final state includes only one negative muon and one positive pion. There is no restriction placed on the number of nucleons in the final state, as we do not attempt to reconstruct these. Due to limitations in the reconstruction, the signal phase space is restricted so that a measurement is only made for events with

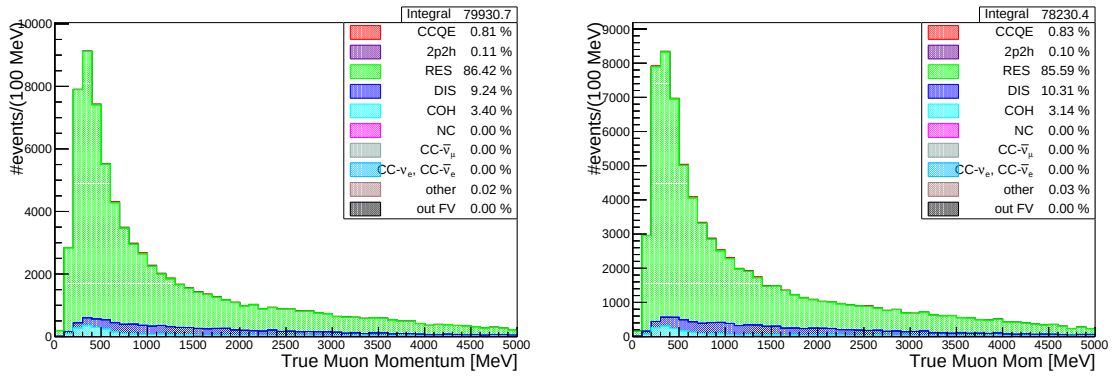
- $p_\mu > 200$ MeV
- $\cos\theta_\mu > 0.3$
- $p_\pi < 1500$ MeV

in the truth space. The reasons for this are discussed in detail in Section 7.2. Further constraints are also applied to individual selection samples, which are discussed in the same section. All plots in Sections 6.1 to 6.4 are drawn without these phase space constraints, so that the full reconstructed phase space can be observed.

5.1.1 Cross-Section Definition

The measurement of the ν_μCC1π^+ cross section on water and hydrocarbon is made using signal interactions occurring within the FGD1 and FGD2 detector fiducial volumes. The choice to measure the cross section in terms of the observed final state topology (CC1π^+) is made in order to avoid dependence on the nuclear model used for Monte Carlo event generation. Measuring an interaction in terms of the true reaction type that took place (CCQE, CCRES, DIS etc.) is hard to do correctly, due to changing topology from final state interactions before the particles exit the nucleus. Instead we choose to measure the interaction in terms of the particles observed in the final state. While a CC1π^+ interaction is predicted in MC to primarily (86.4%) come from resonant interactions (CCRES) creating a single pion, this signal can also be created from Deep Inelastic Scattering (DIS) (9.2%) events where additional pions are reabsorbed before exiting the nuclear medium. A small contribution also comes from coherent pion production events (CCCOH) (3.4%), where the

incoming neutrino interacts coherently with the entire nucleus. Small additional contributions also come from CCQE and 2p2h interactions, although these are mostly negligible. Quoted percentages are for true FGD1 CC1π^+ events prior to any cuts, whilst Figures 5.1a and 5.1b show the full reaction breakdown for the true events in both FGD1 and FGD2, respectively.



(a) FGD1 (b) FGD2

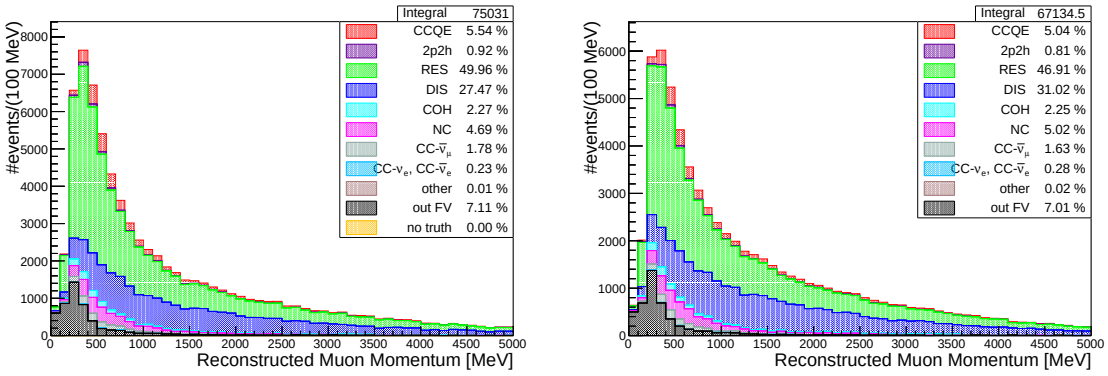
Figure 5.1: True muon momentum distribution for all true CC1π^+ events prior to selection cuts, categorised by true reaction type. Distributions are shown for (a) FGD1 and (b) FGD2 selections.

Along with the described difficulty in accessing the true reaction of an event, it can also be similarly difficult to correctly identify the topology of an event, due to limitations in the reconstruction caused by detector effects. Figures 5.2a and 5.2b show the CC1π^+ samples for the general selection, in FGD1 and FGD2 respectively, for reconstructed MC categorised by the true reaction type of the events. Whilst the primary source of reconstructed CC1π^+ interactions is still RES events, these now only account for 50.0% of the sample. A signal event can also be created from DIS where additional pions interact outside the nucleus or are simply not detected (27.5%), or background resonant interactions where the outgoing pion undergoes charge exchange. Quasi-elastic (CCQE) interactions can also contribute (5.6%), where the outgoing proton is mistaken for a pion due to crossing energy-loss curves, which are used in the particle ID algorithms. Quoted percentages are for the standard FGD1 CC1π^+ selection.

The differential cross section in the i^{th} true bin of variable x is calculated as

$$\left(\frac{d\sigma^{\alpha}}{dx}\right)_i = \frac{N_i^{\text{signal},\,\alpha}}{\epsilon_i^{\alpha}\Phi N_T^{\alpha}\Delta x_i}, \tag{5.1}$$

where $N_i^{\text{signal},\,\alpha}$ is the number of selected CC1π^+ events in the given bin, ϵ_i^{α} is the selection efficiency, Φ is the integrated flux, N_T^{α} is the number of target nucleons within the fiducial volume, and Δx_i is the width of the i^{th} kinematic bin. The index α indicates the target of the interaction, either hydrocarbon (CH) or water (H_2O). Neutrino interactions with other nuclear targets are treated as background. This measurement is made in the four-dimensional space of muon and pion kinematics: p_μ, $\cos\theta_\mu$, p_π and $\cos\theta_\pi$. The full four-



(a) FGD1 (b) FGD2

Figure 5.2: Reconstructed muon momentum distribution for the general CC1π^+ selection, categorised by true reaction type. Distributions are shown for (a) FGD1 and (b) FGD2 selections.

dimensional form of the cross-section definition in Equation (5.1) then becomes

$$\left(\frac{d^4\sigma^\alpha}{dp_\mu d\cos\theta_\mu dp_\pi d\cos\theta_\pi}\right)_{i,j,k,l} = \frac{N^{\text{signal},\,\alpha}_{i,j,k,l}}{\epsilon^\alpha_{i,j,k,l}\Phi N^\alpha_T \Delta p^i_\mu \Delta\cos\theta^j_\mu \Delta p^k_\pi \Delta\cos\theta^l_\pi}. \tag{5.2}$$

In addition to reporting the full four-dimensional cross section, which will be statistically limited due to the number true bins required, we can also integrate over the muon kinematic bins after extraction, to obtain a double-differential result in pion kinematics:

$$\left(\frac{d^2\sigma^\alpha}{dp_\pi d\cos\theta_\pi}\right)_{k,l} = \frac{1}{\Phi N^\alpha_T \Delta p^k_\pi \Delta\cos\theta^l_\pi}\sum_i\sum_j \frac{N^{\text{signal},\,\alpha}_{i,j,k,l}}{\epsilon^\alpha_{i,j,k,l}\Delta p^i_\mu \Delta\cos\theta^j_\mu}. \tag{5.3}$$

This has the advantage of reducing the statistical error on the measurement, whilst still being a novel measurement for T2K.

The measurement is made in terms of the reconstructed outgoing lepton and pion kinematics, since these variables are only smeared by detector effects. This is preferable to making measurements in terms of variables such as four-momentum transfer or neutrino energy, as reconstruction of these also relies on assumptions about the nuclear model; using directly observable variables further reduces model dependency. Finally, the integrated flux Φ is used in the calculation in order to avoid making additional assumptions, in this case with regards to the shape of the neutrino energy spectrum.

5.2 Cross-Section Extraction Strategy

In order to compare the cross-section measurement made in this work to external cross-section models, it is desirable to be able to report the cross section in terms of true kinematic variables. However, the measurement is made on the reconstruction level, in terms of detector level variables. The process used to translate from reconstructed to truth space,

and extract the associated cross section, is detailed in this section.

5.2.1 Unfolding

The cross-section measurement presented is an example of a counting experiment, in which binned event rates are used for calculation. Comparing results to theory predictions in this case is non-trivial, as several effects create differences between the data and the prediction [183]:

- *Detector Effects* – Detector inefficiencies and resolution can cause inaccuracies in reconstruction of kinematic variables. This in turn can cause events to be incorrectly migrated into neighbouring bins, or not be counted at all. This is generally referred to as *smearing*.
- *Statistical Fluctuations* – As a counting experiment, the observed number of events obeys Poissonian statistics. This introduces a statistical uncertainty into the measurement.

The process of moving from the reconstruction level variables and deconvolving these smearing effects, to get back to truth-level comparisons, is known as *unfolding* [183]. In order to do this, we wish to find the unfolding matrix $\mathbf{U}$ which maps the reconstruction level variables R_j to the truth level variables, T_i. This can in general be done by equating $\mathbf{U}$ to the inverse of the smearing matrix, $\mathbf{S}^{-1}$, such that

$$R_j = \sum_{i}^{\text{true bins}} S_{ij} T_i. \tag{5.4}$$

This is achieved using a binned maximum likelihood fit, for which the method is fully described in the following sections. Whilst unfolding via the inverse smearing matrix $\mathbf{S}^{-1}$ is theoretically easy, in practice it can be very error-prone. It is known as an 'ill-posed' problem, in that many possible solutions can exist; a solution which overpopulates the bin T_i can simply compensate by underpopulating bins T_{i-1} and T_{i+1}, with the population of the associated reconstructed bin remaining unchanged. As bin-to-bin migration can happen in either direction, the T_{i-2} and T_{i+2} bins must then also be overpopulated, and so on. This leads to an oscillatory behaviour between bins, which is clearly non-physical. In order to deal with this, regularisation is applied in the unfolding, which is discussed further in Section 5.3.

5.2.2 Binned Maximum Likelihood Method

To determine the number of signal events in a given true kinematic bin, a binned maximum likelihood fit is used, varying a series of parameters which weight the signal in the true bins from Monte Carlo until they best match the data in reconstructed bins. This is done by maximising the likelihood function

$$\mathcal{L} = \mathcal{L}_{\text{stat}}(\vec{x};\vec{\theta}) \times \mathcal{L}_{\text{syst}}(\vec{x};\vec{\theta}) \times \mathcal{L}_{\text{reg}}(\vec{x};\vec{\theta}), \tag{5.5}$$

which returns the set of parameters $\vec{\theta}$ which best describe the data $\vec{x}$. In practice, it is often more convenient to work with the negative of the log likelihood, as minimising this is equal to maximising the log likelihood. According to Wilks' theorem [184], in the limit of large statistics, the minimum of $-2\ln\mathcal{L}$ follows a χ^2 distribution with a number of degrees of freedom equal to the number of bins minus the number of unconstrained fit parameters, allowing Equation (5.5) to be rewritten as

$$\chi^2 = \chi^2_{\text{stat}} + \chi^2_{\text{syst}} + \chi^2_{\text{reg}} \tag{5.6}$$

$$\approx -2\ln\mathcal{L}_{\text{stat}} - 2\ln\mathcal{L}_{\text{syst}} - 2\ln\mathcal{L}_{\text{reg}}. \tag{5.7}$$

As described above, the fact this is a counting experiment leads $\mathcal{L}_{\text{stat}}$ to take the Poissonian likelihood form, where in general the likelihood for a data set $\vec{y}$ to be described by the set of parameters $\vec{n}$ is given by

$$\mathcal{L}_P(\vec{y};\vec{n}) = \prod_i \exp(-y_i)\frac{y_i^{n_i}}{n_i!}, \tag{5.8}$$

where n_i is the number of events in the i^{th} bin, and y_i is the number of events predicted by the model in that bin. Defining the likelihood ratio test value λ as the ratio of the given likelihood to the likelihood to obtain the true values of the n_i if there were no error, and taking $\vec{n}$ as the maximum likelihood estimate of the true values [185], the likelihood ratio value becomes

$$\lambda = \frac{\mathcal{L}_P(\vec{y};\vec{n})}{\mathcal{L}_P(\vec{n};\vec{n})} \tag{5.9}$$

$$= \prod_i \exp(-y_i + n_i)\left(\frac{y_i}{n_i}\right)^{n_i}. \tag{5.10}$$

Applying Wilks' theorem to this, we obtain the χ^2 approximation in the log-likelihood form:

$$\chi^2_P \approx -2\ln\lambda \tag{5.11}$$

$$= -2\ln\left[\prod_i \exp(-y_i + n_i)\left(\frac{y_i}{n_i}\right)^{n_i}\right] \tag{5.12}$$

$$= 2\sum_i \left(y_i - n_i + n_i\ln\frac{n_i}{y_i}\right), \tag{5.13}$$

which in the specific form of this problem can be written as

$$\chi^2_{\text{stat}} = 2\sum_j^{\text{reco bins}} \left(N_j^{\text{exp}} - N_j^{\text{obs}} + N_j^{\text{obs}}\ln\frac{N_j^{\text{obs}}}{N_j^{\text{exp}}}\right), \tag{5.14}$$

where N_j^{exp} is the number of events expected in the j^{th} reconstructed bin from Monte Carlo prediction, and N_j^{obs} is the number observed in data. The systematic term, χ^2_{syst}, takes the

form

$$\chi^2_{\text{syst}} \approx -2 \ln \mathcal{L}_{\text{syst}} \tag{5.15}$$

$$= (\vec{p} - \vec{p}^{\,\text{prior}})^T (\mathbf{V}^{\text{prior}}_{\text{cov}})^{-1} (\vec{p} - \vec{p}^{\,\text{prior}}), \tag{5.16}$$

where $\vec{p}$ is the vector of systematic fit parameters, $\vec{p}^{\,\text{prior}}$ is the vector of prior values, and $\mathbf{V}^{\text{prior}}_{\text{cov}}$ is the related covariance matrix. This allows prior knowledge of theory and experimental data to be encoded into the fit, and to act as a penalty term to the χ^2. The final term in Equation (5.6), χ^2_{reg}, is the χ^2 value due to the regularisation applied in the fit, which is described in further detail in Section 5.3.

5.2.3 Template Parameter Unfolding

To unfold the reconstructed information and retrieve the true, non-smeared variables, the maximum likelihood method described above is applied by assigning free template parameters to each bin in the true distribution. In order to perform a simultaneous extraction, two such parameters are assigned, for signal interactions on different targets:

$$N^{\text{CH, sig}}_{i,\text{ true}} = c_i N^{\text{CH, sig MC}}_{i,\text{ true}} \tag{5.17}$$

$$N^{\text{H}_2\text{O, sig}}_{i,\text{ true}} = o_i N^{\text{H}_2\text{O, sig MC}}_{i,\text{ true}} \tag{5.18}$$

Here $N^{\text{CH, sig}}_{i,\text{ true}}$ and $N^{\text{H}_2\text{O, sig}}_{i,\text{ true}}$ are the predicted number of events in the i^{th} true bin for interactions on hydrocarbon and water respectively, and $N^{\text{CH, sig MC}}_{i,\text{ true}}$ and $N^{\text{H}_2\text{O, sig MC}}_{i,\text{ true}}$ are the events simulated in Monte Carlo within that same bin. The template parameters c_i and o_i, defined separately for the different targets, are allowed to vary without any prior constraint, so as to avoid model bias, and cause changes in the predicted number of events accordingly. In total, the number of signal events predicted in the i^{th} true bin is written as

$$N^{\text{sig}}_{i,\text{ true}} = c_i N^{\text{CH, sig MC}}_{i,\text{ true}} + o_i N^{\text{H}_2\text{O, sig MC}}_{i,\text{ true}}. \tag{5.19}$$

To translate the number of signal events in a true bin i to reconstructed bin j, the detector smearing matrix $\mathbf{S}$ is applied, such that the total number of reconstructed signal events can be written as

$$N^{\text{sig}}_{j,\text{ reco}} = \sum_{i}^{\text{true bins}} \left(c_i N^{\text{CH, sig MC}}_{i,\text{ true}} + o_i N^{\text{H}_2\text{O, sig MC}}_{i,\text{ true}} \right) S_{ij}. \tag{5.20}$$

This can then be modified in order to add the contribution to the reconstructed signal from true background events as

$$N^{\text{sig}}_{j,\text{ reco}} = \sum_{i}^{\text{true bins}} \left(c_i N^{\text{CH, sig MC}}_{i,\text{ true}} + o_i N^{\text{H}_2\text{O, sig MC}}_{i,\text{ true}} + \sum_{b}^{\text{bkg}} N^{\text{bkg MC}}_{ib,\text{ true}} \right) S_{ij}, \tag{5.21}$$

where the index b runs over all backgrounds. Rather than assign template weights to the background contributions, these are constrained by simultaneously fitting sideband regions

rich in the typical background events to the analysis. Finally, the variation of systematic parameters is included in addition to that of the template parameters, giving the full form of the number of reconstructed events in the j^{th} bin as

$$N^{\text{sig}}_{j,\text{ reco}} = \sum_{i}^{\text{true bins}} \left[c_i \left(\prod_{a}^{\text{model}} \omega(a,\vec{x}) \right) N^{\text{CH, sig MC}}_{i,\text{ true}} + o_i \left(\prod_{a}^{\text{model}} \omega(a,\vec{x}) \right) N^{\text{H}_2\text{O, sig MC}}_{i,\text{ true}} + \sum_{b}^{\text{bkg}} N^{\text{bkg MC}}_{ib,\text{ true}} \left(\prod_{a}^{\text{model}} \omega(a,\vec{x}) \right) \right] S_{ij} d_j \sum_{n}^{E_\nu \text{ bins}} \omega_{in} f_n, \tag{5.22}$$

which is used as the number of observed events, N^{obs}_j, in Equation (5.14) to calculate χ^2_{stat}. Here three different sets of systematic parameters are included:

- **Detector Systematic Parameters** – The uncertainty on the number of events in the j^{th} reconstructed bin due to detector effects within ND280 is included via the multiplicative factor d_j. The matrix is generated by throwing toy experiments to obtain weights. More detail on these parameters is given in Section 7.1.1.

- **Cross-Section Model Uncertainties** – Uncertainties in the modelling of both signal and background are included via the set of weights $\omega(a,\vec{x})$, where a is the theoretical model parameter being altered and $\vec{x}$ represents the true kinematics, reaction mechanism and target nucleus. Parameter values are varied around the nominal values, calculated from fits to external data, and the effect of changing the underlying model parameters is propagated using spline functions generated from the `T2KReWeight` package [186, 187]. More detail on these parameters is given in Section 7.1.2.

- **Neutrino Flux Parameters** – Uncertainty in the flux of the neutrino beam is included via the parameters f_n and weights ω_{in}. The flux parameters f_n describe the flux uncertainty in a true bin as a function of the neutrino energy bin n, whilst ω_{in} are the weights which map the neutrino energy bins n to the corresponding true analysis bins i. Altering the flux parameters causes a change in the number of events in a true energy bin, which can correspond to multiple true kinematic bins. The covariance matrix for this is supplied by the T2K Beam group. More detail on these parameters is given in Section 7.1.3.

To perform the likelihood fit described, the `Super-xsLLhFitter` software package [188] is used, an extension of the fitting software used for earlier T2K cross-section analyses. This employs the `Minuit2` minimiser [189] using the MIGRAD algorithm in order to find the minimum of the log-likelihood defined in Equation (5.7). The HESSE algorithm is also used to calculate the Hessian matrix (matrix of second order partial derivatives) around the best-fit points, which inverted gives the covariance matrix for the fit parameters. The output of the fit is the vector of parameters which minimise the log-likelihood, along with the covariance matrix describing their post-fit error values and correlations. Using the post-fit template and nuisance parameters, the number of selected signal events in the i^{th} true

bin is calculated as

$$N_i^{\text{sig}} = \sum_j^{\text{reco bins}} \left[c_i \left(\prod_a^{\text{model}} \omega(a, \vec{x}) \right) N_{i,\text{ true}}^{\text{CH, sig MC}} + o_i \left(\prod_a^{\text{model}} \omega(a, \vec{x}) \right) N_{i,\text{ true}}^{\text{H}_2\text{O, sig MC}} \right] \times (S_{ij})^{-1} d_j \sum_n^{E_\nu \text{ bins}} \omega_{in} f_n, \tag{5.23}$$

which is used to calculate the cross section as in Equation (5.1).

5.3 Regularisation

The process of unfolding can be described as an 'ill-posed' problem, in that many solutions to the problem can exist. As described in Section 5.2.1, the fit can easily raise the number of events in a given true bin (or vice versa), as long as it compensates by lowering the number of events in the neighbouring bins. This gives many solutions that have very similar χ^2 values, and in practice results in non-physical results with strong anti-correlations between bins.

This analysis, following the practice of previous [119] and on-going T2K analyses, employs the use of Tikhonov regularisation [190] in order to deal with this effect. Regularisation adds an additional constraint to the fit requiring that, between neighbouring bins, the number of events changes smoothly. This is achieved via a further penalty term to the total χ^2,

$$\chi^2_{\text{reg}} \approx -2 \ln \mathcal{L}_{\text{reg}} \tag{5.24}$$

$$= \lambda \sum_i^{N-1} (\eta_i - \eta_{i+1})^2, \tag{5.25}$$

where λ is the strength of regularisation, which is a tunable parameter in the fit framework, and η_i (η_{i+1}) is the general template parameter for the i^{th} ($i^{th}+1$) true bin. For the analysis described, η can be either of the template parameters c_i or o_i, but regularisation is not applied between the two. This penalty term causes an increase in the χ^2 for large differences between neighbouring template parameters, biasing the result towards the shape of the MC input. The regularisation parameter λ is chosen such that the contribution of χ^2_{reg} to the total χ^2 is the smallest of the three components.

For the intended 4-dimensional differential cross section in muon and pion kinematics, extra care must be taken in applying regularisation between neighbouring template bins, as there is no reason that the last bin in pion kinematics for a given muon kinematic slice should be constrained by the first bin in pion kinematics for next muon kinematic slice. To account for this in the fit, regularisation is only applied between adjacent pion momentum bins.

5.4 Cross-Section Calculation and Error Propagation

The fit method described in Section 5.2.3 returns the set of best-fit parameters which minimises the log-likelihood, along with a covariance matrix describing the post-fit uncertainties on each parameter, and the correlations between them. Using the best-fit number of signal events, Equation (5.1) can be used to calculate the final cross-section result, but analytically propagating the associated errors through the calculation is infeasible, due to the large number of parameters and correlations between them. Instead, a series of Monte Carlo toy throws of the final fit parameters is used, alternated slightly each time in line with the correlations provided from the post-fit covariance matrix. The variance of the each toy cross section about the nominal value is then used to calculate the uncertainty in each bin.

The post-fit covariance matrix $\mathbf{\Sigma}$ is first Cholesky decomposed [191] into the lower triangular matrix $\mathbf{L}$ and its conjugate transpose $\mathbf{L}^*$ as

$$\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^* \tag{5.26}$$

using ROOT's `TDecompChol` package [192]. A random variation of the fit parameters can then be generated by multiplying $\mathbf{L}$ with a vector of random numbers Gaussian distributed around zero with a width of one:

$$\vec{\theta_t} = \vec{\theta} + (\mathbf{L} \times \vec{r_t}), \tag{5.27}$$

where $\vec{\theta_t}$ is the set of toy throw parameters, $\vec{\theta}$ is the set of best fit parameters, and $\vec{r_t}$ is the set of Gaussian distributed random numbers. This process is repeated for many toy throws, and the new set of toy parameters in each throw is used to calculate the associated cross section. Finally, the toy variations of the cross section are used to calculate the cross-section covariance matrix:

$$V_{ij} = \frac{1}{N}\sum_t \left[\left(\frac{d\sigma}{dx}\right)_{i,t} - \left(\frac{\bar{d\sigma}}{dx}\right)_i\right]\left[\left(\frac{d\sigma}{dx}\right)_{j,t} - \left(\frac{\bar{d\sigma}}{dx}\right)_j\right] \tag{5.28}$$

where t is the index over toys from 1 to N, and $\frac{\bar{d\sigma}}{dx}$ is the best fit cross-section value. In using the HESSE algorithm to determine the covariance between template parameters, an assumption is made that the parameter uncertainties are Gaussian distributed. To ensure that this assumption is sound, a series of fake data studies are carried out in validation of the fitter framework, before unblinding to data.

5.5 Data and MC Statistics Summary

Table 5.I summarises the protons on target (POT) for the ND280 data taking periods used in this analysis, along with the associated MC production used. Sand events – neutrino interactions that occur upstream of ND280 in the sand of the detector surroundings – are also included. These are simulated separately from the regular magnet MC, and the same set of sand MC (run 3a) is used for each run.

ND280 Run	MC POT $\times 10^{20}$	Data POT $\times 10^{20}$	Sand POT $\times 10^{20}$
Run 2w	4.75887	0.43329	10.7297
Run 2a	3.64059	0.35934	10.7297
Run 3a	16.2411	1.58174	10.7297
Run 4w	16.5211	1.64277	10.7297
Run 4a	18.2013	1.78271	10.7297
Run 8w	14.6736	1.58053	10.7297
Run 8a	34.7093	4.14909	10.7297
Total	108.74586	11.52947	

Table 5.I: Summary of real data and Monte Carlo production POT (protons on target) used in the described analysis. MC and data release are both prod6T.

This analysis uses all forward-horn current (FHC) runs: T2K runs 2+3+4+8. The Monte Carlo simulated data used is the updated production 6T (prod6T) MC, generated using ND280 software version `nd280v11r31p43` and `NEUT_D` (`NEUT 5.4.0`). Full details of the production re-spin used and its changes in comparison to the previously used prod6B can be found in the updated version of the numuCCMultiPiAnalysis technical note, T2K-TN-407 [193].

The analysis is performed on modified versions of the official prod6T flattrees (`nd280Highland2 v2r45`). They were regenerated using `nd280Highland2 v2r49`, with a change in the provided parameters file to save individual FGD hit information, which is required for the Michel electron kinematic reconstruction performed in Section 6.4. The package used to perform the selection and systematics analysis described in Sections 6.2 to 6.4 and 7.1 is the developed numuCC1PiAnalysis, using the `nd280Highland 2.61` framework.

Chapter 6

CC1π^+ Event Selection

This chapter describes the development of the signal and control sample selections used for event identification. The work done to access regions of low momentum pion kinematics is also discussed in detail.

6.1 Standard ν_μCCMultiPi Selection

This analysis builds on one of the standard selection inputs to the T2K oscillation analysis, which is described in detail in T2K-TN-199 [194], and further discussed with updates and newer MC production (prod6T) in T2K-TN-407 [193]. Firstly, a ν_μCC-inclusive sample is selected from forward horn current (FHC) runs, by requiring the highest momentum negative track in the event to be a μ^- candidate. Where events reach this stage in the selection, they are searched for charged and neutral pions, and are then split into three distinct topologies:

- **ν_μCC0π** – Events containing no charged or neutral pions are placed into this branch. This sample is mostly made up of CCQE and 2p2h interactions.
- **ν_μCC1π^+** – Events containing exactly one positive pion, but no negative or neutral pions, go into this branch. No limit is placed on the number of baryons. This sample is largely made up of resonant or coherent interactions.
- **ν_μCCOther** – All selected events that do not fall into the above topology definitions get placed into this branch. This sample is largely made up of resonant and DIS interactions.

The selection sequence that is applied to the data in order to obtain these samples is described below. Selections for FGD1 and FGD2 are run in parallel.

6.1.1 Event Quality Cut

This cut first checks that the data is usable. ND280 subruns are assigned data quality flags based on the operational status of sub-detectors at the time of data taking. Data quality flags are required to be good to pass this stage. The events are also required to be in time with the beam trigger.

6.1.2 Track Multiplicity Cut

Events are required to have at least one reconstructed track crossing one of the TPCs.

6.1.3 Track Quality and Fiducial Volume Cut

All tracks that originate in either FGD1 or 2 and have good quality in the TPCs are split into positive and negative groups, ordered by their momentum. A good quality TPC track is defined as having over 18 vertical hit clusters in the TPC, in order to reject short tracks where the reconstruction is less reliable. The highest momentum negative track is selected as the muon candidate, and required to originate from a vertex contained within the FGD fiducial volume (FV). The FGD1 FV is defined by the coordinate limits $|x| < 874.510$ mm, $|y - 55| < 874.510$ mm, and $136.875 < z < 446.955$ mm. The cuts in x and y are chosen to reflect the outer boundaries of the central 182 scintillator bars in the X and Y layers. As each layer contains a total of 192 bars, this equates to excluding the five outermost bars on each side from the FV. The 55 mm offset in y is due to the fact that the XY modules are displaced 55 mm upwards from the centre of the ND280 coordinate system. The cut in z removes the first upstream XY module from the FV volume definition, leaving the remaining 14 modules. In FGD2, the FV is defined as $|x| < 860.000$ mm, $|y - 55| < 860.000$ mm, and $1483.750 < z < 1807.375$ mm. In contrast to the FGD1 FV, this definition only removes the first X layer from the FV; removing the whole XY module would leave the most upstream water module unusable.

6.1.4 External FGD Veto Cut

The second highest momentum TPC track in the event is selected to be the veto track. If the starting position of the veto track is more than 150 mm upstream of the main track's starting position, then the event is rejected. In the FGD2 selection, events are also removed if any tracks are found with a start position in FGD1. This removes mis-reconstructed events that likely originated outside the fiducial volumes of the FGDs (OOFV).

6.1.5 Broken Track Veto

This cut is introduced in order to reject events with incorrectly reconstructed tracks. Where a muon candidate track may originate in the FGD and escape into the TPC, it is possible for the reconstruction to break this track into two separate components: one track fully-contained within the FGD FV, and a second starting in the last few layers of the FGD and exiting into the TPC. In this case, the second track would be classed as the muon candidate. In order to reject such events, whose initial momentum and start position will likely be incorrectly reconstructed due to the breaking of the track, it is required that events with at least one fully-contained FGD track must start at least two layers upstream from the downstream edge of the FGD.

6.1.6 Muon PID Cut

This applies the μ^- TPC particle identification (PID) criteria to the selected μ^- candiate, based on the dE/dx distribution of the track. This is done by using two separate selection criteria, which cut on the differing track hypotheses:

$$\mathcal{L}_{\text{MIP}} = \frac{\mathcal{L}_\mu + \mathcal{L}_\pi}{1 - \mathcal{L}_p} > 0.8 \text{ if } p < 500 \text{ MeV/c}, \tag{6.1}$$

$$\mathcal{L}_\mu > 0.05. \tag{6.2}$$

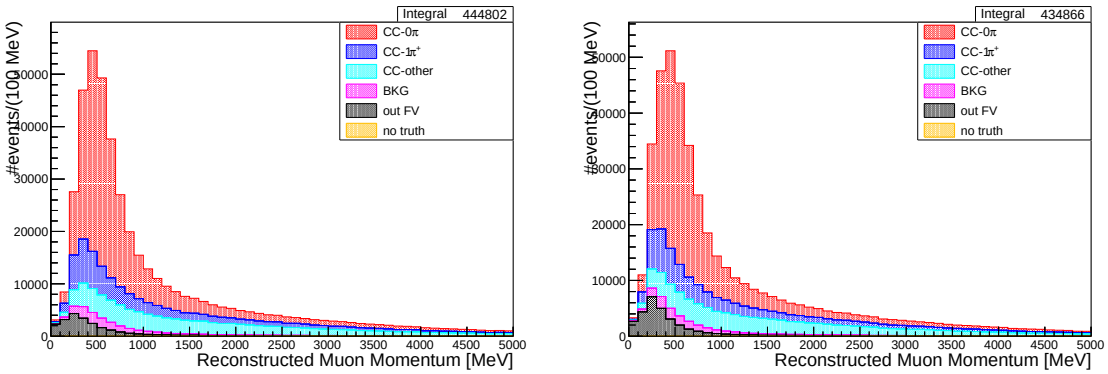
The first condition on the MIP-likelihood is applied to reject electrons that have been selected as the highest momentum negative track, whilst the second is used to remove leading pions and protons. The likelihood function for the i^{th} particle, $\mathcal{L}_i$, is given by

$$\mathcal{L}_i = \frac{\exp\left(\frac{-\delta_i^2}{2}\right)}{\sum_\ell \exp\left(\frac{-\delta_i^2}{2}\right)}, \tag{6.3}$$

where the pull of the i^{th} particle, δ_i, is defined as

$$\delta_i = \frac{dE/dx^{\text{measured}} - dE/dx_i^{\text{expected}}}{\sigma_{(dE/dx^{\text{measured}} - dE/dx_i^{\text{expected}})}}, \tag{6.4}$$

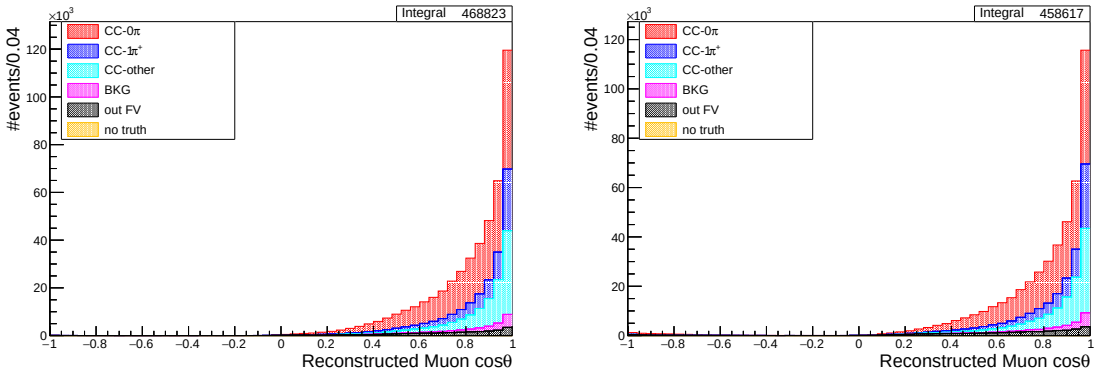
for the particle hypotheses $i = \mu^-$, p, e^-, $\pi^\pm$, and σ is the width of the distribution. The estimate of dE/dx is found using the truncated mean of the energy deposited in the TPC by the crossing particle, where a defined fraction of the total hit clusters is used[8], in order to obtain a better resolution track dE/dx measurement. Events passing this cut correspond to a CC-inclusive selection, the predicted muon momentum and angular distributions for which are given in Figures 6.1 and 6.2, respectively.



(a) FGD1 (b) FGD2

Figure 6.1: Muon candidate momentum for a CC inclusive selection, in (a) FGD1 and (b) FGD2. Events are categorised by the true topology of the interaction.

[8] Full details of the calculation for the truncation factor used can be found in T2K-TN-001 [195].



(a) FGD1 (b) FGD2

Figure 6.2: Muon candidate cosine of the angle to the neutrino direction for a CC inclusive selection, in (a) FGD1 and (b) FGD2. Events are categorised by the true topology of the interaction.

6.1.7 Single Positive Pion Cut

At this point in the selection process, the CC-inclusive sample is branched into the three sub-samples described at the beginning of Section 6.1, based on pion multiplicity in the event. To obtain a CC1π^+ sample, we require that the event has a secondary track found to be a positive pion, and that no other pions are present. The positive pion can be identified from three possible sources:

- *TPC* – The pion is reconstructed in the TPC using TPC PID methods.
- *Michel electrons* – Pions contained in the FGD decay, and are identified by the subsequent decay electrons.
- *Isolated FGD pions* – Fully contained FGD pions with no Michel decay, identified using FGD PID methods.

The number of positive pions in the event (required to be one to pass this cut) is then defined as the sum of the TPC and total FGD π^+ candidates,

$$N^{\pi^+}_{\mathrm{TPC}} + N^{\pi^+}_{\mathrm{FGD}}, \tag{6.5}$$

where

$$N^{\pi^+}_{\mathrm{FGD}} = \begin{cases} N^{\pi^+}_{\mathrm{ME}}, & \text{if } N^{\pi^+}_{\mathrm{ME}} \geq 1 \\ 1, & \text{if } N^{\pi^+}_{\mathrm{ME}} = 0 \text{ and } N^{\pi}_{\mathrm{isoFGD}} \geq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{6.6}$$

The TPC pion PID assignment is based on the highest probability particle method. To ensure the assignment is the most probable, the likelihood and pull calculations in

Equations (6.3) and (6.4) are followed by computing the ratio of the likelihood of a given particle type to the sum of all likelihoods,

$$P_i = \frac{\mathcal{L}_i}{\sum_\ell \mathcal{L}_\ell}, \tag{6.7}$$

where i is the given particle type, and ℓ represents a positive pion, proton or positron in the case of positive TPC tracks, or a negative pion or electron in the negative case. The particle is then tagged with the type that has the highest probability.

A significant fraction of the pions produced in the interactions being searched for will not have sufficient energy to reach a TPC, and instead will stop within the FGD volume. In this case, there are two ways the pions may be identified. If the pion decays within the FGD, then it will almost certainly[9] follow the well understood decay process

$$\pi^+ \longrightarrow \mu^+ + \nu_\mu, \tag{6.8}$$

where the muon itself decays via Michel decay as

$$\mu^+ \longrightarrow e^+ + \nu_e + \bar{\nu}_\mu. \tag{6.9}$$

Decayed pions can be identified via their Michel electron decay products, which are found by searching for hits within delayed FGD time bins. An FGD time bin is defined as a cluster of hits in the FGD from tracks which are passing through at a similar time. Individual hits are ordered in time and then sorted into separate 'time bins' according to the time between subsequent hits, starting from the beginning. If the time difference between hits is less than 100 ns, then hits are placed into the same bin. If the time difference is greater than 100 ns, then a new bin is created, and the latter hit added into the next bin. This is shown diagrammatically in Figure 6.3.

Along with separating bunches, the FGD time bins can be used to tag the presence of Michel electrons. As the muon lifetime ($\sim$2.2 μs) is several orders of magnitude longer than the spacing between time bins, Michel electrons can usually be identified by the additional delayed time bins that they produce. Time bins are considered 'delayed' if the first hit in that bin occurs more than 100 ns outside of one of the eight bunch windows in each beam spill. To be classed as containing Michel electrons, an event must have at least one delayed out-of-bunch FGD time bin. A cut is then placed on the total number of hits in the delayed time bins. This cut was studied for production 6 in T2K-TN-104 [196], defining the cut separately for FGD1 and FGD2, due to their differing amounts of active mass. These cut values were determined by examining the number of delayed FGD hits in a CC-inclusive sample, broken down by true particle causing the delayed time bin. The obtained distributions for FGD1 and FGD2 are shown in Figure 6.4. The study was performed using `highlandv0r27` on production 6A Run 4 files; total MC POT is 4.2×10^{20}, total data POT is 1.35×10^{20}.

[9] Proceeds with a branching fraction of $\Gamma = (99.98770 \pm 0.00004)\%$ [19].

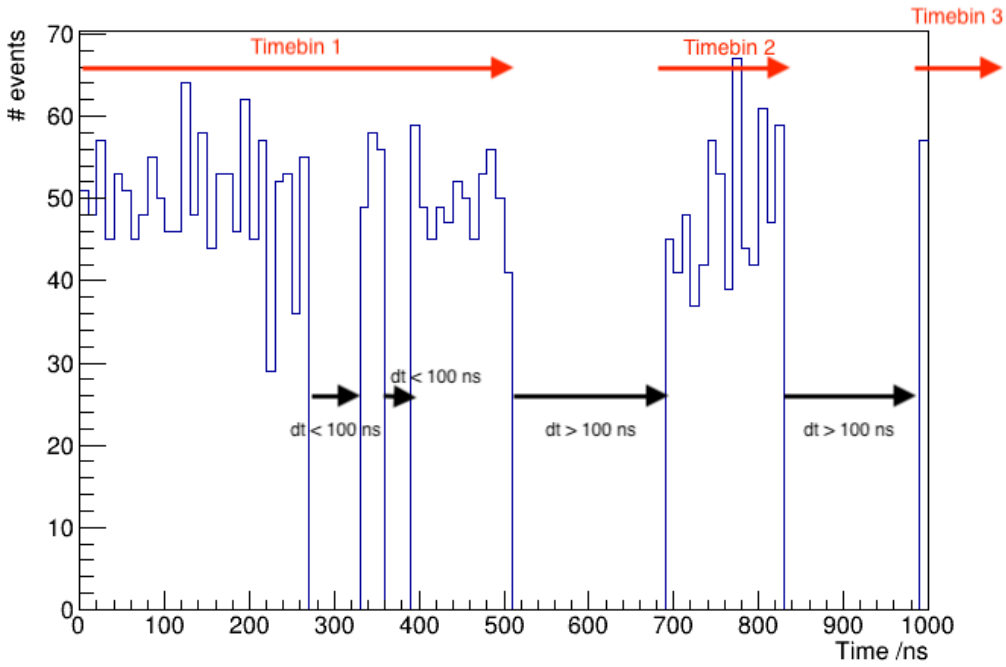


Figure 6.3: Schematic showing FGD time binning, where FGD hits are ordered by time. If there is a gap of less than 100 ns with no hits before the next hit occurs, events stay in the same FGD time bin. If 100 ns passes with no hits, a new bin is created for the next available hits. Event distribution shown is randomly generated for illustrative purposes.

The study shows that the separation of distribution shape between Michel and non-Michel events is much clearer in FGD1 than in FGD2, making it easier to choose a cut that separates the signal from the lower-hit backgrounds. However, the driving reason for the cuts chosen was the large MC low-hit background in both FGDs, which is not observed in data. Whilst the data–MC agreement matches well for Michel electrons at higher numbers of hits, background from neutrons and other low-hit sources seems to be over-simulated. Since this is not well understood, cuts should not be placed within this region of disagreement. In FGD1, a delayed time bin is required to have at least seven hits in order to be considered a Michel electron candidate. In FGD2, at least six hits are required.

If a CC-inclusive event is found to have a Michel electron candidate, then this is used to imply the presence of a positive pion. In selecting CC1π^+ events, the number of pions from either TPC tracks or Michel electrons is required to be *exactly* one. The method used for estimating the kinematics of the pion when it is tagged by the presence of a Michel electron is detailed fully in Section 6.4.

The final allowed pion identification possibility is for those events where the produced pion is fully contained within the FGD volume, but no decay via the Michel chain is detected. In order to identify these positive pions, all events with FGD-only tracks and no identified Michel electrons have the following requirements applied:

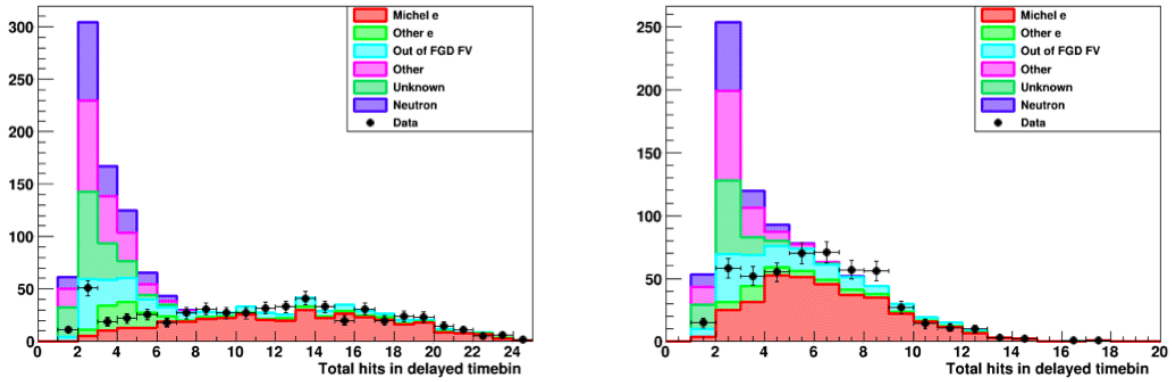


Figure 6.4: Hit distribution for FGD delayed time bins for FGD1 (*left*) and FGD2 (*right*). Plots are categorised by true particle causing the delayed time bin, and are normalised to data POT. Plots taken from T2K-TN-104 [196].

1. Isolated FGD track must be in the same time bunch as the muon candidate.

2. Only one FGD segment associated, in order to avoid broken tracks.

3. Track must be fully contained within either the FGD1 or 2 FV.

4. The absolute value of the cosine of the angle to the neutrino direction, $|\cos\theta_\pi|$, must be greater than 0.3. Higher angle isolated tracks are found to have reconstruction and systematic uncertainty problems, and so are removed from the selection here [197].

5. The calculated FGD pion pull value, $\delta_\pi^{\mathrm{FGD}}$, is required to be in the range $-2.0 < \delta_\pi^{\mathrm{FGD}} < 2.5$. This range is chosen to avoid electron-like pulls at lower values, and proton-like pulls at higher values [194]. Figure 6.5, taken from T2K-TN-152 [198], shows the pion pull distribution for contained tracks in FGD1, prior to adding the cuts in step 4. The distribution is categorised by the true particle type of the isoFGD track.

The pull itself is calculated in a similar way to that of the TPC particle pulls in Equation (6.4), instead using the total deposited energy in the FGD:

$$\delta_i^{\mathrm{FGD}} = \frac{E^{\mathrm{measured}} - E_i(x^{\mathrm{measured}})}{\sigma_i(x^{\mathrm{measured}})}. \tag{6.10}$$

Here, the expected total energy E_i, and its expected resolution σ_i, are both functions of the measured track length, x^{measured}. The expectations for each particle hypothesis i, where $i \in \{\mu, \pi, p\}$, are obtained by performing fits to the true track range against deposited energy, and validated using both MC simulation and real data. More detailed information can be found in [156] and T2K-TN-103 [199]. If a non-zero number of tracks are tagged as isolated FGD pions, with no other mesons present, then the event is classed as CC1π^+.

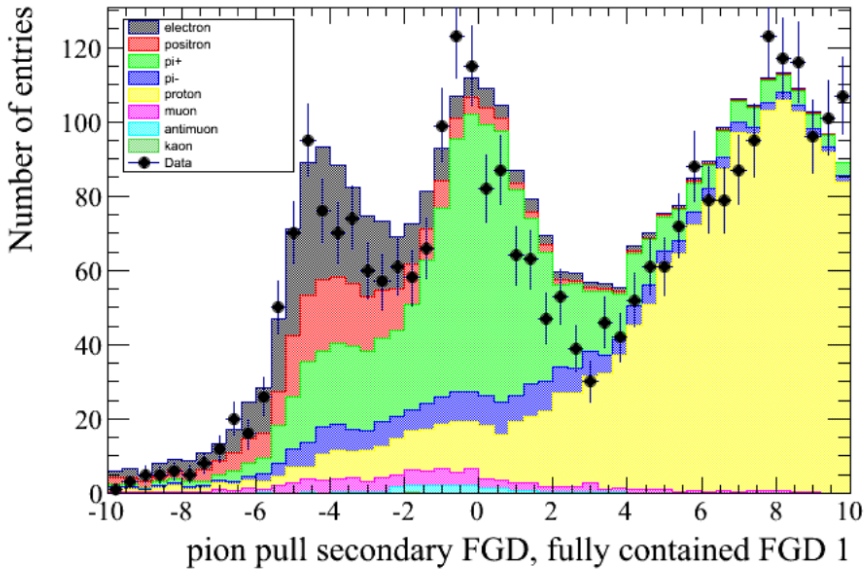


Figure 6.5: Pion pull distribution for contained tracks in FGD1, prior to placing cuts on the cosine of the angle to the neutrino direction, taken from [198].

6.2 Development of a ν_μCC1π^+Selection

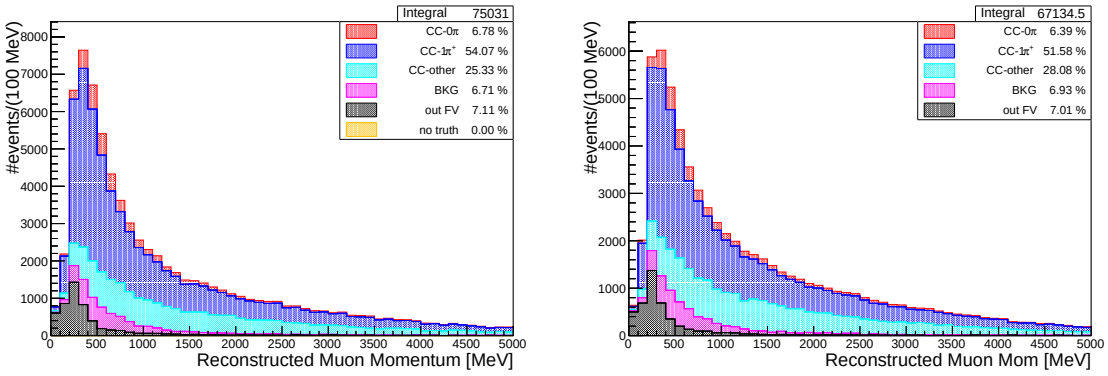
The standard ν_μCC multiple pion selection provides a good base from which to make a CC1π^+ measurement, and the momentum distribution for the muon candidate in both FGDs is given in Figure 6.6, showing the purity of the sample. However, in order to better measure the signal process in question, we should further investigate the behaviour of these events in the detector. In an ideal world, we would create a completely pure sample by cutting out all background processes that could be mistaken for signal. In reality this is not possible, as we also require a reasonable number of events in order to make a measurement that is not severely limited by statistical error. This section describes how the standard selection is modified to optimise background rejection. To judge the effectiveness of cuts, the efficiency and purity after each cut are checked, where the efficiency ϵ of the selection is defined as the fraction of selected signal events to the number of true signal events prior to any cuts,

$$\epsilon = \frac{N^{\mathrm{CC1}\pi^+}_{\text{true, selected}}}{N^{\mathrm{CC1}\pi^+}_{\text{true, pre-cuts}}}, \tag{6.11}$$

and the purity ρ is defined as the fraction of selected events which are true signal,

$$\rho = \frac{N^{\mathrm{CC1}\pi^+}_{\text{true}}}{N^{\mathrm{CC1}\pi^+}_{\text{selected}}}. \tag{6.12}$$

The usefulness of the cuts presented are judged on these two variables, and generally optimised for a maximal value of the chosen figure of merit, $\epsilon \times \rho$.



(a) FGD1 (b) FGD2

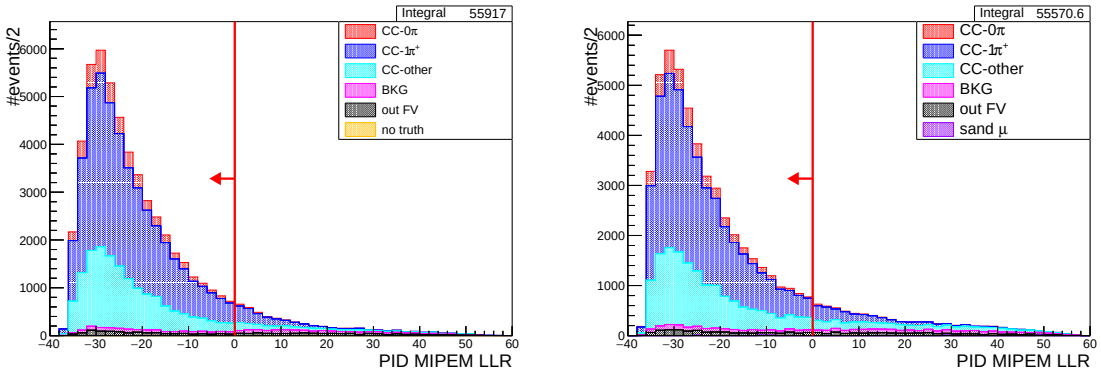
Figure 6.6: Momentum distribution for the muon candidate in the CC1π^+ branch of the standard numuCCMultiPiSelection, in (a) FGD1 and (b) FGD2. Events are categorised by true topology. Despite technically being available in standard multiple pion selection, the ECal π^0 veto (Section 6.2.2) is not applied here, since it is not included in the standard oscillation analysis input sample.

6.2.1 ECal PID

The Tracker ECals provide several discriminating PID variables that can be utilised to further reduce background acceptance in the signal sample. The PID stage of the reconstruction consists of two steps. First, the low-level PID variables are calculated. These describe the circularity, cluster hit charge deviation, layer charge truncated max ratio and dE/dx front-back ratio of tracks and showers in the ECal. A full description of these variables can be found in T2K-TN-111 [200], which are then used as inputs to calculate three log-likelihood ratios. These high-level ECal variables can be used to improve particle identification.

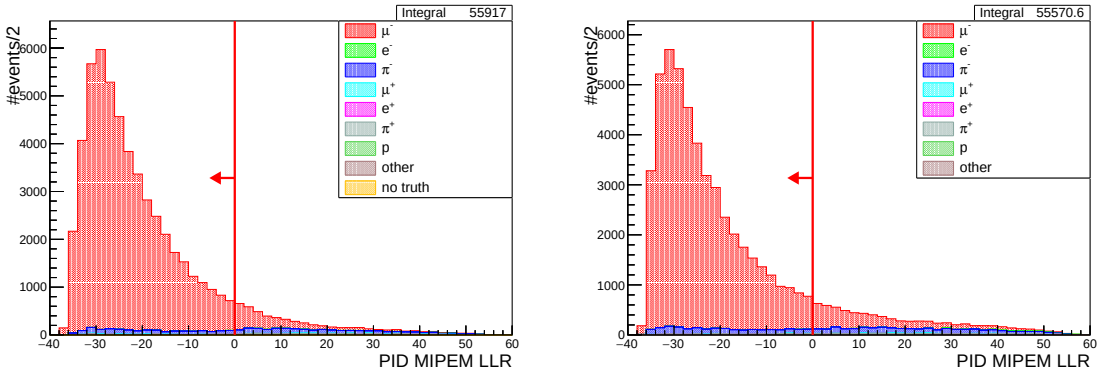
For all events where the main track has an ECal segment, the PID MIPEM value is used as a discriminator, which separates minimum ionising particles from electromagnetic shower-like particles. Of all the events that reach this step in the selection, 30.1% of FGD1 main tracks don't have an ECal segment, compared to only 23.1% for FGD2. This is as expected, since tracks from interactions in FGD2 are more likely to have ECal segments due to the proximity to the DsECal. ECal objects with a MIPEM value of below 0 will be MIP-like, whilst values above will be shower-like. A cut is placed at 0, to reject all events with μ^- candidates categorised as shower-like. The MIPEM distributions of events categorised by their true topology are shown in Figures 6.7a and 6.7b, whilst Figures 6.8a and 6.8b give the same distributions in terms of the true identity of the μ^- candidate. If the main track does not have an ECal segment, then it is allowed to pass this cut without any restrictions imposed. Where a main track has more than one associated segment, only the first segment is checked; this occurs for less than 1% of events which reach this cut.

Figure 6.7 shows that the majority of the CCOther background leaves a MIP-like signature in the ECal, and so will not be affected by the cut. However, there is still a large tail extending out beyond 0, which stays at a roughly constant rate whilst the CC1π^+ signal



(a) FGD1 (b) FGD2

Figure 6.7: Muon candidate MIPEM distribution for all events with an ECal segment that pass the CC1π^+ cut in (a) FGD1 and (b) FGD2. Distribution is categorised by true event topology, with the intended cut at 0 shown. Events left of the cut line will pass. No distinction on pion detection method is made here.

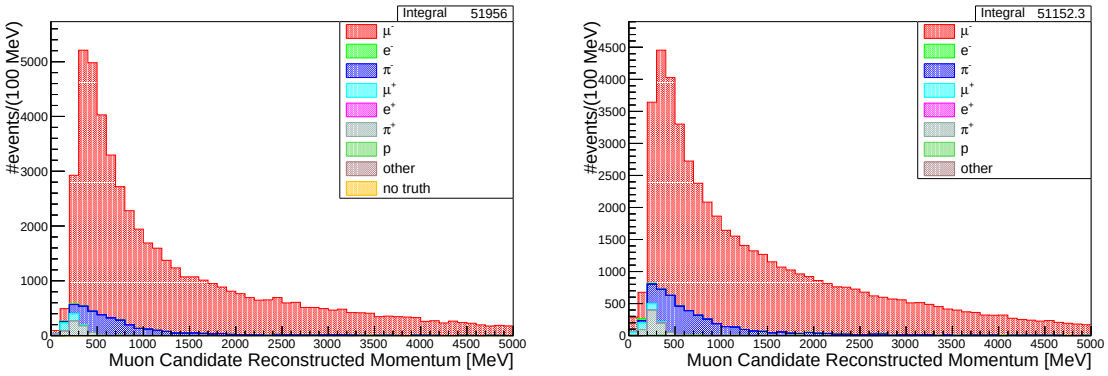


(a) FGD1 (b) FGD2

Figure 6.8: Muon candidate MIPEM distribution for all events with an ECal segment that pass the CC1π^+ cut in (a) FGD1 and (b) FGD2. Distribution is categorised by true particle ID, with the intended cut at 0 shown. Events left of the cut line will pass. No distinction on pion detection method is made here.

contribution decreases more rapidly. The background (BKG) topology, here representing NC or $\bar{\nu}_\mu$ processes, is also evenly distributed in MIPEM, and so this cut removes roughly half of those background events. The main use of this cut however can be seen in Figure 6.8. Examining the true particle ID of the μ^- candidate shows that there is a ~8% contribution from true π^- being mistakenly selected. This misidentification primarily occurs at lower momentum values, as seen in Figure 6.9, and is most likely due to the similarity of muon and pion energy loss in the TPC. The π^- background appears to be evenly distributed in MIPEM, and so cutting out shower-like tracks will remove over half of this background.

Of course, placing a cut at 0 is not necessarily the most optimal way of cutting, even



(a) FGD1 (b) FGD2

Figure 6.9: Momentum distribution for the muon candidate, for all events that pass the CC1π^+ cut where the muon candidate has an ECal segment, for (a) FGD1 and (b) FGD2. Distribution is categorised by true muon candidate particle ID, showing that the π^- background occurs mostly at low momentum values. No distinction on pion detection method is made here.

though this is the defined boundary between MIP-like and shower-like tracks. As part of the selection development for this analysis, a study was done to determine the optimal value at which to cut, maximising the figure of merit $\epsilon \times \rho$. This determined that the optimal value at which to place a cut was 12. Further details of this study are given in Appendix B. However, this also introduces additional difficulties. At the time of writing, the detector systematic uncertainty on the MIPEM variable has only been assessed for a cut placed at 0, and so would require recalculation. Secondly, for the control samples inspected in the original construction of this variable, data–MC agreement was seen to be slightly worse at higher values than at 0. Therefore, it seemed that the small gain in $\epsilon \times \rho$ achieved by optimisation of the cut was not worth the introduction of non-standard and likely larger systematic uncertainties.

6.2.2 ECal π^0 Veto

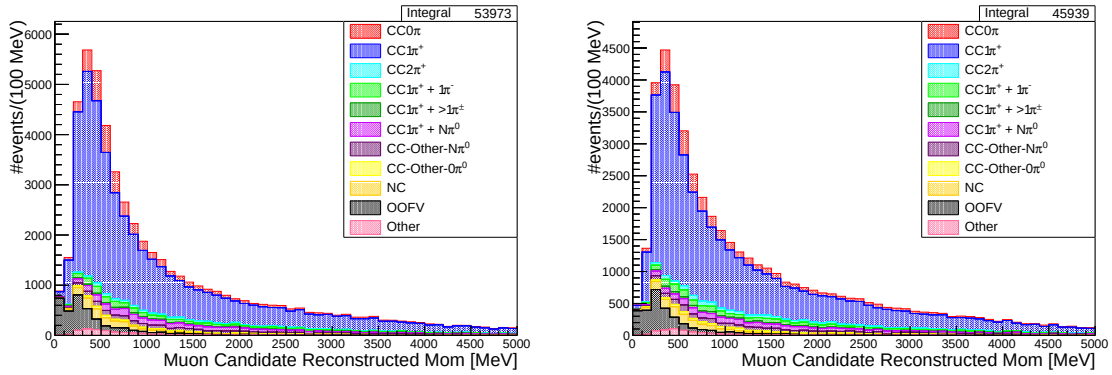
Neutral pions have a mean lifetime of $(8.52 \pm 0.18) \times 10^{-17}$ s [19], and so any produced in neutrino interactions within the detector will decay very quickly to two photons, with a branching ratio of ∼99% [19]. Whilst it is possible for the decay photons to convert into e^+e^- pairs which can be identified in the TPCs, the probability of this happening is low. Instead, the decay photons are most likely to interact in the ECals, causing electromagnetic showers. This cut is used to veto events containing reconstructed π^0 candidates, which are identified under the following requirements:

- The ECal contains an isolated object (no associated tracks or segments in any other sub-detector) in the same time bunch as the μ^- candidate.
- The most energetic isolated ECal object has an electromagnetic energy of at least 30 MeV, and has a likelihood compatible with that of an electromagnetic shower (MIPEM

< 0).

- The most upstream layer hit from the object is within the first five layers of the ECal.

The production 5 version of this cut also included two additional requirements, which propagated the muon candidate and all positive TPC objects to the ECal surface. If the distance between either of these positions and the isolated ECal object was found to be less than 700 mm, the isolated ECal was not considered to be a π^0 candidate. This was included in order to not incorrectly reject events where the muon or positive TPC object reached the ECal, but the ECal object was not correctly associated to the global track. An investigation into whether this should be retained for production 6, documented more thoroughly in T2K-TN-392 [201], found that there were arguments for including it. However the efficiency gain was minimal enough that, for simplicity in not having to recalculate related detector systematics, it was decided not to use this part of the cut in this analysis. For the standard version of the veto that is applied, the distributions in muon momentum of events passing the cut are shown in Figures 6.10a and 6.10b. All events failing the veto are shown in Figures 6.11a and 6.11b.

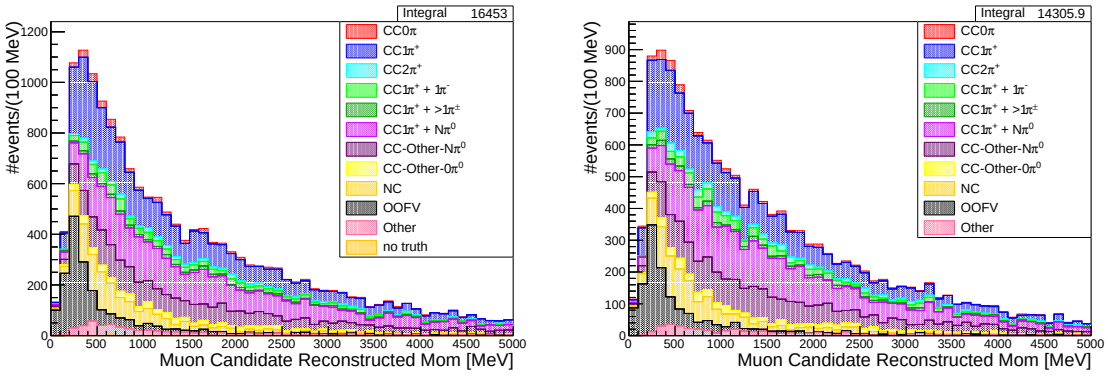


(a) FGD1 (b) FGD2

Figure 6.10: Muon candidate reconstructed momentum for all events passing the π^0 cut in (a) FGD1 and (b) FGD2. Events are categorised by topology, where the CCOther category is broken down into more detailed topology definitions based on the type of additional pions present. Distributions are shown for total CC1π^+ sample.

6.2.3 Barrel ECal Time of Flight Cut

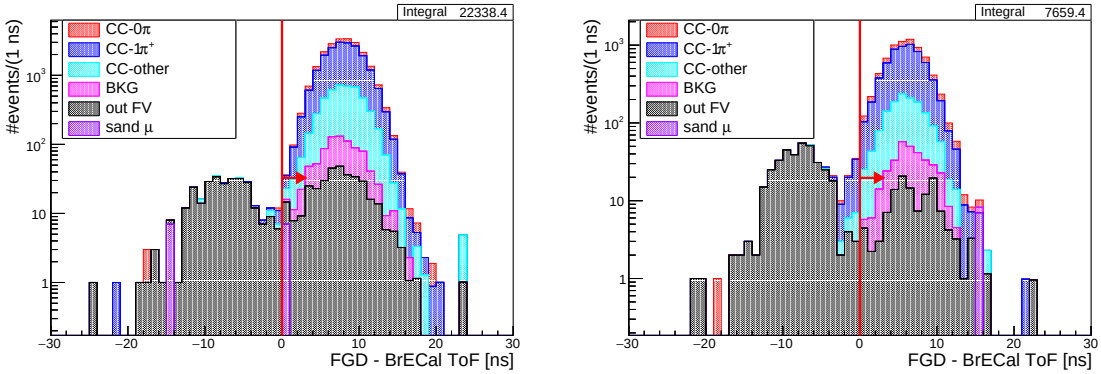
Timing information in ND280 is often used during the reconstruction and analysis stages. In this analysis, the time of flight values between the interaction FGD and the ECals were examined. When considering backward-going or high-angle tracks, the definition of ToF used allows it to take negative values (see Chapter 4). Since this analysis is performed for interactions with forward-going muons ($\cos\theta_\mu > 0.3$), the ToF from either FGD to an ECal, where the μ^- candidate does reach one, should be positive. ToF distributions between FGD1 and FGD2, and the Barrel ECal, are shown in Figures 6.12a and 6.12b, respectively.



(a) FGD1 (b) FGD2

Figure 6.11: Muon candidate reconstructed momentum for all events failing the π^0 cut in (a) FGD1 and (b) FGD2. Events are categorised by topology, where the CCOther category is broken down into more detailed topology definitions based on the type of additional pions present. Distributions are shown for total CC1π^+ sample.

These show that placing a cut at 0 ns and accepting anything with a ToF above gives a small, yet clear, benefit. Nearly 100% of the events with a negative ToF are ones where the interaction takes place out of the FGD FV. Removing these events gives a small increase in purity, for almost no efficiency loss. All events at this stage of the selection which do not have an ECal segment automatically pass the cut; it is applied only if a ToF value can be calculated. At this stage of the analysis, 36% of FGD1 events have at least one segment in the BrECal, compared with 19% of events from FGD2.



(a) FGD1 (b) FGD2

Figure 6.12: Time of flight distribution for the μ^- candidate (main track), between (a) FGD1 and Barrel ECal, and (b) FGD2 and Barrel ECal. Intended cut line at 0 is shown by the arrow. Distribution is shown for total CC1π^+ sample, using a log scale for clarity.

The ToF distributions between each FGD and the Downstream ECal were also checked,

but despite there being more valid[10] ToF values for this detector combination, these distributions were found to be less useful for applying cuts. Even though a significant percentage of events had ECal segments in the DsECal (34% from FGD1 and 62% from FGD2), a negligible amount of these segments displayed negative ToF values. This is due to the fact an OOFV event would have to be travelling upstream from downstream of the DsECal to possess a negative ToF. Instead the main OOFV component appears on the positive side of zero, as can be seen in Figure C.1 in Appendix C. In particular, placing a cut on the ToF to the DsECal after already doing so for the BrECal was found to have almost no effect at all. For this reason, only the BrECal ToF distribution is used for this cut.

It is a known issue that the electronics of the Trip-T detectors (ECals, PØD, SMRDs) and the FGDs are susceptible to time slips between the SCMs on each detector and the main clock module they connect to, which are in part responsible for recording hit times (Chapter 4). Although the reconstruction attempts to correct for this in the data, it is not 100% effective, and timing information remains particularly bad in run 8. For analyses that rely heavily on time of flight information (such as the full-angular acceptance measurements where timing information is required for directional inference), a time data quality cut is introduced. This is similar to the event quality cut described in Section 6.1.1, and is applied at the beginning of the selection to remove any events flagged as having bad timing information. However, as roughly 30–40% of run 8 is deemed to have bad data quality, applying this would result in a 15–20% loss of statistics from the very beginning of the selection. This is clearly not ideal, particularly when only a maximum of 36% of events which make it to this final cut will have the time of flight information used.

Instead, we choose to use the time of flight information without first applying the time quality cut. A study conducted in the course of a ν_eCC-inclusive cross-section measurement (T2K-TN-277 [202]) concluded that the ToF systematic error for 'bad ToF' data should be increased by a factor of 3, and this is the method employed here. As the systematic uncertainty on time of flight resolution is not expected to be the dominant source of error in this analysis, this is preferable to losing a large amount of statistics at the beginning of the selection. Further details of the systematic treatment are given in Section 7.1.1.22.

6.2.4 FGD Layer Reconstruction

FGD1 is composed of layers of polystyrene scintillator $(C_8H_8)_n$, aligned in alternating directions in order to allow 3D tracking of charged particles passing through the detector volume. Whilst the active volume of FGD2 is also composed of scintillator bars aligned in alternating directions, each XY module is interleaved with water layers. This allows neutrino interactions to take place on water, which is of course the target medium in the far detector. Since the FGD2 water layers are inactive, it is impossible to directly reconstruct a neutrino interaction occurring on water. If an interaction occurs in a water layer, the first hit from it should generally appear in the X layer directly downstream of the water layer. Therefore, selecting events with an initial position in an X layer of FGD2 should give a sample of events where the majority of interactions occurred on oxygen. However, this

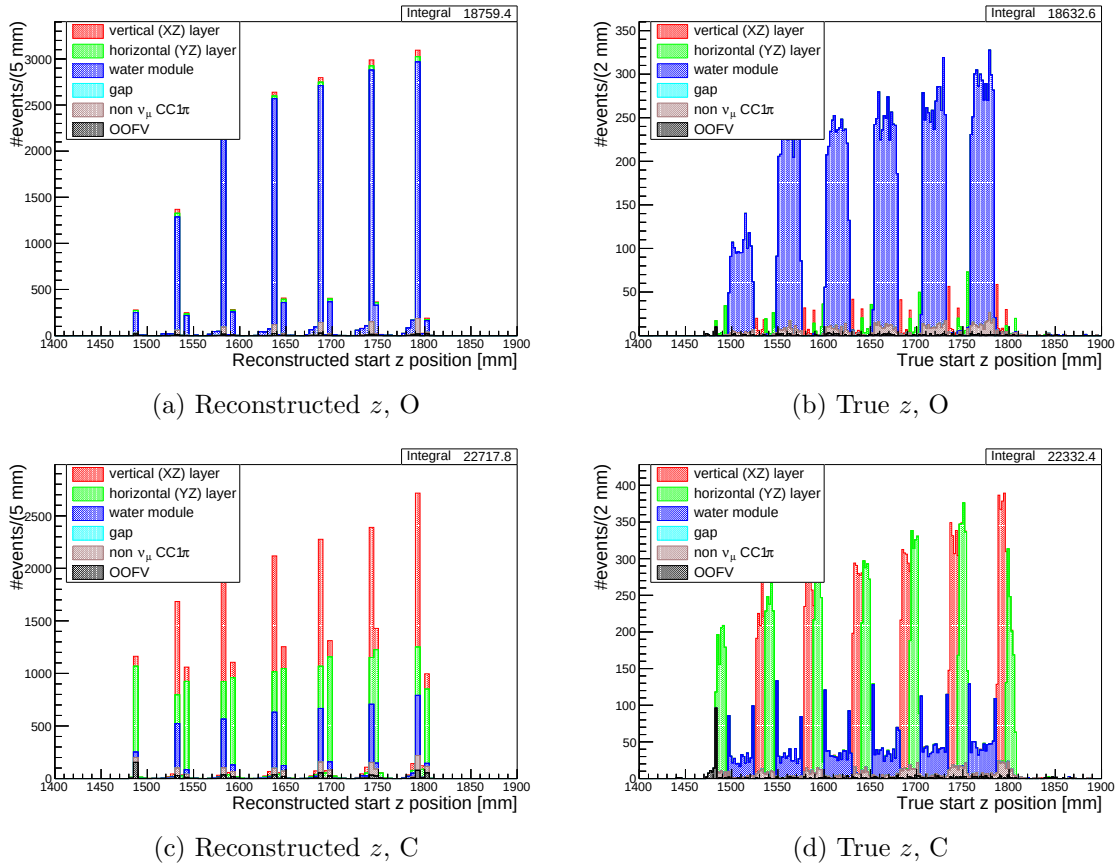
[10]Invalid values occur when the μ^- candidate does not reach an ECal, and so no ToF can be calculated.

sample will of course also include events where the interaction happened within an X layer itself, and hence occurred on carbon. Similarly, tracks where the first hit is in a Y layer will predominantly be from interactions on carbon, but a small background of interactions on oxygen will contaminate the sample. This contamination comes from two main sources: either oxygen interactions that have been backward migrated into the preceding Y layer, or backward-going tracks where the first hit from an oxygen interaction does indeed occur in a Y layer. Figures 6.13a and 6.13b show the reconstructed and true z coordinates of the initial position of the muon candidate, respectively, for oxygen interactions, and are categorised by which part of the FGD the interaction truly took place in. It can be seen that whilst the majority of interactions take place within the water layers, most of these are detected in the next X layer and have their vertex reconstructed there, due to the water layers being inactive. The remainder are mostly detected within a Y layer, and are caused by backward migrated tracks, placing the interaction vertex further upstream than it should be. A more detailed description of backward migration along with the systematic treatment for it is given in Section 7.1.1.6. A small component of the oxygen interactions can also be seen in Figure 6.13b originating from within the X and Y layers. This comes from a small number of interactions occurring on oxygen nuclei that are not part of the water target.

Similarly, Figures 6.13c and 6.13d show the same information, but for interactions on carbon nuclei. As expected, the majority of the interactions happen within the X or Y FGD layers themselves, as well as the initial track position occurring there. There is also a significant portion of interactions that happen on carbon within a water layer, with their vertex being reconstructed in an X layer. However, the largest fraction of these events have their true vertex at the very edge of the water layer. This is expected due to the additional components of the water module, such as the module components themselves. A small amount of events backward migrated into a Y layer are seen, but this is a much smaller effect than is observed for oxygen interactions.

In order to effectively assess the muon and pion kinematics for the described selection, each signal or control sample is separated into FGD1, FGD2x and FGD2y samples, based on the z coordinate of the initial track position. This is achieved by finding the reconstructed vertex position of the event, and comparing with the known FGD geometry. When extracting the cross section on oxygen, the carbon and oxygen contributions are extracted simultaneously, with the FGD1 and FGD2y samples acting as control samples to the FGD2x signal sample. Figure 6.14 shows the reconstructed and true z coordinates of the vertex position for both the FGD2x (Figures 6.14a and 6.14b) and FGD2y (Figures 6.14c and 6.14d) total signal samples. As expected, the FGD2x sample contains events primarily occurring within a water layer, with some additional events from an X layer and, to a lesser extent, a Y layer. Conversely, the FGD2y sample is mostly formed from interactions with first hits reconstructed in a Y layer, with a smaller contribution from water and X layer events.

For completeness, Figure 6.15 shows the reconstructed and true initial z position for each of the total CC1π^+ samples in FGD1, FGD2x and FGD2y. The FGD2 sample is split into FGD2x and FGD2y layers based on the active layer that the vertex is reconstructed in; in the few cases where the vertex is reconstructed in a water layer, the closest active layer is



(a) Reconstructed z, O

(b) True z, O

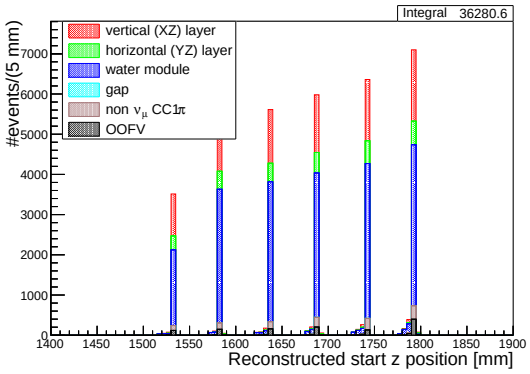
(c) Reconstructed z, C

(d) True z, C

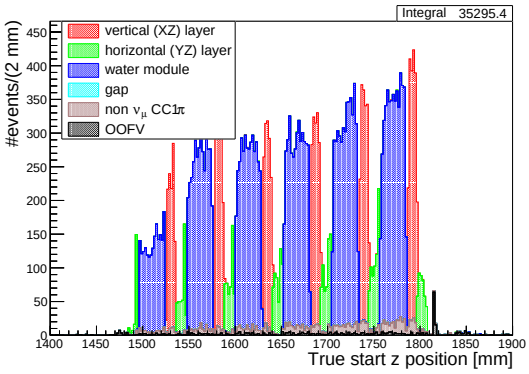
Figure 6.13: Top: Reconstructed track start z coordinate for the muon candidate in the total FGD2 CC1π^+ selection, for true signal interactions on oxygen. (a) shows the reconstructed position, while (b) shows the true position of the reconstructed object. Bottom: First hit z coordinate for the muon candidate in the total FGD2 CC1π^+ selection, for true signal interactions on carbon. (c) shows the reconstructed position, while (d) shows the true position of the reconstructed object. Category shows the true FGD layer type that the interaction happened in. Sample plotted is the total CC1π^+ signal sample.

assigned. These are drawn using a category to show the breakdown of which events are true CC1π^+ interactions, along with what nuclei the interaction happens on. In FGD1 the large majority of interactions occur on carbon nuclei, with some background on hydrogen, which is to be expected from the use of hydrocarbon plastic scintillator. It is also interesting to note the small amount of signal interactions on oxygen, which must come exclusively from other materials since FGD1 does not contain water modules. The majority of the true signal interactions on oxygen come from the FGD2x sample; however, there is also a significant amount of interactions on carbon in this sample (24.4% true CC1π^+, 37.8% total), due to interactions truly happening in an X layer. The FGD2y sample is, as expected, mostly carbon interactions, but with a significant contribution from oxygen interactions backward migrated from a water layer, or occurring on oxygen within an XY module (9.9% true

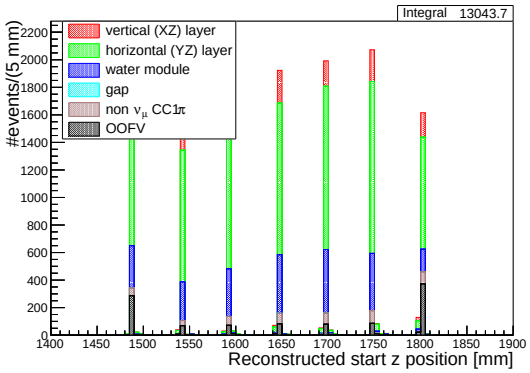
CC1π^+, 17.0% total). This sample is also the smallest, with roughly 13000 events in the analysed MC, compared to 56000 in FGD1. It is clear from this that when organising the binning scheme for the cross-section analysis, particularly for the reconstructed samples, the FGD2y samples will drive the decisions on binning, since low statistics issues will appear here first.



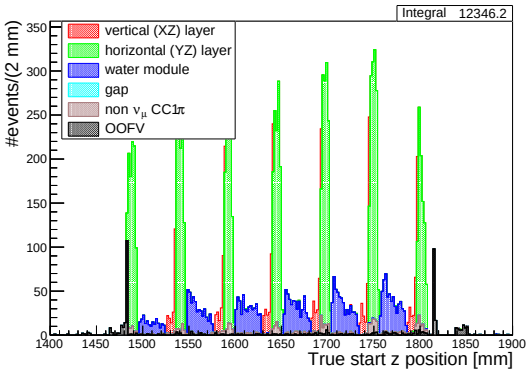
(a) Reconstructed z, X layer



(b) True z, X layer



(c) Reconstructed z, Y layer

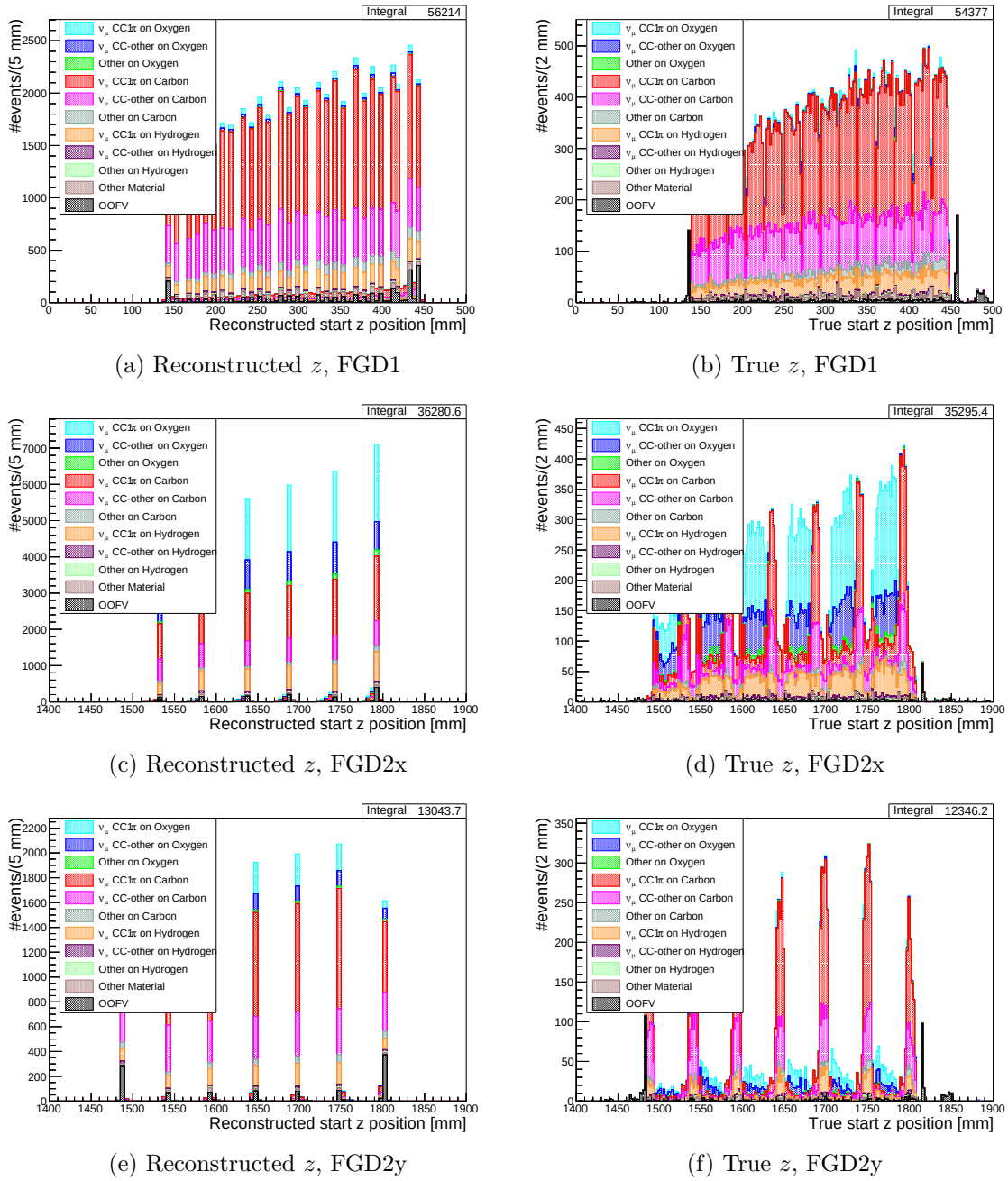


(d) True z, Y layer

Figure 6.14: Top: Reconstructed track start z coordinate for the muon candidate in the total FGD2 CC1π^+ selection, for events reconstructed in an FGD2 X layer. (a) shows the reconstructed position, while (b) shows the true position of the reconstructed object. Bottom: First hit z coordinate for the muon candidate in the total FGD2 CC1π^+ selection, for events reconstructed in an FGD2 Y layer. (c) shows the reconstructed position, while (d) shows the true position of the reconstructed object. Category shows the true FGD layer type that the interaction happened in. Sample plotted is the total CC1π^+ signal sample.

6.2.5 Summary of Signal Selection

As mentioned in Section 6.1.7, the CC1π^+ selection is branched into four signal samples: the total CC1π^+ sample; CC1π^+ using TPC pions (TPC); CC1π^+ using pions from Michel decay in the FGD (ME); and CC1π^+ using isolated FGD pions (isoFGD). The first of these is used simply for the purpose of checking the overall sample kinematics, whilst it is the latter



(a) Reconstructed z, FGD1

(b) True z, FGD1

(c) Reconstructed z, FGD2x

(d) True z, FGD2x

(e) Reconstructed z, FGD2y

(f) True z, FGD2y

Figure 6.15: Track start z coordinate in reconstruction and truth, for FGD1, FGD2x and FGD2y samples, categorised by the true topology and nuclear target of the interaction. Sample plotted is the total CC1π^+ signal sample.

three that will be used for the cross-section fit. The breakdown of the total CC1π^+ sample in terms of the three sub-samples is given in Table 6.I. Expected muon and pion kinematic distributions for the total CC1π^+ sample are presented in Figures 6.16 and 6.17. The reconstructed momentum is found from the global momentum of the selected track, whilst the angle θ is defined as the angle between the outgoing particle and the neutrino direction. The same distributions for the TPC, ME and isoFGD subsamples are shown in Figures 6.18 and 6.19, Figures 6.20 and 6.21, and Figures 6.22 and 6.23, respectively. These plots are

categorised by the true interaction topology, using a more detailed topology definition that breaks down the CCOther background into the different flavours and multiplicity of pions present. The percentage composition of the samples for each true topology is given in Table 6.II. Tables 6.III and 6.IV give the percentage composition of the samples broken down by true reaction type and true μ^- candidate particle type, respectively.

Signal Sample	CC1π^+ Total Composition (%)		
	FGD1	FGD2x	FGD2y
TPC	51.20	65.03	60.37
ME	35.72	26.70	28.94
isoFGD	13.08	8.27	10.69

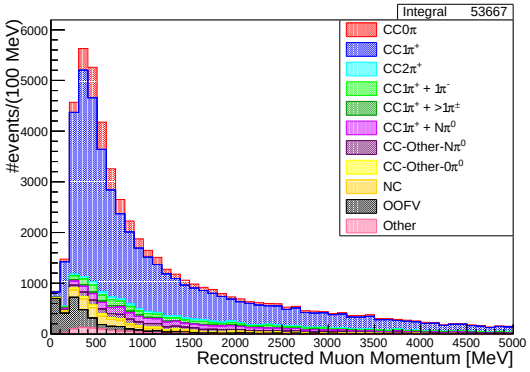
Table 6.I: Total CC1π^+ breakdown in terms of the signal sub-samples that will be used in the fit.

Regarding the total CC1π^+ samples, by the end of the selection described, a signal sample of 65.14% purity is achieved for FGD1, with the FGD2x selection much the same at 65.40%, and FGD2y slightly lower at 59.85%. The total samples are in majority made up of events that go into the TPC samples, where the contribution ranges from 51.20% in FGD1 to 65.03% in FGD2x. Events in these samples have pions that escape into the TPCs, and therefore a much larger range of pion momentum values is observed (Figure 6.19). These pions are almost all forward-going, with a $\cos\theta$ distribution starting at around 0.3, and peaking at 1. The main background to the CC1π^+ TPC samples is events featuring a π^0 that is not reconstructed, with a total contribution to the selected signal of between 9.05% and 10.56%.

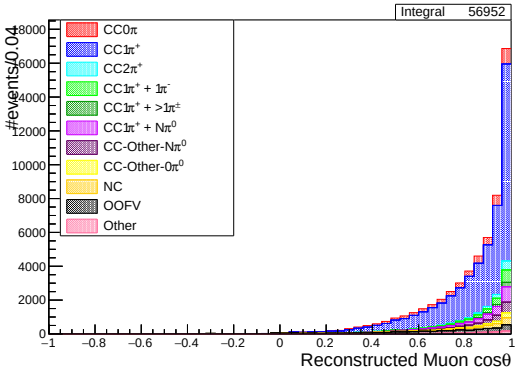
The second largest contribution to the total samples comes from events where the pion is identified in the FGD by the presence of the Michel electron that it decays to. These samples constitute between 35.72% and 26.70% of the total. The fact that these samples rely on Michel electrons to simply infer the presence of pions means that, as standard, no kinematic information would be available for the decayed pion. This shows the need to implement reconstruction, in order to avoid a large reduction in pion statistics. The method for this is detailed in Section 6.4, and the reconstruction is used in order to obtain the kinematics shown in Figure 6.21. As can be seen in Figure 6.20, the major background to the true CC1π^+ interactions in these samples comes from OOFV events, with contributions ranging from 9.16% in FGD1 to 16.55% in FGD2y. This is a particular problem of the Michel electron identification, which relies on delayed hits corresponding to the pion lifetime; hits in the delayed FGD time bin may easily come from OOFV events. However, the fact that the majority of this background is found in the very low muon momentum bins means that these events will not end up being used in the fit, as they occur in regions that will be removed by phase space constraints. Therefore most of the background events of this particular type will not make it into the binned likelihood fit.

The final and smallest contribution to the total CC1π^+ samples are the events from the isoFGD samples, which contribute between 8.27% for FGD2x and 13.08% for FGD1. These are much smaller because the majority of pions with momentum too low to reach

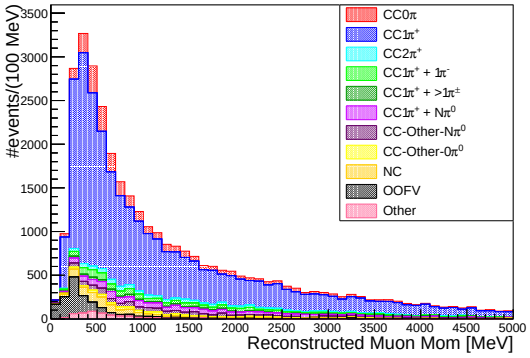
the TPC will decay within the FGD. It is only when the Michel decay cannot be seen that the reconstruction falls back to searching for isolated FGD tracks. Figure 6.23 shows the kinematics of the pions selected in this sample, which are generally very low momentum, with few events having a reconstructed pion momentum of greater than 400 MeV. The $\cos\theta$ distribution can be seen to increase from the most forward or backward-going angles, to peak at $|\cos\theta| \approx 0.3$. The high-angle region in the centre then has very low population due to a restriction placed within the selection, removing events with an initial reconstructed $\cos\theta$ value of between –0.3 and 0.3 due to difficulties in assessing the systematic error in this region. However, a later correction to the vertex placement of isolated FGD tracks causes a small number of events to migrate into this region. This will be dealt with in preparing the samples for input to the fitter by not including the events that get migrated into this region. The major background to true CC1π^+ selection in these samples is CC0π events, where protons produced in the interaction are mistakenly reconstructed as positive pions. This accounts for between 14.86% (FGD1) and 29.63% (FGD2y) of the selected samples.



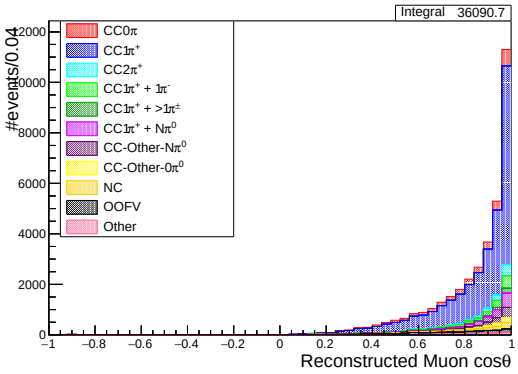
(a) Muon momentum - FGD1



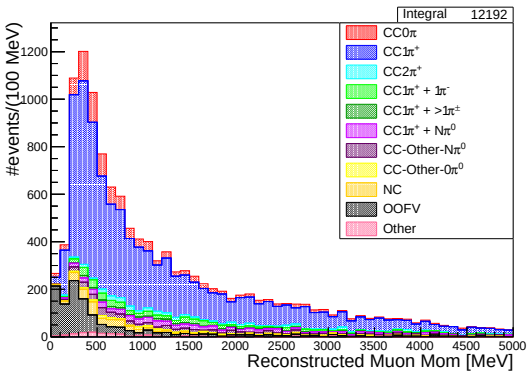
(b) Muon $\cos\theta$ - FGD1



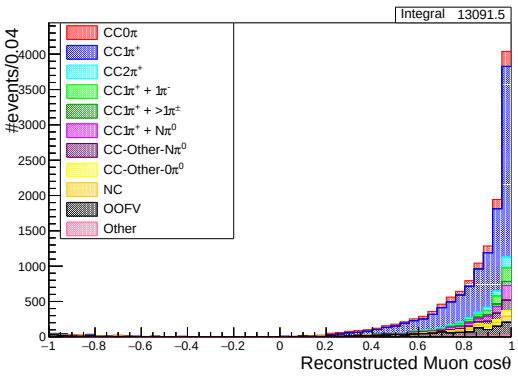
(c) Muon momentum - FGD2x



(d) Muon $\cos\theta$ - FGD2x

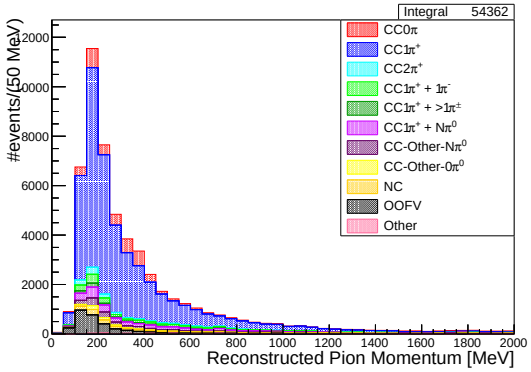


(e) Muon momentum – FGD2y

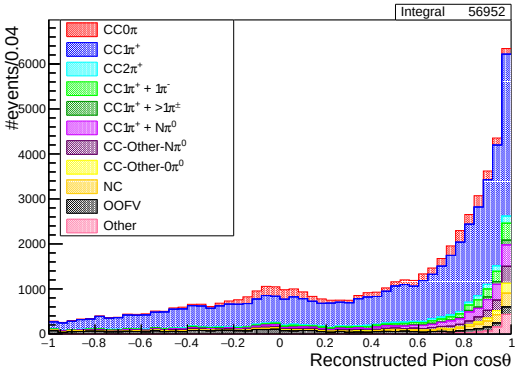


(f) Muon $\cos\theta$ – FGD2y

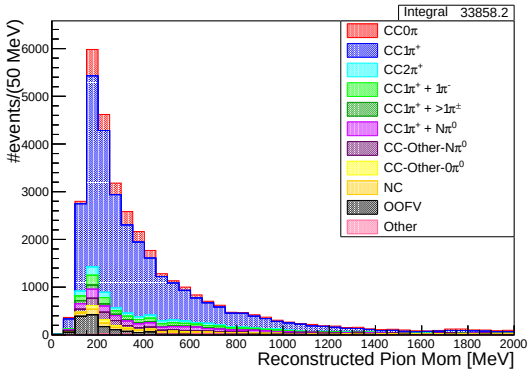
Figure 6.16: Muon kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the total CC1π^+ sample.



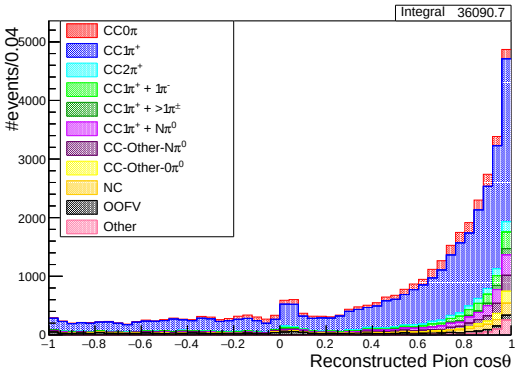
(a) Pion momentum – FGD1



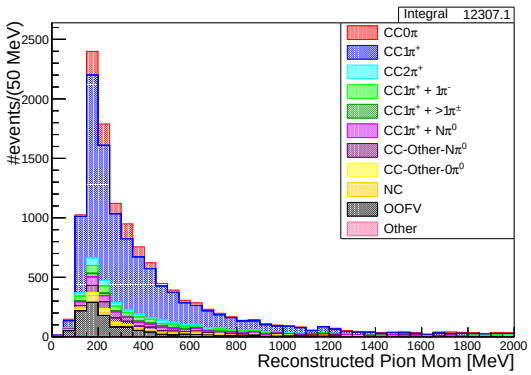
(b) Pion $\cos\theta$ – FGD1



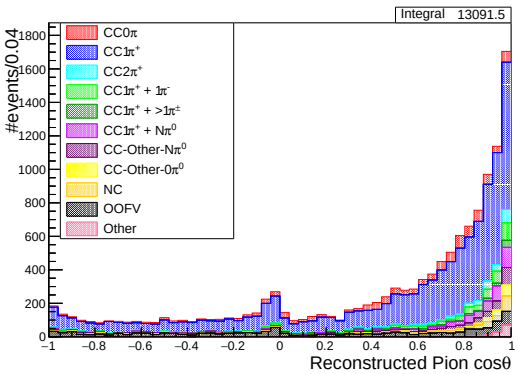
(c) Pion momentum – FGD2x



(d) Pion $\cos\theta$ – FGD2x

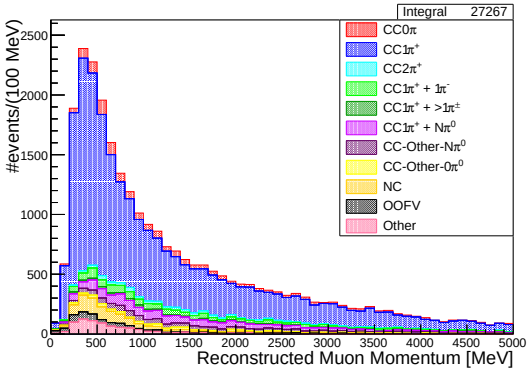


(e) Pion momentum – FGD2y

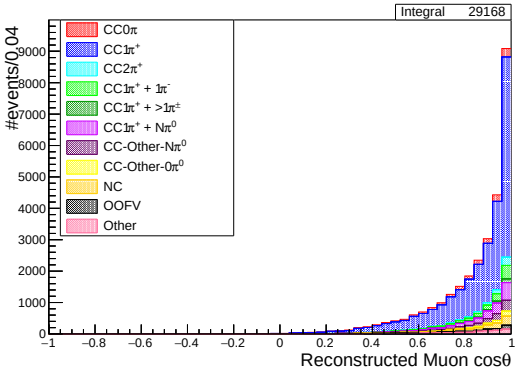


(f) Pion $\cos\theta$ – FGD2y

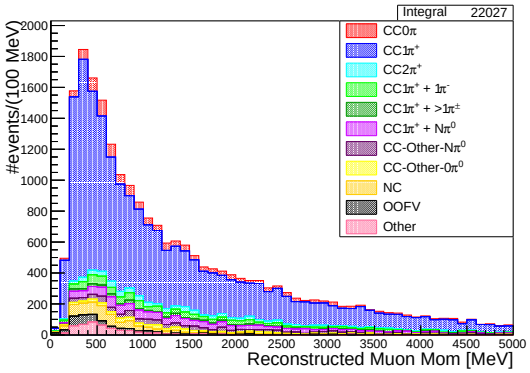
Figure 6.17: Pion kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the total CC1π^+ sample.



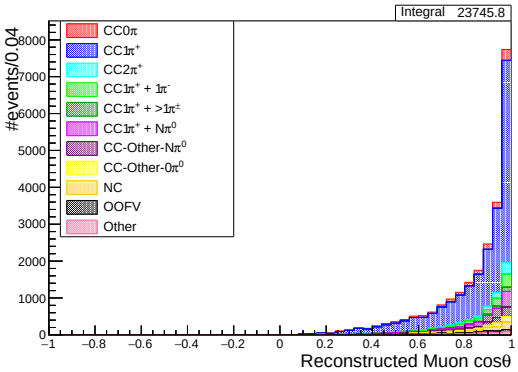
(a) Muon momentum – FGD1



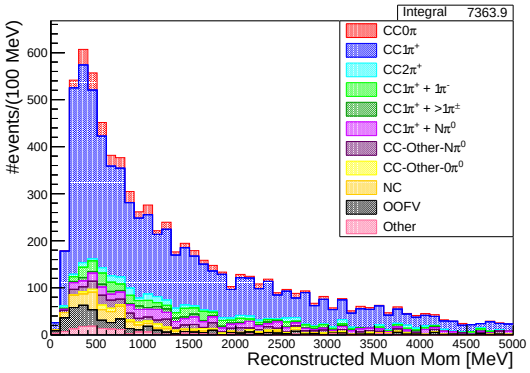
(b) Muon $\cos\theta$ – FGD1



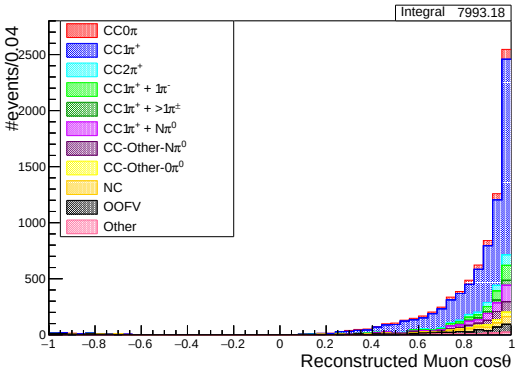
(c) Muon momentum – FGD2x



(d) Muon $\cos\theta$ – FGD2x

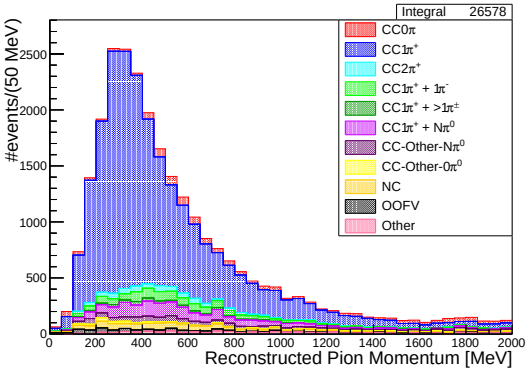


(e) Muon momentum – FGD2y

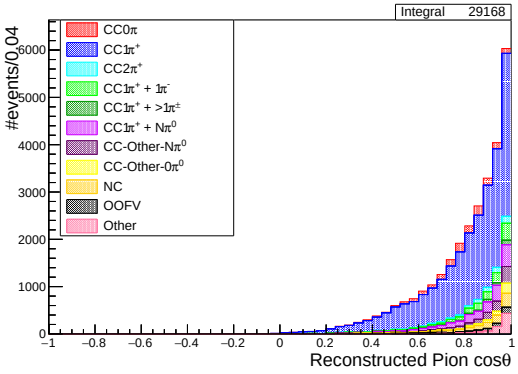


(f) Muon $\cos\theta$ – FGD2y

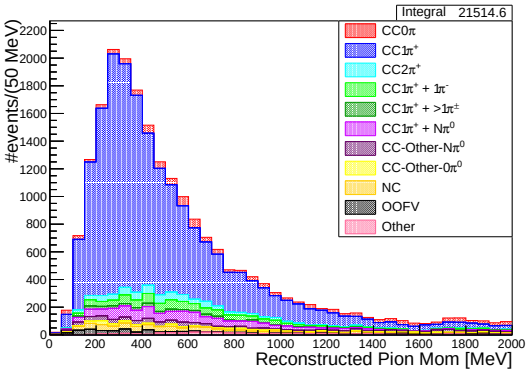
Figure 6.18: Muon kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the TPC CC1π^+ sample.



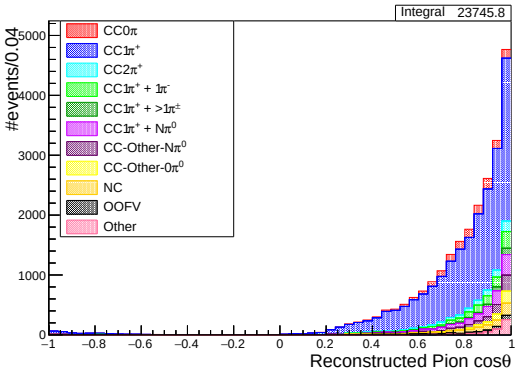
(a) Pion momentum – FGD1



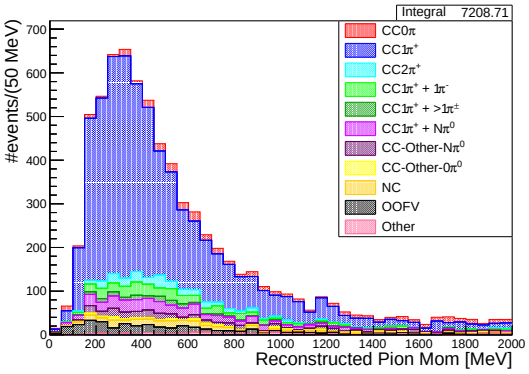
(b) Pion $\cos\theta$ – FGD1



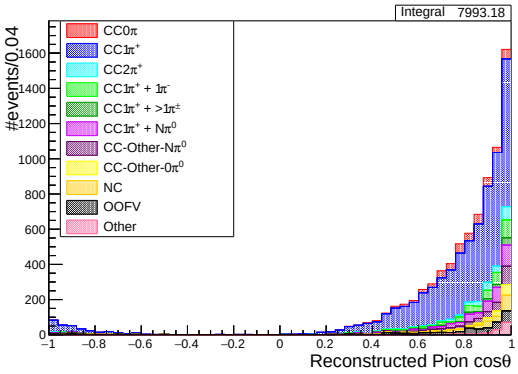
(c) Pion momentum – FGD2x



(d) Pion $\cos\theta$ – FGD2x

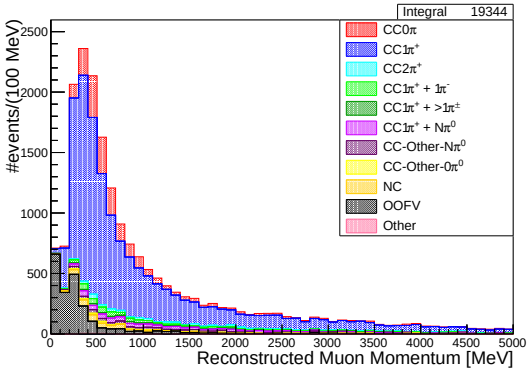


(e) Pion momentum – FGD2y

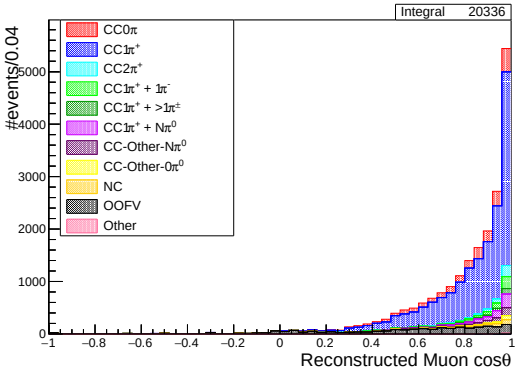


(f) Pion $\cos\theta$ – FGD2y

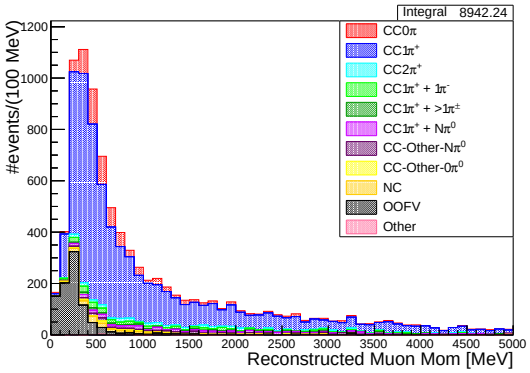
Figure 6.19: Pion kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the TPC CC1π^+ sample.



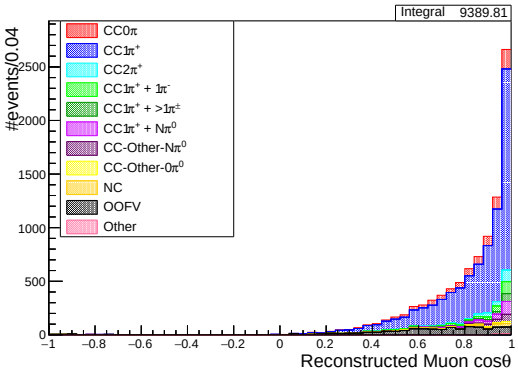
(a) Muon momentum – FGD1



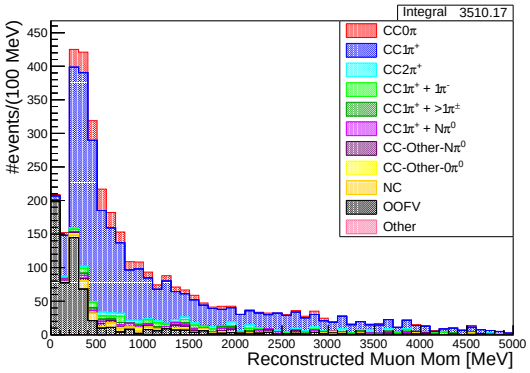
(b) Muon $\cos\theta$ – FGD1



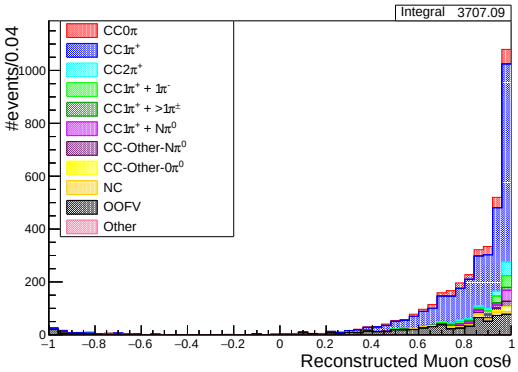
(c) Muon momentum – FGD2x



(d) Muon $\cos\theta$ – FGD2x

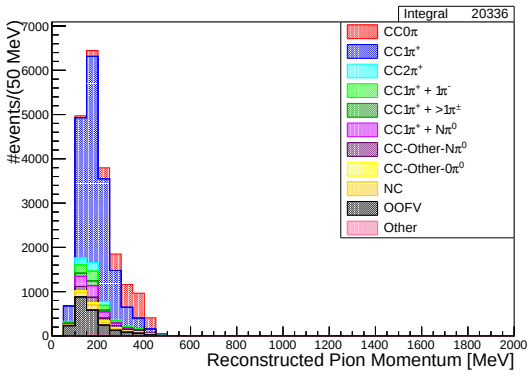


(e) Muon momentum – FGD2y

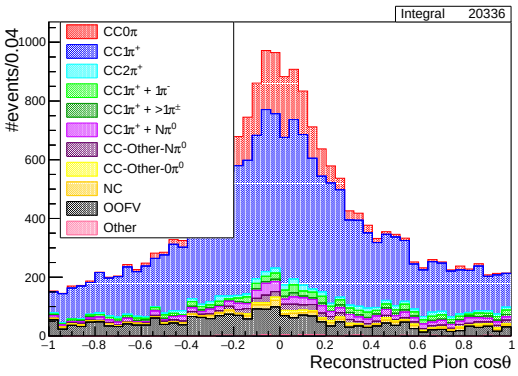


(f) Muon $\cos\theta$ – FGD2y

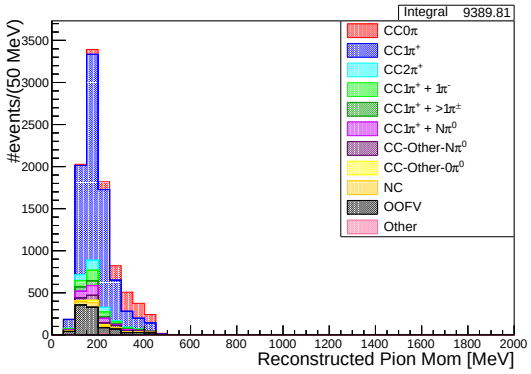
Figure 6.20: Muon kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the Michel electron CC1π^+ sample.



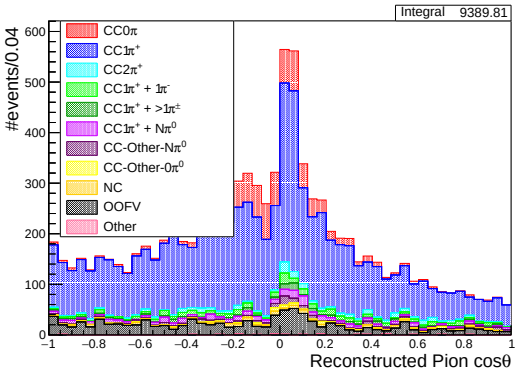
(a) Pion momentum – FGD1



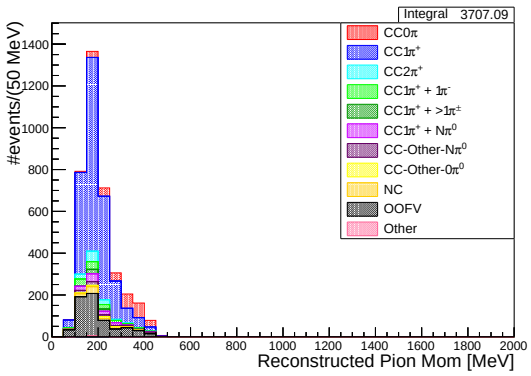
(b) Pion $\cos\theta$ – FGD1



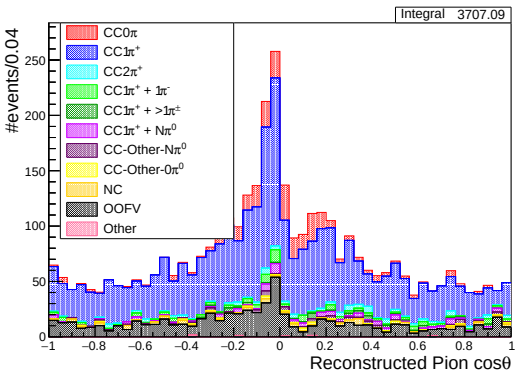
(c) Pion momentum – FGD2x



(d) Pion $\cos\theta$ – FGD2x

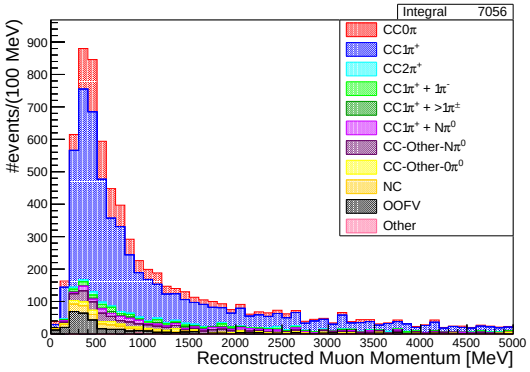


(e) Pion momentum – FGD2y

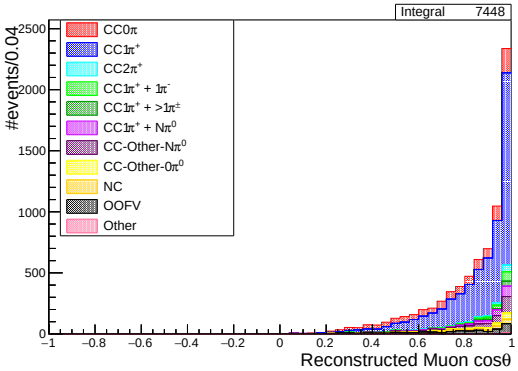


(f) Pion $\cos\theta$ – FGD2y

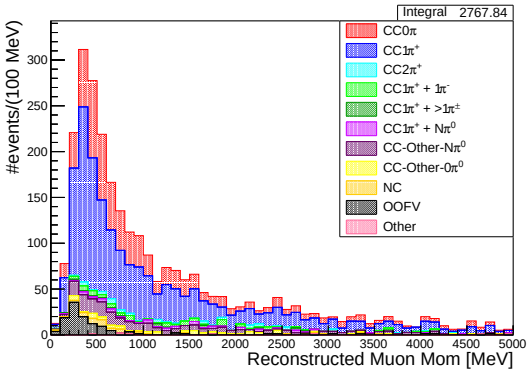
Figure 6.21: Pion kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the Michel electron CC1π^+ sample.



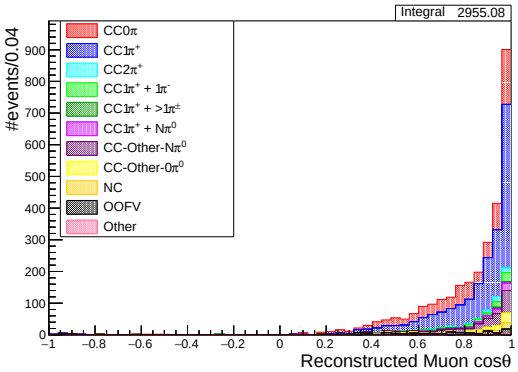
(a) Muon momentum – FGD1



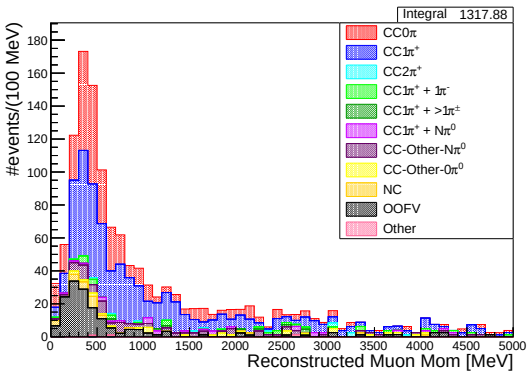
(b) Muon $\cos\theta$ – FGD1



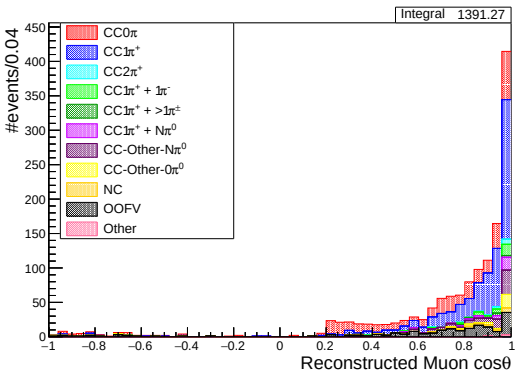
(c) Muon momentum – FGD2x



(d) Muon $\cos\theta$ – FGD2x

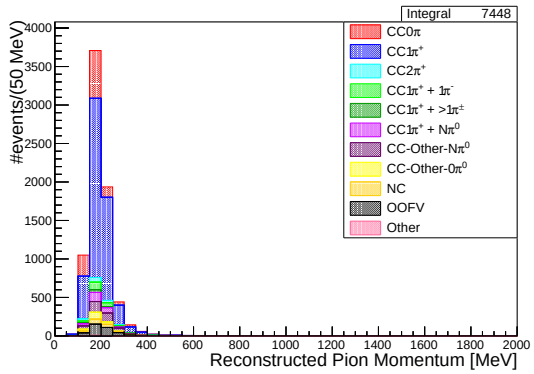


(e) Muon momentum – FGD2y

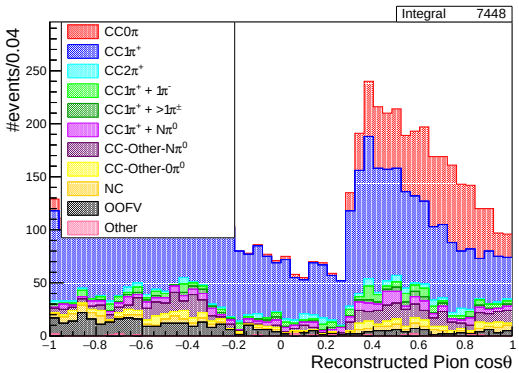


(f) Muon $\cos\theta$ – FGD2y

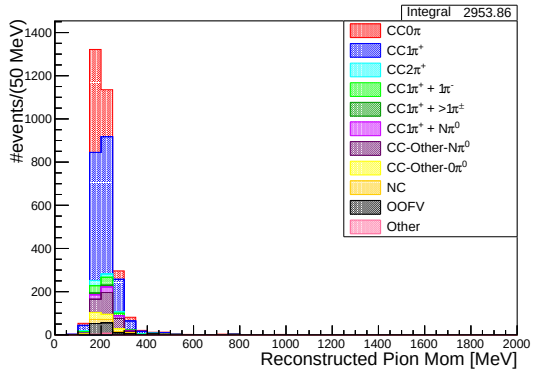
Figure 6.22: Muon kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the isolated FGD CC1π^+ sample.



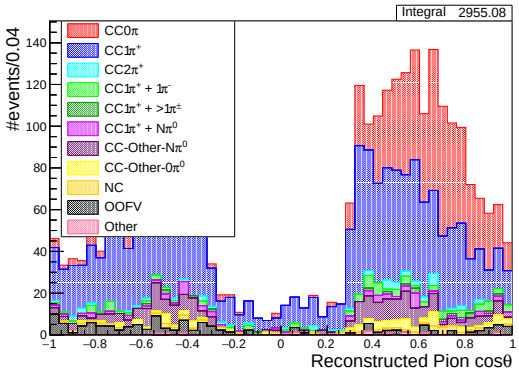
(a) Pion momentum – FGD1



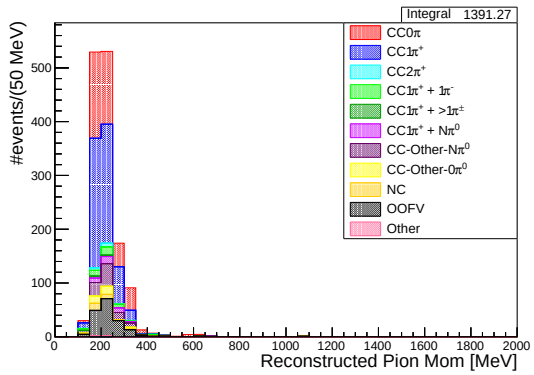
(b) Pion $\cos\theta$ – FGD1



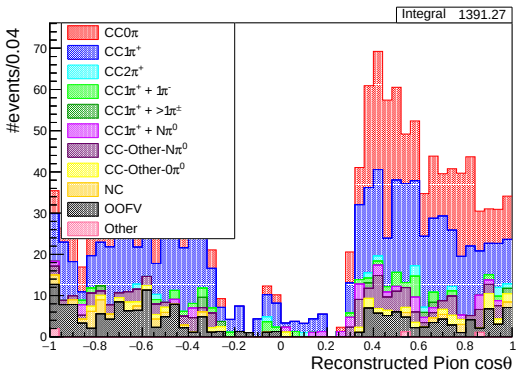
(c) Pion momentum – FGD2x



(d) Pion $\cos\theta$ – FGD2x



(e) Pion momentum – FGD2y



(f) Pion $\cos\theta$ – FGD2y

Figure 6.23: Pion kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the isolated FGD CC1π^+ sample.

Sample	Topology	MC Composition (%)		
		FGD1	FGD2x	FGD2y
CC1π^+ Total	CC0π	8.39	8.29	8.65
	CC1π^+	65.14	65.40	59.85
	CC2π^+	2.42	3.24	3.02
	CC1π^++1π^-	3.81	4.26	4.23
	CC1π^++> 1$\pi^\pm$	1.16	1.42	1.41
	CC1π^++Nπ^0	5.09	4.70	4.61
	CCOther-Nπ^0	3.46	3.32	3.70
	CCOther-0π^0	1.49	1.53	1.69
	NC	2.73	2.72	2.92
	OOFV	4.75	3.63	8.62
	Other	1.56	1.48	1.30
CC1π^+ TPC	CC0π	5.15	5.64	5.34
	CC1π^+	66.17	67.04	62.69
	CC2π^+	2.40	3.57	3.41
	CC1π^++1π^-	4.60	4.76	5.14
	CC1π^++> 1$\pi^\pm$	1.17	1.41	1.57
	CC1π^++Nπ^0	6.63	5.74	5.91
	CCOther-Nπ^0	3.93	3.31	4.16
	CCOther-0π^0	1.72	1.77	1.91
	NC	3.62	3.30	3.61
	OOFV	1.78	1.34	4.31
	Other	2.82	2.12	1.94
CC1π^+ ME	CC0π	10.70	9.18	7.78
	CC1π^+	64.50	66.17	61.26
	CC2π^+	2.59	2.90	2.86
	CC1π^++1π^-	3.11	3.42	3.00
	CC1π^++> 1$\pi^\pm$	1.19	1.65	1.40
	CC1π^++Nπ^0	3.68	2.87	2.62
	CCOther-Nπ^0	2.22	2.09	1.60
	CCOther-0π^0	0.95	0.64	0.93
	NC	1.68	1.49	1.72
	OOFV	9.16	9.34	16.55
	Other	0.22	0.23	0.28
CC1π^+ isoFGD	CC0π	14.86	26.57	29.63
	CC1π^+	62.84	49.95	40.15
	CC2π^+	2.00	1.68	1.19
	CC1π^++1π^-	2.62	2.90	2.32
	CC1π^++> 1$\pi^\pm$	1.05	0.75	0.49
	CC1π^++Nπ^0	2.90	2.10	2.41
	CCOther-Nπ^0	4.97	7.29	6.57
	CCOther-0π^0	2.00	2.40	2.41
	NC	2.09	2.01	2.18
	OOFV	4.44	4.01	12.33
	Other	0.25	0.35	0.33

Table 6.II: True topology breakdown of the total CC1π^+ sample and individual subsamples.

Sample	Interaction	MC Composition (%)		
		FGD1	FGD2x	FGD2y
CC1π^+ Total	CCQE	6.86	6.50	6.89
	2p2h	1.22	1.10	1.14
	RES	59.51	58.52	53.85
	DIS	20.79	23.28	23.02
	COH	2.79	3.03	2.50
	NC	2.65	2.66	2.83
	CC-$\bar{\nu}_\mu$	1.34	1.21	1.03
	CC-ν_e , CC-$\bar{\nu}_e$	0.08	0.06	0.10
	Other	0.01	0.01	0.02
	OOFV	4.75	3.63	8.62
CC1π^+ TPC	CCQE	4.01	4.07	4.05
	2p2h	0.61	0.65	0.50
	RES	55.95	56.91	52.80
	DIS	26.62	27.67	29.37
	COH	4.87	4.25	3.76
	NC	3.52	3.22	3.49
	CC-$\bar{\nu}_\mu$	2.52	1.80	1.58
	CC-ν_e , CC-$\bar{\nu}_e$	0.10	0.07	0.12
	Other	0.01	0.01	0.03
	OOFV	1.78	1.34	4.31
CC1π^+ ME	CCQE	9.34	8.52	6.99
	2p2h	1.70	1.38	1.24
	RES	62.26	63.20	58.85
	DIS	15.24	15.44	14.13
	COH	0.54	0.59	0.38
	NC	1.64	1.45	1.66
	CC-$\bar{\nu}_\mu$	0.08	0.04	0.13
	CC-ν_e , CC-$\bar{\nu}_e$	0.04	0.04	0.06
	Other	0.01	0.00	0.00
	OOFV	9.16	9.34	16.55
CC1π^+ isoFGD	CCQE	11.27	19.57	22.74
	2p2h	2.35	3.76	4.49
	RES	66.03	56.65	46.78
	DIS	12.95	12.78	10.30
	COH	0.72	1.01	0.94
	NC	1.98	2.01	2.10
	CC-$\bar{\nu}_\mu$	0.15	0.21	0.24
	CC-ν_e , CC-$\bar{\nu}_e$	0.10	0.00	0.08
	Other	0.00	0.00	0.00
	OOFV	4.44	4.01	12.33

Table 6.III: True reaction breakdown of the total CC1π^+ sample and individual subsamples.

Sample	Particle	MC Composition (%)		
		FGD1	FGD2x	FGD2y
CC1π^+ Total	μ^-	91.69	92.13	90.81
	e^-	0.18	0.12	0.28
	π^-	5.75	5.27	5.79
	μ^+	0.69	0.69	0.58
	e^+	0.02	0.01	0.07
	π^+	1.50	1.57	2.19
	p	0.15	0.18	0.26
	Other	0.02	0.03	0.03
CC1π^+ TPC	μ^-	91.27	92.82	91.25
	e^-	0.21	0.14	0.26
	π^-	7.98	6.49	7.37
	μ^+	0.04	0.07	0.06
	e^+	0.02	0.01	0.03
	π^+	0.30	0.49	0.79
	p	0.15	0.15	0.20
	Other	0.04	0.05	0.03
CC1π^+ ME	μ^-	91.30	90.04	89.35
	e^-	0.12	0.05	0.31
	π^-	3.36	2.93	3.26
	μ^+	1.84	2.41	1.78
	e^+	0.01	0.00	0.16
	π^+	3.21	4.30	4.76
	p	0.15	0.25	0.34
	Other	0.01	0.01	0.04
CC1π^+ isoFGD	μ^-	94.42	94.79	92.10
	e^-	0.21	0.17	0.33
	π^-	3.45	2.90	3.41
	μ^+	0.12	0.28	0.41
	e^+	0.03	0.00	0.00
	π^+	1.63	1.67	3.40
	p	0.15	0.19	0.34
	Other	0.00	0.00	0.00

Table 6.IV: Leading track true particle ID for the total CC1π^+ sample and individual subsamples.

6.3 Control Samples

In order to accurately measure the CC1π^+ cross section, we wish to have as pure a CC1π^+ sample as is practically possible. Unfortunately, due to reconstruction effects and detector inefficiency, it is impossible to ever obtain a 100% pure signal sample; some backgrounds will always remain since they cannot be differentiated from true signal via detector observables. It can be seen from Tables 6.II and 6.III that the selection has two major backgrounds. Roughly half of these background events are caused by interactions where the true topology involved multiple pions in the final state; these are most likely Deep Inelastic Scattering (DIS) interactions, where several pions were produced but the only one reconstructed was a π^+, causing the event to be tagged as signal. In the standard multi-pi selection these events all fall under the 'CCOther' category, but here are split into more detailed topology definitions. The other major background is from CC0π events, where the reconstructed pion is likely a true proton. The major backgrounds can be split into three separate groups:

- Events containing multiple *charged* pions, where only one π^+ is reconstructed.
- Events containing neutral pions.
- CC0π interactions, where the reconstructed π^+ is a true proton.

Table 6.II suggests that the first two backgrounds are the most significant across the sum of the samples, whereas the CC0π background predominantly enters from the isoFGD sample; in the FGD2 selection, between 26 and 30% of the signal sample is in fact true CC0π. This likely comes from the fact that PID is more accurate within the TPCs than FGDs, and so protons are more easily mistaken for pions in the isoFGD sample. These substantive backgrounds in the signal sample show a clear need for control samples enriched in the principal backgrounds. Based on this, three possible control samples for this selection were identified, primarily based on reconstructed topology:

1. *CC1*π^+ *+ N*$\pi^\pm$ – Require one positively charged pion, plus any number of additional charged pions.
2. *CC1*π^+ *+ N*π^0 – Require one positively charged pion, plus any number of neutral pions.
3. *CC0*π *+ Np FGD* – CC-inclusive events with no identified pions present.

The control samples are described in detail in the following sections.

6.3.1 CC1π^+ + N$\pi^\pm$

The first of the described control samples is formed to address the multiple charged pion background observed in the signal samples (cyan, light, and dark green categories in Figures 6.16 and 6.17). For the total CC1π^+, this accounts for 7.39% of the FGD1 sample, 8.92% of the FGD2x sample, and 8.66% of the FGD2y sample. Whilst this background is observed across all of the three pion detection samples, it occurs most within the TPC

sample. In order to select this control sample, the CCOther selection branch is used. This is the catch-all sample for the CCinclusive events which fail both the CC0π and CC1π^+ cuts. Several further cuts are then applied to select this background sample:

- *CCOther Cut* – Events failing CC0π or CC1π^+ cuts get migrated to this sample.
- *Nπ^+ Cut* – Applying a similar π^+ selection criterion to the one described in Section 6.1.7, positive pions in the event are identified. The only difference here is that this sample does not use Michel electrons to tag pions. The event passes this cut if *at least* one positive pion is present. Where multiple pions are present, the highest momentum candidate is selected as that whose kinematics will be measured.
- *ECal π^0 Veto* – This cut, as described in Section 6.2.2, is applied in order to remove π^0 photon candidates from the ECals.
- *TPC π^0 Veto* – Similar to the ECal version, this cut removes events based on reconstructed positron and electron candidates in the TPCs coming from π^0 decay. This is applied automatically as part of the standard single positive pion cut in the signal selection (Section 6.1.7), but has to be applied separately here.

This series of cuts should select a sample of events with at least one positive pion, and any non-zero number of additional charged pions, since events with *only* one π^+ would not have failed the CC1π^+ cut. In the case where there are multiple positive pions in the sample, the highest momentum one is taken as the leading pion; it is this pion whose kinematics will be measured. The exclusion of kinematics for pions from Michel electrons has very limited effect on the sample, since the highest momentum π^+ generally reaches the TPC. The muon and pion kinematic distributions for the CC1π^+ + N$\pi^\pm$ control sample, hereafter referred to as CS1, are presented in Figures 6.24 and 6.25. This control sample achieves a multiple pion purity of 45–49% for the three FGD samples, with a true CC1π^+ contamination of roughly 13%. Although we ideally want as little true signal in the control samples as possible, completely removing it is not feasible, and this is judged to be a reasonable amount to accept.

6.3.2 CC1π^+ + Nπ^0

The second control sample is formed to address the background coming from π^0-containing events, where the the neutral pion cannot be reconstructed. This is represented by the light and dark purple categories in Figures 6.16 and 6.17. In the total CC1π^+ branch, this accounts for 8.55% of the FGD1 sample, 8.02% of the FGD2x sample, and 8.31% of the FGD2y sample. Similarly to the backgrounds described in Section 6.3.1, these specific backgrounds are most prevalent in the TPC sample. Again, the sample is formed by first taking the CCOther selection branch, and applying a subsequent series of cuts, as follows:

- *CCOther Cut* – Events failing CC0π or CC1π^+ cuts get migrated to this sample.
- *1π^+ Cut* – A slightly modified version of the same standard single positive pion cut from the main selection (Section 6.1.7) is applied here, to select events with only

one reconstructed positive pion. As in Section 6.3.1, the only difference is that pions identified via Michel electrons are not used.

- *Reverse π^0 Veto Cut* – This cut is used to require *at least* one neutral pion in the event. The TPC and ECal π^0 vetoes from Section 6.3.1 are reversed, rejecting the event if no π^0 objects are reconstructed.

These cuts aim to select a sample of events with a single positive pion, and any non-zero number of neutral pions. However, due to reconstruction difficulties, it is fair to assume other events will mistakenly be selected, where particles such as a π^- or proton may inadvertently be reconstructed as a π^+. In this case this is advantageous, as it also allows access to the CCOther-Nπ^0 background in a phase space where the reconstructed π^+ is incorrect, the same as appears in the signal sample. As in Section 6.3.1, the exclusion of pions using the Michel tag doesn't have a large effect on statistics, as these are higher energy events where the π^+ will generally reach the TPC, and means we don't rely on kinematic reconstruction that hasn't been completely verified in the control regions, instead restricting that to its own signal sample where it can be treated with proper care. The muon and pion kinematic distributions for the CC1π^+ + Nπ^0 control sample, hereafter referred to as CS2, are presented in Figures 6.26 and 6.27. The control sample is able to achieve a combined Nπ^0 purity of between 46% and 50%, with a true CC1π^+ contamination slightly higher than that of CS1 at 13–17%.

6.3.3 CC0π + Np FGD

This control sample is designed to address issues in the CC1π^+-isoFGD signal sample, where a large part of the selected signal in fact has a true CC0π topology. For the FGD2x and y samples, this topology contributes as much as 27–30% of the total events, with the FGD1 contribution being somewhat less at ∼15%. This can be seen visually in Figures 6.22 and 6.23. This background seems to come from the misidentification of protons which only have isolated FGD tracks in the reconstruction, and incorrectly get reconstructed as positive pions. In order to form this control sample, the CC0π selection branch is taken, with the following series of cuts applied:

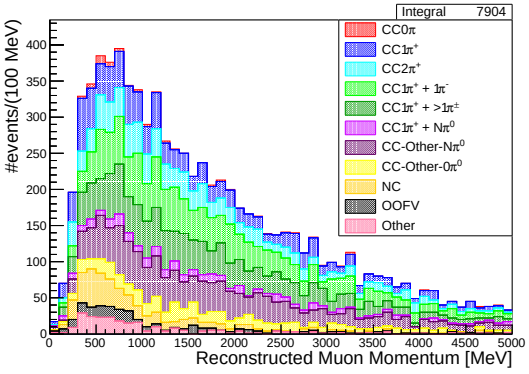
- *ECal π^0 Veto* – This cut, as described in Section 6.2.2, is applied in order to remove π^0 photon candidates from the ECals.

- *FGD Proton Cut* – Using the same FGD PID methods as are used for the pion identification, the proton multiplicity for the event is found. If the event has zero TPC protons and non-zero FGD protons, it is tagged as having FGD topology, and passes the cut. Any other proton topology will be removed from the sample.

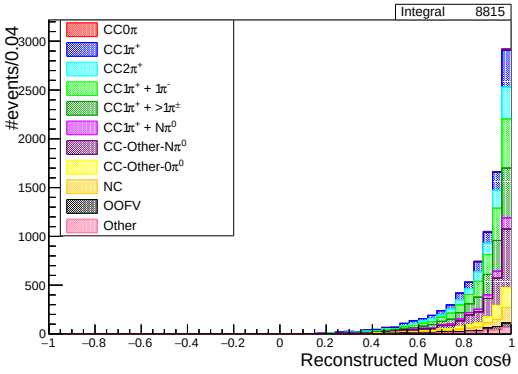
The kinematics for this control sample are presented in Figures 6.28 and 6.29. Since this control sample necessarily does not include a pion, the 'pion' kinematics presented are for the selected proton, but recalculated under the FGD pion hypothesis, thus mimicking the behaviour that the sample is designed to account for. This was considered to be the best way to fit the control sample into the 4-dimensional binning scheme employed in this

analysis, since the CC0π background to the CC1π^+ isoFGD signal sample has its protons mistakenly reconstructed as pions. In FGD1, this control sample has a purity of 83.63%, with a true CC1π^+ contamination of only 8.42%. This represents a much purer control sample than is achieved in the FGD2x and y samples, which have CC0π purity of 65.68% and 61.61%, respectively. This reflects the fact that the FGD1 CC1π^+ isoFGD sample has less CC0π background than the two FGD2 samples do, and is likely caused by more precise reconstruction due to the lack of water layers in FGD1.

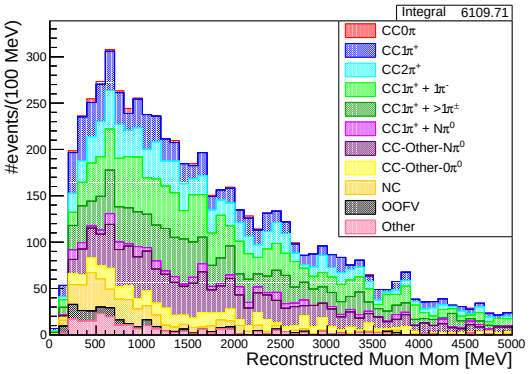
For further comparison, Appendix D provides the true invariant mass W and energy transfer Q^2 for each of the signal and control samples.



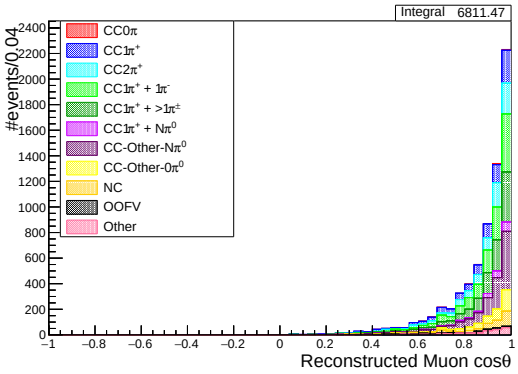
(a) Muon momentum – FGD1



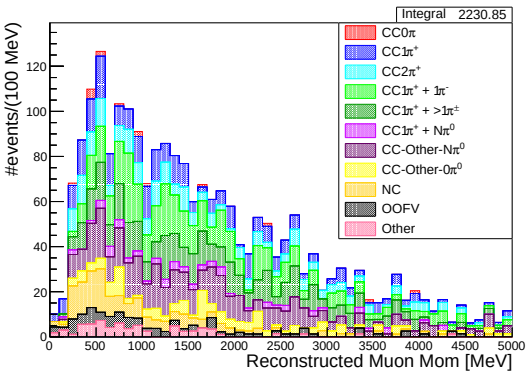
(b) Muon $\cos\theta$ – FGD1



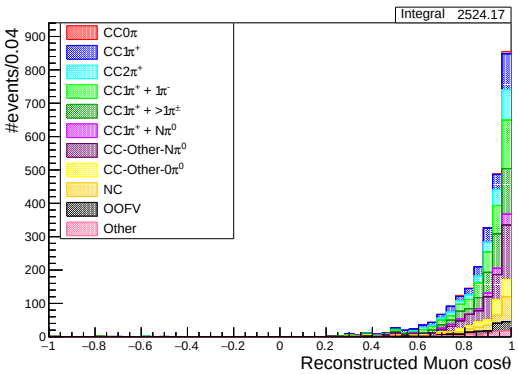
(c) Muon momentum – FGD2x



(d) Muon $\cos\theta$ – FGD2x

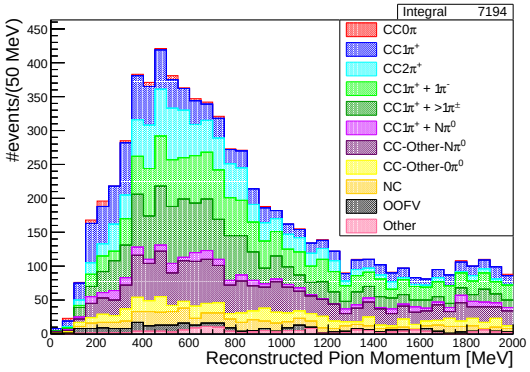


(e) Muon momentum – FGD2y

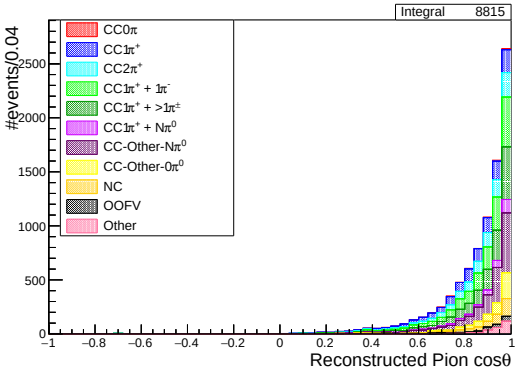


(f) Muon $\cos\theta$ – FGD2y

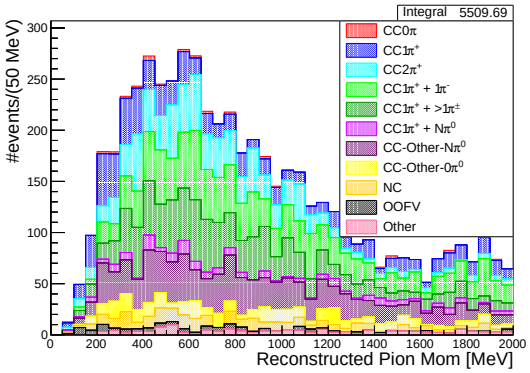
Figure 6.24: Muon kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the CC1π^+ + N$\pi^\pm$ control sample (CS1).



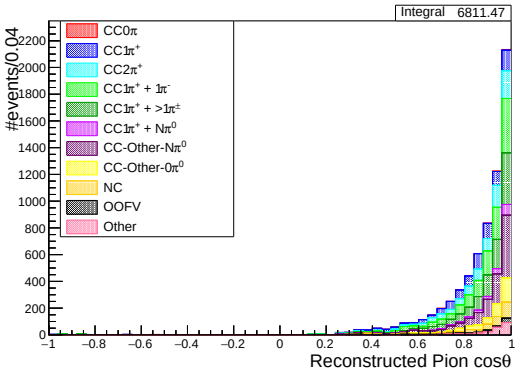
(a) Pion momentum – FGD1



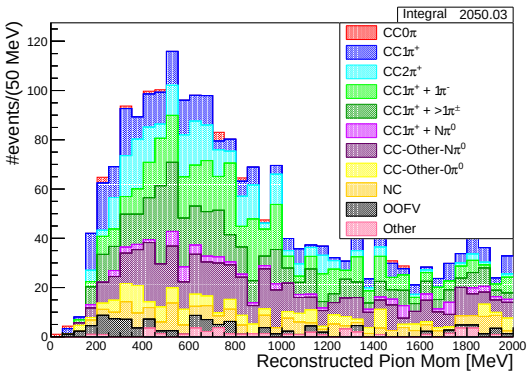
(b) Pion $\cos\theta$ – FGD1



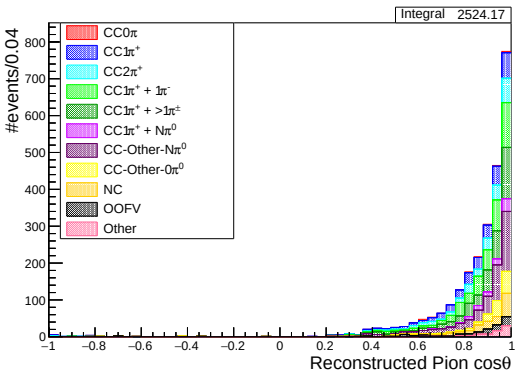
(c) Pion momentum – FGD2x



(d) Pion $\cos\theta$ – FGD2x

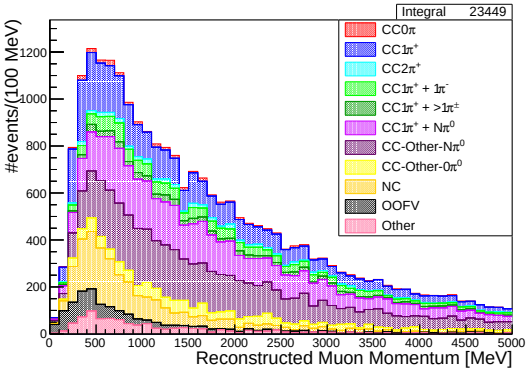


(e) Pion momentum – FGD2y

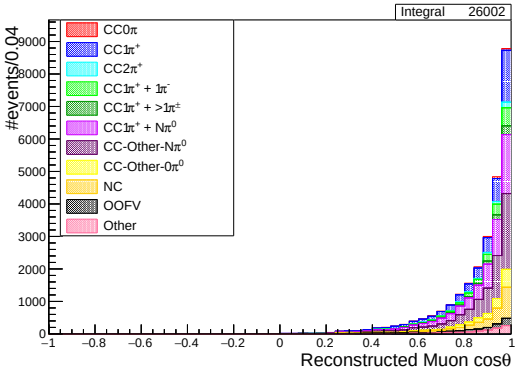


(f) Pion $\cos\theta$ – FGD2y

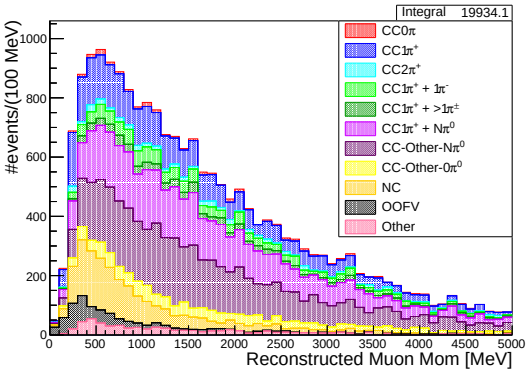
Figure 6.25: Pion kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the CC1π^+ + N$\pi^\pm$ control sample (CS1).



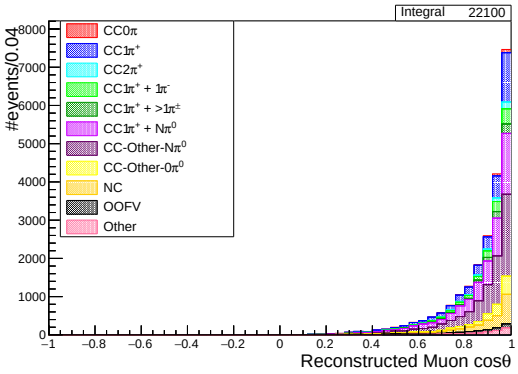
(a) Muon momentum – FGD1



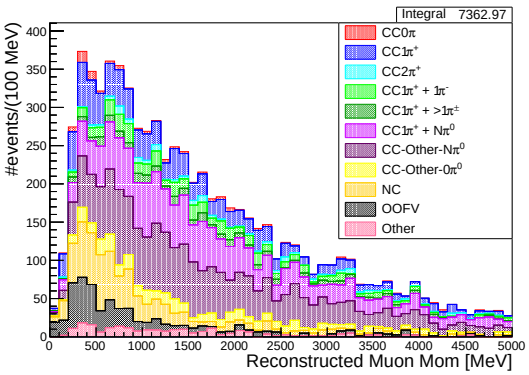
(b) Muon $\cos\theta$ – FGD1



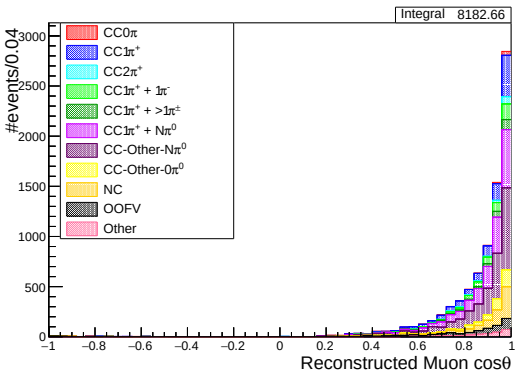
(c) Muon momentum – FGD2x



(d) Muon $\cos\theta$ – FGD2x

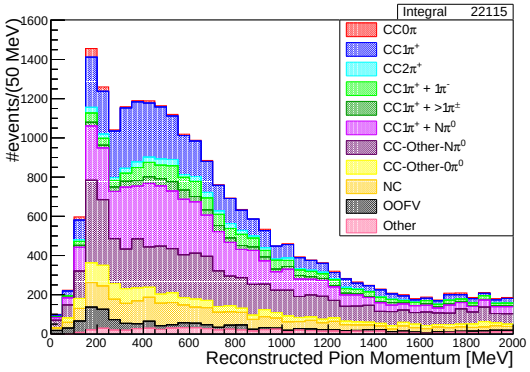


(e) Muon momentum – FGD2y

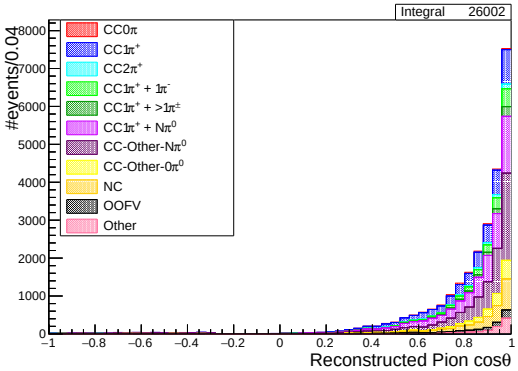


(f) Muon $\cos\theta$ – FGD2y

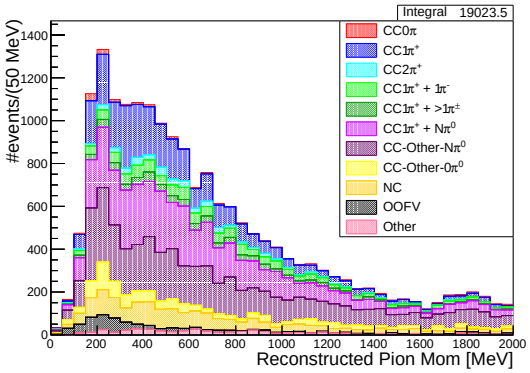
Figure 6.26: Muon kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the CC1π^+ + Nπ^0 control sample (CS2).



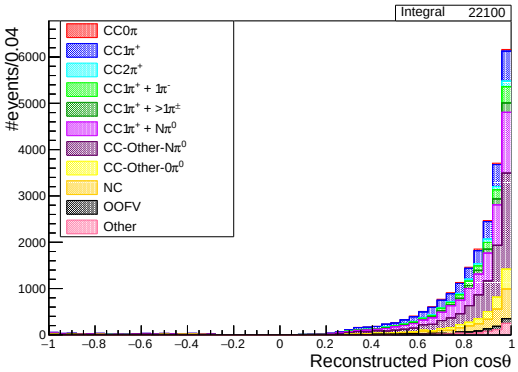
(a) Pion momentum – FGD1



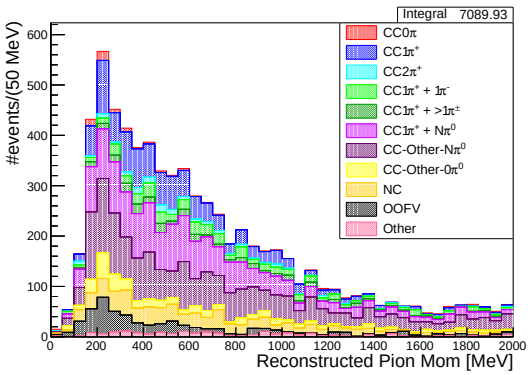
(b) Pion $\cos\theta$ – FGD1



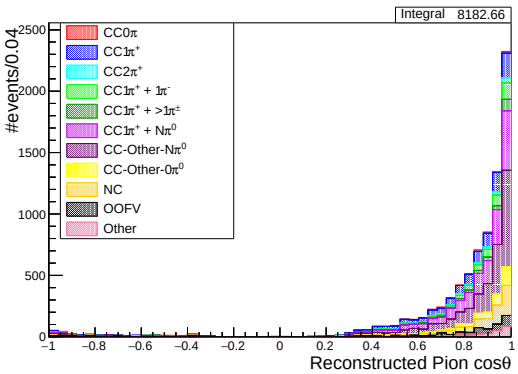
(c) Pion momentum – FGD2x



(d) Pion $\cos\theta$ – FGD2x

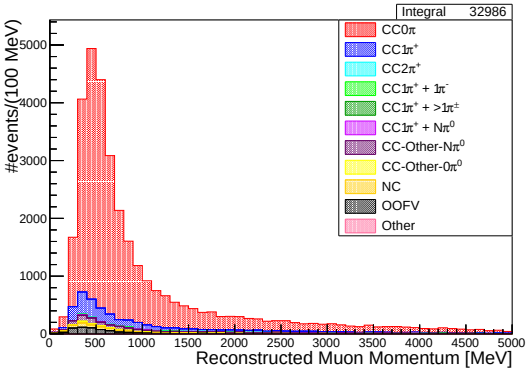


(e) Pion momentum – FGD2y

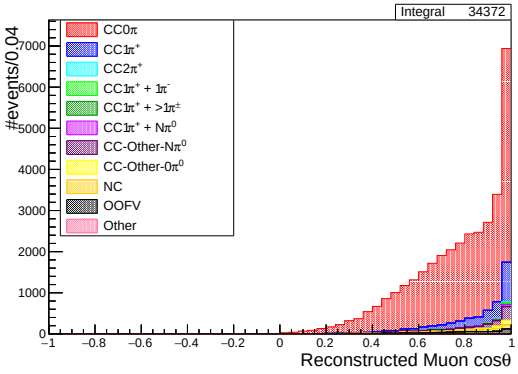


(f) Pion $\cos\theta$ – FGD2y

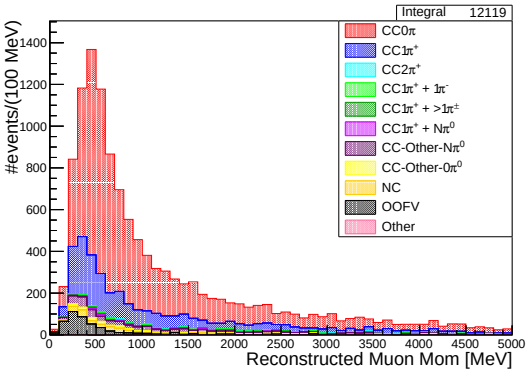
Figure 6.27: Pion kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the CC1π^+ + Nπ^0 control sample (CS2).



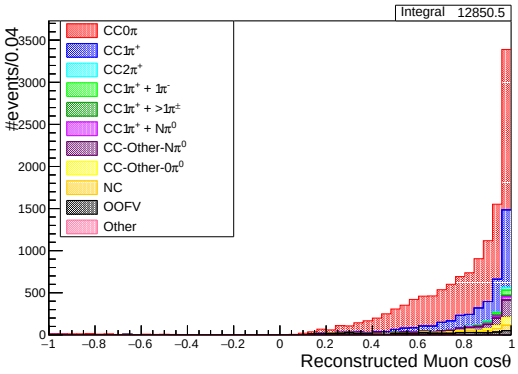
(a) Muon momentum – FGD1



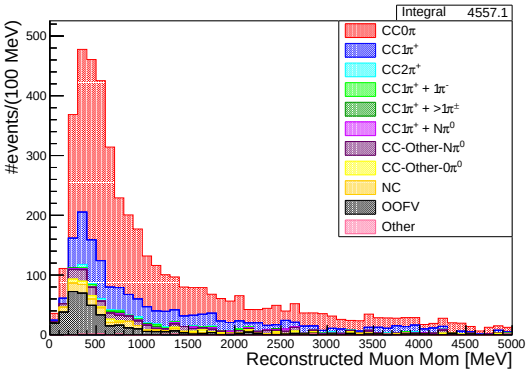
(b) Muon $\cos\theta$ – FGD1



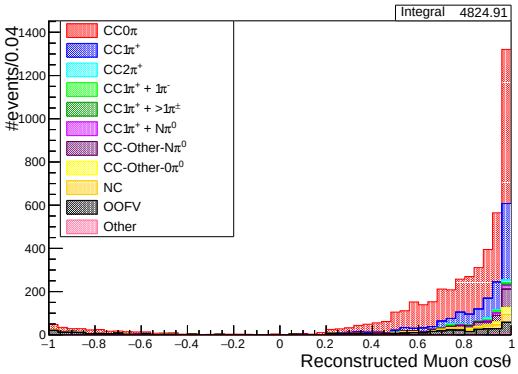
(c) Muon momentum – FGD2x



(d) Muon $\cos\theta$ – FGD2x

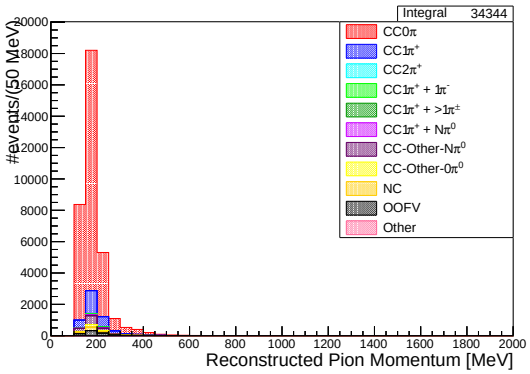


(e) Muon momentum – FGD2y

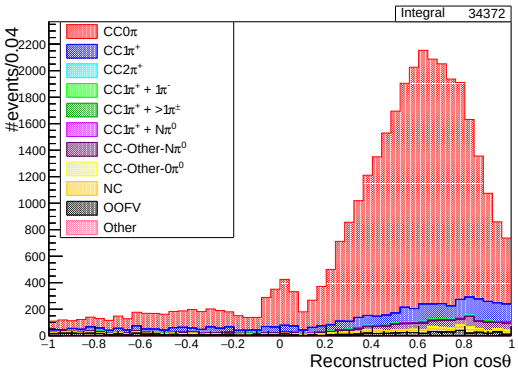


(f) Muon $\cos\theta$ – FGD2y

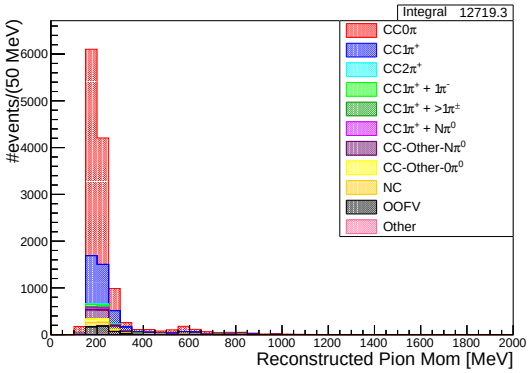
Figure 6.28: Muon kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the CC0π + Np control sample (CS3).



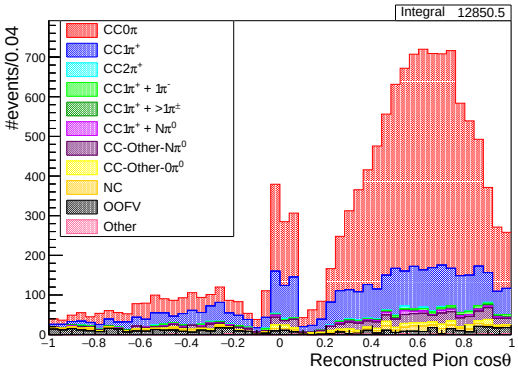
(a) 'Pion' momentum – FGD1



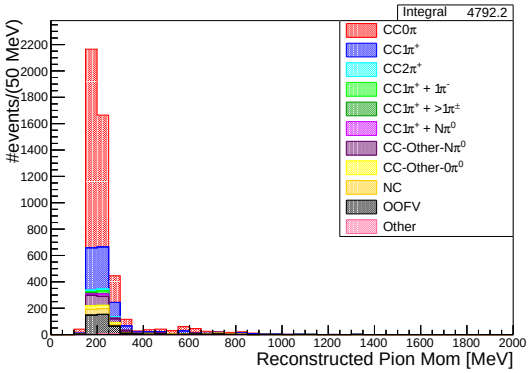
(b) 'Pion' $\cos\theta$ – FGD1



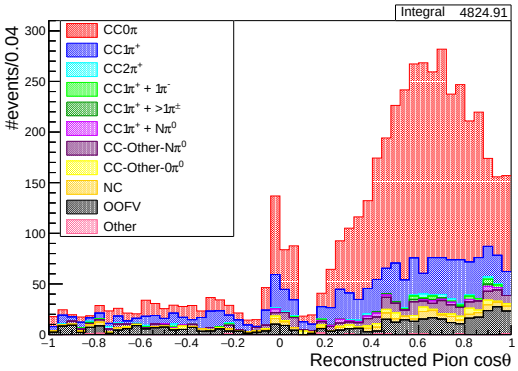
(c) 'Pion' momentum – FGD2x



(d) 'Pion' $\cos\theta$ – FGD2x



(e) 'Pion' momentum – FGD2y



(f) 'Pion' $\cos\theta$ – FGD2y

Figure 6.29: 'Pion' kinematic distributions in FGD1, FGD2x and FGD2y, categorised by true topology, for the CC0π^+ + Np control sample (CS3). Kinematics are for the selected isoFGD proton, but reconstructed under the pion hypothesis, since these are the topology of events that form a background in the signal samples.

Sample	Topology	MC Composition (%)		
		FGD1	FGD2x	FGD2y
$CC1\pi^+ + N\pi^\pm$ (CS1)	$CC0\pi$	0.82	0.75	0.91
	$CC1\pi^+$	13.30	12.82	13.57
	$CC2\pi^+$	10.76	11.73	8.88
	$CC1\pi^+ + 1\pi^-$	20.33	20.45	19.67
	$CC1\pi^+ + > 1\pi^\pm$	17.69	16.83	16.84
	$CC1\pi^+ + N\pi^0$	4.00	3.34	3.49
	$CCOther\text{-}N\pi^0$	17.04	17.34	17.57
	$CCOther\text{-}0\pi^0$	6.55	6.98	6.28
	NC	4.86	5.23	5.76
	OOFV	1.97	1.58	4.47
	Other	2.67	2.97	2.56
$CC1\pi^+ + N\pi^0$ (CS2)	$CC0\pi$	1.61	1.30	1.36
	$CC1\pi^+$	17.17	15.31	13.61
	$CC2\pi^+$	1.96	2.30	2.09
	$CC1\pi^+ + 1\pi^-$	5.77	5.63	5.39
	$CC1\pi^+ + > 1\pi^\pm$	2.94	3.28	3.37
	$CC1\pi^+ + N\pi^0$	22.01	23.38	21.36
	$CCOther\text{-}N\pi^0$	24.29	26.69	26.43
	$CCOther\text{-}0\pi^0$	5.56	6.03	5.78
	NC	8.94	8.54	9.26
	OOFV	6.57	4.85	8.66
	Other	3.12	2.69	2.69
$CC1\pi^+ + Np$ (CS3)	$CC0\pi$	83.63	65.68	61.61
	$CC1\pi^+$	8.42	20.59	18.66
	$CC2\pi^+$	0.14	0.47	0.62
	$CC1\pi^+ + 1\pi^-$	0.36	0.98	0.88
	$CC1\pi^+ + > 1\pi^\pm$	0.04	0.33	0.30
	$CC1\pi^+ + N\pi^0$	0.35	0.85	0.91
	$CCOther\text{-}N\pi^0$	2.84	4.52	4.68
	$CCOther\text{-}0\pi^0$	0.95	1.48	1.41
	NC	1.24	1.53	2.13
	OOFV	2.00	3.45	8.60
	Other	0.03	0.12	0.20

Table 6.V: True topology breakdown of the three sideband samples.

Sample	Interaction	MC Composition (%)		
		FGD1	FGD2x	FGD2y
$CC1\pi^+ + N\pi^\pm$ (CS1)	CCQE	0.79	0.61	0.77
	2p2h	0.06	0.13	0.08
	RES	13.44	12.07	12.12
	DIS	76.74	78.59	74.98
	COH	0.47	0.21	0.42
	NC	4.63	5.01	5.62
	CC-$\bar{\nu}_\mu$	1.64	1.57	1.41
	CC-ν_e , CC-$\bar{\nu}_e$	0.23	0.22	0.14
	Other	0.02	0.02	0.00
	OOFV	1.97	1.58	4.47
$CC1\pi^+ + N\pi^0$ (CS2)	CCQE	1.33	1.01	1.24
	2p2h	0.23	0.11	0.21
	RES	18.47	17.36	15.81
	DIS	62.09	66.34	63.27
	COH	0.71	0.66	0.38
	NC	8.06	7.65	8.24
	CC-$\bar{\nu}_\mu$	1.64	1.11	1.15
	CC-ν_e , CC-$\bar{\nu}_e$	0.88	0.89	1.02
	Other	0.01	0.03	0.00
	OOFV	6.57	4.85	8.66
$CC1\pi^+ + Np$ (CS3)	CCQE	61.10	44.48	40.48
	2p2h	10.91	9.50	9.68
	RES	21.86	34.33	32.53
	DIS	2.77	6.34	6.13
	COH	0.09	0.26	0.27
	NC	1.18	1.42	2.05
	CC-$\bar{\nu}_\mu$	0.03	0.10	0.17
	CC-ν_e , CC-$\bar{\nu}_e$	0.06	0.11	0.08
	Other	0.00	0.00	0.00
	OOFV	2.00	3.45	8.60

Table 6.VI: True reaction breakdown of the three sideband samples.

Sample	Particle	MC Composition (%)		
		FGD1	FGD2x	FGD2y
CC1π^+ + N$\pi^\pm$ (CS1)	μ^-	79.83	76.95	76.78
	e^-	0.25	0.18	0.32
	π^-	17.76	19.91	19.55
	μ^+	0.16	0.22	0.19
	e^+	0.02	0.07	0.05
	π^+	0.69	1.15	1.22
	p	0.91	0.99	1.59
	Other	0.39	0.54	0.31
CC1π^+ + Nπ^0 (CS2)	μ^-	77.14	78.46	74.93
	e^-	1.26	1.21	1.70
	π^-	17.06	15.98	18.40
	μ^+	1.00	0.75	0.65
	e^+	0.23	0.16	0.22
	π^+	2.47	2.45	2.81
	p	0.56	0.68	0.90
	Other	0.28	0.31	0.39
CC1π^+ + Np (CS3)	μ^-	97.64	96.23	94.12
	e^-	0.11	0.17	0.37
	π^-	1.78	2.49	3.86
	μ^+	0.03	0.09	0.11
	e^+	0.03	0.04	0.15
	π^+	0.36	0.74	1.18
	p	0.05	0.21	0.18
	Other	0.00	0.01	0.03

Table 6.VII: Leading track true particle ID for the three sideband samples.

6.4 Pion Kinematic Reconstruction from Michel Electrons

As described in Section 6.1.7, the single positive pion signal can be detected through three different methods, depending on where in the detector the pion candidate reaches. For TPC and isoFGD pions, which are directly detected through the ionisation they cause in the relative sub-detector, methods of reconstructing the momentum and direction of the particle are readily available. However, for pions whose presence is inferred from their Michel electron decay chain (Equations (6.8) and (6.9)), no such kinematic information is available. Since the Michel detected pions make up approximately 30–35% of the total sample (Table 6.I), losing such a large amount of potential information is regrettable, particularly in the interesting low momentum regime in which these pions exist.

6.4.1 Truth Studies

In order to establish the viability of reconstructing the kinematic information of the pions detected via Michel electrons, a series of truth studies were performed. These investigate the possibility of reconstructing pion kinematics based on the geometry of the subsequent Michel decay chain. Two assumptions are made, which are first verified in these truth studies. They are that:

1. the greater the momentum of the pion, the further from the neutrino vertex the Michel electron will be produced, and

2. the direction of the outgoing pion will affect the radial position of the Michel electron production.

A diagram of the interaction chain, from neutrino interaction through to Michel electron production, is given in Figure 6.30, with necessary variables labelled. The separation between the initial neutrino interaction vertex and Michel electron production vertex is labelled as d, whilst the angle that the vector $\vec{d}$ makes with respect to the initial neutrino direction is labelled θ. The outgoing pion angle is labelled as ϕ. To verify the above assumptions, correlations in the MC truth information were searched for between pion kinematics and the associated Michel electron geometry:

1. Correlation between pion momentum p_π, and the separation between neutrino vertex and Michel electron vertex, d.

2. Correlation between the pion angle ϕ, and Michel electron vertex angle θ.

In order to select the events for this truth study, the standard CC1π^+ branch from the multiple pion selection (Section 6.1) is used. For all events that pass the CC1π^+ cut, the list of true particles from the vertex is searched for electrons or positrons in the FGD. The parentage tree of these particles is checked, to identify which electrons or positrons are second-generation daughters of pions. This gives a sample of pions which, in truth, decay to muons and finally Michel electrons.

Figure 6.31 gives the 2D histogram showing the relationship between true separation d and initial pion momentum p_π, where each column is normalised by its highest bin content,

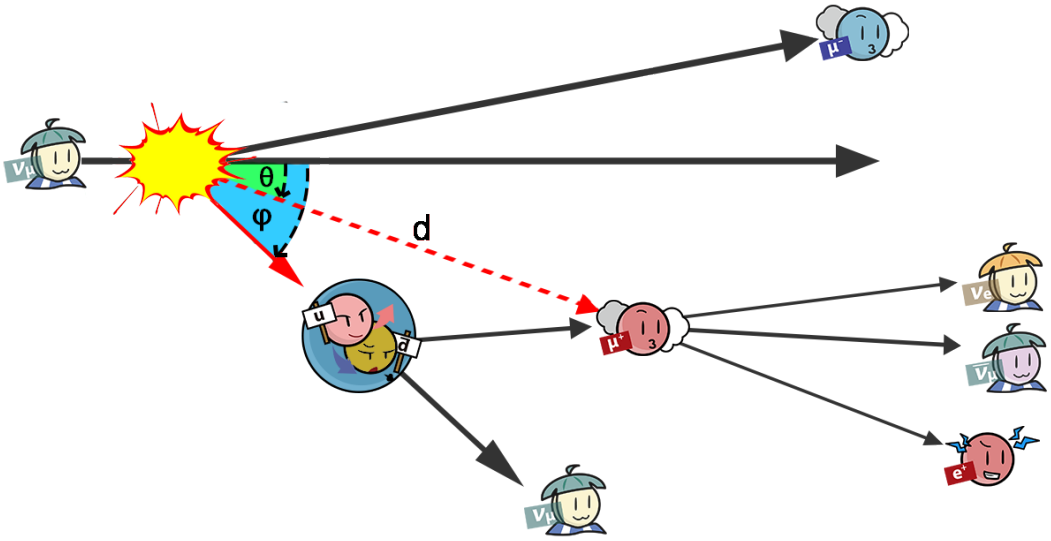


Figure 6.30: Geometric diagram of the pion decay chain to a Michel electron. The initial neutrino interaction is shown in yellow, with the Michel electron production vertex at the μ^+ end position. Vertex separation d is defined as the distance between these two points. ϕ denotes the pion angle to the neutrino direction, whilst θ labels the angle of the separation vector $\vec{d}$ with respect to the neutrino direction. Diagram created using illustrations from [203].

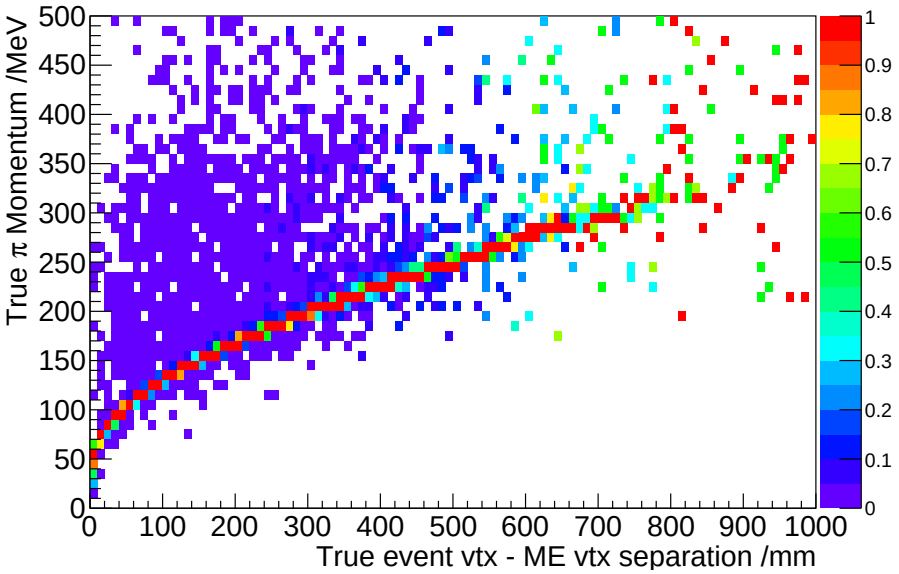


Figure 6.31: Relationship between true Michel electron vertex separation d, and initial pion momentum p_π, for true CC1π^+ events which decay via the Michel chain. Plot is shown for events in FGD1, and bin content is normalised to the highest value in each set of x bins. Plot shows clear correlation between true variables.

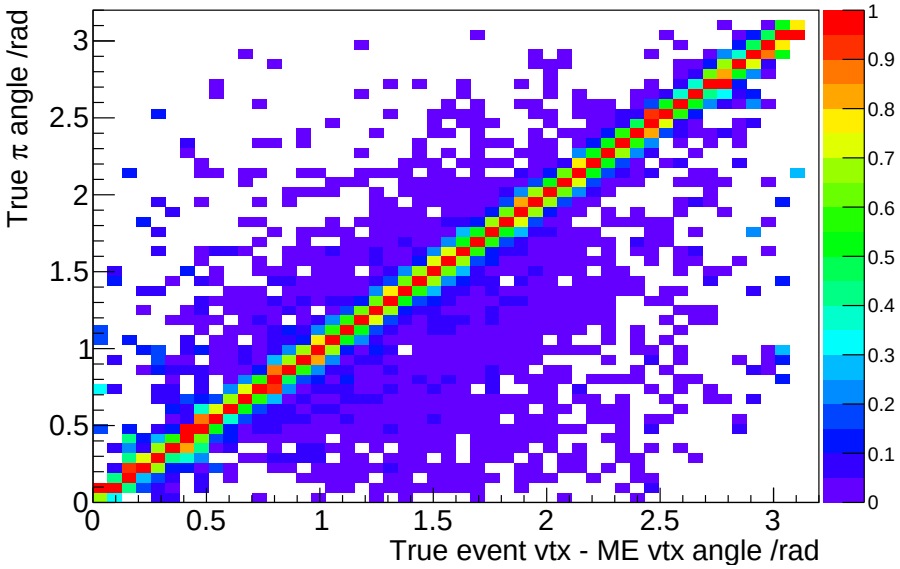


Figure 6.32: Relationship between true angle to Michel vertex θ and initial pion direction ϕ, for true CC1π^+ events which decay via the Michel chain. Plot is shown for events in FGD1, and bin content is normalised to the highest value in each set of x bins. Plot shows clear correlation between true variables.

for ease of viewing. This shows a clear correlation between the true values, although a significant spread in values is also observed above the correlation line, below a vertex separation of ∼500 mm. This suggests that using the vertex separation in order to reconstruct pion momentum should theoretically be viable, although some variation in the results will be expected. Above 500 mm there are very few statistics available, but as each FGD only has a depth (in the beamline direction) of ∼300 mm [158], extrapolation into this region is not expected to cause issues in the reconstruction. However, to ensure that this lack of statistics doesn't cause incorrect reconstruction and in turn bias the measurement, it was decided that events from the Michel electron sample which have a reconstructed momentum of 300 MeV or greater would be excluded from the selected sample when used in the fitting procedure. Again, this should have a minimal effect on the overall sample, as very few events have a true momentum value above this level.

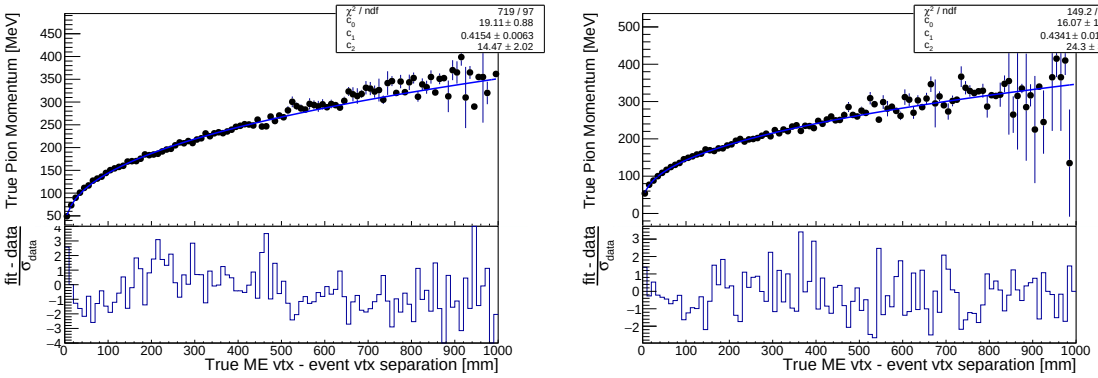
Figure 6.32 shows the corresponding plot for angle, measured in radians. Again this shows a clear correlation between true pion outgoing angle and Michel electron geometry, this time with a clear linear correspondence. As with the momentum comparison there is some spread on the distribution, skewed towards the true value of θ being marginally greater than ϕ.

From both of these truth studies, it can be seen that using the true Michel electron geometry should ideally give a way of reconstructing pion kinematics, before the convolution of detector effects is included. For angular reconstruction, the one-to-one correlation shows that the pion angle can be directly taken as the Michel angle. For momentum, however, the correlation must be fitted to find the function relating the two true quantities. This

is done separately for both FGD1 and FGD2, as the inclusion of water layers in FGD2 could lead to different propagation by any of the involved particles. Each of the momentum histograms for true CC1π^+ events is collapsed into a profile plot along the x axis, where each point shows the mean of the true pion momentum for the corresponding binned value of separation. Both distributions are fitted with the function

$$p_\pi = c_0 \times d^{c_1} + c_2, \tag{6.13}$$

and the most likely fit parameters found by minimising the Pearson's χ^2 value of the fit. The fit profile histograms are given in Figure 6.33, whilst a summary of the obtained fit parameters is shown in Table 6.VIII. As expected, a difference outside of error values is found in the fit parameters between FGD1 and 2, showing the need for separate reconstruction.



(a) FGD1 (b) FGD2

Figure 6.33: Profiled histograms for true CC1π^+ events that decay to Michel electrons in (a) FGD1 and (b) FGD2. Each point shows the mean value of the true pion momentum for each true separation bin. Profiles are fit with the function in Equation (6.13) to find the relationship between the true values.

Fit Parameter	FGD1	FGD2
c_o [MeV/mm]	19.11 ± 0.88	16.07 ± 1.31
c_1	0.4154 ± 0.0063	0.4341 ± 0.0110
c_2 [MeV]	14.47 ± 2.02	24.3 ± 3.5

Table 6.VIII: Optimal fit parameters for the conversion between true Michel vertex separation d and pion momentum p_π, using the function given in Equation (6.13).

6.4.2 Michel Electron Position Reconstruction

Although the truth studies performed show that it is theoretically possible to retrieve pion kinematic information based on the geometry of a Michel electron event, the ability to do this with reconstructed information also has to be verified. This relies on being able to

accurately reconstruct the position at which the Michel electron is produced from the decay muon. As explained in Section 6.1.7, the presence of Michel electrons is identified via the filling of delayed FGD time bins consistent with the lifetime of a charged pion. The reconstructed sample used for these studies is again the standard multiple pion CC1π^+ selection, restricted to those events which are identified through the presence of Michel electrons, and therefore do not currently have pion kinematic information. Within `highland2`, accessing the Michel electron candidate object within the event allows for each of the individual associated FGD hits to be retrieved. Using these, it is possible to estimate the position at which the Michel electron decay occurs. However, the detector geometry and non-event signals impose significant limitations on this.

Firstly, it has already been shown that these event selections are not 100% pure. Figure 6.4 showed the particle breakdown for Michel electron candidates before applying a cut on the total number of hits in a delayed time bin, and although the cuts remove a large part of the background, some of the FGD hits used will not in fact be from Michel electrons. The remaining background is primarily composed of non-Michel electrons, stray neutrons and OOFV events. This means that the position information taken from one of the FGD hits may be completely unrelated to hits left by a Michel electron, if there even was one at all. Naturally this cannot be fully avoided, and is a common problem in all forms of reconstruction. A possible way of mitigating this effect would be to take the charge-weighted average of the position of all FGD hits which contribute to the reconstructed Michel electron. However, this introduces a different complication. In order to estimate the separation between the Michel vertex and primary neutrino interaction vertex accurately, the reconstructed position of the Michel vertex has to be as precise as possible. Averaging over all of the candidate hits, even if they all come from a true Michel electron, inherently moves the reconstructed vertex position away from its true value, as the the electron travels away from the vertex. An illustration of this effect is provided for clarity in Figure 6.34.

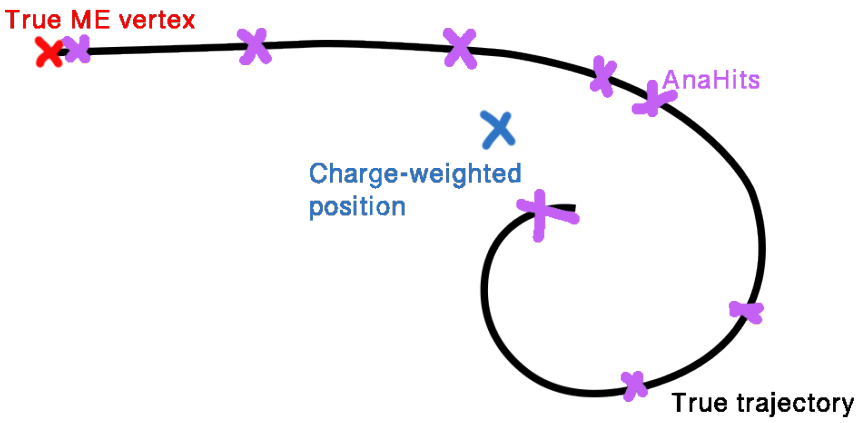


Figure 6.34: Diagram illustrating how using the charge-weighted average position over all candidate Michel electron hits inherently shifts the reconstructed vertex position away from the truth, shown in 2D for simplicity.

Since this effect could significantly affect the reconstructed vertex separation, and the

Michel electron purity (Figure 6.4) is seen to be relatively high anyway, no average over all available FGD hits is performed. Instead, ordering the hits by hit time and taking the earliest should ideally give the hit closest to the true vertex. In order to verify this, the relationship between the time of the first FGD hit from a Michel electron candidate against the time of the true Michel electron production is given in Figure 6.35. Aside from a mild scattering of points, a strong linear correlation between the two variables is seen. This suggests that using the first produced hit contributing to a Michel candidate can in general be suitable for estimating the Michel vertex, since the times roughly match when removing the offset due to the global beam time.

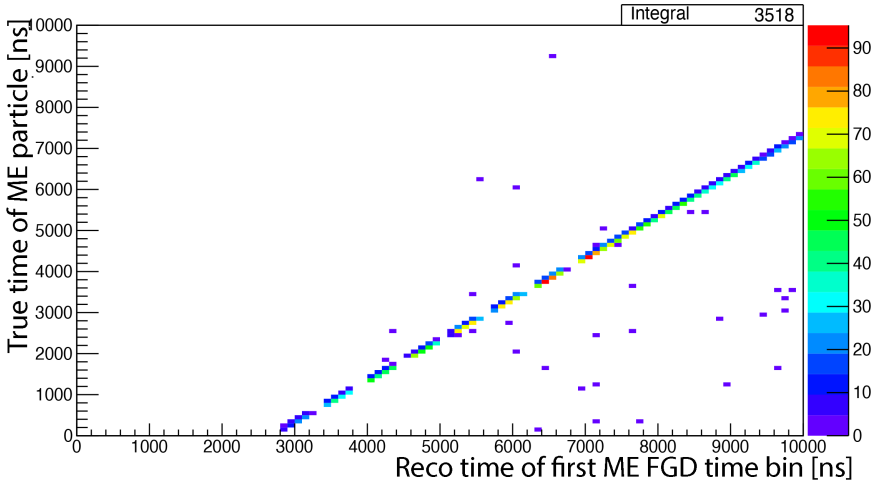


Figure 6.35: Reconstructed time of the first FGD hit contributing to the candidate Michel electron, plotted against time that the true Michel electron is produced. Statistics shown are for runs 2-4 only. A strong linear correlation is observed between the two variables, suggesting the first reconstructed hit when ordered by time is suitable to use. The offset in time is understood to be due to the global detector time, where the beam spill starts at a global time of 2800 ns.

The second limitation of estimating the Michel vertex position from the available FGD hits is due to the geometry of the FGDs. Figure 6.36 shows the FGD structure, which is composed of vertically-stacked layers of scintillator bars, followed by horizontally-stacked layers. Each pair of layers forms an XY module, which is then repeated multiple times, and in FGD2 interleaved with water layers. Although the crossing bar structure of the FGDs allows for precise tracking when a particle crosses many FGD layers, examining individual hits gives less information. For example, if a hit occurs in an X layer, which of the 192 scintillator bars is hit provides a y coordinate, but no x coordinate is available. Similarly, hits occurring in a Y layer provide an x coordinate, but no information on the associated y coordinate. No information is given directly on the z coordinate of a hit, but this can be estimated based on which numbered FGD layer the hit occurs in.

The complication this introduces is that if we take the first hit in time associated to the Michel electron candidate, only an x or y coordinate for the position will be available, not

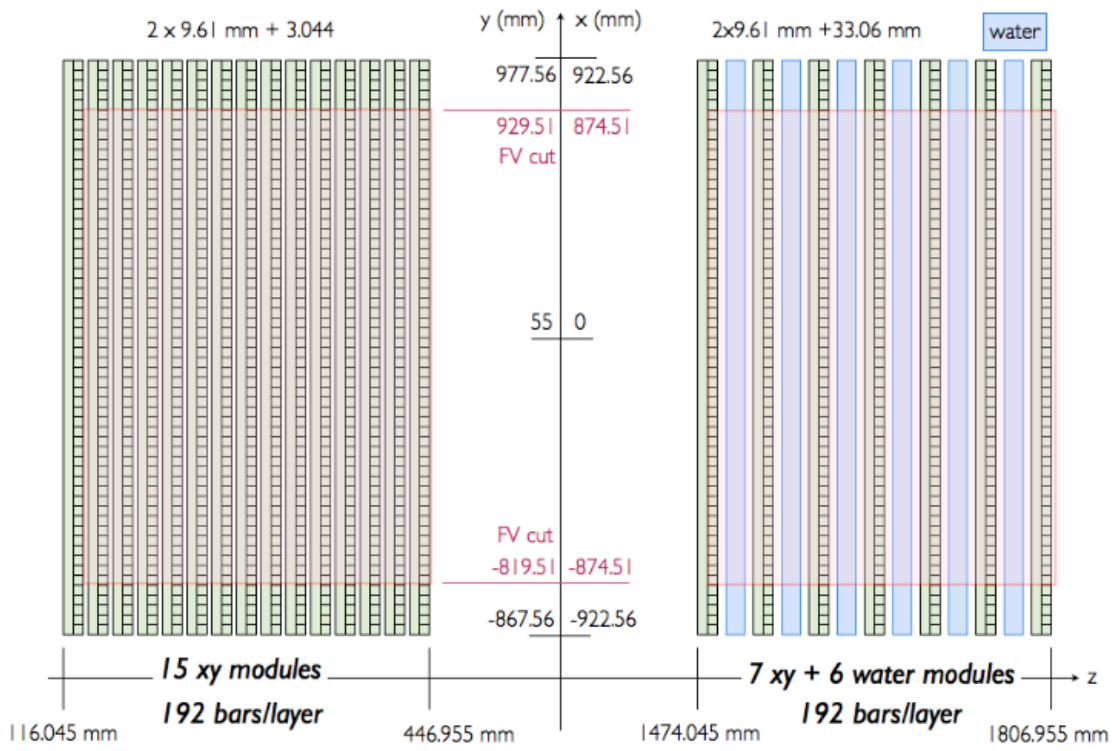
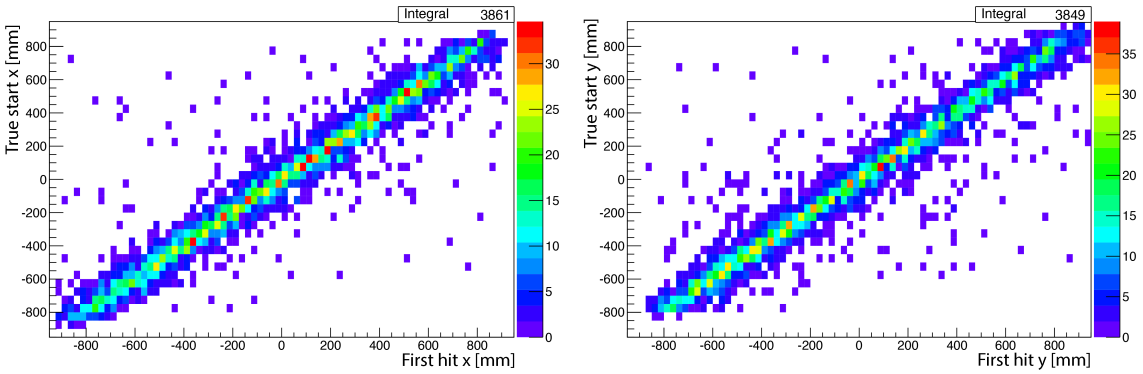


Figure 6.36: Cross-sectional schematic of both FGDs, showing the layered bar structure of the individual XY modules [204].

both. In order to get around this, we modify the vertex estimation to search the available FGD hits (again ordered by time), and take the two hits where an x and a y coordinate are first available. Figure 6.37 shows the truth against reconstruction comparisons for both x and y coordinates, where the reconstructed coordinate comes from the first *available* value in the list of hits.



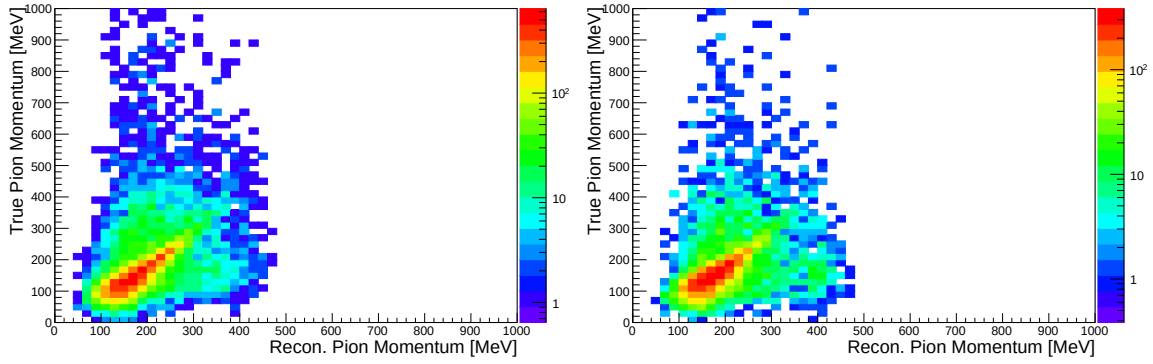
(a) x (b) y

Figure 6.37: Comparison between true Michel vertex coordinate and reconstructed FGD hit with the first *available* values for (a) x and (b) y coordinates. Statistics shown are for runs 2-4 only.

As expected, this introduces some uncertainty into the vertex reconstruction. Since a hit in an X layer may not occur directly after one in a Y layer (or vice versa), by the time one does occur the candidate particle may have significantly moved away from the true Michel vertex. This causes the off-diagonal values observed. However, Figure 6.37 shows that generally using the first available x or y coordinate gives an accurate estimation of the true vertex coordinate. To take this 'hybrid' estimation across multiple hits into account for the reconstructed z coordinate, the average of the z coordinate from each of the two hits is used.

6.4.3 Pion Kinematic Reconstruction

Using the tools and relationships developed in Sections 6.4.1 and 6.4.2, the reconstruction of pion kinematics has been implemented into the ν_μCC1π^+ analysis, for the Michel electron signal branch along with the sideband samples where necessary. As detailed above, the momentum estimation is performed separately for FGD1 and 2, in order to take into account the differences in target medium. For the sample of selected CC1π^+ events which are identified via Michel electrons, with no further cuts applied (the general numuCCMultiPi selection 1π^+ branch), Figure 6.38 shows the comparison of true and reconstructed pion momentum values, specifically for true CC1π^+ events. True pions are taken as the grandparents of the true Michel electrons, identified using the same technique in Section 6.4.1. This shows a clear linear correlation between true and reconstructed values. As expected there is a smearing to the distribution which results in a large number of off-diagonal elements, particularly at higher values of momentum. This is likely due to the fit being less reliable at higher values of momentum, caused by the lower statistics, as can be seen in Figure 6.33.

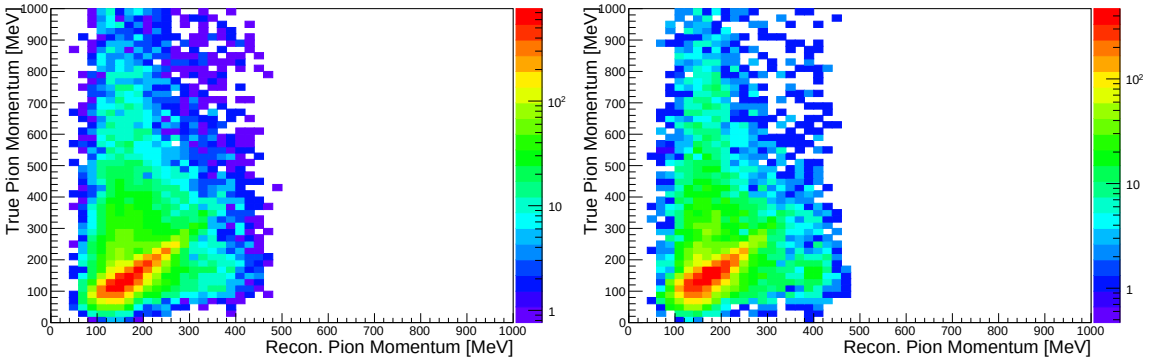


(a) FGD1 - true CC1π^+ events

(b) FGD2 - true CC1π^+ events

Figure 6.38: Pion momentum reconstruction comparison to true momentum values for (a) FGD1 and (b) FGD2, for all true CC1π^+ topology events in the selected ME pion sample. Event distribution is drawn for events passing the general 1π^+ cut.

For comparison, Figure 6.39 shows the same truth against reconstruction plots, but for all selected events in the CC1π^+ from Michel electrons sample, regardless of true event



(a) FGD1 - all events

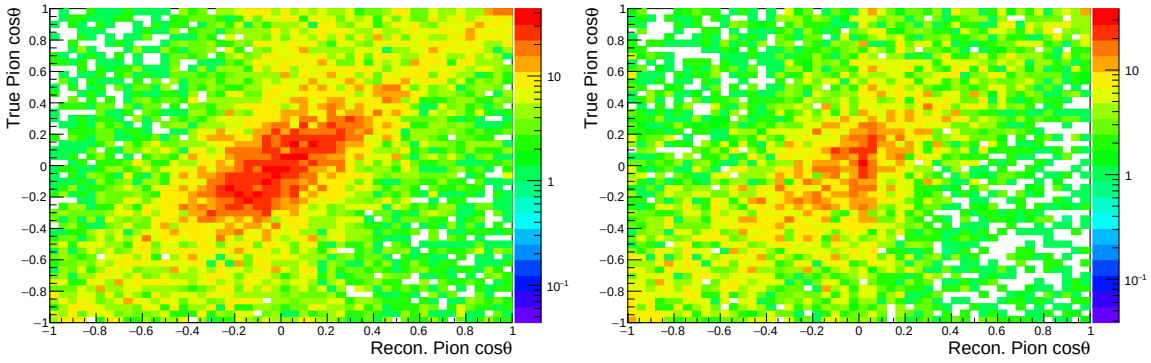
(b) FGD2 - all events

Figure 6.39: Pion momentum reconstruction comparison to true momentum values for (a) FGD1 and (b) FGD2, for the full selected ME pion sample. Event distribution is drawn for events passing the general $1\pi^+$ cut.

topology. Here the off-diagonal elements are more pronounced, with a clear band of events where the pion momentum is reconstructed in the 100–200 MeV region, regardless of its true value. This shows that the reconstruction breaks down with background events, presumably due to being developed with true signal events, whereas the major background is composed of CCOther events.

Similarly, Figures 6.40 and 6.41 show the truth against reconstruction relationships for the cosine of the pion angle to the neutrino direction, with the former for only true CC$1\pi^+$ interactions, and the latter for all selected events. Again both event distributions are those which pass the general CC$1\pi^+$ cut, and are identified via Michel electrons. As with momentum, although a clear linear correlation is observed there is a very large spread, with many events where the $\cos\theta$ value is reconstructed completely wrong. This is due to the limitations in the Michel vertex reconstruction, which will have more of an effect on the angular reconstruction than that of momentum, since momentum relies only on the distance between the two points rather than the exact location of the Michel vertex. As with momentum, the incorrect reconstruction is also more prevalent for background events than signal, because in background topologies there are more opportunities for the reconstruction to go wrong.

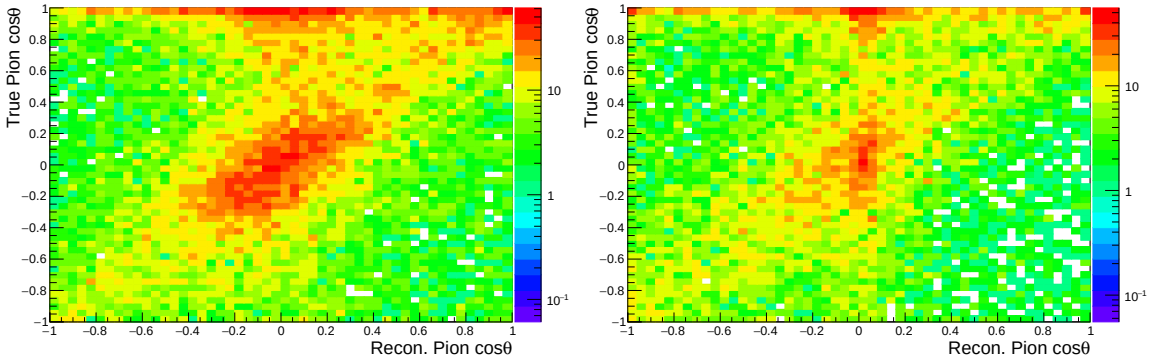
Figures 6.42 and 6.43 show the kinematic truth against reconstruction plots for the Michel signal sample of the developed selection (Section 6.2), after all selection cuts have been applied. This sample has a signal purity of ∼65%, and as usual the major background is CCOther, signifying the presence of additional pions. Figure 6.42 shows that the momentum reconstruction in this sample of interest is generally good, with the majority of events appearing on the diagonal. As expected there is some spread to the data, but this is unavoidable due to the limitations of the Michel vertex position reconstruction caused by the FGD geometry. It is interesting to note that with the inclusion of the SuperFGD as part of the upcoming ND280 upgrade, this technique should become more accurate due to 3D cube-like structure of the SFGD scintillator [205]. In Figure 6.43 it can be seen that



(a) FGD1 - true CC1π^+ events

(b) FGD2 - true CC1π^+ events

Figure 6.40: Pion $\cos\theta$ reconstruction comparison to true $\cos\theta$ values for (a) FGD1 and (b) FGD2, for all true CC1π^+ events in the selected sample. Event distribution is drawn for events passing the general 1π^+ cut.



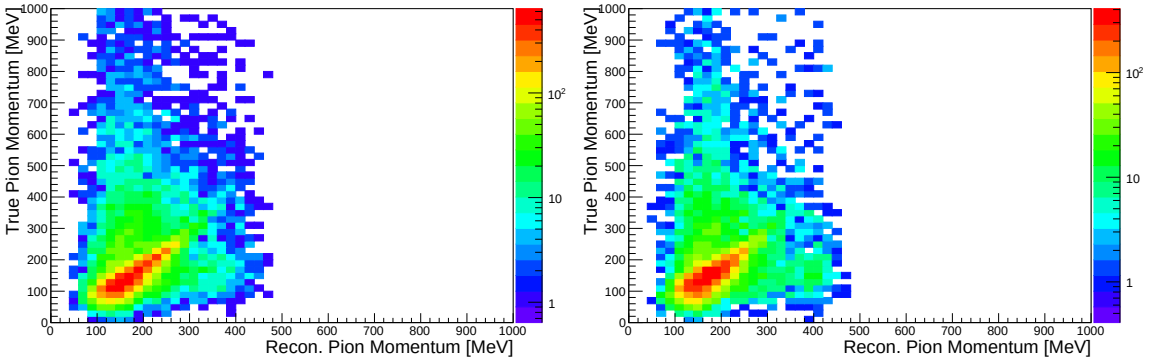
(a) FGD1 - all events

(b) FGD2 - all events

Figure 6.41: Pion $\cos\theta$ reconstruction comparison to true $\cos\theta$ values for (a) FGD1 and (b) FGD2, for the full selected sample. Event distribution is drawn for events passing the general 1π^+ cut.

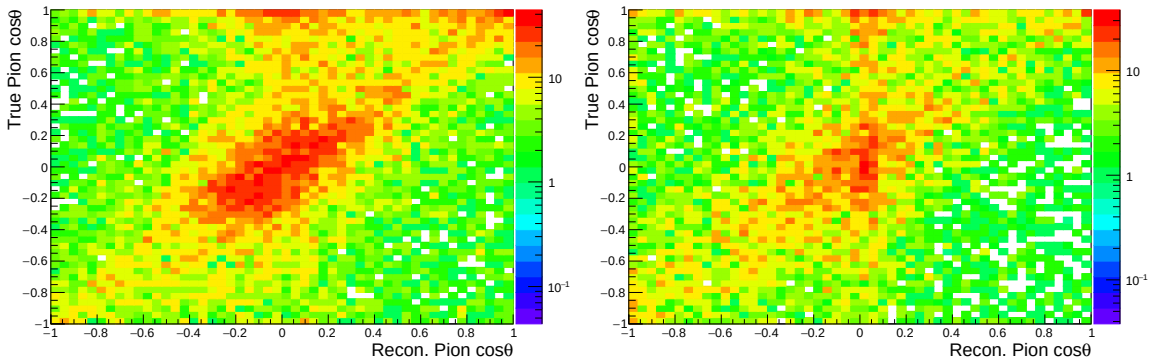
the although there is a general correlation, agreement between true and reconstructed $\cos\theta$ values is worse than for momentum, with a large spread across the full range of values. This again shows that the angular reconstruction is far more susceptible to inaccuracies in the Michel vertex reconstruction, and that the resolution in this region is not good.

As an additional comparison, it is interesting to look at the distribution of the normalised residual values between the true and reconstructed momentum, $(p^{\text{true}} - p^{\text{reco}})/p^{\text{true}}$, for the selected sample in the CC1π^+ Michel electron branch. These are given in Figure 6.44 for both FGD1 and FGD2, with both categorised by the true event topology. In both cases the distributions peak at values of approximately -0.1, meaning that while the reconstruction is generally quite accurate, there is a small bias due to systematically overestimating the momentum in comparison to the true value. The offset peak between residual values of 1.0–1.5 in both FGDs comes from the cases where a pion is mistakenly reconstructed, but no



(a) FGD1 (b) FGD2

Figure 6.42: Pion momentum reconstruction comparison to true momentum values for (a) FGD1 and (b) FGD2, for the selected CC1π^+ from Michel sample after all developed selection cuts are appied.

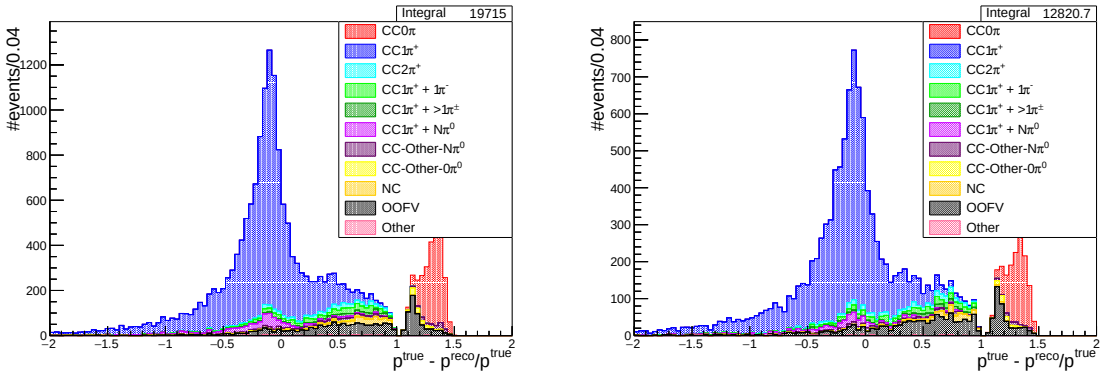


(a) FGD1 (b) FGD2

Figure 6.43: Pion $\cos\theta$ reconstruction comparison to true $\cos\theta$ values for (a) FGD1 and (b) FGD2, for the selected CC1π^+ from Michel sample after all developed selection cuts are applied.

true pion exists – hence this region being mostly populated by true CC0π events. In these cases the true pion momentum value takes the default null value of -999, which results in a normalised residual of greater than 1.0.

Although the kinematic reconstruction of pions from Michel electrons is not perfect, this is the first time that this pion phase space has been accessible. The very low momentum region is particularly interesting as pions with such low momentum are below the Cherenkov threshold for Super-K. In order to make use of this reconstruction whilst taking into account the resolution, binning in this region will have to be coarse, and should prioritise momentum over angle, as the momentum reconstruction works better. As mentioned in Section 6.4.1, the Michel electron sample will be restricted to events with momentum less than 300 MeV, so that the only kinematic region used is the one where the reconstruction can be trusted. It can be seen from Figure 6.42 that adding this restriction only discards a minimal number



(a) FGD1 (b) FGD2

Figure 6.44: Normalised residual values between the true and reconstructed momentum, $(p^{\text{true}} - p^{\text{reco}})/p^{\text{true}}$, in (a) FGD1 and (b) FGD2. Distributions shown are for the selected CC1π^+ from Michel electron sample after all developed selection cuts are applied, and are categorised by true event topology.

of events, as the majority have a reconstructed momentum in the 100–250 MeV region. Full details of the binning choices for this sample and the rest of the analysis are given in Section 7.2.

6.4.4 Reconstruction Uncertainty Evaluation

In order to make use of this reconstruction, it is necessary to assess the systematic uncertainties associated with the reconstruction method. To do this, a similar method to the momentum by range systematic assessment in Section 7.1.1.8 is employed.

6.4.4.1 Control Sample Selection

We first form a control sample of stopping pions in FGDs 1 and 2, so that the data–MC discrepancy can be evaluated separately for each. This is done by applying the following cuts to the MC or data events:

- *Event Quality Cut* – See Section 6.1.1
- *Track Quality and Candidate Available Cut* – All reconstructed tracks that use a TPC are sorted according to momentum. The track quality cut (Section 6.1.3) is applied to each, along with requiring the charge of the track to be positive. Of these tracks, the highest momentum one is set as the main track. If after this no main track exists, the event fails the selection.
- *Pion Main Track Cut* – For all candidate tracks, TPC particle ID is applied, giving a likelihood for each of three particle hypotheses: p, e^-, $\pi^\pm$. The highest likelihood method is employed, and particles which are most likely to be a pion pass this cut.

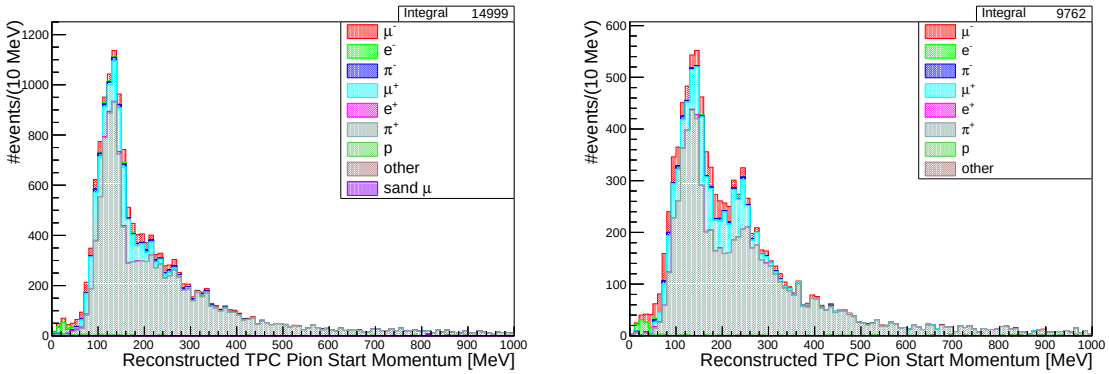
- *FGD Stopping Layer Cut* – At this point the selection splits into two branches, one for each FGD. For tracks that enter the FGD detector volume, the position of the end of the track is found. If the track ends in the first 10 layers of the FGD, the event passes the selection. If there is no FGD segment, or if the track stops downstream of the 10^{th} FGD layer, the event fails the cut.
- *FGD Michel Electron Object Cut* – In each branch, the corresponding FGD is checked for Michel electron candidates, which are filled as described in Section 6.1.7. If a candidate exists, the event passes the cut.

This provides two separate control samples, one for each FGD, where a pion-like track stops in the first few layers of the FGD, and decays to a Michel electron candidate. The TPC track is required in order to utilise the TPC PID methods, along with obtaining an accurate estimation of the track momentum prior to entering the FGD. The pion momentum reconstruction method described in Sections 6.4.2 and 6.4.3 is applied to the Michel electron candidate, in order to provide an estimate for the momentum of the pion in the FGD. As the event vertex this time does not exist within the FGD, we take the first available FGD segment in the region of interest, and define the separation d as the distance between the start position of the FGD segment and the Michel electron vertex. Doing this makes it possible to compare the momentum of the pion candidate measured from the TPC to the estimation by range from the Michel electron in the FGD. The reduced volume of interest requirement on the FGD is added to prevent too much change in momentum between the TPC and FGD. It is expected that the momentum distributions will be slightly different, since it is known that the momentum by Michel range estimation is not perfect, and some energy will be lost as the particle traverses the TPC, although the low density of the TPC gas will minimise this.

In applying the pion momentum reconstruction method to the Michel electrons from events selected as part of this sample, an additional restriction is also added. As detailed in Section 6.4.1, the truth studies used as a basis for this reconstruction show few events above a true momentum of 300 MeV, where the few that do exist are very scattered with no clear relationship between vertex separation and momentum. The fits performed to these relationships in Figure 6.33 for both FGDs show unreliable fitting in this region, due to the low statistics and therefore high statistical errors. For this reason it was decided that the Michel electron signal sample for the cross-section measurement described would be restricted to events below a reconstructed momentum of 300 MeV, where the reconstruction is more reliable. In order to allow the systematic treatment being developed to migrate events in and out of the accepted momentum range, but avoid extrapolating outside of the region that the momentum – separation relationship is fit for, only events with a separation of 1000 mm or less are used in this control sample. This will cause a cut off in momentum values at around 350–400 MeV, which is outside of the range of interest for the signal sample.

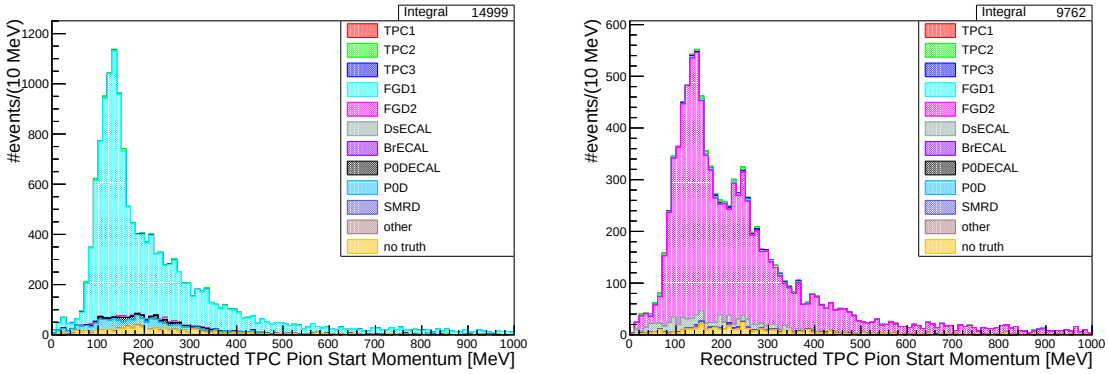
The developed control sample selection is first applied to the Monte Carlo on its own, using all FHC MC available, which is detailed in Table 5.I. Figures 6.45a and 6.45b show the pion candidate TPC reconstructed momentum for the FGD1 and FGD2 control sample

selections, respectively. Both distributions are categorised by true particle type for the pion candidates, which are found to be overwhelmingly positive pions. The major background to the pion track selection comes from positive muon tracks, and to a lesser extent negative muons where the sign of the charge has been misreconstructed. Figure 6.46 shows the



(a) FGD1 (b) FGD2

Figure 6.45: TPC track momentum from the selected pion candidate in the described control sample, in (a) FGD1 and (b) FGD2, categorised by true particle type. TPC momentum is taken from the start of the reconstructed track.



(a) FGD1 (b) FGD2

Figure 6.46: TPC track momentum from the selected pion candidate in the described control sample, in (a) FGD1 and (b) FGD2, categorised by true detector in which the particle track ends. TPC momentum is taken from the start of the reconstructed track.

same distributions, but categorised by the detector in which the selected track truly ends, which for each selection should ideally be the appropriate FGD. The distribution in Figure 6.46a gives the track end breakdown for the FGD1 selection. As expected the majority of selected pions do end in FGD1 as required, but a small background is present from tracks ending in the PØD, likely coming from incorrectly reconstructed, backward-going tracks. For FGD2 (Figure 6.46b), again the majority of events stop as required in FGD2, with a small background from tracks which continue through the FGD and end in the DsECal.

Looking instead at the momentum distributions shown in Figures 6.45 and 6.46, it can be seen that the distributions sharply peak between 100 and 200 MeV, which is the expected range for the low momentum pions that will decay to Michel electrons within the detector. The FGD1 distribution peaks more sharply, however, with FGD2 exhibiting a clear double peaked structure. This is particularly obvious when focusing on the momentum distribution of the true pions in Figure 6.45b, and looking at this in Figure 6.45a also shows a slight double peaked structure in FGD1. To investigate the cause of this double peaked structure in both FGDs, the TPC pion momentum distribution was drawn for each cut in the control sample selected. Doing so found that the structure first appears after applying the pion-like PID cut to the TPC main track, where the distribution is shown for the FGD1 selection in Figure 6.47. This shows that, rather than being a physical double peak in the momentum

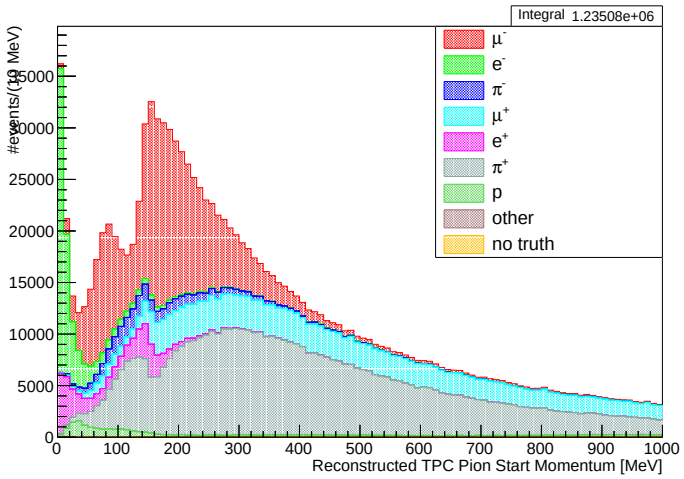
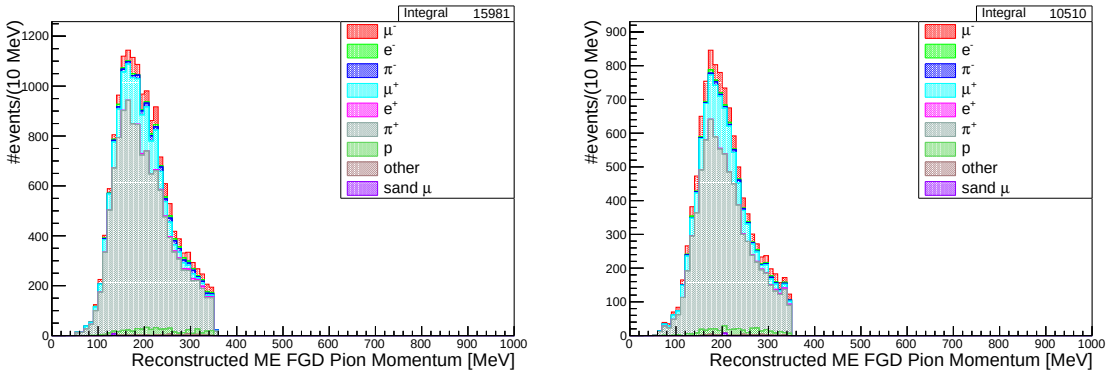


Figure 6.47: TPC track momentum from the selected pion candidate, for events which pass the pion-like TPC PID cut on the main track, categorised by true particle type. A dip in the true pion distribution can be seen at ∼150 MeV, which corresponds to the crossing of the muon and pion energy-loss curves used in the PID algorithm.

distribution, this structure is actually a dip in the expected smooth curve, cause by an inefficiency in the PID method. The dip region appearing at ∼150 MeV after the PID cut suggests that events are being removed incorrectly due to the overlapping muon and pion energy-loss curves in this momentum range, which is an unavoidable feature of the PID algorithm.

Applying the developed algorithm to the Michel electron information in the selected samples gives the reconstructed FGD momentum distributions presented in Figures 6.48a and 6.48b for FGD1 and FGD2, respectively. Here it can be seen that the obtained momentum distributions are overall of similar shape to that of the TPC momentum; however, some specific features are not replicated. Firstly, the peak of the momentum distribution occurs in both FGDs at a slightly higher value than in the TPC, particularly in FGD2, suggesting that the Michel electron algorithm is slightly over-estimating the momentum. The algorithm used also seems unable to properly reproduce the double peaked structure

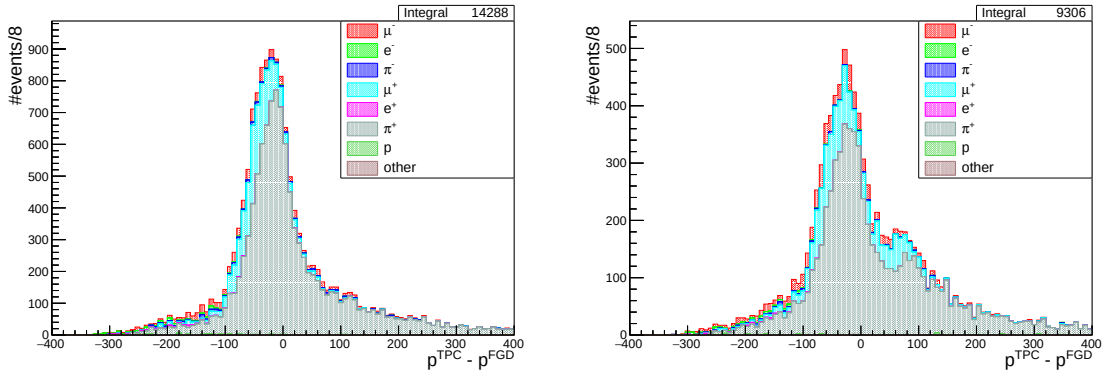


(a) FGD1 (b) FGD2

Figure 6.48: Reconstructed FGD momentum for the selected pion candidate, using the Michel electron reconstruction algorithm, in (a) FGD1 and (b) FGD2, categorised by true particle type. A cut off in values is seen at about 350 MeV, due to the 1000 mm upper limit on the vertex separation distance.

previously observed in the TPC momentum. The sharp cut-off in momentum at around 350 MeV is a result of the 1000 mm separation limit imposed on the reconstruction, so as to not extrapolate outside of the region that the truth studies were initially fitted for.

In order to compare the reconstruction on an event-by-event basis, Figures 6.49a and 6.49b show the difference between the TPC and FGD reconstructed momentum of the pion candidate, for FGD1 and FGD2 respectively. The distribution for FGD1 shows a peak at slightly below zero, meaning that the ME reconstruction generally slightly overestimates the momentum in comparison to the TPC value, as was inferred from the FGD plots previously. It is interesting to note however that, looking at the underlying distributions for the true



(a) FGD1 (b) FGD2

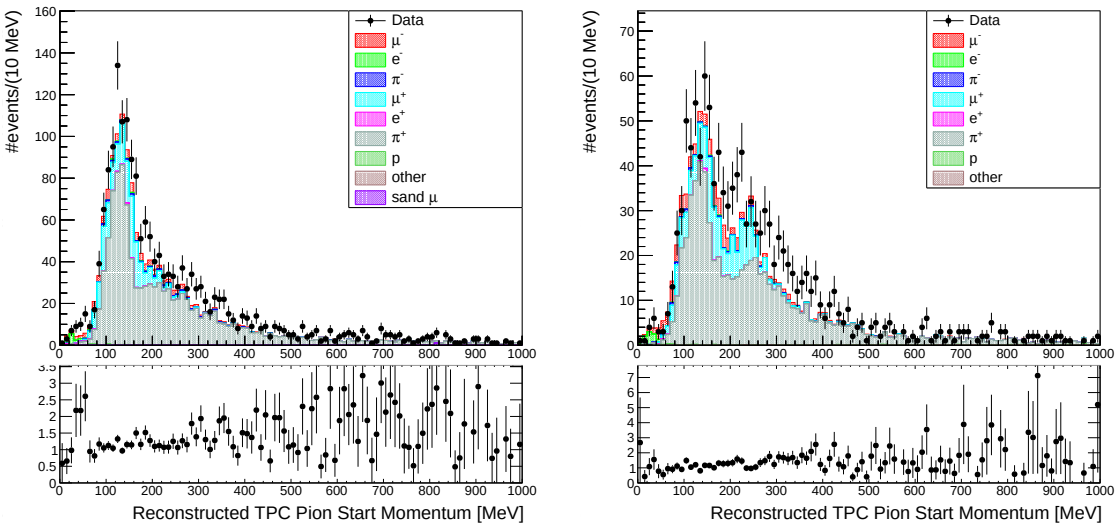
Figure 6.49: Difference between TPC and FGD momentum, $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$, in (a) FGD1 and (b) FGD2, categorised by true particle type.

particle ID, the momentum distribution for true pions is slightly closer to zero than for the total distribution, which is instead pulled to a more negative value by the muon background.

Overall, the distribution is mostly symmetric; however, a longer tail exists out towards positive values caused by the long tail in TPC momentum and the cut off in FGD momentum due to the vertex separation constraint. The distribution for FGD2 exhibits slightly different behaviour: while the peak is still near zero, it is at a slightly more negative value than for FGD1. The major difference is the lack of symmetry, where a smaller local maximum appears at a $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$ value of ∼100 MeV, caused by the inability of the Michel pion reconstruction algorithm to reconstruct the secondary peak in TPC pion momentum. Since this secondary peak is an unavoidable consequence of the TPC PID inefficiency, there is no feasible way of removing it from the selection. Instead, the effect of this was taken into account by increasing the width of the function used to fit this distribution, which will be used to compare the responses of the control sample in both data and MC.

6.4.4.2 Data–Monte Carlo Comparisons

The developed control sample is needed to understand how any differences in MC and data manifest, and to quantify these potential differences in order to develop a systematic error treatment to cover them. The control sample selection is applied to the FHC data set detailed in Table 5.I, and compared to the MC response for each relevant variable distribution. Figure 6.50 shows the TPC track momentum distribution for the selected pion candidate, for both FGD1 and FGD2, with the underlying MC distribution compared to the data points in black. The first thing that is immediately obvious when inspecting these



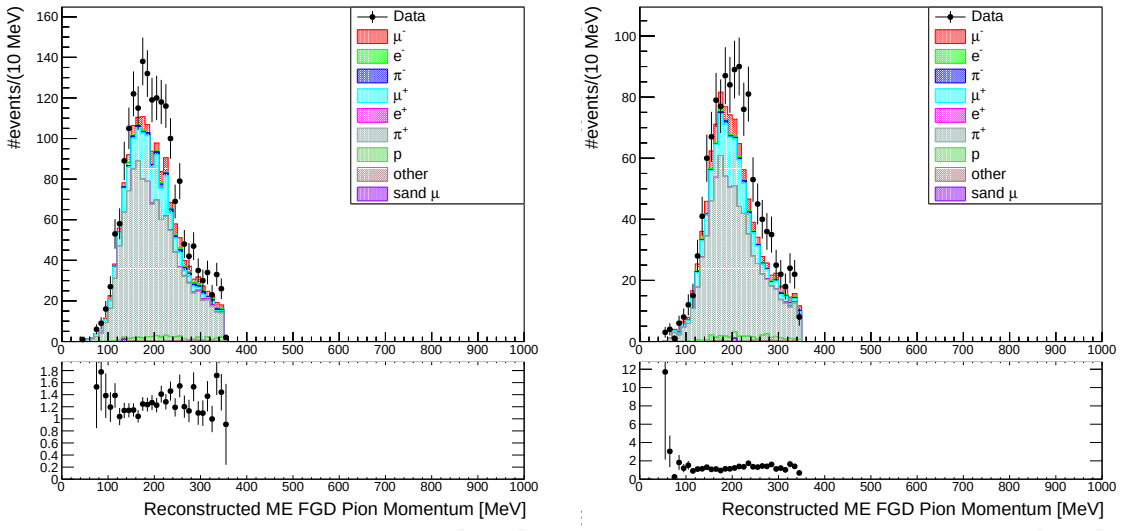
(a) FGD1 (b) FGD2

Figure 6.50: TPC track momentum from the selected pion candidate in the described control sample, in (a) FGD1 and (b) FGD2, categorised by true particle type. Black points show variable distribution in data. TPC momentum is taken from the start of the reconstructed track.

comparisons is the excess of data events over MC. Comparing the integrated number of events, there is found to be a data excess of 25% in both FGD control samples. This excess

does not occur in a particular region: instead, there is seen to be a constant excess across the whole momentum region plotted. This can even be seen in the FGD2 comparison in Figure 6.50b – despite the relatively low statistics once the MC is scaled to data POT, the secondary peak can still be seen in both data and MC. To check whether this disagreement was being introduced by a feature of the selection, the TPC momentum distributions for all previous cut values were plotted for both data and MC. In doing this, it was found that the data–MC discrepancy exists prior to any cuts, and therefore it is not the selection cuts themselves that are causing the discrepancy. It was also noted that prior to any cuts the excess is still seen to be a constant factor across the range of values, rather than being localised in any specific region. It therefore appears that this is a normalisation issue which, although not ideal, is also not particularly surprising for a non-standard event selection such as this control sample. Hence, it was decided that the discrepancy would be treated as a normalisation issue, and accounted for in the subsequent parametrisation of the $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$ residual distributions.

Figure 6.51 gives the MC and data comparisons for the FGD reconstructed momentum from Michel electrons distribution. Again, a clear excess of approximately 25% is observed in the data distributions, which mostly follows the same shape as the underlying MC distributions for both FGDs. There is a very slightly different shape in the peaks of the MC

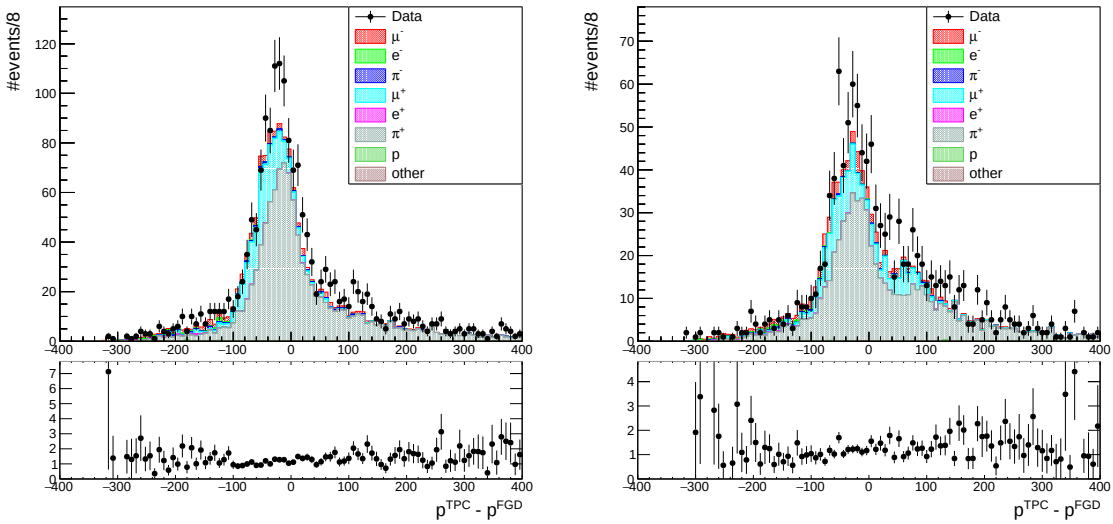


(a) FGD1 (b) FGD2

Figure 6.51: Reconstructed FGD momentum for the selected pion candidate, using the Michel electron reconstruction algorithm, in (a) FGD1 and (b) FGD2, categorised by true particle type. Black points show variable distribution in data. A cut off in values is seen at about 350 MeV, due to the 1000 mm upper limit on the vertex separation distance.

and data distributions in FGD2, as data is seen to peak at a marginally higher momentum value than in MC. This is likely due to incorrect modelling of background events in the MC which, as discussed in Sections 6.4.2 and 6.4.3, can cause issues in the reconstruction of the Michel electron vertex position and thus the momentum reconstruction.

Finally, Figure 6.52 shows the data–MC comparisons for the difference between the TPC and FGD momentum measurements, $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$, in both FGD1 and FGD2. Once again, a clear excess of approximately 25% is observed in the data, which is distributed evenly across the range of values shown in the plots. In Figure 6.52a, the excess in data shows a



(a) FGD1 (b) FGD2

Figure 6.52: Difference between TPC and FGD momentum, $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$, in (a) FGD1 and (b) FGD2, categorised by true particle type. Black points show variable distribution in data.

small secondary peak at around 100 MeV, which is also possibly present in the MC, but to a slightly lesser degree. Figure 6.52b displays the much more obvious peak, both in data and MC, and it is reassuring to see this secondary peak occurring at roughly the same residual momentum value in both.

In order to obtain a systematic uncertainty, it is necessary to parametrise the data and MC distributions for both FGDs, so that the differences in each can be quantitatively assessed. Examining the shapes of the residual momentum distributions by eye, it was decided that a Lorentzian distribution function would likely fit best. In this case we use the three-parameter Lorentzian distribution,

$$f(x; x_0, \gamma_0, I) = \frac{I}{\pi\gamma_0 \left[1 + \left(\frac{x - x_0}{\gamma_0}\right)\right]}, \tag{6.14}$$

where x_0 is the location of the maximal value, γ_0 is the full width at half maximum, which characterises the width of the distribution, and I is the height of the peak. The addition of I is required both for the fact that the distributions are not normalised to unity, and to account for the excess observed in data. Initial tests of fitting the data and MC distributions found that, although particularly successful at identifying the peak values, the Lorentzian function could not fully describe the shapes. This is due to the fact that the described probability density function is symmetric, and the obtained momentum residual distributions show some skew towards positive values, particularly in FGD2.

In order to account for this, we follow the procedure used in [206], and modify Equation (6.14) to an asymmetric form by replacing γ_0 with the sigmoidally varied γ,

$$\gamma(x, a) = \frac{2\gamma_0}{1 + \exp\left[a\left(x - x_0\right)\right]}, \tag{6.15}$$

where a is an asymmetry parameter that is also allowed to vary within the fit. Negative values of a will skew the distribution towards positive residual values, whereas a positive a will skew it the opposite way.

Figure 6.53 shows the fits obtained for the FGD1 data and MC distributions. The MC fit (solid blue line) is found to reproduce the shape of the MC histogram quite well. A χ^2/Ndf of 6.03 is found for the fit; despite this, the agreement between the MC and the fit is visually good. In terms of best fit parameters, a MC central value of $x_0 = -25.03 \pm 0.52$ MeV is found; as expected, the momentum reconstructed by the FGD is a slight over-estimation in comparison to the momentum measured by the TPC. The fit value of the full width at half maximum is found to be $\gamma_0 = (85.38 \pm 1.20)$ MeV, which represents the width of the distribution of values. A slight skew towards positive values, represented by a small negative value of a, is also observed. Comparing to the data fit (solid red line), a peak value slightly

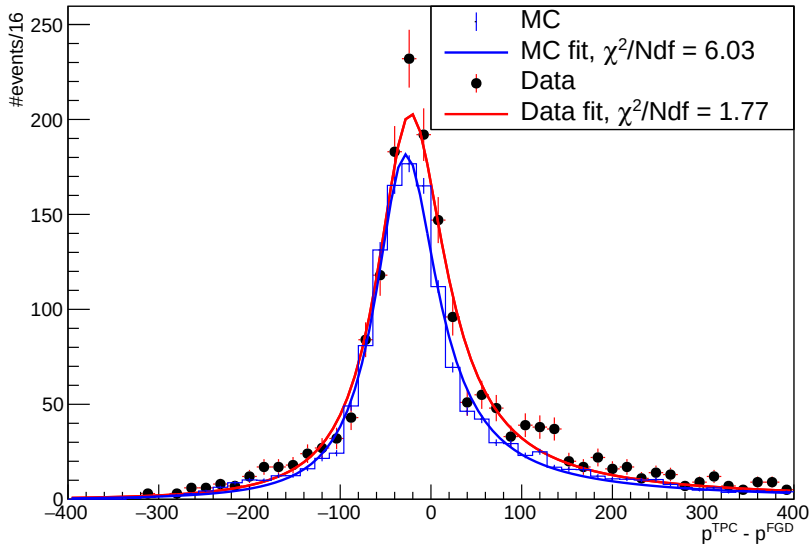


Figure 6.53: Fits to both MC (blue) and data (red) distributions for the FGD1 momentum residual, $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$. Both distributions are fitted with the asymmetric Lorentzian distribution in Equations (6.14) and (6.15), using the MIGRAD minimiser.

closer to zero is found in data, with $x_0 = -19.43 \pm 1.41$ MeV. However, the more accurate reconstruction for the values at the peak position is counteracted by the larger width of the overall distribution, with the best fit parameter for the full width at half maximum being $\gamma_0 = 93.57 \pm 4.15$. The asymmetry parameter is found to be within error of the value found for the MC. All fit parameters and their associated errors for the FGD1 residual values are

presented in Table 6.IX.

Parameter	Data		MC	
	Value	Error	Value	Error
x_0	-19.43	1.41	-25.03	0.52
γ_0	93.57	4.15	85.38	1.20
I	29743.5	734.8	24260.1	219.7
a	-0.0055	0.0009	-0.0059	0.0003

Table 6.IX: Fit parameters for the fitting of an asymmetric Lorentzian function to the distribution of $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$ in FGD1, performed for both MC and data.

Figure 6.54 shows the equivalent fits to the data and MC distributions for the FGD2 control sample. Again, the asymmetric Lorentzian function is found to fit better to data than MC, with χ^2/Ndf values of 1.40 and 6.78, respectively. In both data and MC, the peak values are found to be slightly closer to zero than for FGD1, with values of -16.14 ± 2.58 and -22.68 ± 0.84 MeV, respectively. This puts the peaks at roughly 2–3 MeV closer to zero than for FGD1, although the FGD1 and 2 values for data still agree within error. This is somewhat surprising, as it was expected that the water layers in FGD2 and the double peak in the TPC distribution of the FGD2 control sample would cause greater disagreement in TPC and FGD momentum measurements. Naturally, the presence of the secondary peak in the FGD2 distributions causes a large increase in the values of the full width at half maximum, which was required in order to take into account the lower overall resolution caused by this. In both data and MC, which for FGD2 agree within error, the width is found to be roughly 30 MeV higher when compared with FGD1. The secondary peak is also accounted for in the asymmetry parameter a, which is larger for FGD2 than FGD1. These values are also found to agree within the errors given, which is consistent with the behaviour in FGD1.

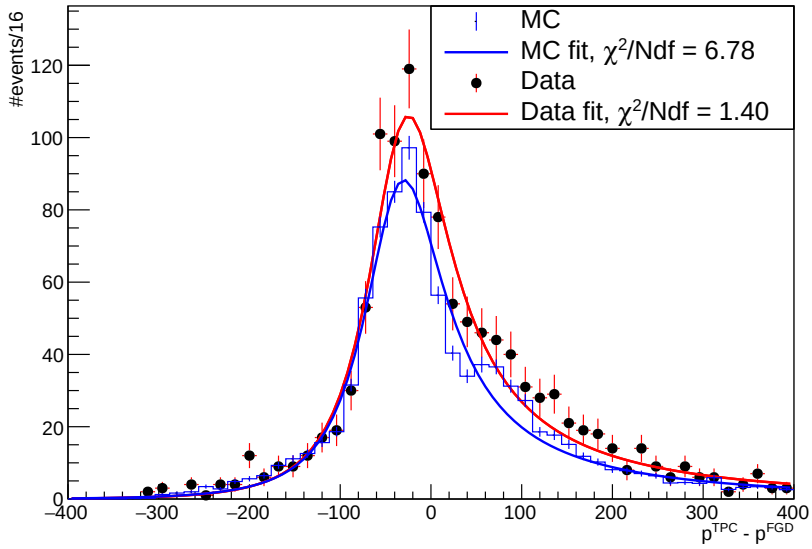


Figure 6.54: Fits to both MC (blue) and data (red) distributions for the FGD2 momentum residual, $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$. Both distributions are fitted with the asymmetric Lorentzian distribution in Equations (6.14) and (6.15), using the MIGRAD minimiser.

Parameter	Data		MC	
	Value	Error	Value	Error
x_0	-16.14	2.58	-22.68	0.84
γ_0	123.30	6.49	118.39	2.22
I	20119.1	673.6	16164.7	186.4
a	-0.0095	0.0016	-0.0085	0.0004

Table 6.X: Fit parameters for the fitting of an asymmetric Lorentzian function to the distribution of $p^{\mathrm{TPC}} - p^{\mathrm{FGD}}$ in FGD2, performed for both MC and data.

6.4.4.3 Systematic Uncertainty Propagation

The final step in calculating the uncertainty on the pion momentum reconstruction from its Michel electron decay chain is to propagate the differences observed between data and MC in the control sample back through the signal sample. This is done as somewhat of a hybrid form between a variation and weight-like systematic. In the general `highland` systematics package (discussed fully in Section 7.1.1), variation systematics are calculated by applying a small variation to a nominal kinematic value based on the bias and resolution seen in control samples, and then the effect of this change as the event goes through the selection is assessed. As the Michel electron reconstruction is performed near the end of the analysis, and only if certain conditions are met, this is not feasible. Instead the systematic is built as a weight-like one, which is evaluated at the end of the selection. Instead of applying

a weight to each event, however, the weight is set to exactly one, and the momentum value calculated by the described algorithm is varied. This allows potential changes to the momentum reconstruction due to the finite resolution to be observed, without major changes to the analysis procedure.

First, we define the absolute error on the momentum measurement, which sums up the contributions from the differences between mean value x_0 and distribution width γ_0 in data and MC, along with their associated error values:

$$\sigma^2_{abs} = \left(x_0^{\text{data}} - x_0^{\text{MC}}\right)^2 + \sigma^2_{x_0^{\text{data}}} + \sigma^2_{x_0^{\text{MC}}} + \left(\gamma_0^{\text{data}} - \gamma_0^{\text{MC}}\right)^2 + \sigma^2_{\gamma_0^{\text{data}}} + \sigma^2_{\gamma_0^{\text{MC}}}. \tag{6.16}$$

Normalising this by the true momentum of the selected pion in the event then gives the relative resolution, σ_{rel}, which is used in the systematic uncertainty evaluation. The toy variation α, drawn from a Gaussian distribution, is applied with the relative resolution as

$$p^{\text{new}} = (1 + \sigma_{\text{rel}}\alpha) \times \left(p^{\text{nom}} + x_0^{\text{MC}} - p^{\text{true}}\right) + p^{\text{true}}, \tag{6.17}$$

and finally the addition of the bias is corrected for, to obtain the momentum varied according to the resolution,

$$p^{\text{new}}_{\text{corr}} = p^{\text{new}} - x_0^{\text{MC}}. \tag{6.18}$$

To assess the relative error contribution caused by this variation in the momentum, the selection is run using 500 'toy' throws, where a different, randomly drawn, Gaussian variation is applied to each. This is again done for the full FHC MC described in Table 5.I. This yields the relative error distributions shown in Figure 6.55, which at this point are plotted for the three FGD samples (FGD1, FGD2x and FGD2y) that are obtained in the sample selection. The reconstructed pion momentum distributions for these samples are shown in the same binning scheme for reference in Figure 6.56. The pion momentum range is shown up to 400 MeV but, as discussed in Section 6.4.1, the signal sample will be restricted to 300 MeV and below when included in the cross-section fit, as the reconstruction is more reliable in the higher statistics, lower momentum region. As described in Section 6.4.4.1, the control sample is valid for events up to ∼350 MeV (∼1000 mm vertex separation), in order to allow the systematic treatment to migrate events in and out of the 200–300 MeV bin. The bin with the largest error is seen to be the 0–100 MeV one for each of the FGD samples, which is caused by the very low statistics observed in this momentum range. Above this, the relative error averages out to approximately 2%, with FGD2y error always the highest; this is again expected due to the number of events in the sample.

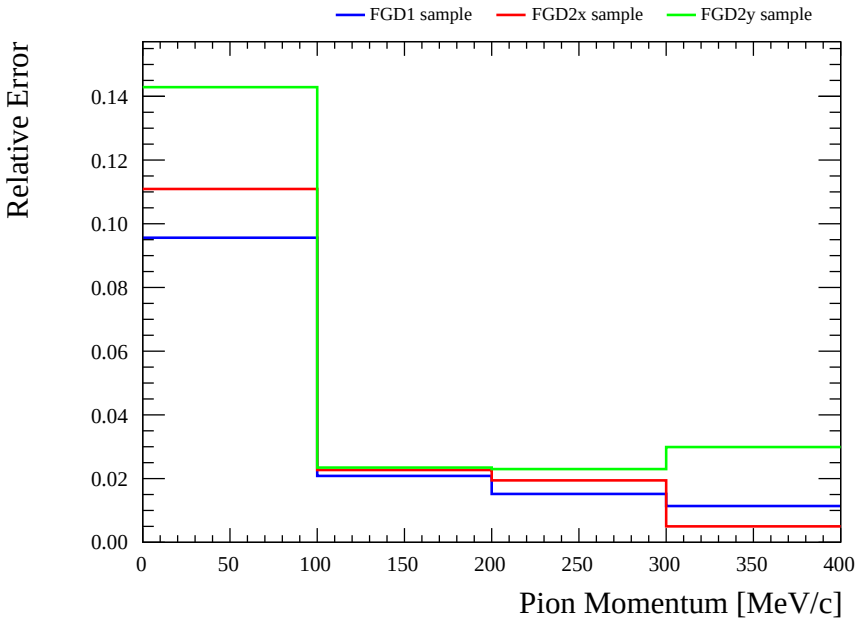
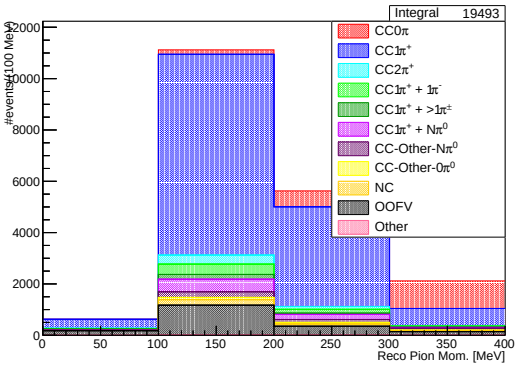
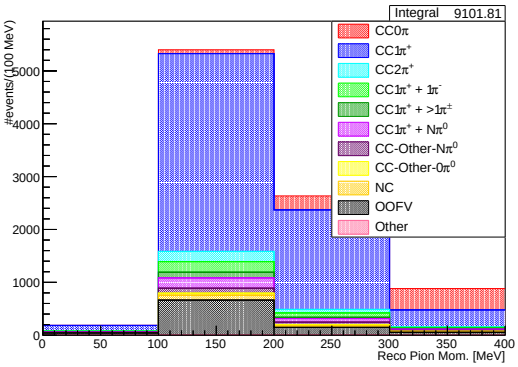


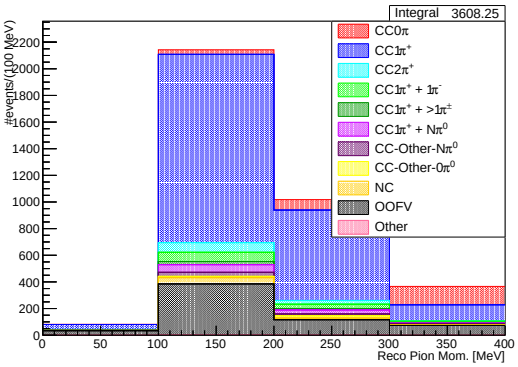
Figure 6.55: Relative error values for the FGD pion momentum from Michel electron reconstruction systematic, plotted as a function of the reconstructed pion momentum. Distributions shown are for the CC1π^+ ME sample, plotted separately for FGD1 (blue), FGD2x (red) and FGD2y (green).



(a) FGD1



(b) FGD2x



(c) FGD2y

Figure 6.56: Reconstructed pion momentum in CC1π^+ ME sample for (a) FGD1, (b) FGD2x and (c) FGD2y. Distributions are given in the same binning as Figure 6.55 for clarity, but the 300–400 MeV bins will not be included in the fit signal samples.

Chapter 7

Systematic Uncertainty, Efficiency and Binning Studies

7.1 Systematic Uncertainties

7.1.1 Detector Systematic Uncertainties

In using different varieties of reconstruction and detector-derived variables, detector related uncertainties have to be taken into account as part of the measurement. The majority of these systematic uncertainties have been thoroughly assessed and used in previous analyses. For the additional ECal cuts and pion reconstruction that were added into this analysis, less established or even entirely new systematics had to be implemented. A summary of the included systematic uncertainties is given in Table 7.I.

The systematic uncertainties can generally be split into two separate types. The first set, known as variation systematics, deal with the uncertainty associated with the scale or resolution of reconstructed variables, such as particle momentum measurements. To assess the effect of these systematic errors, small variations are applied to reconstructed variables. These variations are based on observed differences between data and MC in specific control samples, and scaled by a random number usually drawn from a Gaussian distribution of mean zero and width unity. Systematic propagation is achieved by running the selection multiple times for a number of 'toy throws', and the effect of the applied variations to the measured variables calculated.

The second set are known as weight systematics, which are further divided into efficiency and normalisation systematics. The former of these is associated with the uncertainty due to reconstruction or detection efficiencies, whilst the latter is a normalisation effect for a number of observed events for certain processes, such as OOFV events or sand muons. Rather than having to rerun the selection for every toy to observe the effect, these systematics are propagated via weights that alter the total number of events in the respective sample. Weights for efficiency systematics are propagated by applying the ratios of efficiencies between data and MC in control samples, which are assumed to be the same for the selected sample, again scaled by random numbers drawn from a Gaussian distribution. Weights for normalisation systematics are calculated simply by taking the related weight uncertainty,

and scaling it by the random Gaussian number.

In order to assess the impact of the described detector uncertainties on the measurement, the signal and control sample selections from Sections 6.2 and 6.3 were re-run, with all uncertainties listed in Table 7.I, along with the newly developed one described in Section 6.4.4.3. The selection is run using 500 toy throws, where each toy event is run separately with randomly varied parameters, allowing the contribution from each individual systematic error to be accurately assessed.

Systematic	T2K TN Reference	Type
B-Field Distortions	TN-061	Variation
ECal PID	TN-279	Efficiency
ECal π^0 Veto Pileup	TN-270	Normalisation
FGD Hybrid Tracking Eff.	TN-223	Efficiency
FGD PID	TN-223	Variation
FGD Pion Mom from ME Resol	TN-417	Variation-like
FGD2 Backward Migration	TN-368	Efficiency
Michel Electron Eff.	TN-104	Efficiency
Momentum Range Resol.	TN-216	Variation
OOFV Background	TN-098	Normalisation
Pile-Up	TN-152	Normalisation
Pion Secondary Interactions	TN-125	Normalisation
Proton Secondary Interactions	TN-216	Normalisation
Sand Background	TN-077	Normalisation
TPC–ECal Matching	TN-279	Efficiency
TPC–FGD Matching	TN-075	Efficiency
TPC Charge ID Eff.	TN-229	Efficiency
TPC Cluster Eff.	TN-234	Efficiency
TPC Momentum Resolution	TN-222	Variation
TPC Momentum Scale	TN-081	Variation
TPC PID	TN-221	Variation
TPC Tracking Eff.	TN-163	Efficiency
ToF Resolution	TN-245	Variation

Table 7.I: Summary of detector systematic uncertainties included in the analysis.

7.1.1.1 B-Field Distortions

Distortions in the UA1 magnetic field surrounding the ND280 tracker result in TPC field distortions, which are corrected for in two ways, described in full detail in T2K-TN-061 [207]. The main correction for the distortions is applied at the reconstruction level, using the B field map created from measurements of the magnetic field within the ND280 basket. The position of each hit cluster is drifted back to the ionisation point, and the field map applied in order to correct the y and z positions of the clusters. This correction itself is referred to as the B field correction.

A second correction, known as the empirical correction, is based on measurements from the TPC laser system, where a laser is used to illuminate aluminium dots on the cathode.

The emitted photo-electrons drift to the anode and are recorded by the Micromegas. This is done for each dot, using measurements with B field both off and on, in order to isolate the effect of the distortions. The expected and measured positions are compared in order to empirically derive a distortion map as a function of drift distance. This empirical correction is then applied instead of the nominal correction, in order to get the systematic uncertainty for the B field distortions.

7.1.1.2 ECal PID

ECal PID methods are employed in the muon MIPEM cut described in Section 6.2.1. In order to estimate the systematic uncertainty on this cut, the selection efficiencies in data and MC for a cut placed at a MIPEM value of 0 are compared, for high purity electron, muon and proton control samples. Full details of the control sample selection and data–MC differences can be found in T2K-TN-279 [208], whilst the final data–MC efficiency differences for the control samples are summarised in Table 7.II.

Sample	Module	Data – MC efficiency
Sand muon	DsECal	$0.6 \pm 0.1\%$
Cosmic muon	BrECal	$1.9 \pm 0.1\%$
e^+e^-	DsECal	$-1.2 \pm 0.9\%$
	BrECal	$-1.9 \pm 2.1\%$
Proton	DsECal	$-1.8 \pm 0.7\%$
	BrECal	$1.3 \pm 0.9\%$

Table 7.II: Data–MC efficiency differences for ECal PID MIPEM control samples, from T2K-TN-279 [208].

7.1.1.3 ECal π^0 Veto Pileup

Uncertainty in the ECal π^0 veto cut can come from two separate cases: where an event does not have its true π^0 reconstructed, and so is not removed from the sample, or when a false π^0 is reconstructed and the event is incorrectly removed. For all events that reach the π^0 veto cut, a weight is applied as:

$$w_e = 1 + \alpha \delta e_{\text{cat}}, \tag{7.1}$$

where δe_{cat} is equal to $P(\text{pass}|\pi^0)$ if the event contains a π^0, or $P(\text{removed}|\text{no } \pi^0)$ if the event does not contain any π^0s, and α is the random toy variation drawn from the Gaussian distribution. Full details of these calculations are given in T2K-TN-270 [209].

7.1.1.4 FGD Hybrid Tracking Efficiency

The goal of the FGD-contained hybrid track study in T2K-TN-223 [210] is to calculate the efficiency with which fully contained FGD tracks are reconstructed, in the presence of a long FGD-TPC matched muon candidate track. In order to do this, 'hybrid' files containing additional simulated FGD tracks are created for both data and MC samples.

This involves taking a CC0π selection with a muon candidate crossing a TPC, and selecting out $\sim$1000 vertices without FGD-contained tracks. At each vertex position, 100 positive pions and protons are generated (as separate samples) using the `GEANT` Particle Gun. The hit information from the generated pions or protons is inserted into the reconstruction-level file along with the original muon candidate vertex. Finally, these hybrid files are reprocessed using FGD-only reconstruction. The separate pion and proton reconstruction efficiencies are calculated as the ratio of events with at least one isoRecon object to the total number of events. The MC and data efficiencies and errors are then plotted as a function of the cosine of the angle between the true pion/proton direction and the direction of the reconstructed muon candidate, where the latter comes from the original data or MC reconstruction.

7.1.1.5 FGD PID

Non-interacting particles which stop in the FGD volume lose all of their energy through ionisation, allowing the energy deposit and particle range to be translated into a mass. As described in Section 6.1.7, the normalised difference between the measured and expected energy loss is used to form particle pull distributions, which are used by the PID method.

In order to assess the systematics associated with the FGD PID, pull distributions can be modelled by a Gaussian function. These are checked using control samples where particles cross the TPC, allowing for TPC PID, and then stop in the FGD volume. A set of toy experiments can be generated with the FGD PID measurements varied, and used to provide the systematic errors on the PID, for two particle hypotheses – muon and proton. Due to the small mass difference between muons and pions, it is assumed that the pion PID systematic error can reasonably be covered by that of the muon, since the energy deposition will be similar. Full details of the analysis and control samples used to assess these systematics is given in T2K-TN-223 [210].

7.1.1.6 FGD2 Vertex Backward Migration

Backward migration is defined as the effect where an interaction vertex is reconstructed upstream from its true position. This is often caused by incorrect reconstruction of backward-going tracks. For example, a forward-going muon and a backward-going proton may be reconstructed together as a single muon track, placing the reconstructed interaction vertex further upstream, at the proton's end position. The effect of this vertex migration is particularly important to understand for this analysis due to the splitting of the FGD2 sample into individual X and Y layers. This splitting is used to estimate the number of interactions on water and hydrocarbon, and migration between samples could cause a large effect in the eventual fit.

In order to estimate the number of backward migrated tracks, the distance between the z coordinate of the first hit from the muon candidate and the z coordinate of the track fit in the first hit layer is used. The 'fit' position is defined by taking the start point of the track along with its direction, and using this to estimate the z coordinate of the first hit. The uncertainty is then estimated from this by comparing the fractions of migrated

and non-migrated tracks between data and MC. Full details of the analysis are given in T2K-TN-368 [211].

7.1.1.7 Michel Electron Systematics

The original treatment of the Michel electron systematics is fully described in T2K-TN-104 [196], T2K-TN-128 [212] and T2K-TN-152 [198], with an update for production 6B added in T2K-TN-212 [213]. The systematic errors on the Michel electron selection arise from the detection efficiency and purity, which are treated separately.

Tagging efficiency: In order to assess the differences in Michel electron tagging efficiency between data and MC, FGD-triggered cosmic samples are used. In FGD1, the sample selection looks for tracks stopping in the FGD1 FV, with one good quality track required in TPC2 and none allowed in TPC1. In order to remove electron contamination, the stopping distance in the FGD is compared with the TPC2 measured momentum of the track, and events consistent with a muon selected. For the FGD2 sample, cosmic tracks stopping in FGD2 are selected, with no tracks in TPC3. PID based on the FGD2 momentum-by-range is then applied in order to increase muon purity. Using these samples, the Michel electron cut (Section 6.1.7) is applied, in order to calculate the detection efficiency. The detection efficiency in FGD1 was found to be $(56.5 \pm 0.9)\%$ for MC, and $(56.4 \pm 0.2)\%$ for data runs 2–4. In FGD2, the detection efficiency was $(41.4 \pm 0.7)\%$ for MC, and $(42.8 \pm 0.1)\%$ for data runs 2–4.

Purity: Particles that enter from outside the FGDs are capable of leaving energy depositions that can mimic the signature of Michel electrons, causing them to pass the cut; these are considered external background to the Michel cut. The external background rate in the FGDs is measured in both data and MC, using empty beam spills with no FGD activity during beam bunch periods. This has to be done separately for each run, as the overall external background rate increases with beam power. Typical false identification rates for data are $(1.8 \pm 0.05) \times 10^{-3}$ in FGD1 and $(1.3 \pm 0.05) \times 10^{-3}$ in FGD2, and for MC $(0.7 \pm 0.009) \times 10^{-3}$ and $(0.4 \pm 0.006) \times 10^{-3}$ (full tabulated values taken from T2K-TN-212 [213]). The false identification rate is seen to be significantly higher in data than MC due to the fact that cosmics are not simulated as part of either the beam or sand MC.

7.1.1.8 Momentum by Range Resolution

For particle tracks that do not reach the TPC, momentum can be estimated from the range of the particle as it crosses the other sub-detectors. In this analysis, this method is used specifically for pion tracks which do not reach the TPC or decay to Michel electrons, and instead leave only an isolated FGD track. To assess the systematic error associated with this momentum by range reconstruction, a control sample is formed of tracks which have a TPC component, and stop in a detector of interest (FGD, BrECal or SMRD). For these tracks, the TPC component is used to provide the momentum of the particle, for which the data–MC discrepancies have been separately evaluated and understood. Applying the

momentum by range algorithm in the stopping detector, the calculated momentum by range can be compared to the TPC momentum. The remaining data–MC discrepancy can then be attributed to the error in range momentum calculation. In order to evaluate the systematic on this, normalised difference between TPC and range momentum for data and MC is fitted with a Gaussian function, and the results combined as

$$\sigma^2_{\text{syst}} = (\mu_{\text{data}} - \mu_{\text{MC}})^2 + \sigma^2_{\mu_{\text{data}}} + \sigma^2_{\mu_{\text{MC}}} + (\sigma_{\text{data}} - \sigma_{\text{MC}})^2 + \sigma^2_{\sigma_{\text{data}}} + \sigma^2_{\sigma_{\text{MC}}}. \quad (7.2)$$

The tabulated results of the Gaussian fitting for each possible stopping detector are given in Table 7.III, reproduced from T2K-TN-216 [214], where full details of the calculations are given. The error is then propagated by taking each toy throw and using it to smear the

Detector	$\cos\theta$	μ_{data}	μ_{MC}	σ_{data}	σ_{MC}
FGD1	−1.0–1.0	-0.0112 ± 0.0016	0.0010 ± 0.0005	0.0532 ± 0.0015	0.0567 ± 0.0005
BrECal	−1.0–0.6	-0.1267 ± 0.0091	0.1255 ± 0.0018	0.1926 ± 0.0079	0.1890 ± 0.0016
BrECal	0.6–1.0	-0.1303 ± 0.0114	0.0993 ± 0.0023	0.2240 ± 0.0095	0.2056 ± 0.0021
SMRD	−1.0–0.6	-0.0609 ± 0.0198	0.0457 ± 0.0027	0.2194 ± 0.0171	0.1850 ± 0.0027
SMRD	0.6–1.0	-0.0629 ± 0.0154	0.0342 ± 0.0027	0.2899 ± 0.0154	0.2584 ± 0.0026

Table 7.III: Gaussian fit parameters for the normalised difference between TPC and FGD range momentum, for data and MC.

resolution, by the formula

$$p^{\text{new}}_{\text{range}} = (1 + \sigma_{\text{resol}}\alpha) \times (p^{\text{nom}}_{\text{range}} - p_{\text{true}}) + p^{true}, \quad (7.3)$$

where α is the random toy variation.

7.1.1.9 OOFV Background

Out of fiducial volume (OOFV) events are defined as those whose interactions are reconstructed as originating from within the FGD FV, but whose true vertices are actually outside. The FGD fiducial volume definition is given in Section 6.1.3. As described in T2K-TN-098 [215], the OOFV background is split into nine categories, using the true information contained within a MC sample of CC-inclusive FGD events.

Rate Uncertainty: The rate uncertainty for OOFV occurring within the tracker dead material is set to 0%, as it is considered that the scaling of the cross section with A for material other than carbon is well understood [213]. For interactions on heavier nuclei outside the tracker, production 6 data and MC are used to compute the interaction rate in the PØD, BrECal and SMRD. Due to the rate of interactions in the magnet or electronics being harder to measure, the relative difference of the rate prediction of events from `NEUT` and `GENIE` was taken to be the uncertainty for these. The uncertainties are summarised in Table 7.IV.

Background Origin	Rate Uncertainty (ν_μ)
PØD	5.1%
ECal	11.6%
SMRD	4.9%
Other	13.6%

Table 7.IV: Rate uncertainties on OOFV events, depending on their origin.

Reconstruction Uncertainty: For OOFV events within the FGD, the uncertainty is closely related to FGD hit efficiency. Between data and MC there is near perfect agreement, and so there is no reconstruction uncertainty placed on this category. Due to the fact that neutral and backward events cannot be distinguished by the reconstruction, these categories also have no uncertainty placed on them. Interactions in the upstream tracker dead material are found to be negligible and thus also receive no uncertainty, whilst the more prevalent interactions in the downstream tracker dead material are assigned an uncertainty of 5%. For the remaining four OOFV categories, the uncertainty is computed from the TPC–FGD incomplete matching rate, described in Section 7.1.1.15. A summary of the contribution from OOFV categories along with their assigned reconstruction uncertainties for both FGDs is given in Table 7.V.

Category	Fraction		Reconstruction Uncertainty	
	FGD1	FGD2	FGD1	FGD2
OOFV inside the FGD	11.1%	11.6%	0%	0%
In tracker upstream	3.1%	7.2%	0%	0%
In tracker downstream	10.3%	8.1%	5%	5%
From neutral parent	22.5%	14.0%	0%	0%
Backward event	17.1%	21.3%	0%	0%
High angle event	10.1%	15.8%	33%	28%
Last module failure	5.5%	8.9%	35%	17%
Double skipped layers	5.8%	4.5%	55%	82%
Hard scattering	14.2%	8.7%	32%	21%

Table 7.V: OOFV category breakdown, along with the associated reconstruction uncertainty, for both FGDs 1 and 2.

7.1.1.10 Pileup

Event pileup refers to non-beam events and sand interactions that happen to overlap with neutrino interactions from beam events, preventing the detector from correctly reconstructing the event. Whilst there are several possibilities that can cause pileup, for the ν_μ analysis the only significant effect comes from sand muons in coincidence with magnet events. A study from T2K-TN-212 [213] shows that, of all the MC events rejected from the FGD1 selection by the upstream TPC1 veto, only $(21.7 \pm 0.7)\%$ were CC interactions, with the rest being external. As the standard MC simulation does not include sand muons and therefore

does not account for events being removed due to sand muon coincidence, a correction has to be made. Since the vertex selection in FGD2 also places a veto on TPC2 activity, the pileup systematic can be assessed in a very similar way for both FGDs, with TPC1 pileup examined for the FGD1 selection, and TPC2 pileup for the FGD2 selection.

The procedure for calculating this correction is outlined fully in [198], where the numbers were evaluated for each data run separately, also splitting by the PØD water in/out status. The number of TPC1 (TPC2) events N_s^{TPC} in a separate sand file is counted, relating to a fixed POT, POT_s. The data intensity $I_d = POT/N_{\mathrm{spills}}$ is found from the data sample, and used the calculate the effective number of spills. Dividing this by the number of bunches per spill N_b in the data set gives a correction to the number of TPC1 (TPC2) events per bunch:

$$C_s = \frac{N_s^{\mathrm{TPC}} \times I_d}{POT_s \times N_b}. \tag{7.4}$$

The number of selected events in MC can then be reweighted with the following factor:

$$w_c = (1 - C_s). \tag{7.5}$$

Given that there is a 10% uncertainty in the sand muon simulation, along with possible differences between the actual and simulated beam intensity, a systematic uncertainty arises on this pileup contribution. This is evaluated through a data–MC comparison of the number of TPC1 (TPC2) events per bunch, with the sand muon contribution added to the MC. This is calculated as

$$\Delta_{\mathrm{data:MC}} = C_d - (C_{MC} + C_s), \tag{7.6}$$

where

$$C_d = \frac{N_d^{\mathrm{TPC}}}{N_{\mathrm{spills}} \times N_b} \text{ and } C_{MC} = \frac{N_{\mathrm{MC}}^{\mathrm{TPC}} \times I_d}{POT_{\mathrm{MC}} \times N_b}.$$

In order to avoid double counting uncertainties, the procedure is to choose the larger number out of $\Delta_{\mathrm{data:MC}}$ or $0.1 \times C_s$, and apply this as the pileup uncertainty, σ_{pileup}. This systematic is then propagated as a normalisation error as

$$w_{\mathrm{pileup}} = 1 + \alpha\sigma_{\mathrm{pileup}}, \tag{7.7}$$

where α is a random variation of the toy variable.

7.1.1.11 Pion Secondary Interactions

Pion secondary interactions (SI) are defined as the interactions that a pion undergoes outside of the nucleus that it was produced in. Since the model used for these interactions is found to disagree with the external data, a systematic has to be applied, based on the observed data-MC differences. The full treatment of this is discussed in T2K-TN-125 [216]. However

in the move to production 6T, the standard `GEANT` model originally used for modelling interactions was switched for the newer `NEUT` cascade model. Since this model tunes the MC based on data, a reduction in the systematic errors for pion SI was observed [193].

7.1.1.12 Proton Secondary Interactions

This is similar to the equivalent for pion SI described in Section 7.1.1.11. Each proton within the volume of interest (VOI) is taken into account, where the VOI for FGD1(2) is the FGD1(2) volume plus the material between the active volume of the FGD1(2) and its downstream TPC, TPC2(3). All true primary protons that interact within the VOI are propagated through in steps of 0.1 mm, in order to find the probability of interaction. By varying the interaction cross section at each step based on the data uncertainty, the final weight is calculated, with full details given in T2K-TN-216 [214].

7.1.1.13 Sand Background

Sand muons are produced when beam neutrinos interact in the sand surrounding the detector. These can enter the detector and leave tracks which may be mistaken for neutrino interactions occurring in the FGDs. This is modelled by generating a specific set of sand MC, as described in T2K-TN-077 [217]. In order to assess the contribution from sand muon events, the standard CC multipion selection for FGD1 and 2 is applied to the sand MC. The systematic error on the number of sand events is evaluated by also selecting a sand enriched sample of events with tracks entering through the upstream wall of the PØD. Both of these MC rates, normalised to POT, are summed and compared to the rate observed in data. The data–MC discrepancy is found to be $\sim$10%, and so a 10% systematic uncertainty is applied to the number of predicted sand events.

7.1.1.14 TPC–ECal Matching

Similarly to the TPC–FGD matching, any difference between MC and data in the efficiency of matching track components between a TPC and ECal module will result in a different number of events passing a selection. Since several of the newer cuts in this selection require the muon track to have an ECal component, it is necessary to understand the uncertainty associated with any data–MC discrepancies. In order to assess this, six control samples are selected for tracks consistent with electrons, positrons, muon, anti-muons and protons, which leave the TPC and enter the BrECal or DsECal. The matching efficiency is calculated based on the number of selected tracks containing an ECal segment:

$$\epsilon = \frac{\text{Track passing selection with ECal segment}}{\text{Track passing selection}} \tag{7.8}$$

To determine the systematic uncertainty, the difference between data and MC efficiencies is calculated, whilst the uncertainties are defined by the statistical uncertainty on those

efficiencies combined in quadrature:

$$\sigma = \epsilon_{\text{data}} - \epsilon_{\text{MC}} \tag{7.9}$$

$$\delta\sigma = \sqrt{\delta\epsilon^2_{\text{data}} + \delta\epsilon^2_{\text{MC}}} \tag{7.10}$$

The full unbinned and binned efficiencies and systematics for each particle hypothesis and ECal are given in T2K-TN-279 [208].

7.1.1.15 TPC–FGD Matching

The TPC–FGD matching efficiency characterises how well the reconstruction is able to match TPC and FGD tracks together. Incorrect matching between detectors can lead to improper vertex reconstruction with an incorrect position. The two sets of uncertainties evaluated for this, which are both described in detail in T2K-TN-075 [218], are summarised here.

Basic matching efficiency: This is the efficiency of matching any of the selected TPC tracks to a hit in the upstream FGD. The systematic uncertainty is taken to be the difference between the matching rate in data and MC for the muon candidate. Since this is insensitive to the FV, this systematic is used to reweight all tracks.

Complete matching efficiency: This is given by the complete matching rate, the count of how often the TPC–FGD matched track is matched with a track in the upstream TPC (leading to a TPC–FGD–TPC matched track). In the case of sand muons, complete matching occurs when the global track starts in the first module of the FGD, outside of the FV z boundary. In the case of cosmic events, the FGD track must start outside of the FGD FV x and y limits, in order to ensure it will be matched to a surrounding detector track. Incomplete matching occurs if a global track starts inside the FGD FV, which can cause background contamination by incorrectly locating the vertex in carbon or water. The systematic is determined from the incomplete matching rate, and is used to reweight OOFV events.

7.1.1.16 TPC Charge ID Efficiency

For this analysis, the updated charge ID algorithms are used, which rely on the global charge identification from a combination of the ND280 subdetectors. The new method has two main goals, and from this two different systematic errors can be drawn. These are:

- probability of swapping local TPC charge identification.
- probability of the global tracking swapping the sign of the charge obtained from the best among the local segments.

The charge sign determination relies mainly on TPC reconstructed information, focusing on the curvature of the track within TPC segments. This depends upon the track length,

number of hits and the spatial resolution at each measured point. Instead of dealing with the complication of three variable dependencies, the momentum error from the tracking fit is used, as it should be unique and have the correct dependencies. The error in the momentum is used to parametrise the the data-MC disagreement for different numbers of TPC segments, for both the local and global charges. The tabulated values along with a full description of the method and calculations used are given in T2K-TN-229 [219].

7.1.1.17 TPC Cluster Efficiency

The TPC cluster efficiency is defined as the probability of finding a group of adjacent single TPC pad hits, or clusters, that corresponds to a single point in the ionised gas trace left by a charged particle crossing the TPC gas volume. The centre of the ionisation is found by forming the hits into clusters in the X (horizontal), or Y (vertical) direction[11]. The distinction between horizontal and vertical clustering is decided upon based on the absolute value of the local angle of the track with respect to the Z direction in the ZY plane. If the value is greater than 55°, horizontal clusters are used, with vertical clusters used otherwise.

A difference in the TPC clustering efficiency between data and MC is found to be the main reason for different fractions of data and MC events passing the TPC track quality cut, which requires events to have over 18 associated TPC clusters with hits in. In order to account for this, TN-234 [220] calculates an additional TPC cluster efficiency α:

$$\alpha = \frac{(\epsilon^{\mathrm{MC}} - \epsilon^{\mathrm{data}})}{\epsilon^{\mathrm{MC}}}. \tag{7.11}$$

This allows MC tracks to lose additional clusters due to an inefficiency equal to α, causing event migration until the function matches to data. α is evaluated independently for vertical and horizontal, with fit values found to be $\alpha = 0.0011 \pm 0.0002$ in the vertical, and $\alpha = 0.0007 \pm 0.0001$ in the horizontal. This weighting is then applied to tracks which are just above the threshold required to pass the TPC quality cut described in Section 6.1.3 [193].

7.1.1.18 TPC Momentum Resolution

The study on TPC momentum resolution is done using tracks which cross multiple TPCs, in order to build a fully reconstructed track with sensitivity to the TPC resolution. The goal is to compare the TPC and global momentum resolutions for both data and MC, to find the smearing factor that removes any differences. Taking a track that crosses at least two TPCs allows the difference between the reconstructed momentum of the global track from the two separate TPC segments to be evaluated. The difference in the inverse of the momentum transverse to the magnetic field between the TPC segments ($\Delta 1/p_t$), corrected for energy loss in the intermediate FGD, is seen to have an approximately Gaussian shape centred at zero. The main contributor to the standard deviation of this is the intrinsic resolution of the TPCs involved. Fitting the distribution with a Gaussian in order to find the standard deviation for different kinematic ranges gives an estimate of the momentum

[11] In internal TPC documentation these are referred to as Y and Z, using a local TPC coordinate system which differs from the global ND280 coordinate system.

resolution. This is done for a control sample detailed in T2K-TN-222 [221], where a full description of the method to convert to a systematic uncertainty is also given.

7.1.1.19 TPC Momentum Scale

The momentum scale error is obtained from the B-field measurement, described in T2K-TN-081 [222]. The uncertainty applied in the propagation is 0.57%.

7.1.1.20 TPC PID

The TPC PID is an essential part of the full selection, which allows for estimation of the particle type based on its energy loss in the TPC gas system. This is summarised in Section 6.1.7, with the equation for the pull used for particle discrimination given in Equation (6.4). For this analysis, the TPC-PID systematics are particularly important due to the possible misidentification between muons and pions in the leading track case, an issue which is further explored in the ECal MIPEM cut (Section 6.2.1). Other confusion between pions and protons can also occur in searching for the reconstructed secondary particle in order to classify the CC-inclusive event by a certain exclusive topology, with electron misidentification for muons or pions also having a small effect.

In order to estimate the systematics, separate control samples are built for muons, electrons and protons, extracted directly from beam events. Since pions have an energy loss similar to muons, they are covered by the muon systematics. The systematics are estimated from the data–MC difference, by grouping events in momentum bins and fitting the obtained pull distribution with a Gaussian. It is observed that there is good agreement between data and MC values in the high momentum regions, whilst at low momentum the values diverge. From these distributions, two quantities are calculated for the systematic propagation:

- The difference between the pull mean values in data and MC. This gives an estimation of the systematic bias between the two.
- The ratio between the pull sigma values ($\sigma_{\text{data}}/\sigma_{\text{MC}}$). This is used to estimate the smearing to be applied when producing toys.

For muon candidates, the pull mean difference $|\delta_{\text{data}}-\delta_{\text{MC}}|$ and pull sigma ratio $\sigma_{\text{data}}/\sigma_{\text{MC}}$ are found to be 0.11 and 1.01, respectively. The equivalent values for proton candidates evaluate to 0.49 and 1.21. Full tabulated values and a more detailed description of the control samples used for the systematic error evaluation can be found in T2K-TN-221 [223].

7.1.1.21 TPC Tracking Efficiency

This systematic uncertainty describes the efficiency with which the TPCs are able to successfully reconstruct the tracks from particles which have crossed them. Incorrect reconstruction can lead to the wrong classification of neutrino event topology, migrating events between sub-samples and possibly removing them from the selection all together.

A full description of the method used to evaluate track reconstruction efficiency is given in T2K-TN-163 [224], where both the upstream and downstream detectors (excluding the

FGDs) of the TPC of interest are used to select a single muon track. The tracks in the surrounding detectors are known as the 'reference tracks'. Using geometric cuts, it is checked whether the reference tracks and the track in the given TPC originate from the same particle, based on proximity. Where this is not the case, the TPC is classed as inefficient for that event. The efficiency values are calculated separately for each TPC, for both data and MC, and are given in Table 7.VI. The ratio of the data to MC efficiencies, with statistical error in the control samples accounted for, is then used to calculate a weight to be applied to events. Full values and a more in-depth description of the systematic error analysis procedure can be found in the aforementioned technical note.

ϵ	TPC1	TPC2	TPC3
Data	$99.9^{+0.1}_{-0.1}\%$	$99.7^{+0.2}_{-0.7}\%$	$99.3^{+0.5}_{-0.2}\%$
MC	$99.6^{+0.2}_{-0.3}\%$	$99.5^{+0.3}_{-0.4}\%$	$99.8^{+0.1}_{-0.2}\%$

Table 7.VI: TPC tracking efficiency values for data and MC.

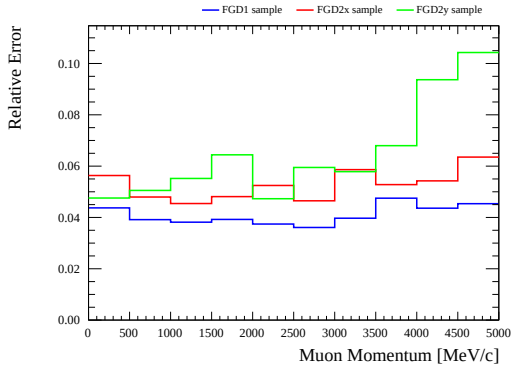
7.1.1.22 ToF Resolution

Time of flight information between the ND280 sub detectors that a track passes through can be particularly useful for determining the direction that the particle that left the track was travelling in. However, the timing information between sub detectors is known to suffer from differences between data and MC (see Chapter 4). The current method used to account for this is to fit measured ToF distributions between each detector combination in data and MC with a sum of two Gaussian distributions, which represent the forward and backward-going peaks. The difference between the mean and standard deviation of each is used to form an additional Gaussian distribution, by which the MC is smeared so that it resembles the data. Full details on the smearing factor calculations can be found in T2K-TN-245 [225]. As a conservative estimate, the associated error is then set to the maximum bias or resolution correction.

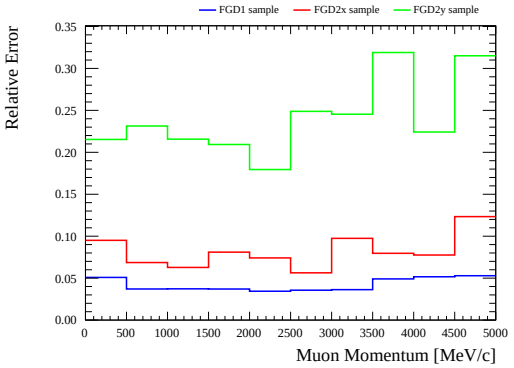
An additional treatment must be applied in order to account for the greater uncertainty caused by cutting on the time of flight information for events from run 8. Due to electronics issues during operation, the quality of the inter-detector timing information in this run was significantly worse than for runs 2–4. As was mentioned in Section 6.2.3, to avoid removing a large number of events which have bad timing quality, regardless of whether the timing information would be used, we instead inflate the uncertainty for events from run 8 which have the FGD–BrECal ToF cut applied. A study performed for the same purposes in T2K-TN-277 [202] found that data deemed to have bad timing quality in run 8 has a resolution roughly three times larger than of that with good timing quality. To take this into account, the variation applied during the systematic propagation is therefore smeared by a factor of three for run 8 events marked as having bad timing data, to replicate the worse resolution for timing measurements, which is not simulated in MC.

7.1.1.23 Detector Systematics Summary

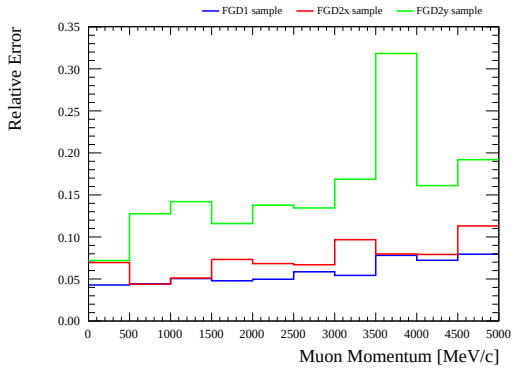
The total contributions from all systematic errors, for each of the selected signal and sideband samples, are presented in Figure 7.1. Several uncertainties more specific to this analysis are presented in the text here: the FGD backward migration systematic is shown in Figure 7.2, the FGD momentum by range resolution for isoFGD pions in Figure 7.3 and the time of flight resolution in Figure 7.4. All other systematic uncertainties are given in Figures E.1 to E.20, in Appendix E. Although the most relevant systematic plot for the FGD pion momentum from Michel electron reconstruction systematic has already been presented and discussed in Figure 6.55, the systematic distributions for all signal and background samples for this uncertainty are included in Figure E.6. For completeness, summaries of the contributions from the individual sources of systematic error to each sample, along with the total systematic error in each, are presented in Table 7.VII (FGD1), Table 7.VIII (FGD2x) and Table 7.IX (FGD2y). Additionally, the pre-fit uncertainty covariance matrix used as an input to the fit is provided in Figure 7.5, whilst the associated correlation matrix is given in Figure 7.6.



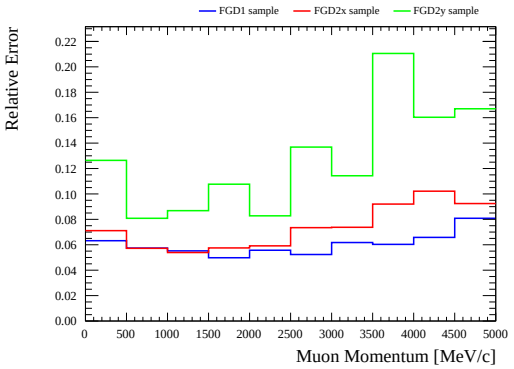
(a) CC1π^+ TPC



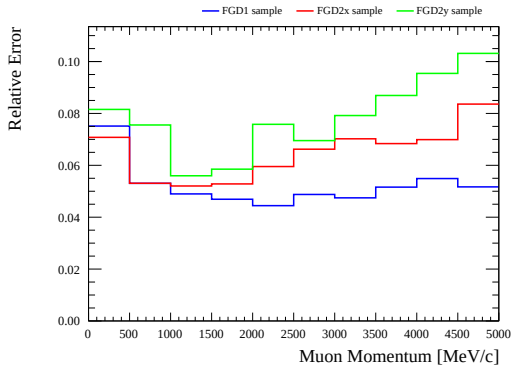
(b) CC1π^+ ME



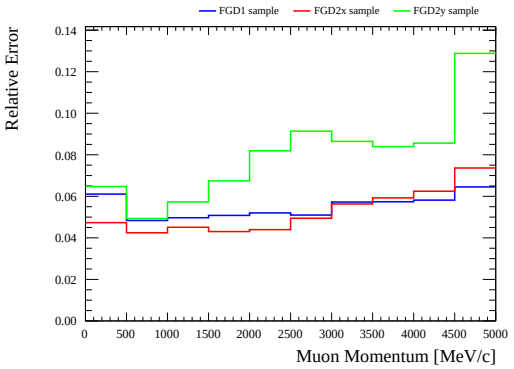
(c) CC1π^+ isoFGD



(d) CS1

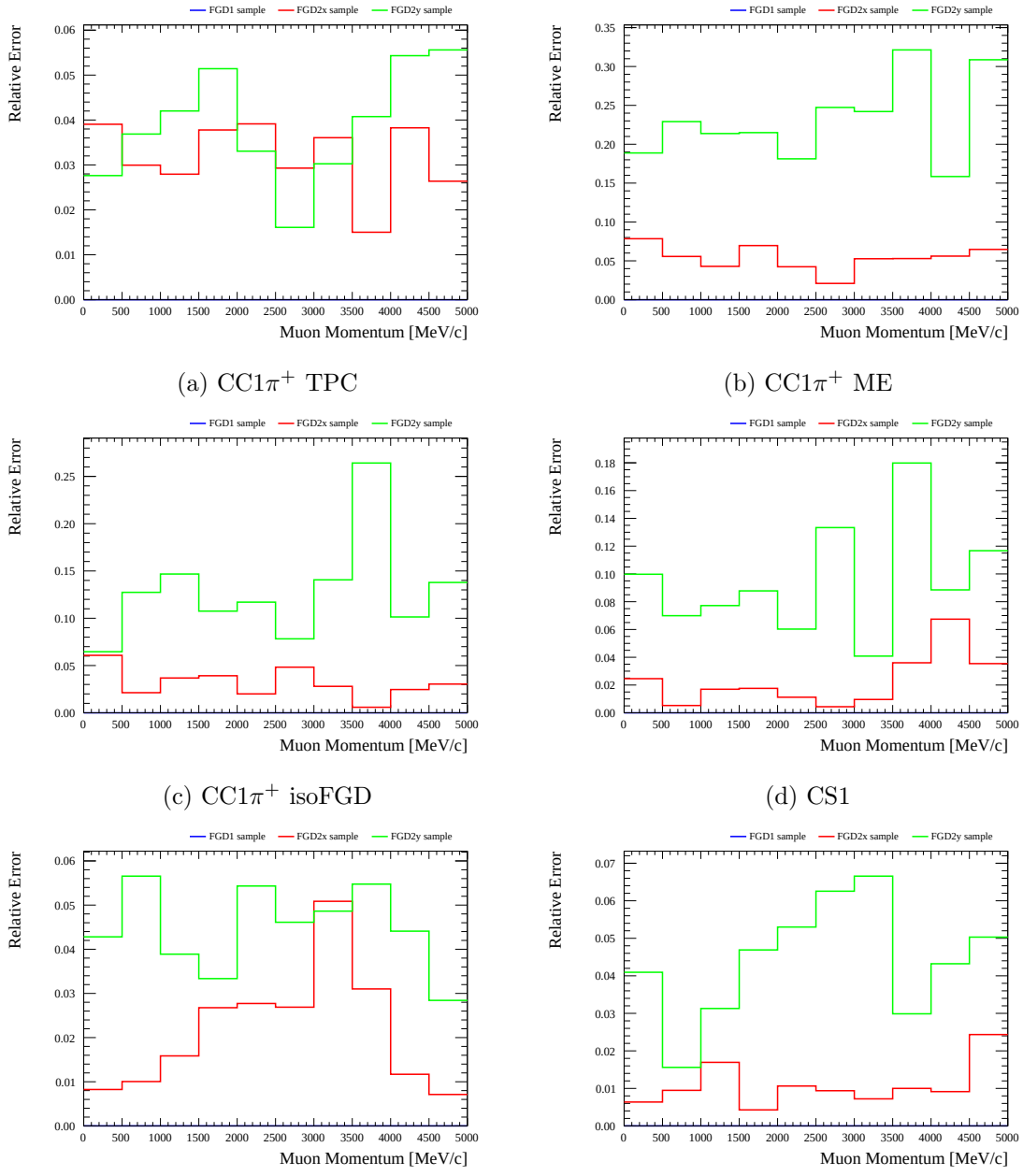


(e) CS2



(f) CS3

Figure 7.1: Relative error values for all systematics, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



(a) CC1π^+ TPC

(b) CC1π^+ ME

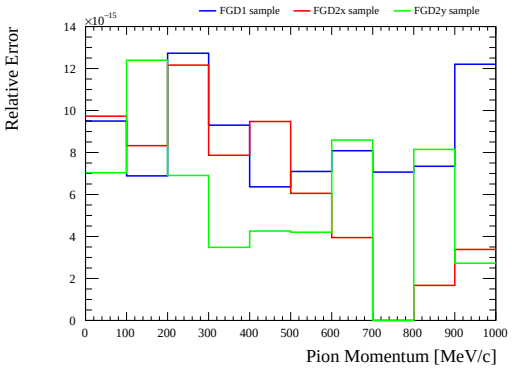
(c) CC1π^+ isoFGD

(d) CS1

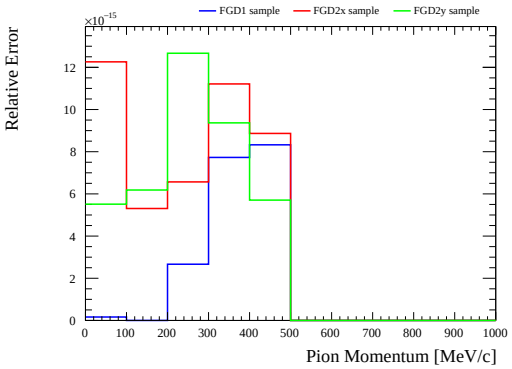
(e) CS2

(f) CS3

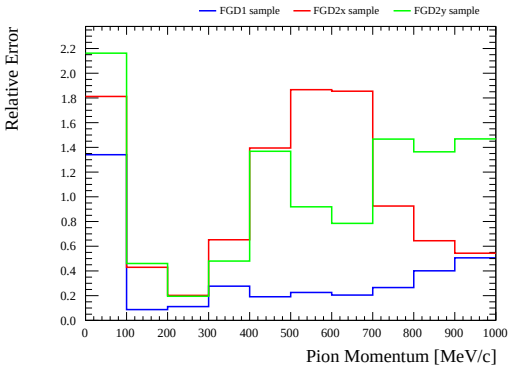
Figure 7.2: Relative error values for the FGD2 backward migration systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



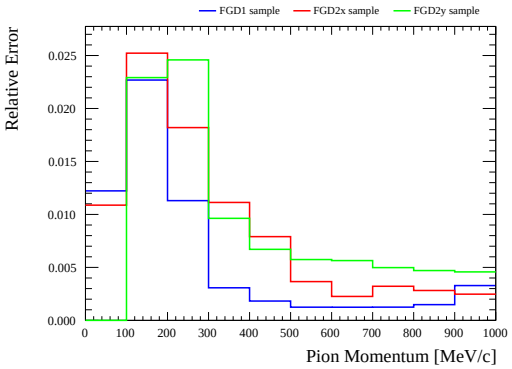
(a) CC1π^+ TPC



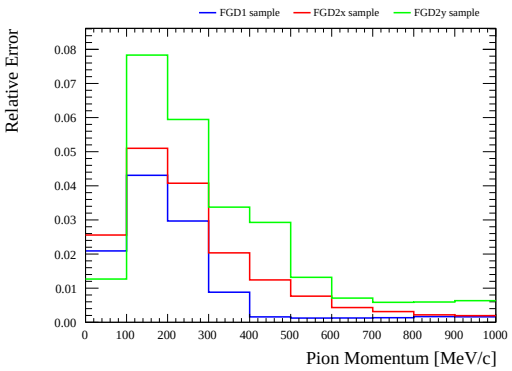
(b) CC1π^+ ME



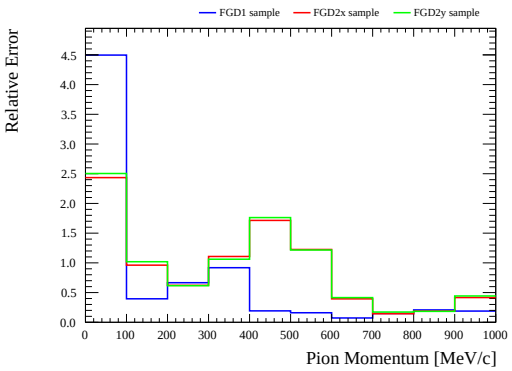
(c) CC1π^+ isoFGD



(d) CS1

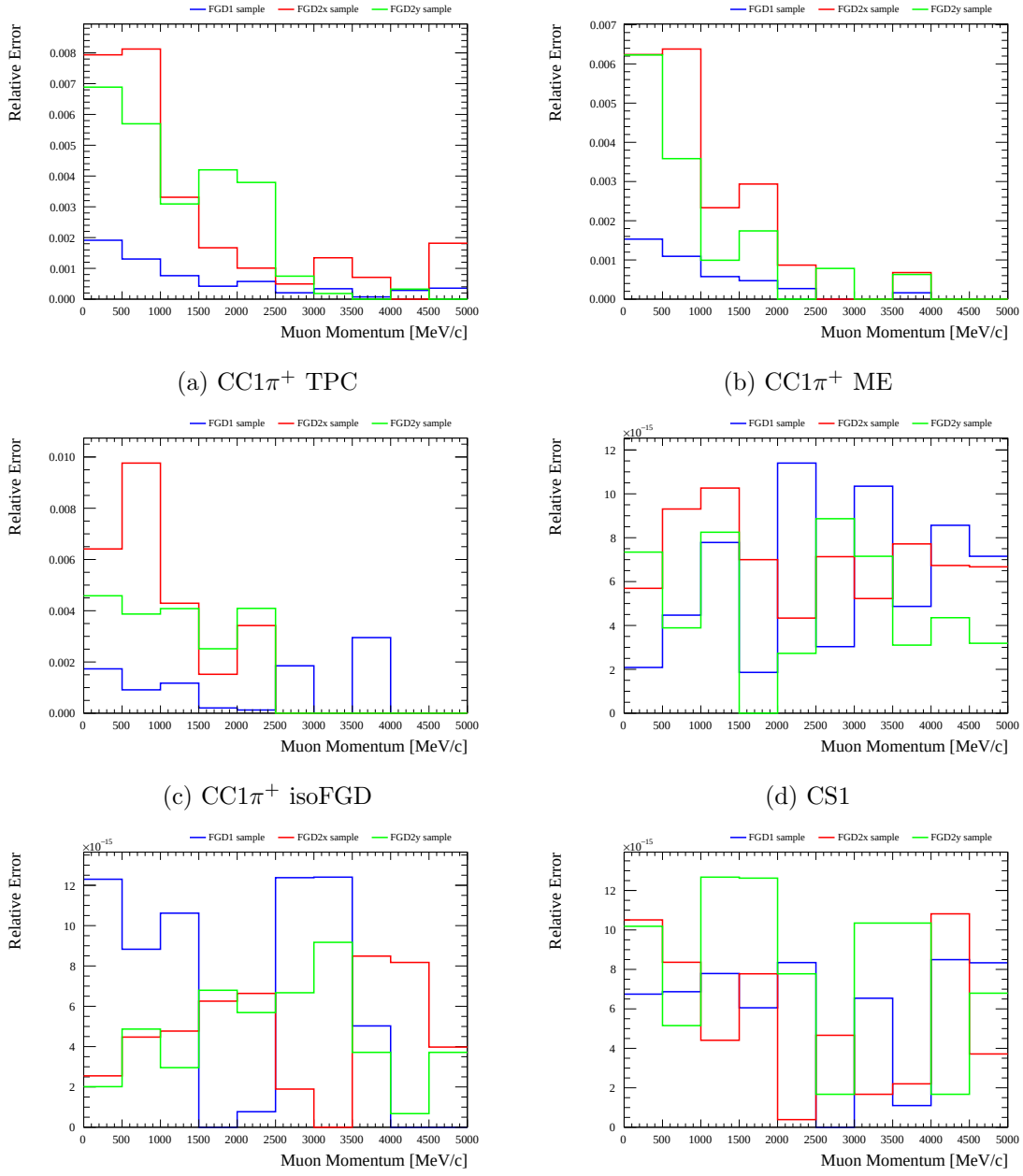


(e) CS2



(f) CS3

Figure 7.3: Relative error values for the isoFGD pion momentum by range resolution systematic, as a function of reconstructed pion momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green). Error values in the TPC and ME samples are expected to be zero, as the momentum by range is not used in these branches.



(a) CC1π^+ TPC

(b) CC1π^+ ME

(c) CC1π^+ isoFGD

(d) CS1

(e) CS2

(f) CS3

Figure 7.4: Relative error values for the time of flight resolution systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).

Systematic	Relative Uncertainty (FGD1) %					
	TPCπ^+	MEπ^+	isoFGDπ^+	CS1	CS2	CS3
B-Field Distortions	0.006	0.009	0.006	0.018	0.020	0.007
ECal PID	2.079	1.812	1.753	3.275	4.856	3.127
ECal π^0 Veto Pileup	0.033	0.033	0.033	0.033	0.683	0.033
FGD Hybrid Tracking Eff.	0.206	0.051	2.693	0.257	0.275	2.427
FGD PID	0.011	0.054	0.426	0.042	0.024	0.460
FGD Pion Mom ME Resol	0.000	2.114	0.405	0.000	0.000	0.000
FGD2 Backward Migration	0.000	0.000	0.000	0.000	0.000	0.000
Michel Electron Eff.	0.380	1.215	0.513	0.106	0.355	0.044
Momentum Range Resol.	0.000	0.000	0.506	0.025	0.060	0.231
OOFV Background	0.101	1.503	0.395	0.183	0.209	0.107
Pile-Up	0.336	0.350	0.335	0.348	0.345	0.333
Pion Secondary Interactions	3.373	2.938	3.849	6.623	3.514	0.858
Proton Secondary Interactions	0.660	0.748	0.114	0.740	0.569	2.695
Sand Background	0.003	0.199	0.000	0.025	0.003	0.004
TPC–ECal Matching	0.985	0.558	0.529	1.342	1.131	0.694
TPC–FGD Matching	0.280	0.088	0.128	0.398	0.405	0.026
TPC Charge ID Eff.	0.293	0.108	0.147	0.243	0.347	0.130
TPC Cluster Eff.	0.021	0.011	0.011	0.025	0.031	0.011
TPC Momentum Resolution	0.017	0.013	0.013	0.037	0.041	0.006
TPC Momentum Scale	0.025	0.029	0.004	0.027	0.007	0.005
TPC PID	1.138	0.698	0.462	1.496	0.622	0.291
TPC Tracking Eff.	0.904	0.410	0.507	0.562	0.166	0.704
ToF Resolution	0.099	0.097	0.103	0.000	0.000	0.000
Total	3.643	3.383	3.685	5.323	5.493	5.432

Table 7.VII: Average systematic uncertainty values for all FGD1 signal and background samples.

Systematic	Relative Uncertainty (FGD2x) %					
	TPCπ^+	MEπ^+	isoFGDπ^+	CS1	CS2	CS3
B-Field Distortions	0.014	0.028	0.009	0.035	0.014	0.010
ECal PID	1.884	1.650	1.653	4.321	5.799	2.219
ECal π^0 Veto Pileup	0.033	0.033	0.033	0.033	0.758	0.033
FGD Hybrid Tracking Eff.	0.041	0.021	0.747	0.049	0.056	0.842
FGD PID	0.018	0.032	0.630	0.096	0.026	0.417
FGD Pion Mom ME Resol	0.000	2.307	0.167	0.000	0.000	0.000
FGD2 Backward Migration	3.486	6.460	3.554	0.302	1.669	0.356
Michel Electron Eff.	0.455	3.066	0.397	0.197	0.439	0.183
Momentum Range Resol.	0.000	0.000	5.719	0.008	0.261	11.766
OOFV Background	0.102	0.864	0.310	0.103	0.160	0.250
Pile-Up	0.356	0.369	0.336	0.371	0.355	0.337
Pion Secondary Interactions	2.280	1.844	1.198	3.321	1.833	1.030
Proton Secondary Interactions	0.754	0.749	0.600	0.888	0.689	3.404
Sand Background	0.018	0.044	0.000	0.000	0.005	0.008
TPC–ECal Matching	0.818	0.542	0.394	1.354	1.196	0.314
TPC–FGD Matching	0.271	0.155	0.195	0.320	0.396	0.039
TPC Charge ID Eff.	0.490	0.292	0.238	0.565	0.257	0.218
TPC Cluster Eff.	0.020	0.010	0.011	0.029	0.030	0.009
TPC Momentum Resolution	0.017	0.017	0.014	0.072	0.047	0.010
TPC Momentum Scale	0.024	0.014	0.011	0.033	0.015	0.010
TPC PID	1.249	1.074	0.491	1.883	0.692	0.372
TPC Tracking Eff.	1.554	0.776	0.799	1.654	1.213	0.762
ToF Resolution	0.483	0.449	0.536	0.000	0.000	0.000
Total	5.148	6.973	3.914	6.363	6.441	4.608

Table 7.VIII: Average systematic uncertainty values for all FGD2x signal and background samples.

Systematic	Relative Uncertainty (FGD2y) %					
	TPCπ^+	MEπ^+	isoFGDπ^+	CS1	CS2	CS3
B-Field Distortions	0.030	0.020	0.038	0.037	0.051	0.011
ECal PID	2.013	1.535	1.551	4.592	5.642	2.437
ECal π^0 Veto Pileup	0.033	0.033	0.033	0.033	0.666	0.033
FGD Hybrid Tracking Eff.	0.044	0.033	0.599	0.055	0.059	0.771
FGD PID	0.043	0.056	0.946	0.159	0.054	0.407
FGD Pion Mom ME Resol	0.000	2.563	0.225	0.000	0.000	0.000
FGD2 Backward Migration	3.173	22.683	10.476	8.207	4.227	3.012
Michel Electron Eff.	0.456	2.944	0.296	0.245	0.423	0.165
Momentum Range Resol.	0.000	0.000	5.219	0.037	0.440	11.956
OOFV Background	0.231	1.250	0.709	0.254	0.378	0.338
Pile-Up	0.368	0.378	0.347	0.364	0.370	0.360
Pion Secondary Interactions	2.647	2.036	1.205	3.276	2.453	1.001
Proton Secondary Interactions	0.764	0.772	0.641	0.877	0.721	3.558
Sand Background	0.000	0.057	0.000	0.000	0.000	0.000
TPC–ECal Matching	0.845	0.498	0.380	1.433	1.177	0.302
TPC–FGD Matching	0.597	0.194	0.474	0.981	0.767	0.084
TPC Charge ID Eff.	0.477	0.156	0.197	0.662	0.235	0.195
TPC Cluster Eff.	0.020	0.010	0.006	0.027	0.030	0.008
TPC Momentum Resolution	0.041	0.034	0.056	0.099	0.113	0.023
TPC Momentum Scale	0.029	0.017	0.033	0.044	0.012	0.011
TPC PID	1.176	1.093	0.739	1.898	0.834	0.416
TPC Tracking Eff.	1.329	0.771	0.667	1.487	0.987	0.631
ToF Resolution	0.419	0.356	0.311	0.000	0.000	0.000
Total	5.086	20.501	9.448	10.433	7.470	5.809

Table 7.IX: Average systematic uncertainty values for all FGD2y signal and background samples.

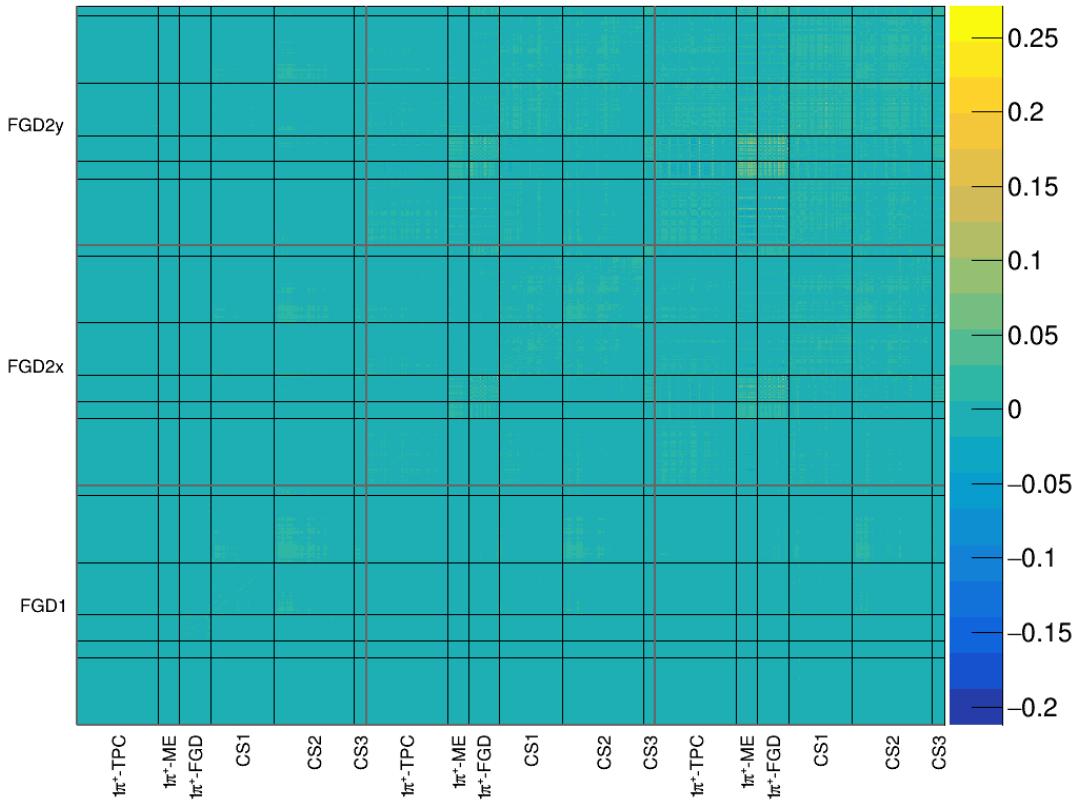


Figure 7.5: Detector systematic uncertainty pre-fit covariance matrix for all 18 samples, with signal and control samples shown on the x-axis and FGD layer sample on the y-axis.

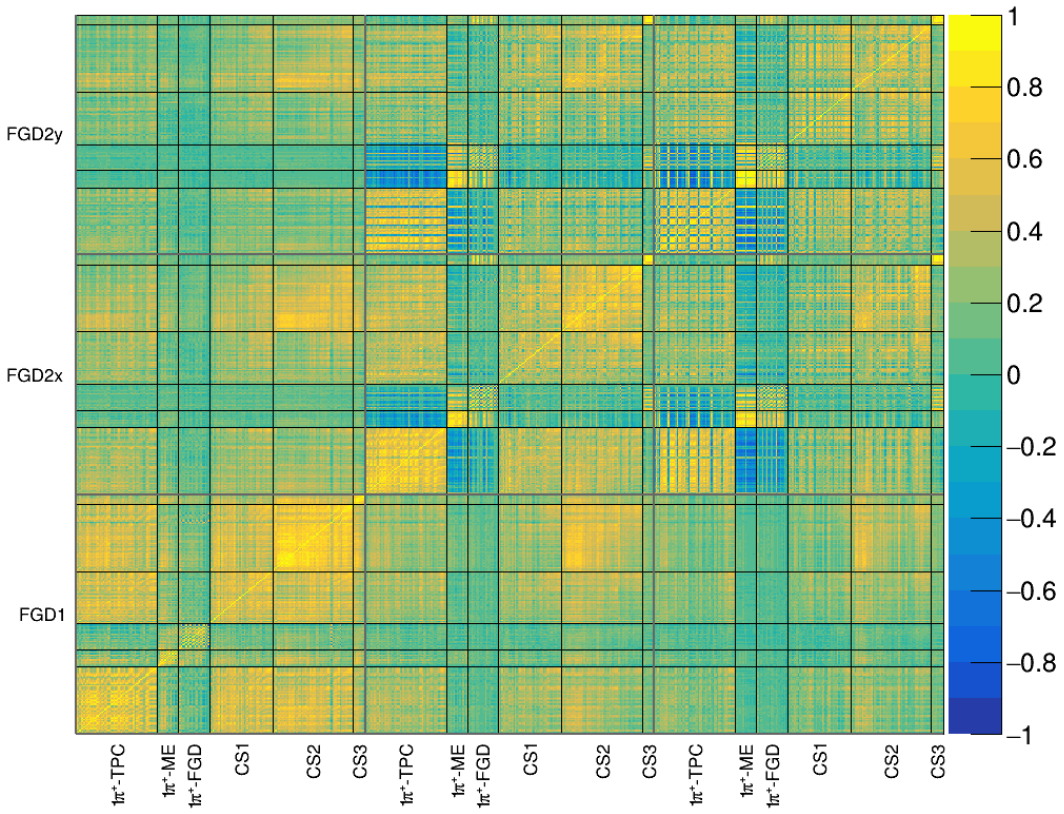


Figure 7.6: Detector systematic uncertainty pre-fit correlation matrix for all 18 samples, with signal and control samples shown on the x-axis and FGD layer sample on the y-axis.

7.1.2 Cross-Section Model Systematic Uncertainties

The Monte Carlo prediction used in the signal and background selection for this analysis is produced using the `NEUT` neutrino interaction generator, using a set of cross-section parameters fixed at certain values. In order to allow the cross-section likelihood fit additional freedom, specific parameter dials are used to vary the underlying parameters, and evaluate the effect on the results of doing so. The cross-section dials used are presented in Table 7.X, where the nominal values are set using recent recommendations from the T2K Neutrino Interaction Working Group (NIWG) [226]. Each of these dials controls a specific parameter or reaction mechanism, allowing a single dial to adjust the number of relevant events whilst leaving unrelated event rates unaffected. The obtained covariance and correlation matrices for the parameters used in this analysis are presented in Figures 7.7 and 7.8, respectively.

To calculate event rate alterations due to changes in the underlying parameter values, the `T2KReWeight` package is used to generate weights to be applied to events. To avoid having to run `T2KReWeight` for every event with every possible combination of parameters, the weights generated by `T2KReWeight` are approximated by sets of response functions, known as *splines*. In order to allow dials the freedom to change event rates within a given true bin (template parameter in the fit), splines for each dial are binned in reduced definitions of molecular target, true reaction, and reconstructed sample (with no distinction between FGD layer), as well as the true binning scheme. Full details on both the template parameter and model systematics binning schemes are given in Section 7.2.2.

Parameter Dial	Dial Type	Nominal Value [GeV]
MACCQE	Shape	1.03 ± 0.06
CA5	Shape	0.96 ± 0.15
MARES	Shape	1.07 ± 0.15
I12RES	Shape	0.96 ± 0.4
CCCOH	Norm	1 ± 1
CCMULTIPI	Norm	1 ± 0.5
CCDIS_NORM	Norm	1 ± 0.035
DIS_BY_CORR	Shape	0 ± 0.5
MULTIPI_BY_CORR	Shape	0 ± 0.5
MULTIPI_XSEC_AGKY	Shape	0 ± 0.5
FSI_PI_ABS	Shape	1.404 ± 0.432
FSI_PI_PROD	Shape	1.002 ± 1.101
FSI_CEX_LO	Shape	0.697 ± 0.305
FSI_CEX_HI	Shape	1.800 ± 0.288
FSI_INEL_LO	Shape	1.069 ± 0.313
FSI_INEL_HI	Shape	1.824 ± 0.859

Table 7.X: Cross-section model systematic parameters used in the analysis, given with their nominal dial values and associated errors.

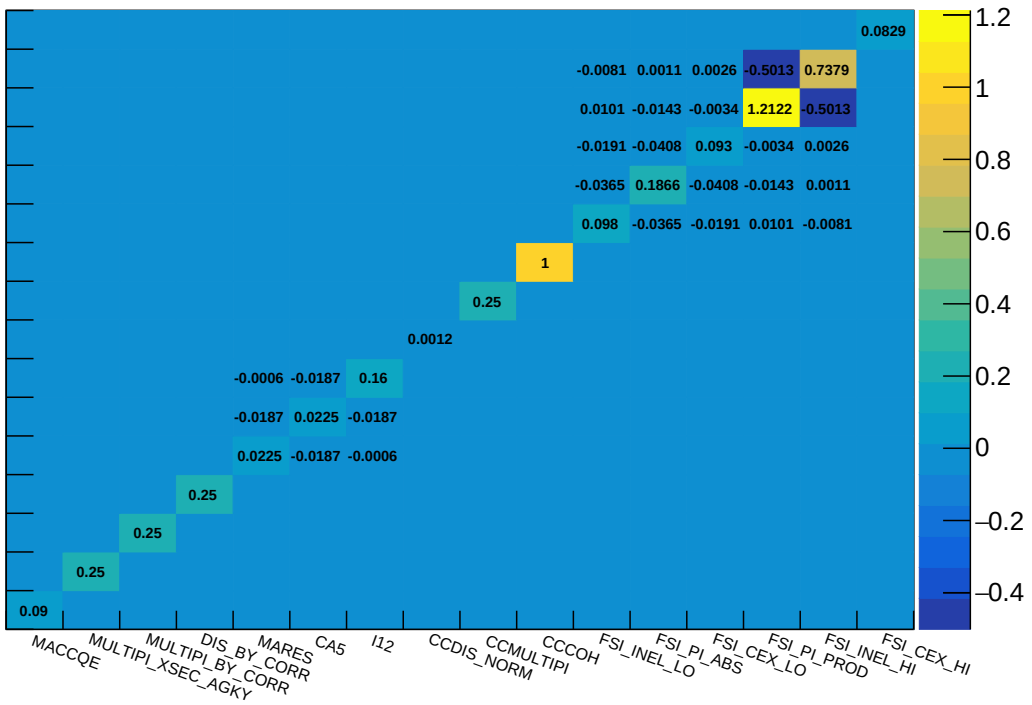


Figure 7.7: Cross-section parameter uncertainty covariance matrix, for the list of dials used in this analysis, given in Table 7.X.

7.1.3 Flux Systematic Uncertainties

In order to calculate the cross section in a model independent way, the calculation in Equation (5.1) is performed using the integrated flux. To take into account the uncertainty on the predicted flux, the covariance matrix is included in the χ^2_{syst} term in Equation (5.16). This covariance matrix is provided by the T2K beam group, and uses the 2020 flux release (`13av7.1`) [227]. The covariance matrix is binned in true neutrino energy, and is produced in two separate forms – one with the standard binning scheme used by the BANFF oscillation analyses, and another with a slightly finer binning.

To make this usable by the analysis however, the matrix and binning scheme must be slightly modified. Figure 7.9 shows the true neutrino energy distribution for the selected events. Whilst the region below 400 MeV is generally low in statistics, the fact that the 0–100 MeV bin contains zero events will be a particular issue. Because of this, the parameter assigned to this bin will do nothing, in which case the likelihood curve will be completely flat. This in turn means the second derivative of the likelihood function will be zero, returning an invalid Hessian matrix and thus a non-convergence in the fit. Due to the lack of any events in this bin, it was deemed acceptable to completely remove the parameter from the fit. The covariance matrix therefore becomes a 19×19 matrix, and the neutrino energy binning starts from 100 MeV. The resultant matrix is shown in Figure 7.10. The energy ranges corresponding to each bin number are given for completeness in Table 7.XI. Here we use the fine binned version of the matrix, which has slightly more bins than the version

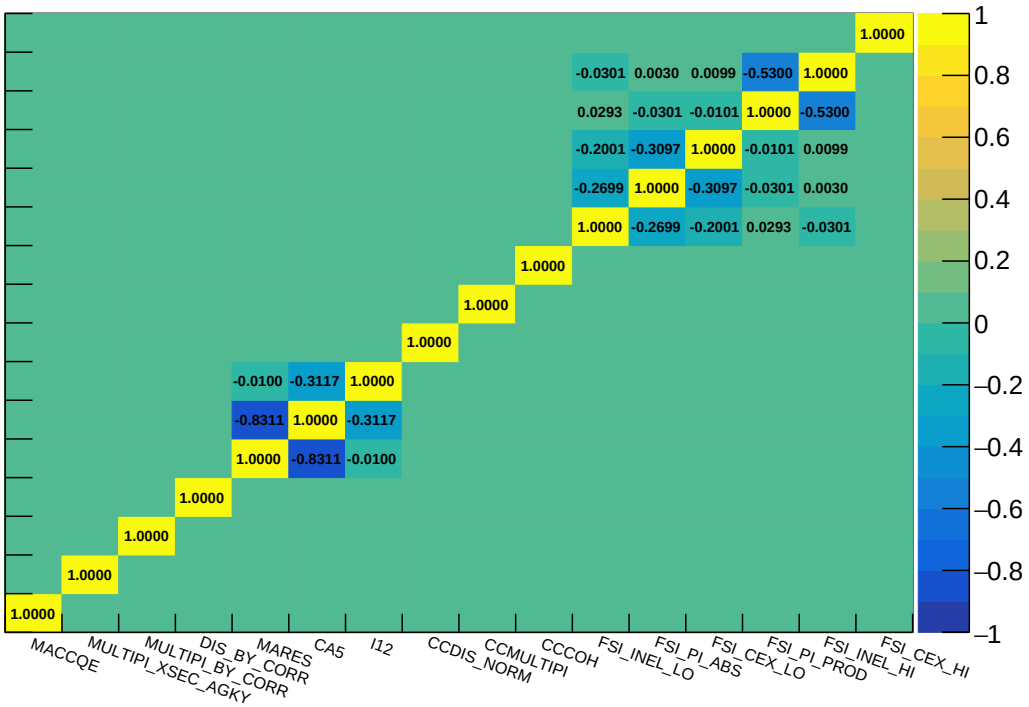


Figure 7.8: Cross-section parameter correlation matrix, calculated from the covariance matrix in Figure 7.7, for the cross-section dials listed in Table 7.X.

used in the oscillation analysis.

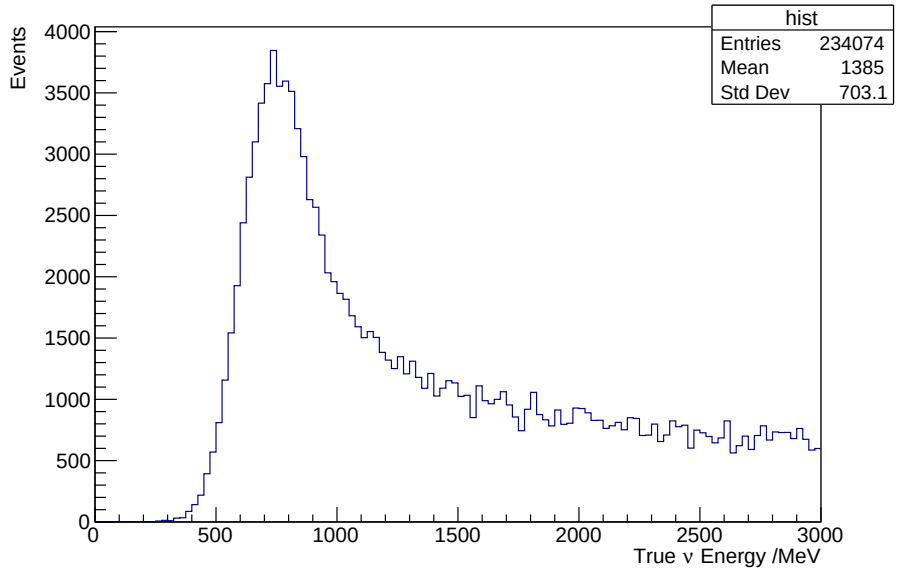


Figure 7.9: True neutrino energy distribution for the selected events from all samples combined. The peak is seen to be slightly higher than the T2K off-axis flux peak, due to the energy required to produce a pion candidate. The 0–100 MeV region contains zero events, and very low statistics are observed up until ∼400 MeV.

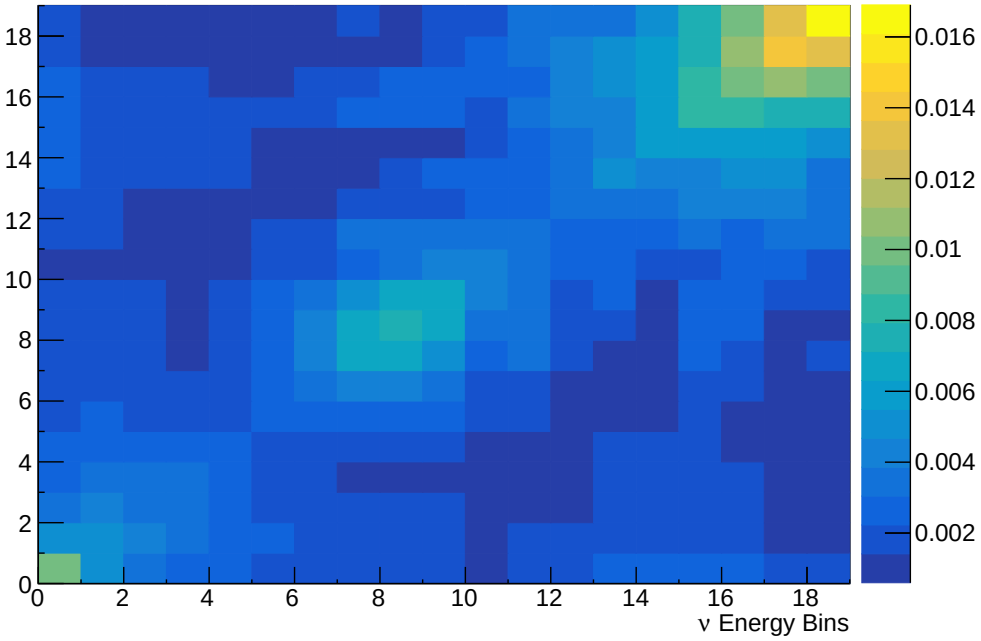


Figure 7.10: Neutrino flux covariance matrix used in the fit, given for ND280 FHC ν_μ runs 1–9. Neutrino energy binning is given in Table 7.XI.

Bin Index	ν Energy [GeV]
0	0.1 – 0.2
1	0.2 – 0.3
2	0.3 – 0.4
3	0.4 – 0.5
4	0.5 – 0.6
5	0.6 – 0.7
6	0.7 – 0.8
7	0.8 – 1.0
8	1.0 – 1.2
9	1.2 – 1.5
10	1.5 – 2.0
11	2.0 – 2.5
12	2.5 – 3.0
13	3.0 – 3.5
14	3.5 – 4.0
15	4.0 – 5.0
16	5.0 – 7.0
17	7.0 – 10.0
18	10.0 – 30.0

Table 7.XI: Binning in neutrino energy used for the flux covariance matrix given in Figure 7.10.

7.1.4 Number of Targets

In order to extract the cross section on a single target nucleon of either hydrocarbon or water, Equation (5.1) involves dividing by the total number of target nucleons within the fiducial volume of the two FGDs considered. The number of target nucleons of element α is calculated as

$$N_t^\alpha = N_A \sum_i^{\text{modules}} m_\alpha^i \frac{N_\alpha}{M_\alpha^A}, \tag{7.12}$$

where i runs over the individual XY scintillator and water modules, m_α^i is the mass of the target element in the fiducial volume of the i^{th} module, $N_A = 6.022 \times 10^{23}$ is the Avogadro constant, and N_t^α is calculated separately for $\alpha =$ C or O. M_α^A is the relative atomic mass of the element, and N_α is the number of nucleons in the element α. In order to account for the presence of additional neutrons from less common isotopes of the two elements, the number of nucleons N_α is set to the weighted average across the possible isotopes, and thus cancels with the relative atomic mass M_α^A. The mass of the i^{th} module can be represented as

$$m_\alpha^i = \rho_\alpha^i V_{FV} = \rho_{\alpha,\text{a}}^i \Delta X_{FV} \Delta Y_{FV} \tag{7.13}$$

where $\rho_{\alpha,\text{a}}^i$ is the areal density of the element α in the i^{th} module, and ΔX_{FV} and ΔY_{FV} are the length of the FGD fiducial volume in the x and y directions, respectively. Due to the symmetry of the detector around the z direction, $\Delta X_{FV} = \Delta Y_{FV} = 174.902$ cm. This finally enables us to rewrite the number of target nucleons as

$$N_t^\alpha = N_A \Delta X_{FV} \Delta Y_{FV} \left(n^{\text{scint}} \rho_{\alpha,\text{a}}^{\text{scint}} + n^{\text{water}} \rho_{\alpha,\text{a}}^{\text{water}} \right), \tag{7.14}$$

where n^{scint} and n^{water} are the total number of XY scintillator and water modules, respectively. As shown in Figure 6.36, the FGD fiducial volume definition involves removing the first full module from FGD1, and the first X layer from FGD2, which in total corresponds to 20.5 XY scintillator modules. Full details of the scintillator module components and how density measurements were made can be found in T2K-TN-091 [228]. FGD2 features 6 water modules in total, all of which are included within the FV. However, several different designs of the water panels exist with slightly varied density measurements, and thus we have to implement them separately for an accurate calculation of target nucleon number. The designs of the different panels, along with the density measurements for each element, are detailed in T2K-TN-198 [229]. A summary of the different panel designs and their positions in FGD2 is presented in Table 7.XII. The areal density measurements of constituent elements and their associated uncertainties are given in Table 7.XIII, where the density measurements for the water panels are made prior to filling with water. Finally, Table 7.XIV gives the total density for oxygen and hydrogen in each water panel within FGD2 after being filled, where the values of all other elements are understood to remain the same.

Using Equation (7.14) and the summarised measurements of the areal density for each

Panel Number	Position in FGD	Design
0	1	Old-style
1	2	Old-style
2	3	Old-style
3	4	Old-style
4	5	Old-style
5	Spare	Old-style
6	6	New-style, thick
7	Spare	New-style, thin
8	Spare	New-style, thick

Table 7.XII: Water module panel designs.

Element	$\rho_{\alpha,\mathrm{a}}^{\mathrm{scint}}$	$\rho_{\alpha,\mathrm{a}}^{\mathrm{water,\ old}}$	$\rho_{\alpha,\mathrm{a}}^{\mathrm{water,\ thick}}$
C	1.849 ± 0.0092	0.422 ± 0.007	0.415 ± 0.008
O	0.0794 ± 0.0048	0.093 ± 0.004	0.091 ± 0.004
H	0.1579 ± 0.0021	0.044 ± 0.001	0.044 ± 0.001
Ti	0.0355 ± 0.0059	–	–
Si	0.0218 ± 0.0043	0.011 ± 0.001	0.011 ± 0.001
N	0.0031 ± 0.0012	–	–
Mg	–	0.007 ± 0.001	0.007 ± 0.001

Table 7.XIII: Elemental composition of the XY and water modules, in g/cm^2, summarised from T2K-TN-091 [228] and T2K-TN-198 [229]. Values are given for empty water panels.

Panel	O	H
0	2.060 ± 0.005	0.292 ± 0.001
1	2.070 ± 0.005	0.294 ± 0.001
2	2.066 ± 0.005	0.293 ± 0.001
3	2.077 ± 0.005	0.295 ± 0.001
4	2.065 ± 0.005	0.293 ± 0.001
6	2.059 ± 0.005	0.292 ± 0.001

Table 7.XIV: Elemental composition for the filled water panels, in g/cm^2, for oxygen and hydrogen, summarised from T2K-TN-198 [229]. All other elemental compositions are the same as prior to filling with water.

constituent element, the number of carbon and oxygen targets can be calculated. The calculation for carbon is simple, as the carbon composition remains the same, regardless of whether the water panels are filled or not. The calculation for oxygen, however, is comparatively more complicated. This is due to damage to the first water layer (panel 0), which was sustained prior to ND280 run 7 (described in Section 3.2.2.3). This resulted in the water panel being drained, and thus in run 8 we treat panel 0 as empty of passive water, making the assumption that the panel was fully drained. In order to account for this, the number of oxygen targets is calculated separately for runs 2–4 and run 8, and the average

number weighted by the collected POT in each state is used,

$$N_t^O = \frac{N_{t,234}^O N_{234}^{\mathrm{POT}} + N_{t,8}^O N_8^{\mathrm{POT}}}{N^{\mathrm{POT}}}, \tag{7.15}$$

where N^{POT} is simply the total collected data POT for all FHC runs. These run-by-run POT values are given in Table 5.I.

In order to calculate the errors on these values, toys are thrown to obtain random values of the areal densities of each element, which are correlated by applying the Cholesky decomposition of the areal density covariance matrices. The covariance matrices themselves are calculated from the errors on the areal density measurements in Tables 7.XIII and 7.XIV, and the supplied correlation matrices, shown in Tables 7.XV to 7.XVII for the scintillator modules, old water panels and new thick water panel, respectively. 10^6 toys are thrown to build the distributions given in Figures 7.11 and 7.12, where the mean is taken as the number of targets, and the RMS value as the uncertainty. Doing this, applying the POT weighting for oxygen from Equation (7.15), and then making the assumption that the number of C (O) atoms is equal to the number of CH (H_2O) molecules, gives:

$$N_t^{\mathrm{CH}} = (7.448 \pm 0.043) \times 10^{29} \rightarrow 0.6\% \text{ error}$$
$$N_t^{\mathrm{H_2O}} = (2.404 \pm 0.032) \times 10^{29} \rightarrow 1.3\% \text{ error}$$

	C	O	H	Ti	Si	N
C	1.000	0.210	0.587	−0.193	−0.161	0.226
O		1.000	0.115	0.830	0.068	−0.033
H			1.000	−0.121	−0.879	0.875
Ti				1.000	0.074	−0.097
Si					1.000	−0.972
N						1.000

Table 7.XV: Correlation coefficients for the areal density of elements in the XY scintillator modules, obtained from T2K-TN-091 [228]. Symmetrical bottom half of the matrix is omitted.

	C	O	H	Mg	Si
C	1.000	0.791	0.976	0.748	0.748
O		1.000	0.697	0.988	0.988
H			1.000	0.678	0.678
Mg				1.000	1.000
Si					1.000

Table 7.XVI: Correlation coefficients for the areal density of elements in the old-style water modules, obtained from T2K-TN-198 [229]. Symmetrical bottom half of the matrix is omitted.

	C	O	H	Mg	Si
C	1.000	0.792	0.957	0.649	0.649
O		1.000	0.720	0.953	0.953
H			1.000	0.647	0.647
Mg				1.000	1.000
Si					1.000

Table 7.XVII: Correlation coefficients for the areal density of elements in the new-style thick water modules, obtained from T2K-TN-198 [229]. Symmetrical bottom half of the matrix is omitted.

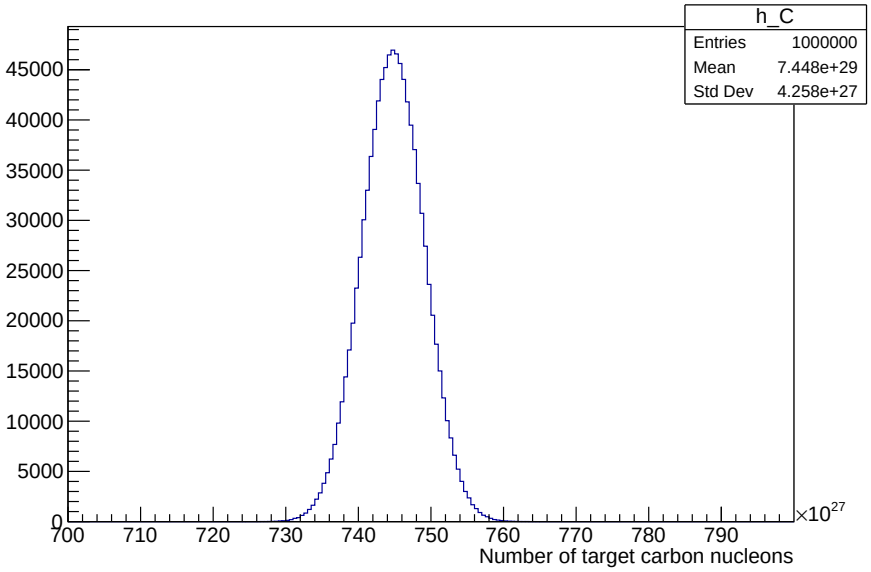
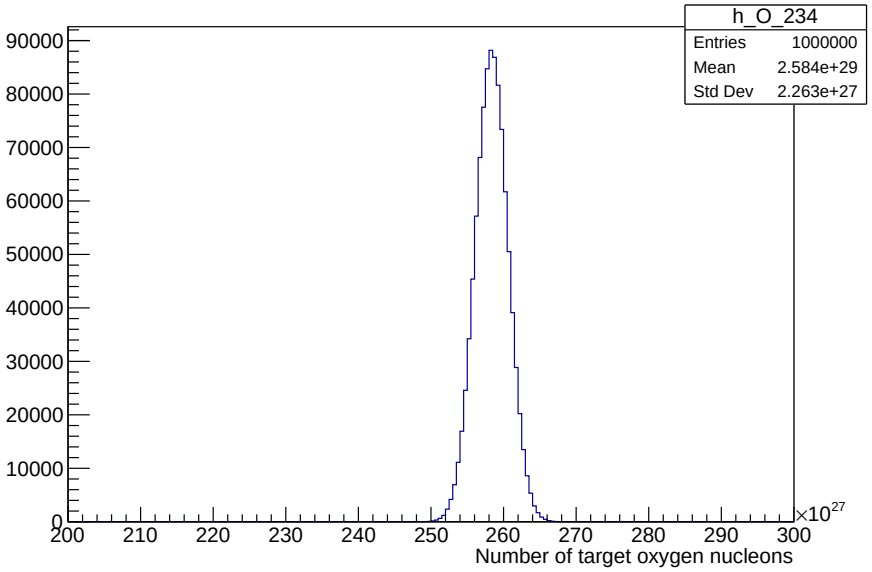
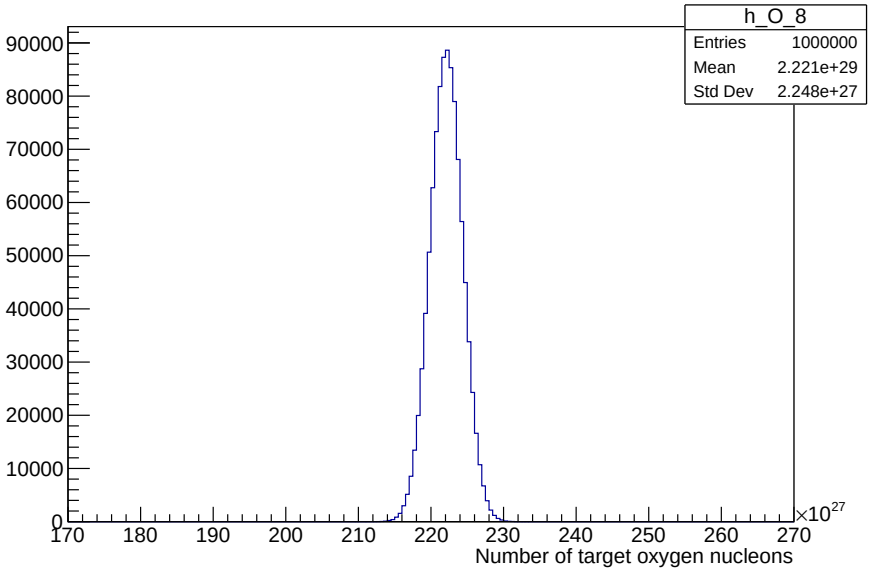


Figure 7.11: Number of carbon target nucleons in the total fiducial volume of both FGDs, generated by throwing areal density values with 10^6 toys.



(a) Runs 2–4



(b) Run 8

Figure 7.12: Number of oxygen target nucleons in the total fiducial volume of both FGDs, for (a) runs 2–4 and (b) run 8, generated by throwing areal density values with 10^6 toys.

7.2 Efficiencies and Binning

The following sections describe the studies used to inform the necessary phase space restrictions to the measurement, and the template binning scheme in which the measurement will be extracted. Reconstructed sample and cross-section spline binning schemes are also discussed.

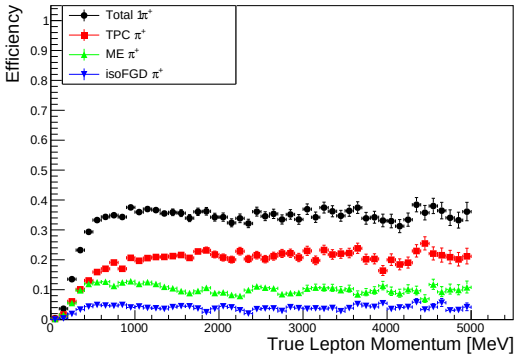
7.2.1 Phase Space Constraints and Efficiency

As discussed, there will always be some areas of our desired kinematic phase space that it is difficult to perform measurements in, due to limitations imposed by the inefficiency of the detector. In some cases, this will mean that measurements in certain regions of phase space cannot be made, as performing efficiency correction in regions with a very low or rapidly changing efficiency relies heavily on assumptions made by the generator model used. Before attempting to use the developed selection to perform a fit to the data, it is important to properly understand where these regions are, and what constraints they will put on the analysis.

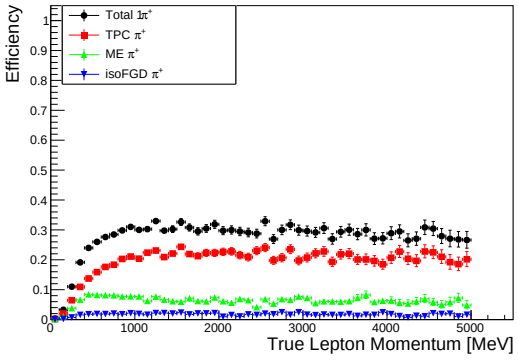
Figures 7.13 and 7.14 show the efficiency of the developed selections as a function of the true muon and pion kinematic distributions of interest, respectively. In each distribution, the efficiency is plotted for the individual signal samples (TPC, ME and isoFGD π^+), along with the total CC1π^+ signal for which the cross section will be measured. In each of these efficiency distributions, no restrictions are placed on the other kinematics of interest.

In both Figures 7.13a and 7.13b it can be seen that the efficiency as a function of true lepton momentum is generally flat for the majority of the range shown, other than the very low momentum values. This is due to the minimum threshold energy required for a muon candidate to be seen by the detector. The switch on in efficiency occurs in the 100–200 MeV region in both FGDs, and by about 400 MeV the efficiency has reached its maximal amount, beyond which it is mostly flat bar statistical fluctuations. This behaviour holds for each of the three sub-samples, where the maximal efficiency come from TPC pions, followed by ME and finally isoFGD pions. In order to avoid the very low momentum region where efficiency is practically zero and the subsequent sharp turn on, we choose to place a phase space constraint on the muon momentum, and do not attempt to measure events with a momentum of less than 200 MeV. Although this doesn't completely exclude the region of changing efficiency, it is also important to balance this requirement with our desire for statistics; in Figure 6.16 it was shown that the peak of the reconstructed momentum distribution occurs around 300–500 MeV, so a higher phase space constraint would remove a large number of events. The binning scheme for the analysis is later chosen in a way that should account for this.

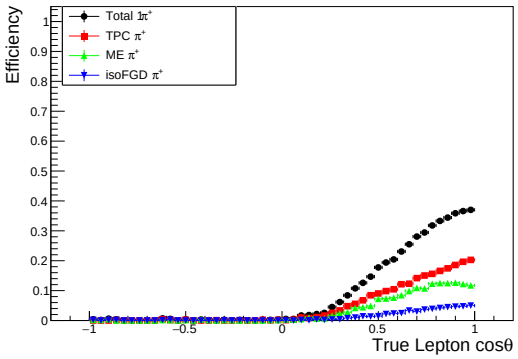
Figures 7.13c and 7.13d give the selection efficiency as a function of the true lepton angle to the neutrino direction. Here the turn on in efficiency is seen to occur at a value of true $\cos\theta_\mu$ of around 0.3, and below this the efficiency is practically zero. This is also observed in Figure 6.16, where there are very few events in the lower region. For backward going tracks ($\cos\theta < -0.3$), the near zero efficiency comes from the detector's inability to reconstruct



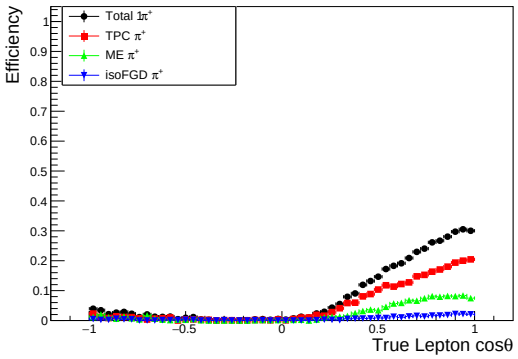
(a) True Lepton Momentum (FGD1)



(b) True Lepton Momentum (FGD2)



(c) True Muon Cosine (FGD1)

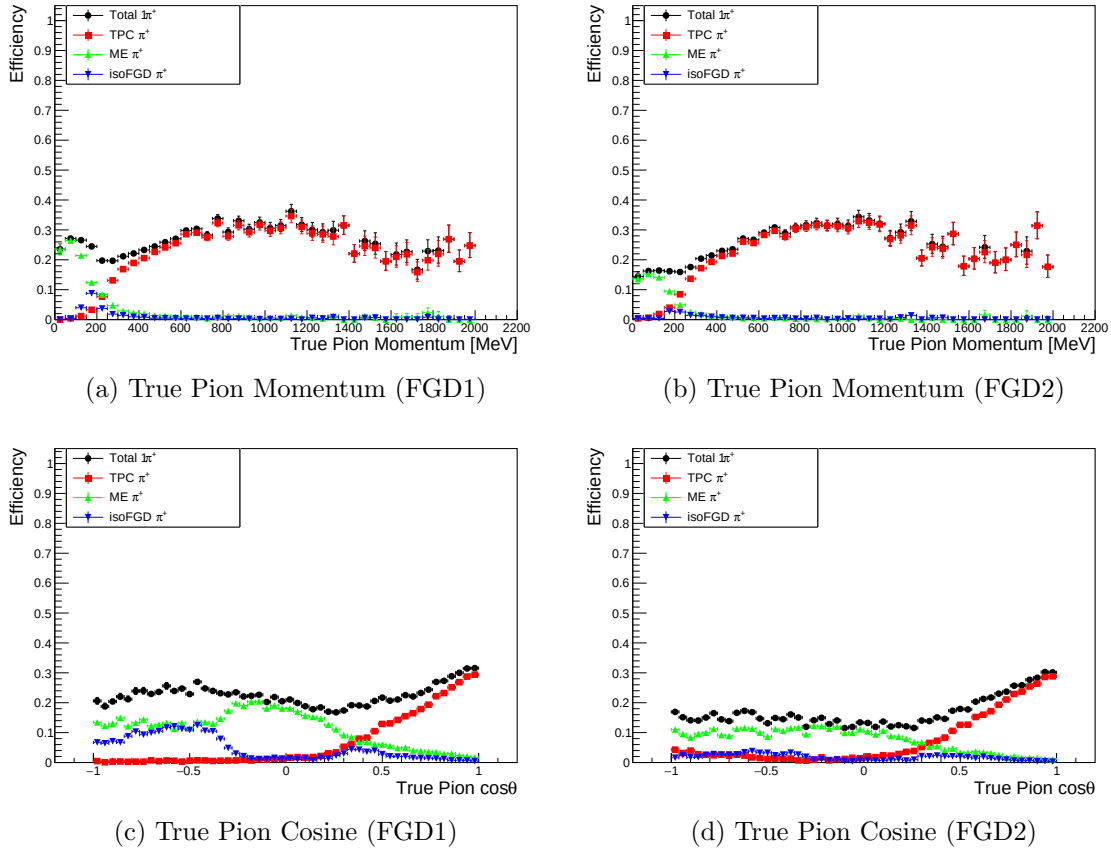


(d) True Muon Cosine (FGD2)

Figure 7.13: CC1π^+ selection efficiency in true muon kinematics for FGD1 and FGD2, as a function of (a)-(b) true lepton momentum, and (c)-(d) true lepton $\cos\theta$. The total efficiency for CC1π^+ events is plotted in black, which is a sum of the individual selection efficiencies for the three different signal samples: TPC, ME or isoFGD π^+. Error bars are statistical.

tracks in this region correctly. The high-angle region has near-zero efficiency because these events are those where the candidate muon generally doesn't enter a TPC, instead going up or down directly into the barrel ECal. Since this selection requires a non-zero number of good-quality TPC tracks (Sections 6.1.2 and 6.1.3), selection efficiency within this high-angle region is naturally expected to be almost zero. Past T2K work has performed data–MC comparisons for the so-called '4π' full angular acceptance of CC-inclusive events [225], with upcoming analyses again working on reconstruction in the backward and high-angle regions. However, since the majority of CC1π^+ interactions feature forward-going muons, and the backward and high-angle regions are relatively impure due to high OOFV backgrounds, this analysis only focuses on forward-going muon events. For that reason, an additional phase space constraint is added – this analysis only attempts to measure the cross section in the region $\cos\theta_\mu > 0.3$.

Arguably, the more interesting efficiency plots are those for the pion kinematics, as they show how the three different pion sub-samples give access to very different regions of true pion phase space. In Figures 7.14a and 7.14b, the pion efficiency can be seen to be



(a) True Pion Momentum (FGD1)

(b) True Pion Momentum (FGD2)

(c) True Pion Cosine (FGD1)

(d) True Pion Cosine (FGD2)

Figure 7.14: CC1π^+ selection efficiency in true pion kinematics for FGD1 and FGD2, as a function of (a)-(b) true pion momentum, and (c)-(d) true pion $\cos\theta$. The total efficiency for CC1π^+ events is plotted in black, which is a sum of the individual selection efficiencies for the three different signal samples: TPC, ME or isoFGD π^+. Error bars are statistical.

completely dominated by the TPC π^+ sample, down to around 200 MeV. Although there is a small underlying peak here from the isoFGD π^+, this is still generally low due to this being the lowest statistics sample of the three. Below 200 MeV, the efficiency is entirely driven by the Michel electron sample, showing the importance of including such events if we wish to access these regions of phase space. At a true momentum of around 1700 MeV in both FGDs, a drop in efficiency is observed. This is understood to be caused by the crossing energy loss curves of pions and protons, which confuses the TPC PID algorithm. Since there is very little that can be done about this, and because this high momentum region is low in statistics anyway, a further phase space constraint is added here, specifying that we do not consider pions of 1500 MeV or above. This avoids having to place additional bin edges around the high momentum efficiency dip, where this would likely lead to very low population bins.

A similar behaviour is seen in the true pion $\cos\theta$ distributions of Figures 7.14c and 7.14d, where the efficiency in the forward-going region is almost entirely driven by the TPC π^+ sample, and the rest of the parameter space dominated by the ME sample. The main

difference can be observed in the isoFGD efficiency, which is very low down to $\cos\theta_\pi \approx -0.3$, but in FGD1 increases below this value, whereas for FGD2 it stays near-zero. The low efficiency in the forward region comes simply from the majority of forward-going pions escaping into the TPC, and so falling into the TPC π^+ sub-sample instead. In the high-angle region, the low efficiency comes from the fact that FGD PID systematics are not well understood here, where particles potentially only cross one or two scintillator bars. Because of this, no attempt is made to reconstruct isoFGD tracks in the high-angle region. Section 6.4.3 concludes that due to the low resolution of the ME pion angular reconstruction, momentum should be prioritised in the binning scheme for the Michel electron samples, which leads us to only bin in momentum for these events. No phase space constraint is applied here, but as the high angle region relies on the ME samples which only have one angular bin, no binning in angle is performed in truth for values of $\cos\theta_\pi$ below 0.5.

Applying these three phase space constraints, the full signal definition for the CC1π^+ measurement becomes events with one negative muon, one positive pion and no other mesons, where

- $p_\mu > 200$ MeV,
- $\cos\theta_\mu > 0.3$,
- $p_\pi < 1500$ MeV.

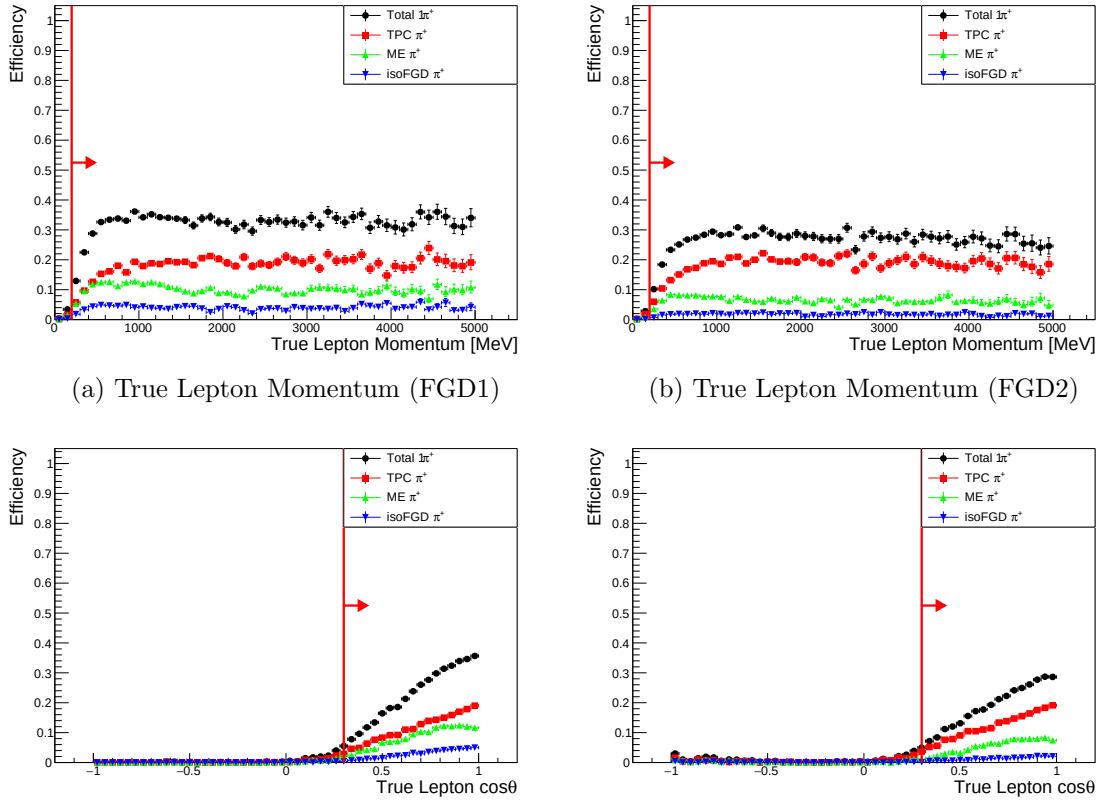
The 1D efficiency in each of the four kinematic variables of interest are presented again in Figures 7.15 and 7.16, where the described phase space constraints are applied to all variables but the one being plotted.

7.2.2 Binning

7.2.2.1 Template Parameter Binning

Deciding on a binning scheme for a cross-section measurement can be particularly difficult, as it involves balancing several, often competing, criteria. Ideally, we would like to bin as finely as possible, so as to better resolve the shape of the cross-section distribution. However, measurements are limited by the number of events counted, and binning schemes should be chosen so that the measurement is not limited by the statistical uncertainty on individual bins. The criteria that should be adhered to when deciding on a binning scheme in which to present a cross-section measurement are as follows:

- Bin edges should be placed such that every bin in the phase space has a sufficient number of events, so that statistical error is not the limiting factor on the measurement. In general, this means having on the order of 100 events in every bin, so that the statistical error is on the order of the other sources of error.
- The size of each bin should be at least as large as the resolution of the detector in that same bin. If bins of smaller size than the detector resolution are used, this can result in large anti-correlations between neighbouring bins.



(a) True Lepton Momentum (FGD1)

(b) True Lepton Momentum (FGD2)

(c) True Lepton Cosine (FGD1)

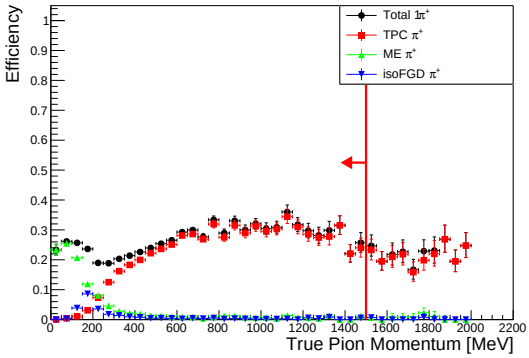
(d) True Lepton Cosine (FGD2)

Figure 7.15: CC1π^+ selection efficiency in true muon kinematics for FGD1 and FGD2, as a function of (a)-(b) true lepton momentum, and (c)-(d) true lepton $\cos\theta$, where phase space constraints are applied to all variable but the one being plotted. The total efficiency for CC1π^+ events is plotted in black, which is a sum of the individual selection efficiencies for the three different signal samples: TPC, ME or isoFGD π^+. Error bars are statistical.

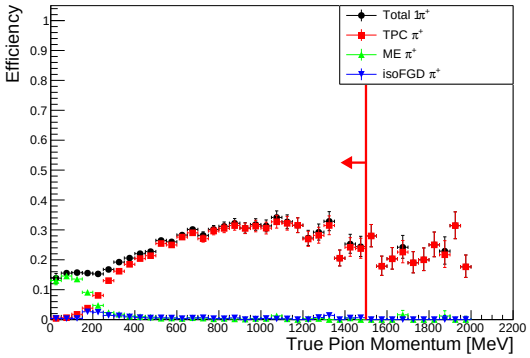
- The transfer matrix between the true and reconstructed distributions should ideally be as diagonal as possible, which can generally be achieved by using similar binning schemes in both true and reconstructed spaces.

- As mentioned previously, selection efficiencies within each bin should ideally be as flat as possible. If the efficiency within a given bin is flat, then the efficiency correction for the events missed due to detector inefficiency is simple. However, if the efficiency within a bin is changing as a function of the underlying truth distributions, then the efficiency correction becomes dependent upon assumptions made by the generator model used.

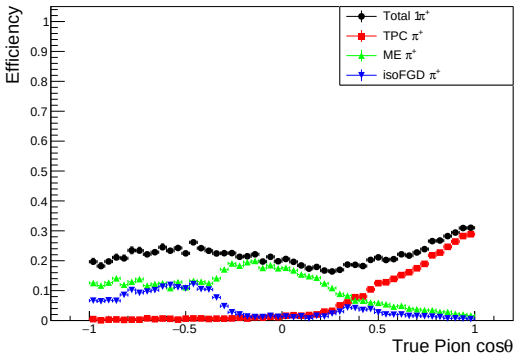
Although not strictly a condition required for a binning scheme, when designing one it is also important to keep in mind exactly what we want to measure. For this analysis in particular, where a measurement is made in terms of four kinematic variables, certain variables of interest should be prioritised in order to account for the finite statistics available.



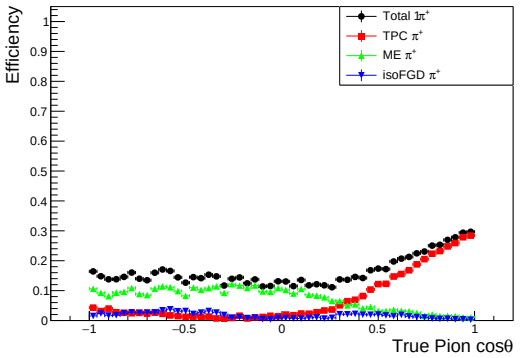
(a) True Pion Momentum (FGD1)



(b) True Pion Momentum (FGD2)



(c) True Pion Cosine (FGD1)

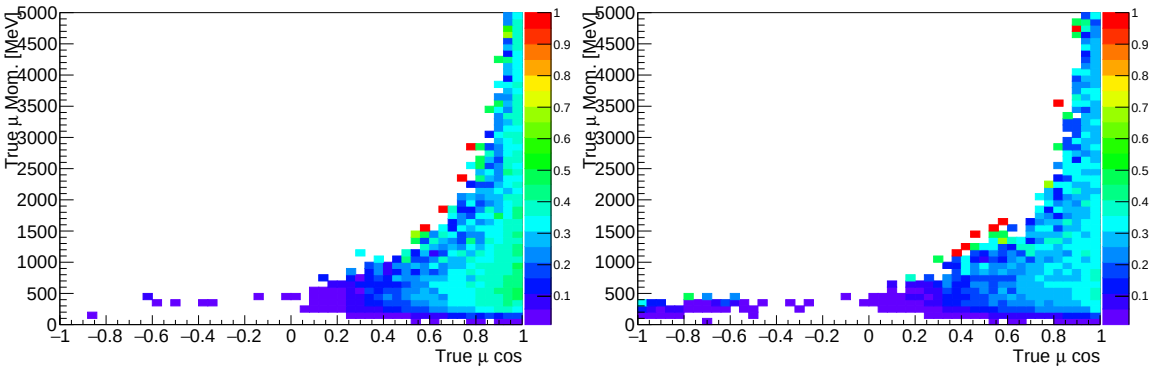


(d) True Pion Cosine (FGD2)

Figure 7.16: CC1π^+ selection efficiency in true kinematics for FGD1 and FGD2, as a function of (a)-(b) true pion momentum and (c)-(d) true pion cos θ, where phase space constraints are applied to all variable but the one being plotted. The total efficiency for CC1π^+ events is plotted in black, which is a sum of the individual selection efficiencies for the three different signal samples: TPC, ME or isoFGD π^+. Error bars are statistical.

Multiple measurements of the muon neutrino CC1π^+ differential cross section exist already as a function of muon kinematics, both on water and carbon. The muon kinematics are therefore deemed to be of lower priority than the pion kinematics, for which the number of measurements is more limited. Cross sections in terms of pion kinematics are of particular interest due to the pions' greater sensitivity to final state interactions, along with the opportunity to study the low pion momentum region which falls below the Cherenkov threshold of Super-K. For this reason, the numbers of bins in muon momentum and cosine of the angle to the neutrino direction are limited as much as possible in line with efficiencies, allowing finer binning in pion kinematics, so as to better resolve the cross-section distribution in those dimensions.

To inform the choice of binning scheme in muon kinematics, Figure 7.17 shows the efficiencies in the two-dimensional space of true muon momentum and cosine of theta, for both FGD1 and FGD2. Both of these plots reinforce the need for the restrictions on the muon phase space detailed in Section 7.2.1 , with negligible efficiency observed in the regions



(a) FGD1 (b) FGD2

Figure 7.17: CC1π^+ selection effieciency as a function of true muon momentum and $\cos\theta$, for (a) FGD1 and (b) FGD2. The phase space constraints discussed in Section 7.2.1 are not applied, and integration over all pion kinematics is implicity performed.

of $\cos\theta_\mu < 0.3$ and $p_\mu < 200$ MeV. This is where the first bin edges will be placed, with events below these edges being retained but moved to out-of-phase-space (OOPS) bins. As a completely arbitrary choice, the restriction on the muon angle is applied first, followed by the momentum. Therefore the few events with $\cos\theta_\mu < 0.3$ **and** $p_\mu < 200$ MeV will fall into the first OOPS bin. After that, the true muon kinematic phase space is divided up into six bins, taking into account the criteria detailed above. In particular, the bin edges are chosen such that each one has sufficient events to have at least six bins in pion kinematics. This ensures that the pion resolution within each muon bin should always be as good as, if not better than, that of the muon.

Once the true phase space has been appropriately divided up in terms of muon kinematics, the same exercise is performed for the pion kinematics. Here the bin edges are chosen primarily based on the distribution of the 2D efficiency in pion momentum and angle, for each possible muon kinematic bin, and then bins merged together where necessary to obtain sufficient statistics. The binning study is performed and optimised based on the number of true hydrocarbon or water interactions in the relevant true kinematic bins, as these are the signals we wish to extract the cross section for. In general, the choice of bin edges is based upon the true interactions on water, as this is the statistically limiting signal. The choice of binning can be conveniently represented by a series of 2-dimensional histograms in true pion kinematics, with one for each true muon kinematic bin.

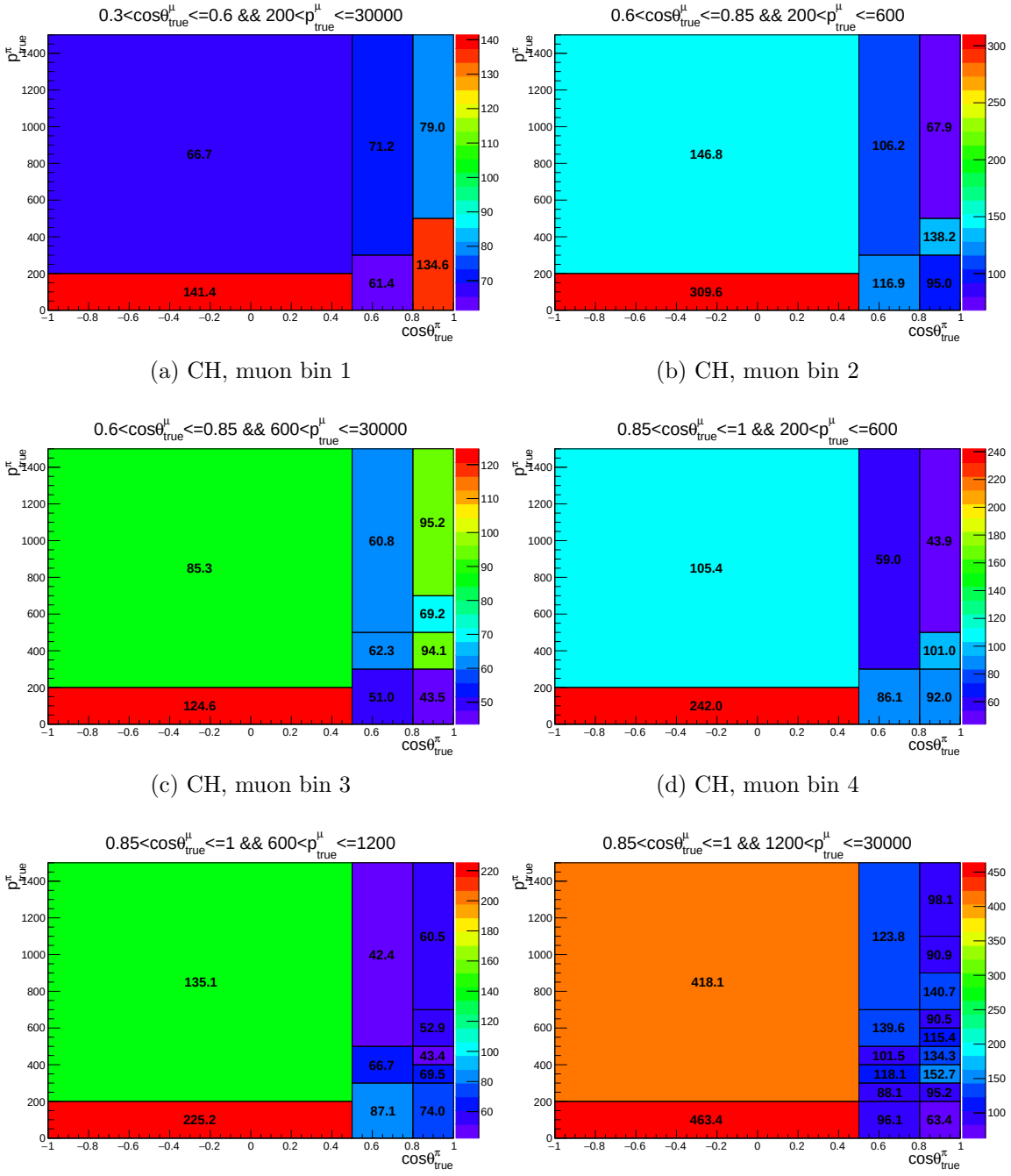
In order to perform efficiency corrections without introducing too much model dependence to the measurement, the cross section will be extracted in fine bins. This of course goes against the first bullet point in the binning criteria above, and will result in high statistical error within each extracted bin. Once extraction is completed, the fine-binned result can be integrated over, in order to report a cross-section measurement in coarser bins but with a reduced statistical error. Instead, the minimum number of required events in a true bin is set to be 10. This is somewhat of an arbitrary choice, and simply serves as

a minimum requirement in order to avoid statistical fluctuations causing any given bin to drop to 0 events. Following a Poissonian distribution

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{7.16}$$

where λ is the expected value and k the number of occurrences, statistical fluctuations in the MC causing a bin with an expected value of $\lambda = 10$ events to drop to 0 should have a probability of 4.5×10^{-5}. Although there are ways of dealing with this, it is best to avoid such cases if possible, and hence this minimum number of events is used in the construction of the pion kinematic binning scheme.

The 2D histograms showing the bin population for the chosen 4D binning scheme are given in Figures 7.18 and 7.19, for true interactions on hydrocarbon and water, respectively. It can be seen from both sets of histograms that as we reach bins with higher values of momentum and $\cos\theta$, more events are present and thus the binning scheme in pion kinematics can be made finer. This is particularly apparent in Figures 7.18e and 7.18f, which despite being binned finely in both angle and momentum always have a sufficient number of events in each bin, due to the true CC1π^+ signal being dominated by forward going TPC pions. Theoretically the binning in the true CH sample could be made even finer, but for simplicity it is preferable for the CH and H_2O signals to use the same binning scheme. The H_2O signal is lower in statistics, with several bins that have $\sim$10 events in when scaled down to data POT. The full binning scheme for both true signals is presented in Table 7.XVIII, and results in 53 signal bins plus 3 OOPS bins per target, for a total of 112 template parameter bins.



(a) CH, muon bin 1

(b) CH, muon bin 2

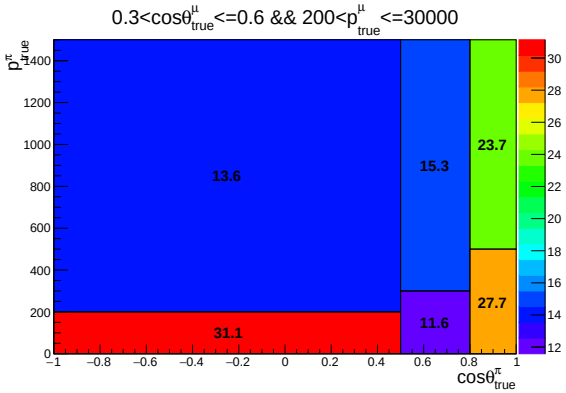
(c) CH, muon bin 3

(d) CH, muon bin 4

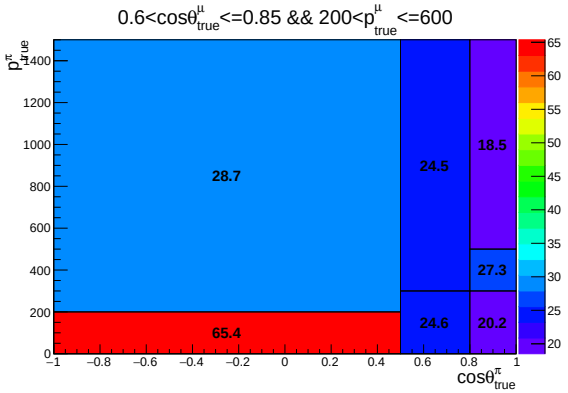
(e) CH, muon bin 5

(f) CH, muon bin 6

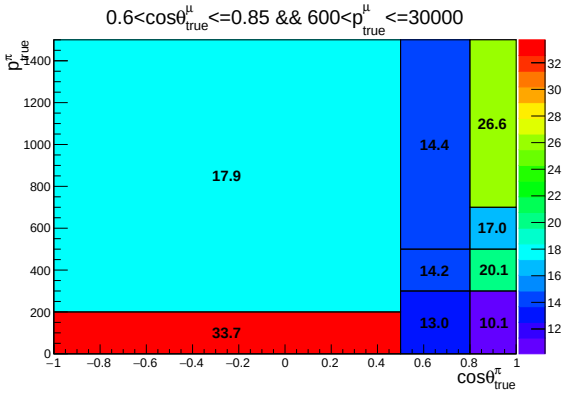
Figure 7.18: 2-dimensional histograms in true pion momentum (y-axis) and $\cos\theta$ (x-axis), for each true muon kinematic bin. Bin content gives the expected number of true events on hydrocarbon for full FHC MC scaled down to data POT.



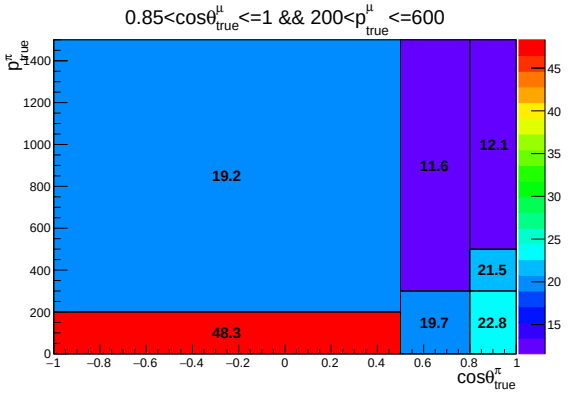
(a) H_2O, muon bin 1



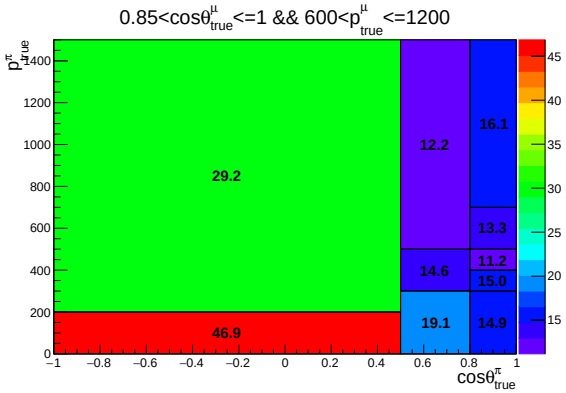
(b) H_2O, muon bin 2



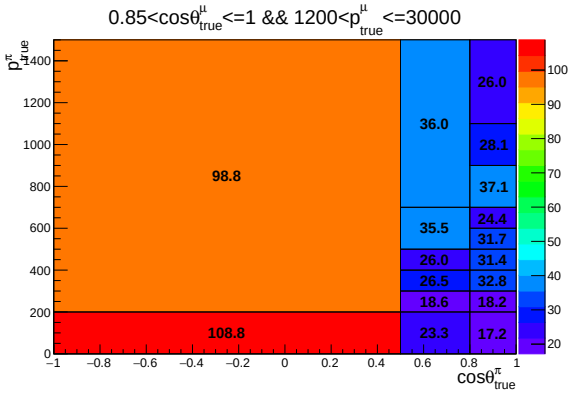
(c) H_2O, muon bin 3



(d) H_2O, muon bin 4



(e) H_2O, muon bin 5



(f) H_2O, muon bin 6

Figure 7.19: 2-dimensional histograms in true pion momentum (y-axis) and $\cos\theta$ (x-axis), for each true muon kinematic bin. Bin content gives the expected number of true events on water for full FHC MC scaled down to data POT.

Bin Number	$\cos\theta_\mu$	p_μ	$\cos\theta_\pi$	p_π
0	–1.0 – 0.3	0 – 30000	–1.0 – 1.0	0 – 30000
1	0.3 – 1.0	0 – 200	–1.0 – 1.0	0 – 30000
2	0.3 – 1.0	200 – 30000	–1.0 – 1.0	1500 – 30000
3	0.3 – 0.6	200 – 30000	–1.0 – 0.5	0 – 200
4	0.3 – 0.6	200 – 30000	–1.0 – 0.5	200 – 1500
5	0.3 – 0.6	200 – 30000	0.5 – 0.8	0 – 300
6	0.3 – 0.6	200 – 30000	0.5 – 0.8	300 – 1500
7	0.3 – 0.6	200 – 30000	0.8 – 1.0	0 – 500
8	0.3 – 0.6	200 – 30000	0.8 – 1.0	500 – 1500
9	0.6 – 0.85	200 – 600	–1.0 – 0.5	0 – 200
10	0.6 – 0.85	200 – 600	–1.0 – 0.5	200 – 1500
11	0.6 – 0.85	200 – 600	0.5 – 0.8	0 – 300
12	0.6 – 0.85	200 – 600	0.5 – 0.8	300 – 1500
13	0.6 – 0.85	200 – 600	0.8 – 1.0	0 – 500
14	0.6 – 0.85	200 – 600	0.8 – 1.0	500 – 1500
15	0.6 – 0.85	600 – 30000	–1.0 - 0.5	0 – 200
16	0.6 – 0.85	600 – 30000	–1.0 – 0.5	200 – 1500
17	0.6 – 0.85	600 – 30000	0.5 – 0.8	0 – 300
18	0.6 – 0.85	600 – 30000	0.5 – 0.8	300 – 1500
19	0.6 – 0.85	600 – 30000	0.8 – 1.0	0 – 300
20	0.6 – 0.85	600 – 30000	0.8 – 1.0	300 – 500
21	0.6 – 0.85	600 – 30000	0.8 – 1.0	500 – 1500
22	0.85 – 1.0	200 – 600	–1.0 – 0.5	0 – 200
23	0.85 – 1.0	200 – 600	–1.0 – 0.5	200 – 1500
24	0.85 – 1.0	200 – 600	0.5 – 0.8	0 – 300
25	0.85 – 1.0	200 – 600	0.5 – 0.8	300 – 1500
26	0.85 – 1.0	200 – 600	0.8 – 1.0	0 – 300
27	0.85 – 1.0	200 – 600	0.8 – 1.0	300 – 500
28	0.85 – 1.0	200 – 600	0.8 – 1.0	500 –1500
29	0.85 – 1.0	600 – 1200	–1.0 – 0.5	0 – 200
30	0.85 – 1.0	600 – 1200	–1.0 – 0.5	200 – 1500
31	0.85 – 1.0	600 – 1200	0.5 – 0.8	0 – 300
32	0.85 – 1.0	600 – 1200	0.5 – 0.8	300 – 500
33	0.85 – 1.0	600 – 1200	0.5 – 0.8	500 – 1500
34	0.85 – 1.0	600 – 1200	0.8 – 1.0	0 – 300
35	0.85 – 1.0	600 – 1200	0.8 – 1.0	300 – 400
36	0.85 – 1.0	600 – 1200	0.8 – 1.0	400 – 500
37	0.85 – 1.0	600 – 1200	0.8 – 1.0	500 – 700
38	0.85 – 1.0	600 – 1200	0.8 – 1.0	700 – 1500
39	0.85 – 1.0	1200 – 30000	–1.0 – 0.5	0 – 200
40	0.85 – 1.0	1200 – 30000	–1.0 – 0.5	200 – 1500
41	0.85 – 1.0	1200 – 30000	0.5 – 0.8	0 – 200
42	0.85 – 1.0	1200 – 30000	0.5 – 0.8	200 – 300
43	0.85 – 1.0	1200 – 30000	0.5 – 0.8	300 – 400
44	0.85 – 1.0	1200 – 30000	0.5 – 0.8	400 – 500
45	0.85 – 1.0	1200 – 30000	0.5 – 0.8	500 – 700
46	0.85 – 1.0	1200 – 30000	0.5 – 0.8	700 – 1500
47	0.85 – 1.0	1200 – 30000	0.8 – 1.0	0 – 200
48	0.85 – 1.0	1200 – 30000	0.8 – 1.0	200 – 300
49	0.85 – 1.0	1200 – 30000	0.8 – 1.0	300 – 400
50	0.85 – 1.0	1200 – 30000	0.8 – 1.0	400 – 500
51	0.85 – 1.0	1200 – 30000	0.8 – 1.0	500 – 600
52	0.85 – 1.0	1200 – 30000	0.8 – 1.0	600 – 700
53	0.85 – 1.0	1200 – 30000	0.8 – 1.0	700 – 900
54	0.85 – 1.0	1200 – 30000	0.8 – 1.0	900 – 1100
55	0.85 – 1.0	1200 – 30000	0.8 – 1.0	1100 – 1500

Table 7.XVIII: Template parameter binning scheme, for hydrocarbon and water signals.

7.2.2.2 Reconstructed Sample Binning

In order to provide the selected signal and control samples as inputs to the fit, binning schemes must also be developed for those samples. These are optimised in the same way as the template parameter binning scheme, but instead working with the number of reconstructed events in each individual sample, of which there are a total of 18. Six different binning schemes are used, one for each of the signal and control sample definitions, where the same signal/control samples in different FGD layers use the same binning scheme for ease. Because of this, binning schemes are optimised according to the event rates in the FGD2y samples, as this is the statistically limiting set. The minimum requirement for the number of events in each bin is also relaxed, as the event rate dropping to 0 in a reconstructed bin poses no issues, as long as the template bin that the given reconstructed bin maps to receives contribution from elsewhere. With the large number of samples used in the fit, there is enough degeneracy in template bin coverage from sample bins to be confident that this is the case. Descriptions of certain constraints for the individual signal and control samples are given here.

CC1π^+ TPC Signal Sample

The binning scheme for the CC1π^+ signal sample where pions are identified in the TPCs is much the same as the template parameter binning, as over half of the identified CC1π^+ interactions come from this sample (see Table 6.I). The major difference between the two is that the signal sample binning scheme does not cover the pion momentum range $0 < p_\pi < 100$ MeV. This is due to the fact that pions with momentum in this region generally fall under the TPC momentum threshold, and are also unlikely to escape the FGD to leave enough hits in the TPC that the event would pass the track quality cut. The full binning scheme for this signal sample is given in Table 7.XIX.

CC1π^+ ME Signal Sample

Despite being the second highest source of signal events, the CC1π^+ signal sample where the presence of a positive pion is inferred from identified Michel electrons uses a very coarse binning scheme. This is due to the low resolution of the developed reconstruction, particularly in the angular reconstruction where it was decided that only one bin in $\cos\theta_\pi$ would be used. This results in each of the six muon kinematic bins containing only two pion bins: one covering the 0–200 MeV p_π region, the second covering 200–300 MeV, with no distinction made for outgoing pion angle. This is the sample that gives access to the low pion momentum region in the true binning scheme that we wish to extract for. As previously discussed, the sample only goes up to 300 MeV in pion momentum due to the range of validity of the fit to true vertex separation against momentum (Section 6.4.1). The full binning scheme is given in Table 7.XX.

CC1π^+ isoFGD Signal Sample

Although the CC1π^+ with isolated FGD pions signal sample contributes the lowest number of events to the total CC1π^+ sample, it is able to be binned slightly finer than the CC1π^+ ME

sample due to the better detector resolution. The same lower pion momentum requirement of 100 MeV from the TPC sample is applied, in order to account for the energy threshold of the FGD and the near-zero efficiency for the FGD sample in this region of true kinematics. An additional restriction is applied to avoid accepting events in the $|\cos\theta_\pi| < 0.3$ region. The events should be removed in the reconstruction, due to a lack of control samples to assess the systematic uncertainty on a momentum by range propagation for these high angle events. However, a later vertex correction makes it possible for events outside of the region to be migrated in, and thus this binning restriction is applied to remove the small number of events that this happens to. The full binning scheme for this sample is given in Table 7.XXI.

CC1π^+ + N$\pi^\pm$ (CS1)

Out of the three control samples used in this measurement, this one has the least number of events, with ∼9000 MC events in the FGD1 control sample. However, the larger range of pion momenta observed in this sample means that a relatively fine binning scheme is still possible, and thus a total of 42 reconstructed bins are used, including the first three OOPS ones. As the majority of the events in this sample come from TPC pions, the binning is made by merging bins from the TPC signal binning scheme, again using the FGD2y sample as the limiting factor. As pions from Michel electrons are excluded from this control sample, a lower limit of 100 MeV is again set on the pion momentum. The full binning scheme is given in Table F.I in Appendix F.

CC1π^+ + Nπ^0 (CS2)

As with the binning scheme for CS1, this binning scheme is again dominated by TPC pion events, and uses much the same bin edges as the TPC signal sample. Some edges are adjusted to account for the overall higher pion momentum range observed for these events. Again, due to the fact Michel electron pions are not tagged, the 100 MeV pion momentum lower bin edge is kept. The full binning scheme is given in Table F.II in Appendix F.

CC0π^+ + N isoFGDp (CS3)

The binning scheme for this control sample is created slightly differently to all others in that, other than the regular upper phase space constraint on pion momentum of 1500 MeV, the binning scheme is two-dimensional in muon kinematics only. The 'pion' kinematics for the phase space constraint are obtained by reconstructing the selected proton momentum under the pion hypothesis, as this control sample is designed to deal with the CC0π background to the CC1π^+ isoFGD sample, where an outgoing proton is mistaken for a pion. The full binning scheme is given in Table F.III in Appendix F.

7.2.2.3 Cross-Section Spline Binning

As was described in Section 7.1.2, in order to approximate the effect of reweighting nominal event rates using cross-section parameter dials, response functions known as splines are used. To ensure that the splines have the ability to alter the event rate within a single true template bin, splines are additionally binned in a reduced definition of the molecular target

on which the interaction occurs, a reduced definition of the true interaction type, and the reconstructed sample type, where no distinction is made between FGD layer. These binning schemes are given in Tables 7.XXII to 7.XXIV.

Bin Number	$\cos\theta_\mu$	p_μ	$\cos\theta_\pi$	p_π
0	–1.0 – 0.3	0 – 30000	–1.0 – 1.0	0 – 30000
1	0.3 – 1.0	0 – 200	–1.0 – 1.0	0 – 30000
2	0.3 – 1.0	200 – 30000	–1.0 – 1.0	1500 – 30000
3	0.3 – 0.6	200 – 30000	–1.0 – 0.5	100 – 1500
4	0.3 – 0.6	200 – 30000	0.5 – 0.8	100 – 300
5	0.3 – 0.6	200 – 30000	0.5 – 0.8	300 – 1500
6	0.3 – 0.6	200 – 30000	0.8 – 1.0	100 – 300
7	0.3 – 0.6	200 – 30000	0.8 – 1.0	300 – 500
8	0.3 – 0.6	200 – 30000	0.8 – 1.0	500 – 1500
9	0.6 – 0.85	200 – 600	–1.0 – 0.5	100 – 1500
10	0.6 – 0.85	200 – 600	0.5 – 0.8	100 – 300
11	0.6 – 0.85	200 – 600	0.5 – 0.8	300 – 1500
12	0.6 – 0.85	200 – 600	0.8 – 1.0	100 – 300
13	0.6 – 0.85	200 – 600	0.8 – 1.0	300 – 500
14	0.6 – 0.85	200 – 600	0.8 – 1.0	500 – 1500
15	0.6 – 0.85	600 – 30000	–1.0 – 0.5	100 – 1500
16	0.6 – 0.85	600 – 30000	0.5 – 0.8	100 – 300
17	0.6 – 0.85	600 – 30000	0.5 – 0.8	300 – 500
18	0.6 – 0.85	600 – 30000	0.5 – 0.8	500 – 1500
19	0.6 – 0.85	600 – 30000	0.8 – 1.0	100 – 300
20	0.6 – 0.85	600 – 30000	0.8 – 1.0	300 – 500
21	0.6 – 0.85	600 – 30000	0.8 – 1.0	500 – 700
22	0.6 – 0.85	600 – 30000	0.8 – 1.0	700 – 1500
23	0.85 – 1.0	200 – 600	–1.0 – 0.5	100 – 1500
24	0.85 – 1.0	200 – 600	0.5 – 0.8	100 – 300
25	0.85 – 1.0	200 – 600	0.5 – 0.8	300 – 1500
26	0.85 – 1.0	200 – 600	0.8 – 1.0	100 – 300
27	0.85 – 1.0	200 – 600	0.8 – 1.0	300 – 500
28	0.85 – 1.0	200 – 600	0.8 – 1.0	500 – 1500
29	0.85 – 1.0	600 – 1200	–1.0 – 0.5	100 – 1500
30	0.85 – 1.0	600 – 1200	0.5 – 0.8	100 – 300
31	0.85 – 1.0	600 – 1200	0.5 – 0.8	300 – 500
32	0.85 – 1.0	600 – 1200	0.5 – 0.8	500 – 1500
33	0.85 – 1.0	600 – 1200	0.8 – 1.0	100 – 300
34	0.85 – 1.0	600 – 1200	0.8 – 1.0	300 – 400
35	0.85 – 1.0	600 – 1200	0.8 – 1.0	400 – 500
36	0.85 – 1.0	600 – 1200	0.8 – 1.0	500 – 700
37	0.85 – 1.0	600 – 1200	0.8 – 1.0	700 – 1500
38	0.85 – 1.0	1200 – 30000	–1.0 – 0.5	100 – 1500
39	0.85 – 1.0	1200 – 30000	0.5 – 0.8	100 – 200
40	0.85 – 1.0	1200 – 30000	0.5 – 0.8	200 – 300
41	0.85 – 1.0	1200 – 30000	0.5 – 0.8	300 – 400
42	0.85 – 1.0	1200 – 30000	0.5 – 0.8	400 – 500
43	0.85 – 1.0	1200 – 30000	0.5 – 0.8	500 – 700
44	0.85 – 1.0	1200 – 30000	0.5 – 0.8	700 – 1500
45	0.85 – 1.0	1200 – 30000	0.8 – 1.0	100 – 200
46	0.85 – 1.0	1200 – 30000	0.8 – 1.0	200 – 300
47	0.85 – 1.0	1200 – 30000	0.8 – 1.0	300 – 400
48	0.85 – 1.0	1200 – 30000	0.8 – 1.0	400 – 500
49	0.85 – 1.0	1200 – 30000	0.8 – 1.0	500 – 600
50	0.85 – 1.0	1200 – 30000	0.8 – 1.0	600 – 700
51	0.85 – 1.0	1200 – 30000	0.8 – 1.0	700 – 900
52	0.85 – 1.0	1200 – 30000	0.8 – 1.0	900 – 1100
53	0.85 – 1.0	1200 – 30000	0.8 – 1.0	1100 – 1500

Table 7.XIX: TPC sample binning scheme, for FGD1, FGD2x and FGD2y samples.

Bin Number	$\cos\theta_\mu$	p_μ	$\cos\theta_\pi$	p_π
0	–1.0 – 0.3	0 – 30000	–1.0 – 1.0	0 – 30000
1	0.3 – 1.0	0 – 200	–1.0 – 1.0	0 – 30000
2	0.3 – 0.6	200 – 30000	–1.0 – 1.0	0 – 200
3	0.3 – 0.6	200 – 30000	–1.0 – 1.0	200 – 300
4	0.6 – 0.85	200 – 600	–1.0 – 1.0	0 – 200
5	0.6 – 0.85	200 – 600	–1.0 – 1.0	200 – 300
6	0.6 – 0.85	600 – 30000	–1.0 – 1.0	0 – 200
7	0.6 – 0.85	600 – 30000	–1.0 – 1.0	200 – 300
8	0.85 – 1.0	200 – 600	–1.0 – 1.0	0 – 200
9	0.85 – 1.0	200 – 600	–1.0 – 1.0	200 – 300
10	0.85 – 1.0	600 – 1200	–1.0 – 1.0	0 – 200
11	0.85 – 1.0	600 – 1200	–1.0 – 1.0	200 – 300
12	0.85 – 1.0	1200 – 30000	–1.0 – 1.0	0 – 200
13	0.85 – 1.0	1200 – 30000	–1.0 – 1.0	200 – 300

Table 7.XX: ME sample binning scheme, for FGD1, FGD2x and FGD2y samples.

Bin Number	$\cos\theta_\mu$	p_μ	$\cos\theta_\pi$	p_π
0	–1.0 – 0.3	0 – 30000	–1.0 – 1.0	0 – 30000
1	0.3 – 1.0	0 – 200	–1.0 – 1.0	0 – 30000
2	0.3 – 0.6	200 – 30000	–1.0 – 0.5	100 – 200
3	0.3 – 0.6	200 – 30000	–1.0 – 0.5	200 – 1500
4	0.3 – 0.6	200 – 30000	0.5 – 1.0	100 – 1500
5	0.6 – 0.85	200 – 600	–1.0 – 0.5	100 – 200
6	0.6 – 0.85	200 – 600	–1.0 – 0.5	200 – 1500
7	0.6 – 0.85	200 – 600	0.5 – 1.0	100 – 1500
8	0.6 – 0.85	600 – 30000	–1.0 – 0.5	100 – 200
9	0.6 – 0.85	600 – 30000	–1.0 – 0.5	200 – 1500
10	0.6 – 0.85	600 – 30000	0.5 – 1.0	100 – 1500
11	0.85 – 1.0	200 – 600	–1.0 – 0.5	100 – 200
12	0.85 – 1.0	200 – 600	–1.0 – 0.5	200 – 1500
13	0.85 – 1.0	200 – 600	0.5 – 1.0	100 – 1500
14	0.85 – 1.0	600 – 1200	–1.0 – 0.5	100 – 200
15	0.85 – 1.0	600 – 1200	–1.0 – 0.5	200 – 1500
16	0.85 – 1.0	600 – 1200	0.5 – 1.0	100 – 1500
17	0.85 – 1.0	1200 – 30000	–1.0 – 0.5	100 – 200
18	0.85 – 1.0	1200 – 30000	–1.0 – 0.5	200 – 1500
19	0.85 – 1.0	1200 – 30000	0.5 – 1.0	100 – 200
20	0.85 – 1.0	1200 – 30000	0.5 – 1.0	200 – 1500

Table 7.XXI: isoFGD sample binning scheme, for FGD1, FGD2x and FGD2y samples.

Bin Index	Target
0	Hydrocarbon
1	Water
2	Other

Table 7.XXII: Reduced target binning for cross-section splines.

Bin Index	Target
0	CCQE
1	RES (CC1π^+)
2	RES (non-CC1π^+)
3	DIS
4	Other

Table 7.XXIII: Reduced reaction binning for cross-section splines.

Bin Index	Target
0	CC1π^+ TPC
1	CC1π^+ ME
2	CC1π^+ isoFGD
3	CS1
4	CS2
5	CS3

Table 7.XXIV: Reconstructed sample binning for cross-section splines.

Chapter 8

Fit Validation

Before using the likelihood fit framework to perform a fit to real ND280 data, it is important to go through a series of tests to understand the response of the fitter is as expected. This chapter presents a series of these tests, which are performed by making modifications to the nominal Monte Carlo simulation. All fitter tests are performed with the MC scaled roughly to data POT, so the effect of statistical fluctuations is realistic.

8.1 Summary of the Fit

Four different types of parameters are included in the fit:

- *Template parameters* – One template parameter is assigned to each of the true kinematics bins, resulting in 59 for hydrocarbon and 59 for water, for a total of 118 template parameters. Template parameters have no prior value or uncertainty applied.

- *Flux parameters* – One flux parameter is assigned for each of the neutrino energy bins in the flux covariance matrix, totalling 19 parameters. The prior values and uncertainties are set according to the provided covariance matrix, given in Section 7.1.3.

- *Cross-section parameters* – One cross-section parameter is assigned to each of the model dials described in Section 7.1.2, where the prior values, uncertainties and correlations are given. A total of 16 cross-section parameters are given to the fit.

- *Detector uncertainty parameters* – One parameter is assigned for each reconstructed sample bin, as described in Section 7.2.2.2, totalling 579 parameters. Each parameter has a prior value of one, with the uncertainties and correlations taken from the covariance matrix in Figure 7.5.

This results in a total of 732 parameters that are input to the fit where, as described previously in Chapter 5, the best fit set of values is found by minimising the associated log-likelihood functions.

8.2 Asimov Fits

An Asimov fit is defined as that where the input Monte Carlo is fitted to itself. In the case of the fitter described, Asimov fits are performed as a first test of the fit framework, in order to ensure that there are no significant issues with the set-up. As the MC is being fitted to itself, the fit should slowly explore the parameter space, where any movement away from the nominal parameter values should result in a raised χ^2 value, before the fitter converges on the nominal parameter values.

8.2.1 Asimov Fit with Nominal Prior Values

The first fit validation performed is a very simple Asimov fit, where the prior values start at nominal, the results for which can be seen in Figure 8.1. For all four sets of fit parameters, the prior and post-fit values are found to be exactly the same, as is expected. Additionally a reduction in error from prior to post-fit is observed in most parameters, particularly in some of the detector uncertainty parameters which have high prior errors. The template parameter plot shows the values for the hydrocarbon parameters (c_i's), followed by water (o_i's), where it can be seen that the relative error on the post-fit values is quite high, particularly for water. This is caused by the fine binning scheme used for efficiency correction, and in collapsing the result down to lower dimensions should reduce.

To assess the impact of each non-template parameter (otherwise known as nuisance parameters) on the fit result, it is additionally useful to plot the parameter pulls and associated fit constraints, pre- and post-fit. The pull of the i^{th} parameter in the fit can be calculated as

$$\text{Pull}_i = \frac{\eta_i^{\text{old}} - \eta_i^{\text{new}}}{\sigma(\eta_i^{\text{old}})}, \tag{8.1}$$

where η_i^{old} and η_i^{new} are the prior and post-fit values for the i^{th} nuisance parameter respectively, and σ is the associated uncertainty. Figure 8.2 shows the pull values and associated prior and post-fit errors for the 20 parameters with the largest post-fit error values, given in order. As would be expected for an Asimov fit, all pull values are seen to be 0. The majority of the large post-fit constraints come from cross-section parameters, mostly relating to DIS and multiple pion interactions, or FSI. The remainder are seen to come from detector parameters, and can mostly be attributed to CS1 in the FGD2y sample.

While the errors on the template parameters are seen to be rather large, particularly for the water parameters, it is still interesting to present this result as the full four-dimensional differential cross section, before integrating over bins. The 4D Asimov cross section using `NEUT` MC is shown in Figures 8.3 and 8.4 for hydrocarbon and water, respectively, as a function of pion momentum. Rows show slices of pion angle, whilst columns give slices of muon kinematics. The result for hydrocarbon shows that, despite the large errors, at forward-going angles for both the muon and pion, the shape of the cross section can be well resolved. Whilst some shape is seen in the water result, the larger errors are clearly more limiting here.

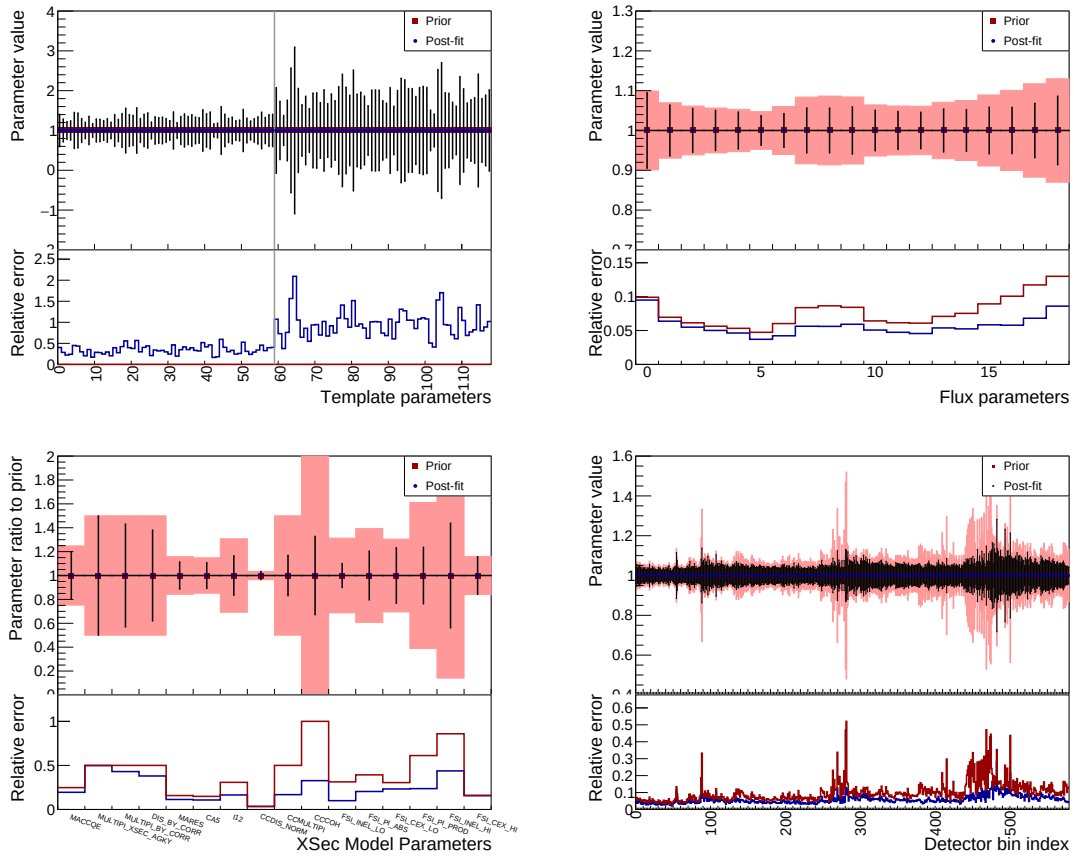


Figure 8.1: Template, flux, cross-section model and detector systematic parameter fit results, for the Asimov fit with nominal prior values.

To report a result with lower statistical error, unfolded bins can be combined before the cross section in each bin is calculated. The approach used for this is to combine bins in muon kinematics, so that a double differential measurement in pion kinematics can be reported. In order to do this efficiently, the choice is made to restrict the phase space to events with $\cos\theta_\mu > 0.6$ (along with the general analysis phase space constraints), so as to not be limited to a coarser binning scheme than is necessary. The binning scheme used for the double differential measurement in pion kinematics is presented in Table 8.I, and MC only Asimov results for both hydrocarbon and water are given in Figure 8.5.

8.2.2 Asimov Fit with Random Prior Values

While the Asimov fit in Section 8.2.1 shows that the fitter will return the known set of best-fit parameters, we are somewhat taking for granted that the full parameter space is being explored. As the fit starts at the best-fit values, it could easily converge at these values post-fit without many iterations. To check whether the fitter is exploring the parameter space as we require, a second type of Asimov fit is performed, where the initial values are randomly set.

This test finds that the fit has no issue in converging on the known best-fit parameters,

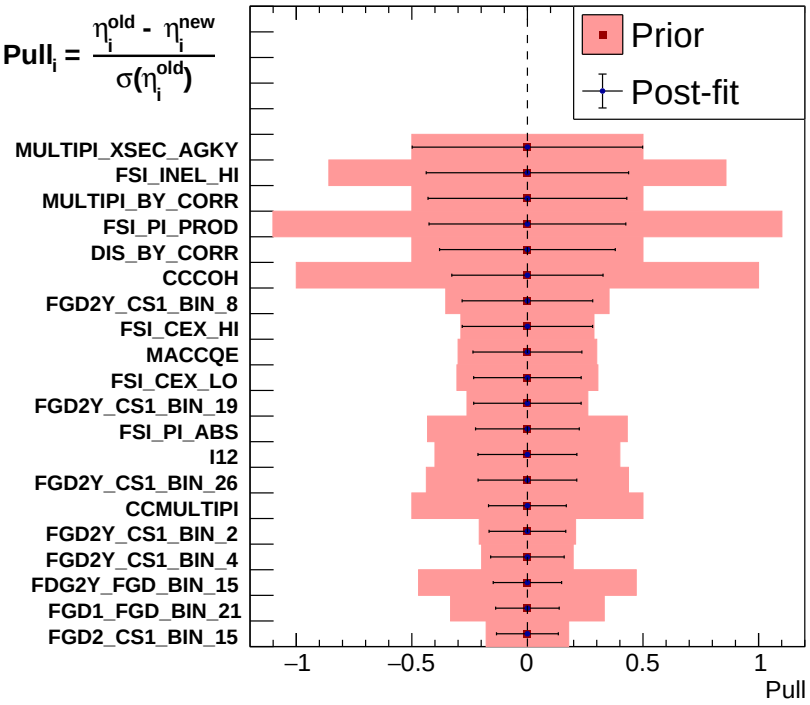


Figure 8.2: Pull values and associated errors for the 20 nuisance parameters with the largest post-fit error, given in order.

Bin Number	$\cos\theta_\pi$	p_π [MeV]
0	$-1 - 0.5$	$0 - 200$
1		$200 - 1500$
2	$0.5 - 0.8$	$0 - 300$
3		$300 - 1500$
4	$0.8 - 1.0$	$0 - 300$
5		$300 - 500$
6		$500 - 1500$

Table 8.I: Pion kinematic bins for reporting a double differential cross-section measurement, in the restricted phase space of $\cos\theta_\mu > 0.6$.

despite starting from random values, and is thus capable of exploring the parameter space as required. As the post-fit parameters are exactly the same as for the initial Asimov fit (Figure 8.1), the plots are not repeated.

8.3 Statistical Fluctuations and Error Validation

The previous Asimov fits described are an important first test for the fit machinery, albeit rather simple. In practice, when applying the fit to real data, the set of best fit parameters will be completely unknown, and the data could potentially have quite different event distributions to what is predicted in the Monte Carlo. In order to check the capacity of the

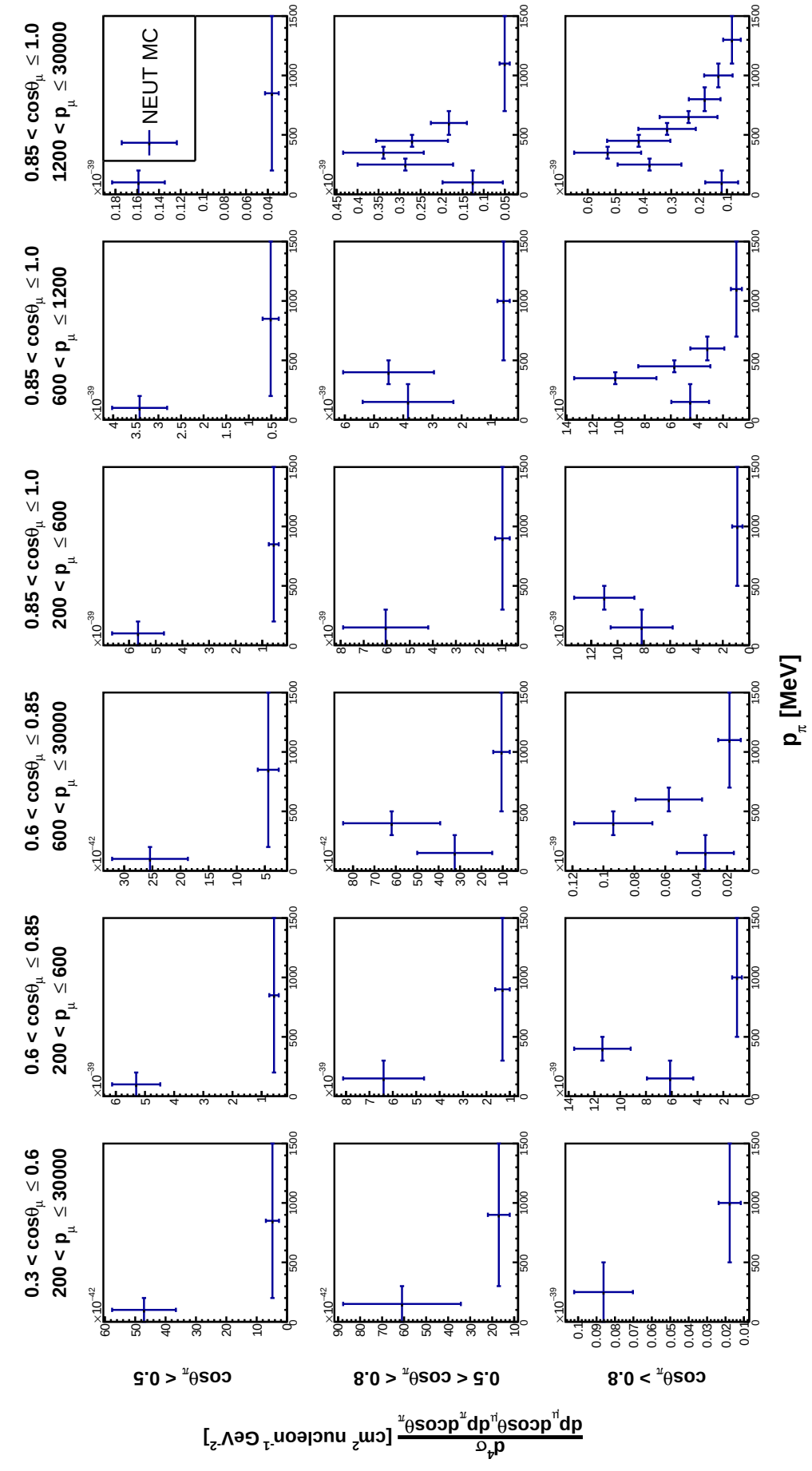


Figure 8.3: 4D differential MC only cross-section result for interactions on hydrocarbon.

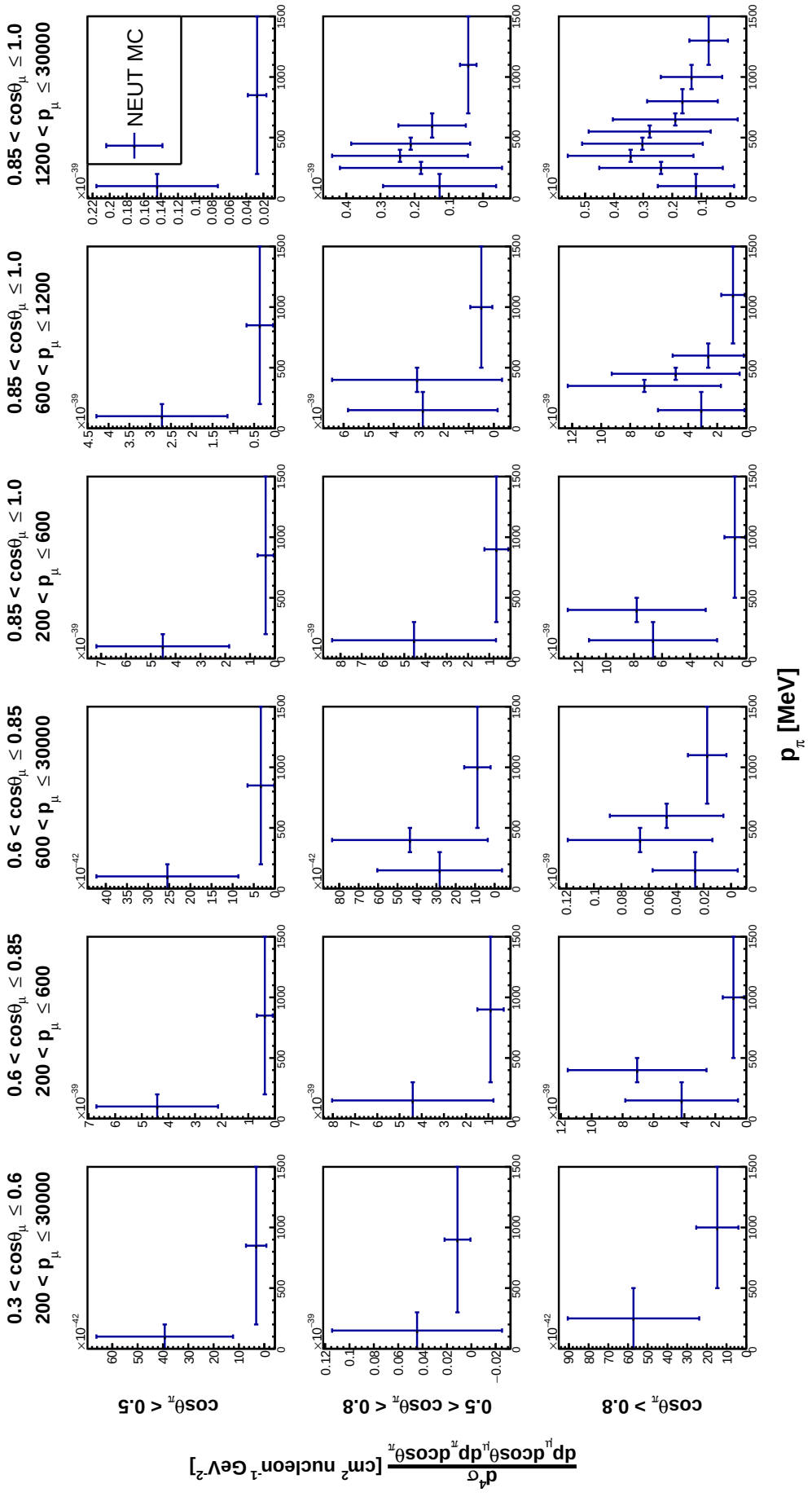
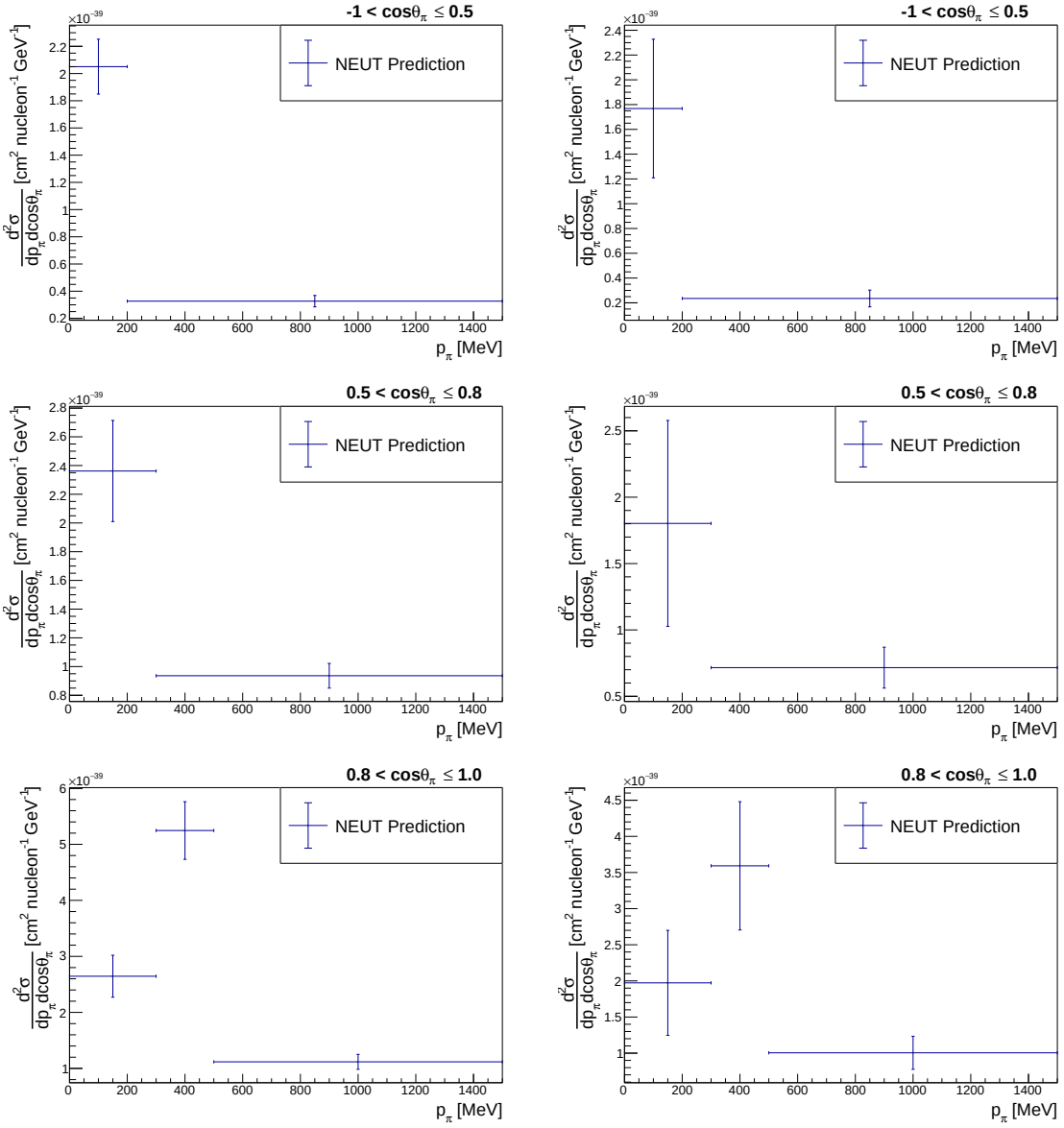


Figure 8.4: 4D differential MC only cross-section result for interactions on water.



(a) CH (b) H_2O

Figure 8.5: 2D differential cross section in pion momentum and $\cos\theta$ for interactions on (a) hydrocarbon and (b) water. Phase space is limited to $p_\mu > 200$ MeV and $\cos\theta_\mu > 0.6$.

fitter to deal with such differences, a series of fake-data studies are performed.

8.3.1 Statistical Fluctuations

As mentioned in Section 7.2.2.1, the possibility of the event rate in a given bin falling to zero can sometimes cause issues with convergence of the fit. Whilst the binning schemes are chosen in order to limit the chance of this happening, it is still possible, and therefore it is important to test the effect of statistical fluctuations on the fit. Fake data is created by applying Poissonian statistical fluctuations to the nominal MC, as would be expected in real data. This is also an important test of the ability of the fitter to converge for cases

where the number of events in a given reconstructed bin is significantly far away from what is expected.

In general, the fit performs fairly well at this, with almost all post-fit values closer to the fake data value than the prior value was. Event distributions for the FGD1 samples are given in Figure 8.6, and the fit parameter values are given in Figure 8.7. As is expected, the post-fit parameter values are somewhat randomly distributed around the nominal, although a slight downward trend overall is observed. Out of the 118 total template parameters, only one is seen to take a best-fit value slightly below 0. The associated fake data cross-section results are shown as a 4D differential measurement on hydrocarbon and oxygen in Figures 8.8 and 8.9, where it can be seen that the cross-section results for fake data (blue bars) generally fluctuate around the nominal MC value (red bands). Finally, results from restricting the muon phase space to $\cos\theta_\mu > 0.6$ and integrating over muon bins to obtain the double differential cross section in pion kinematics are given in Figure 8.10. Integrating over the muon bins causes some of the random fluctuations to average out, giving fake data results where the best-fit points all lie within the error band of the nominal value.

8.3.2 Error Validation

Additional tests using fluctuated fits can be performed in order to understand whether the resultant χ^2 value from a fit suggests that the result can be trusted. To perform this test, 400 fits were run, with both statistical and systematic fluctuations applied. Of these 400, only 14 had fluctuations so large compared to nominal MC that the fit was unable to converge. The distribution of resultant total χ^2 values for the successfully converging fits are shown in Figure 8.11, which is seen to have a mean χ^2 value of 505. This is slightly higher than would initially be expected, as the number of degrees of freedom in the fit would be expected to equal the total number of free parameters, minus the number of template fit parameters. For 579 reconstructed sample bins and 118 template parameters, this would nominally result in 461 degrees of freedom. It is, however, important to remember that a large number of the template parameters are expected to contain very low numbers of events, due to the fine binning scheme used for cross-section extraction. This is particularly true in the water template parameters, which are expected to be lower in statistics anyway. Where template parameter bins are particularly low in events and the assumed Gaussianity within bins breaks down, they will not fully 'count' as a whole degree of freedom, and thus the expected number of degrees of freedom in the fit is likely higher than the naive prediction of 461 anyway.

Following the fitting procedure, all of the converged fits can be used to calculate a series of differential cross-section distributions, in the same way as described previously. For each toy cross section, the χ^2 value of the result with respect to the nominal MC result can be calculated as

$$\chi^2 = (\vec{\sigma}_{\mathrm{MC}} - \vec{\sigma}_{\mathrm{toy}})\, \mathbf{V}_{\mathrm{cov}}^{-\mathbf{1}}\, (\vec{\sigma}_{\mathrm{MC}} - \vec{\sigma}_{\mathrm{toy}})^{\mathbf{T}}\,, \tag{8.2}$$

where $\vec{\sigma}$ is the vector of cross-section measurements for either nominal MC or in each toy,

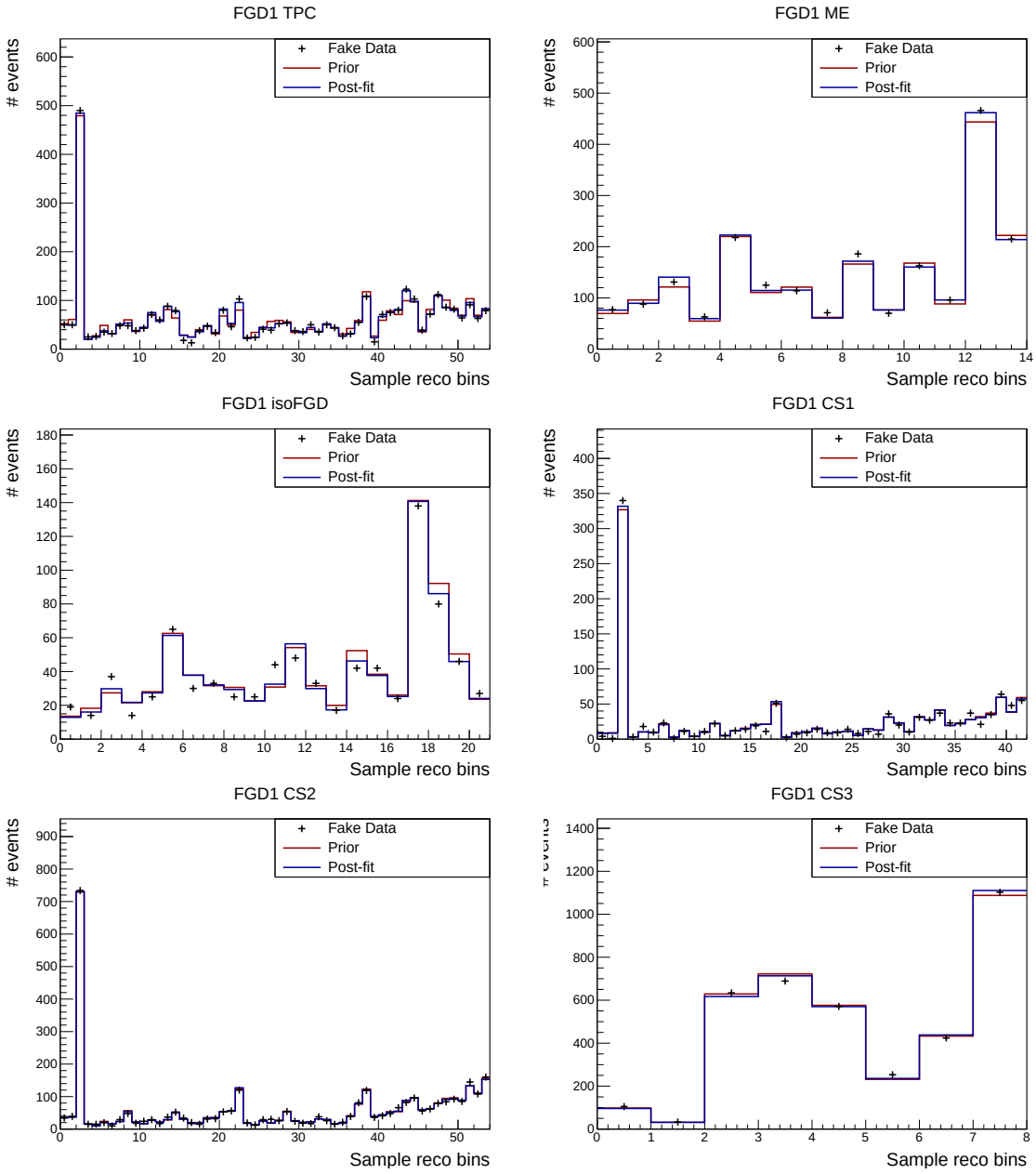


Figure 8.6: Prior (red), post-fit (blue) and fake data (black crosses) reconstructed event distributions for FGD1 samples, for the fit to statistically fluctuated fake data.

and $\mathbf{V}_{\text{cov}}^{-1}$ is the obtained cross-section covariance matrix. Calculating the χ^2 value for each toy, for both the 4D and 2D results, allows distributions of the χ^2 to be built. These are given in Figure 8.12. Both distributions looks as expected, following a χ^2 shape and with mean values close to the expected number of degrees of freedom for the results when approximating to the number of bins (112 in 4D, 14 in 2D). Taking the mean values, we are able to obtain effective numbers of degrees of freedom of 113.9 and 12.8, for the 4D and 2D results respectively.

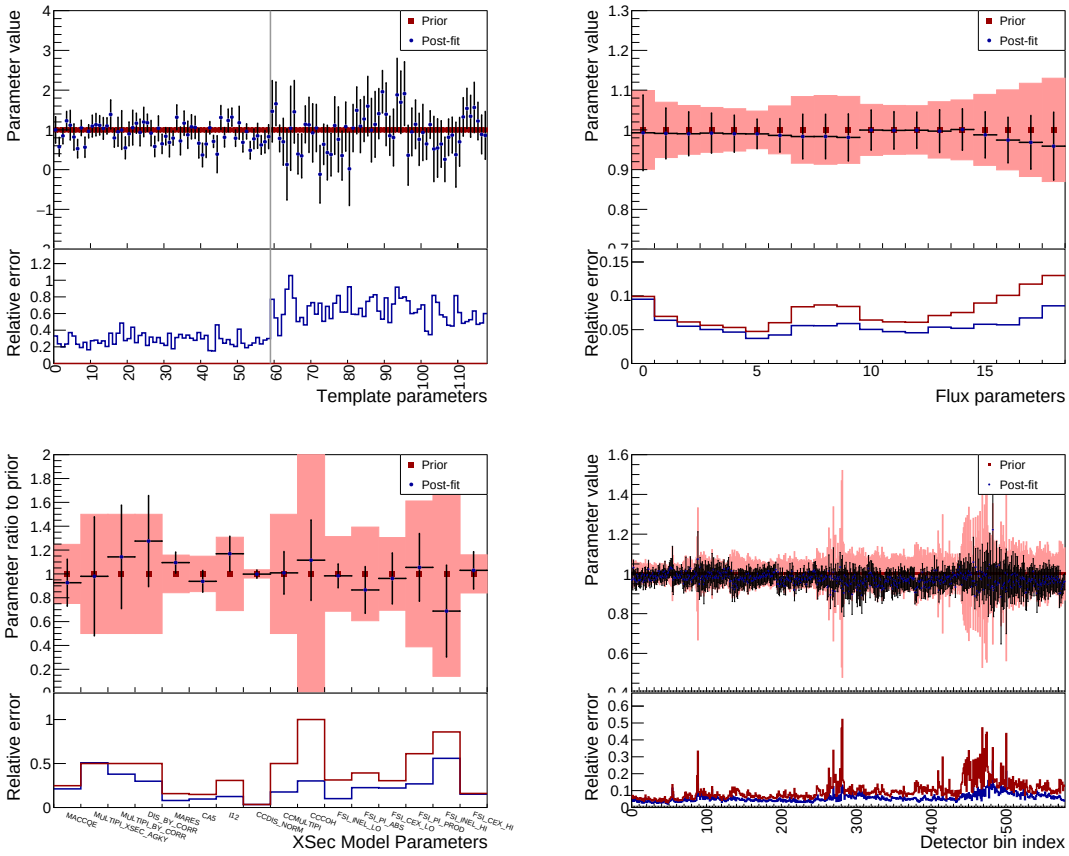


Figure 8.7: Template, flux, cross-section model and detector systematic parameter fit results, for the fit to statistically fluctuated fake data.

8.4 Signal and Model-Enhanced Fake Data

The final set of fake data studies presented are performed to check that the fitter is capable of resolving changes made to the underlying signal, which could occur in data if the MC prediction for either of the signals is not accurate. An additional concern is whether the fit is capable of dealing with changes to systematic parameters, without the rest of the parameter space being adversely affected.

8.4.1 Signal Enhanced Fake Data

This set of fake data is designed to ensure that the fitter can accurately reconstruct changes in the signal cross section. The fake data set is created from the nominal MC, by reweighting all CC1π^+ signal events on H_2O by a factor of 1.2. All other events, including background on H_2O, have their initial weight left at 1.0. The post-fit parameters given in Figure 8.13 show perfect reconstruction of this, with all CH post-fit parameters at their prior value of 1.0, whereas the H_2O parameters have taken post-fit values of 1.2. All systematic parameters also have post-fit values of 1.0, as expected. Propagating this fit through to a cross section gives the results shown in 4D in Figures 8.14 and 8.15 for hydrocarbon and water respectively, and 2D pion kinematics in Figure 8.16. As expected, these show complete agreement to the

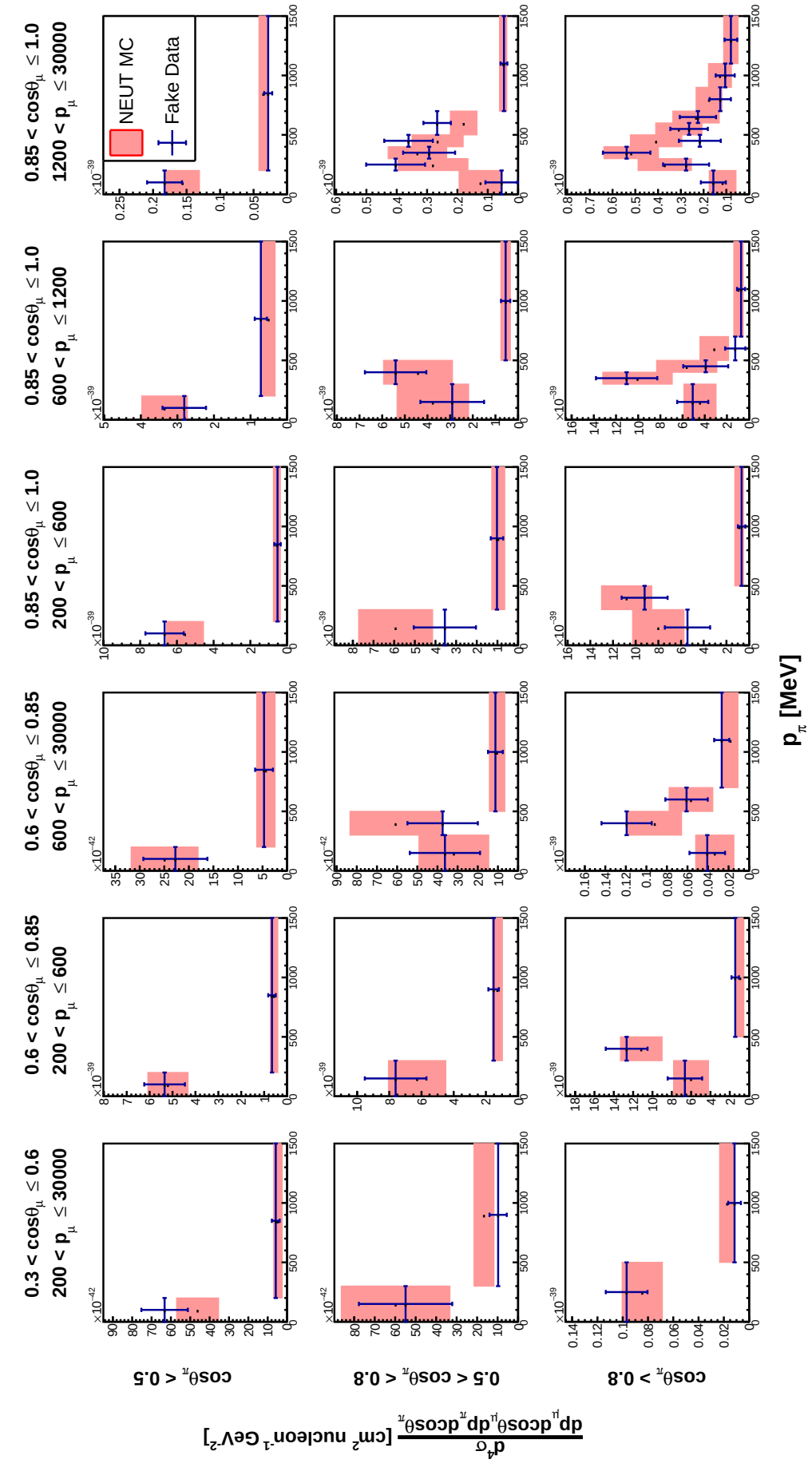


Figure 8.8: 4D differential cross-section result for interactions on hydrocarbon, for statistcally fluctuated MC fake data.

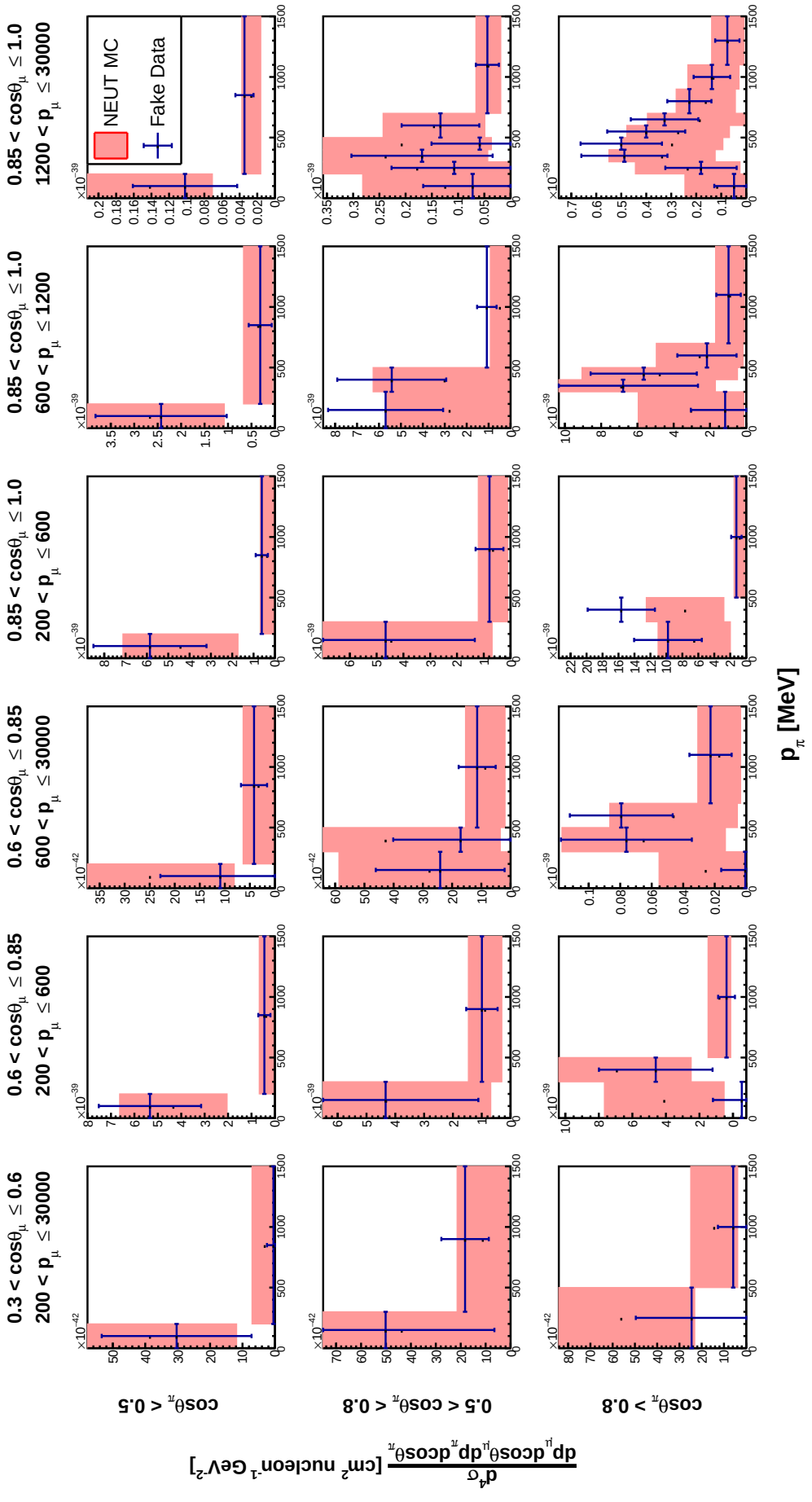
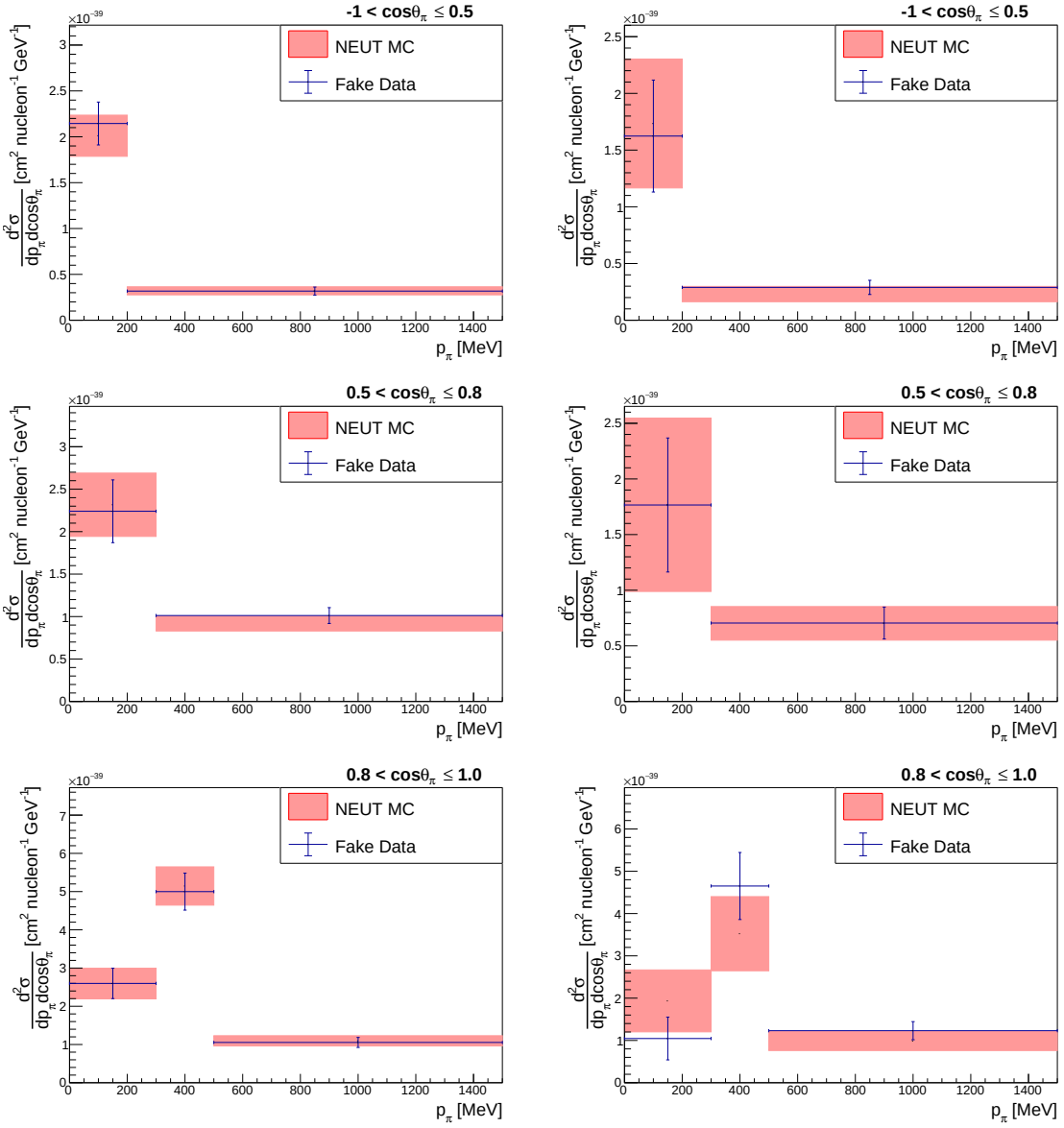


Figure 8.9: 4D differential cross-section result for interactions on water, for statistcally fluctuated MC fake data.



(a) CH (b) H_2O

Figure 8.10: 2D differential cross section in pion momentum and $\cos\theta$ using statistically fluctuated MC fake data, for interactions on (a) hydrocarbon and (b) water. Phase space is limited to $p_\mu > 200$ MeV and $\cos\theta_\mu > 0.6$.

nominal MC cross section for the hydrocarbon case, whereas the water cross-section result has been inflated by 20%.

8.4.2 Model Enhanced Fake Data

This fake data study is designed to ensure that the fitter can handle changes to specific uncertainties within the model, without causing drastic changes to other parameters. The fake data set is created by reweighting events in the nominal MC which proceed via DIS interactions by a factor of 1.2. All other weights remain at nominal.

The fit parameter response to this fake data set can be seen in Figure 8.17, with results

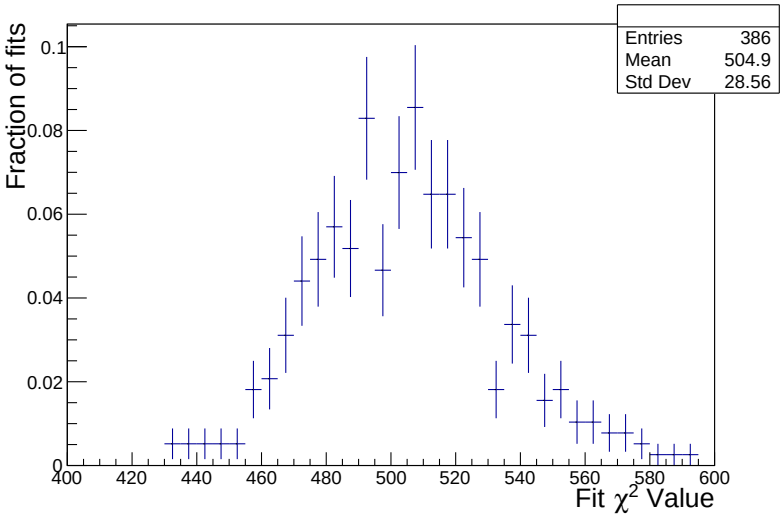
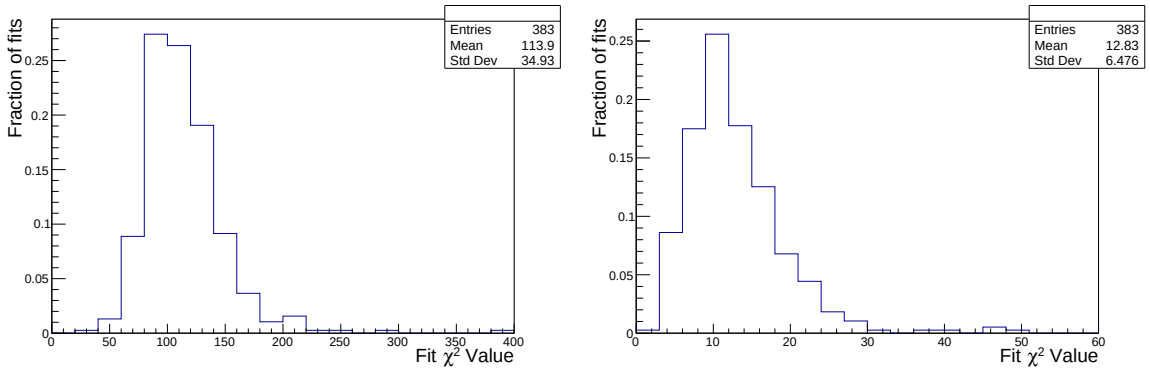


Figure 8.11: Distribution of fit total χ^2 values for 386 successful statistically and systematically fluctuated fits.



(a) 4D (b) 2D

Figure 8.12: Distribution of cross-section χ^2 values, for the (a) 4D and (b) 2D cross-section results.

as expected. The template parameter post-fit values lie either on or very close to 1.0, with a similar structure in the variations for both CH and H_2O. An overall rise in flux parameters is seen, although a greater increase is seen in higher flux bins, corresponding to the higher energy incoming neutrinos needed to facilitate DIS interactions. In the cross-section model parameters, a reduction in the MACCQE dial value is observed, along with increases in the DIS related dials, whilst the rest remain close to their nominal value. It is also interesting to observe the response in the detector systematic parameters. While the TPC parameters remain close to one, there is a slight drop in the ME and isoFGD parameters, the samples of which generally contain lower energy interactions. A rise is then observed in the first two control samples (CC1π^++N$\pi^\pm$ and CC1π^++Nπ^0), which are known from Table 6.VI to originate from predominantly DIS interactions. This is followed by a drop in the CC0π+Np control sample, another generally low energy sample restricted to events where the proton is contained within the FGD. This pattern then repeats across

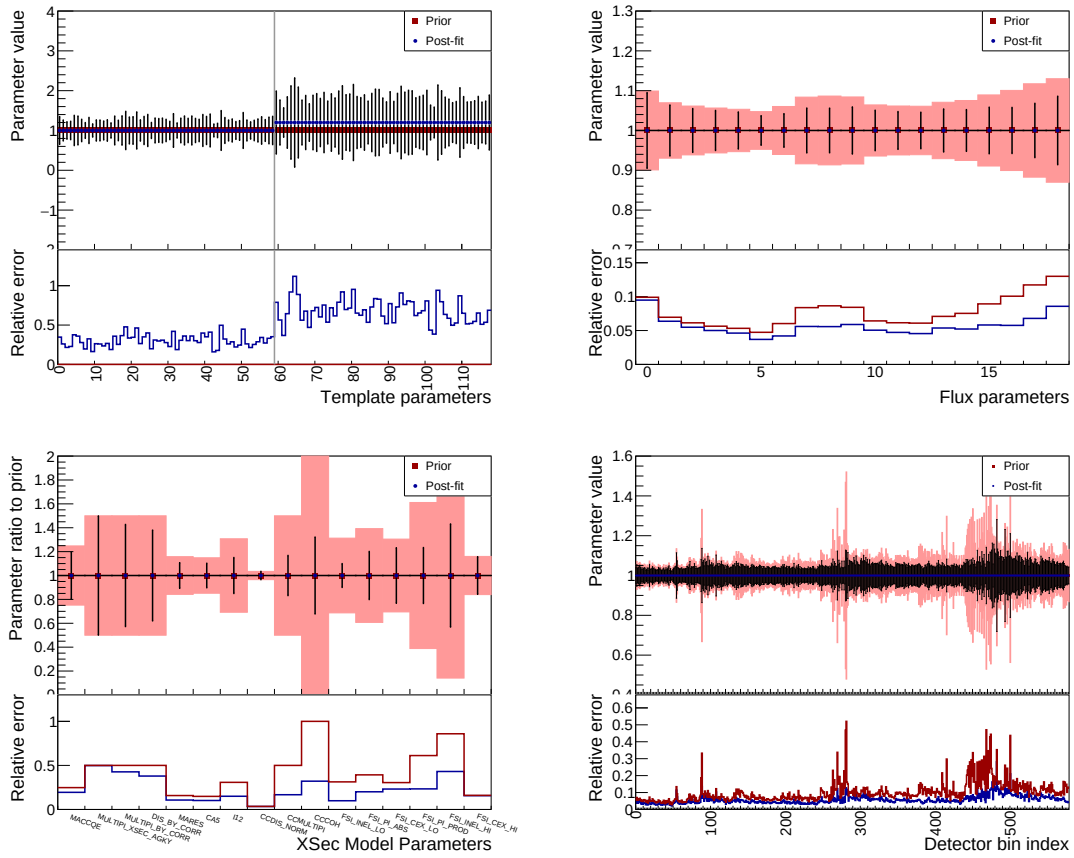


Figure 8.13: Template, flux, cross-section model and detector systematic parameter fit results, for the fit to H_2O signal enhanced fake data.

the three FGD layer samples. This suggests that the fit cannot completely resolve the effect of the additional DIS component by simply changing the cross-section parameter dials, but as the rest of the parameters still lie around one this is not seen as a cause for concern.

Cross-section results for the DIS enhanced fake data sample are given for the full 4D result on hydrocarbon and oxygen in Figures 8.18 and 8.19 respectively, whilst Figure 8.20 gives the double differential result in pion kinematics. As expected from the resultant fit parameters, the obtained fake data cross section is very similar to the NEUT nominal; increasing the flux of incoming neutrinos will increase the number of events, without a change to the interaction cross section.

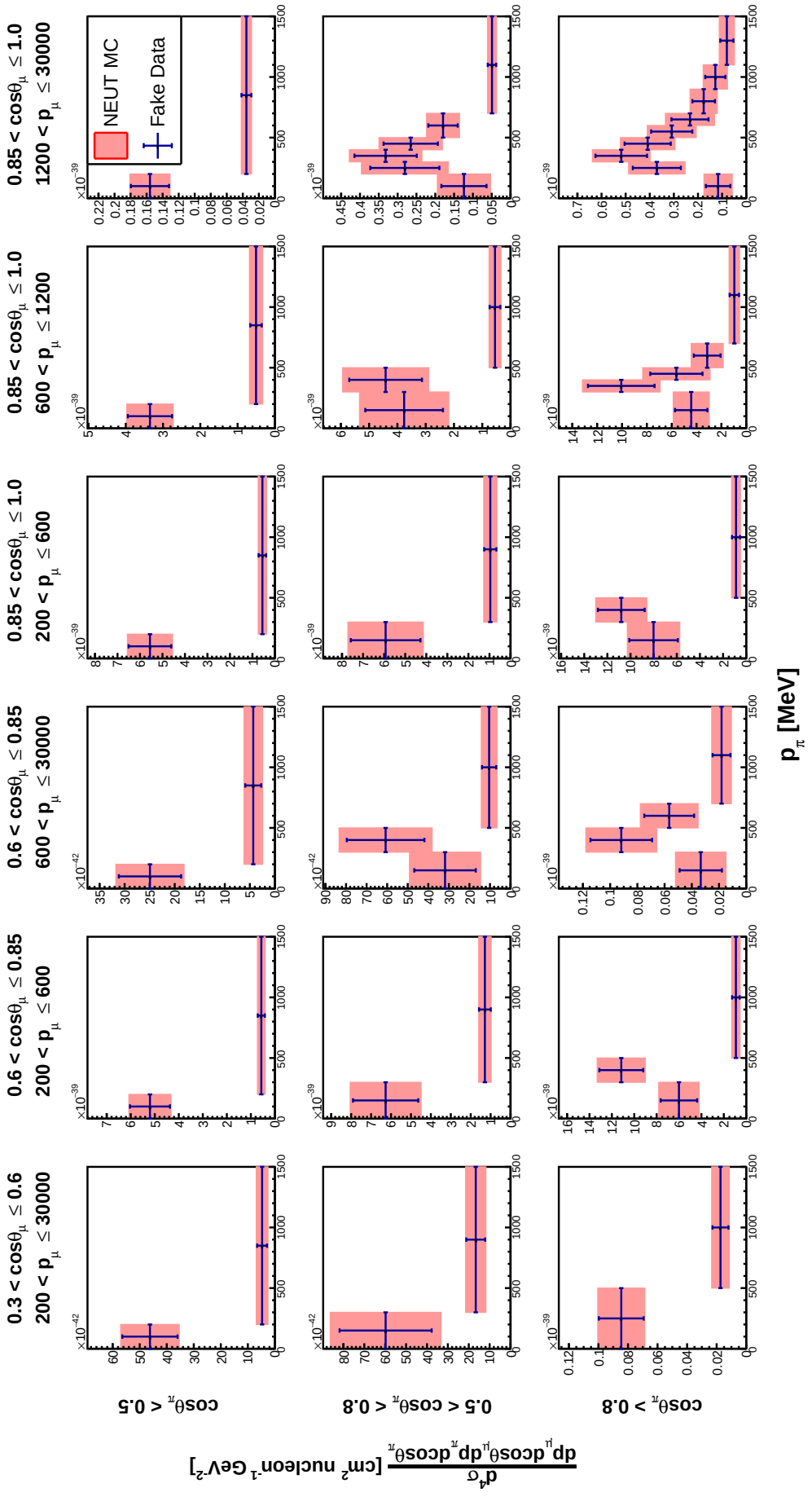


Figure 8.14: 4D differential cross-section result for interactions on hydrocarbon, for H_2O signal enhanced fake data.

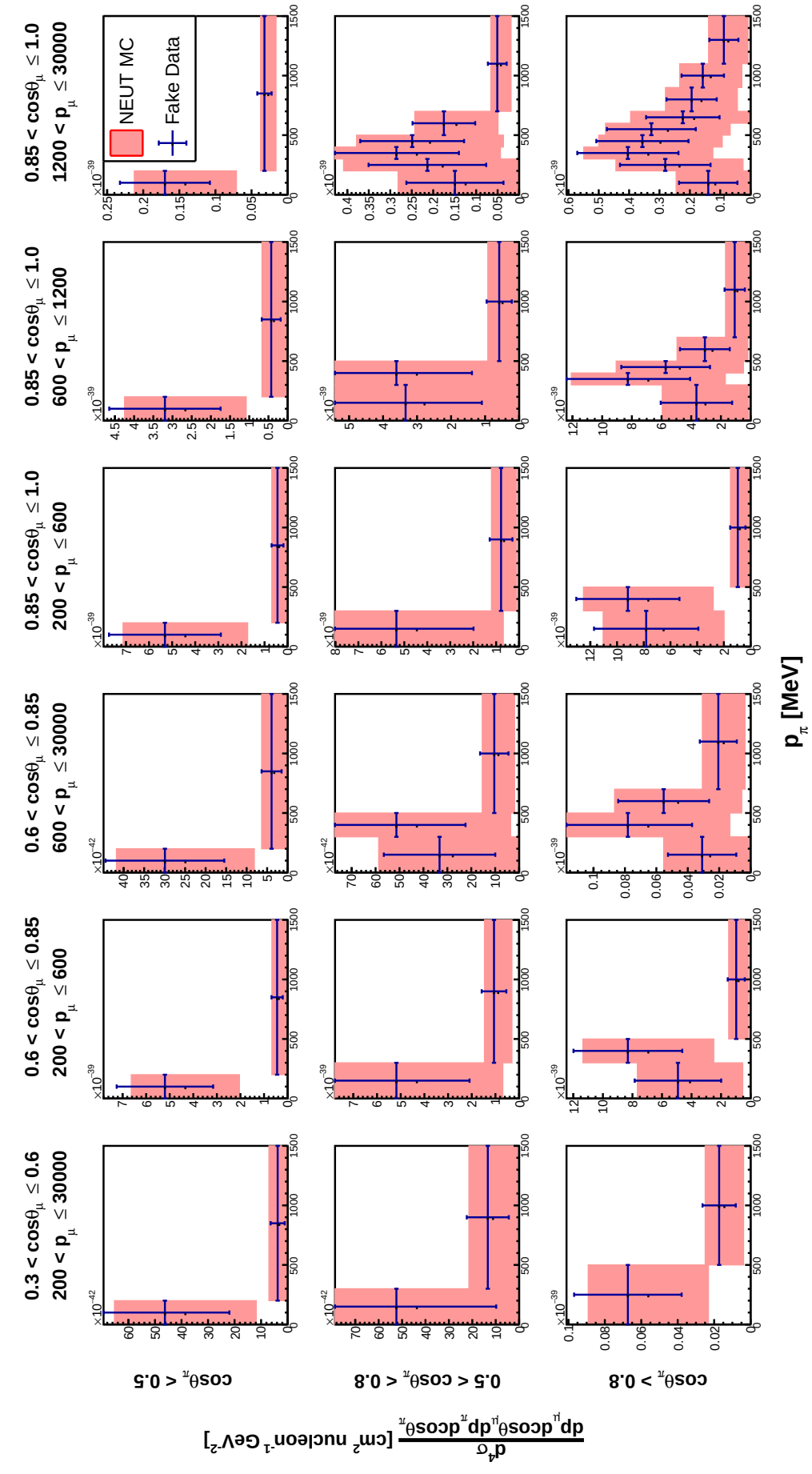
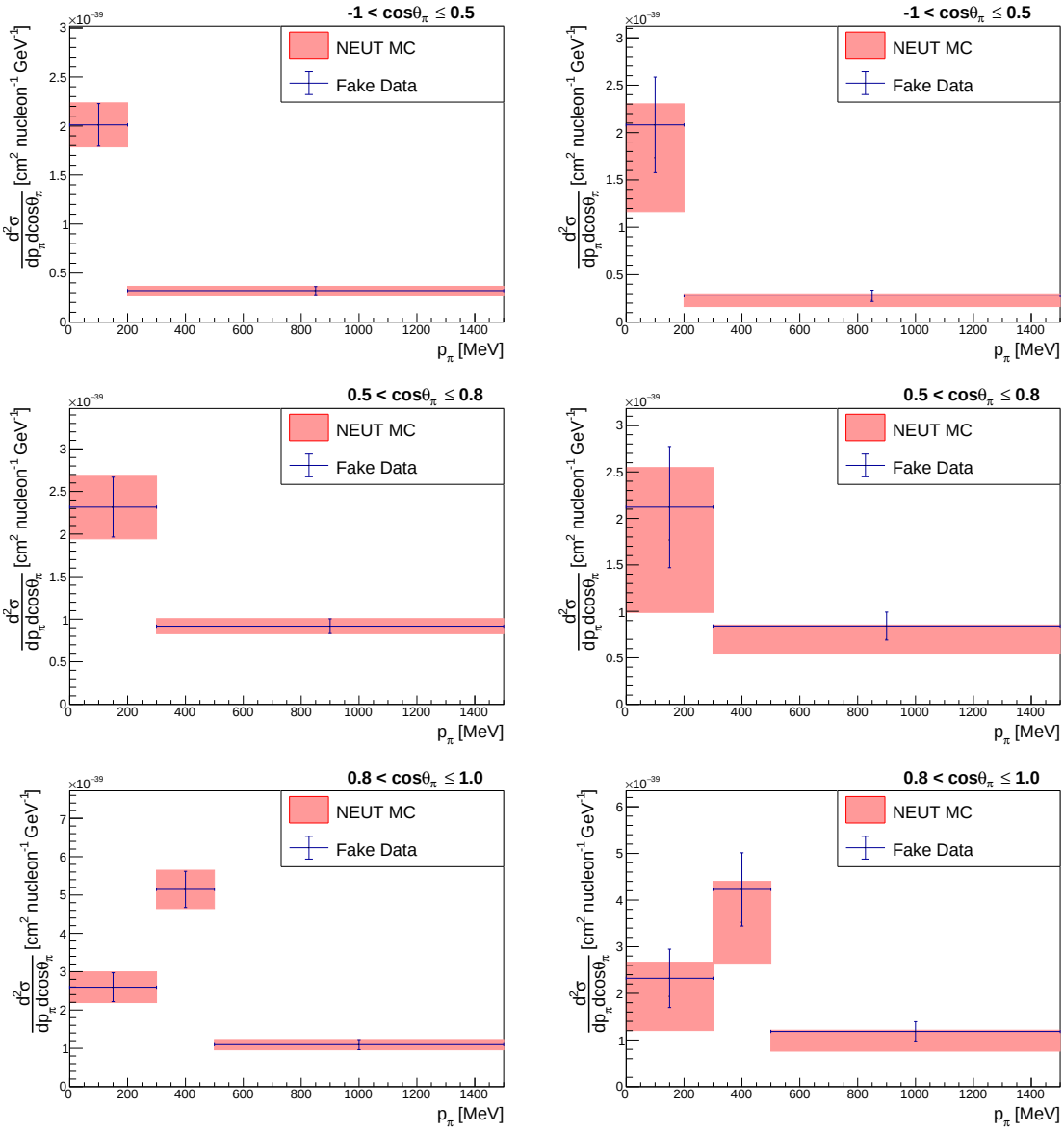


Figure 8.15: 4D differential cross-section result for interactions on water, for H_2O signal enhanced fake data.



(a) CH (b) H_2O

Figure 8.16: 2D differential cross section in pion momentum and $\cos\theta$ using H_2O signal enhanced fake data, for interactions on (a) hydrocarbon and (b) water. Phase space is limited to $p_\mu > 200$ MeV and $\cos\theta_\mu > 0.6$.

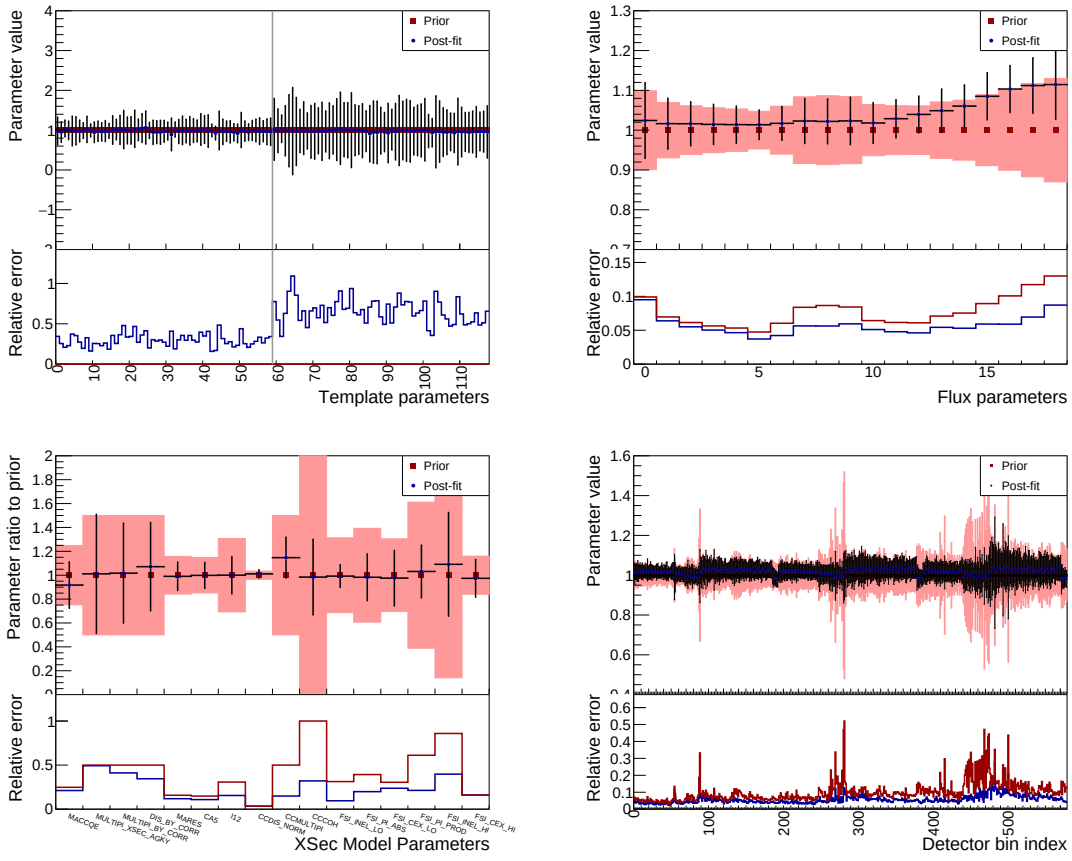


Figure 8.17: Template, flux, cross-section model and detector systematic parameter fit results, for the fit to DIS enhanced fake data.

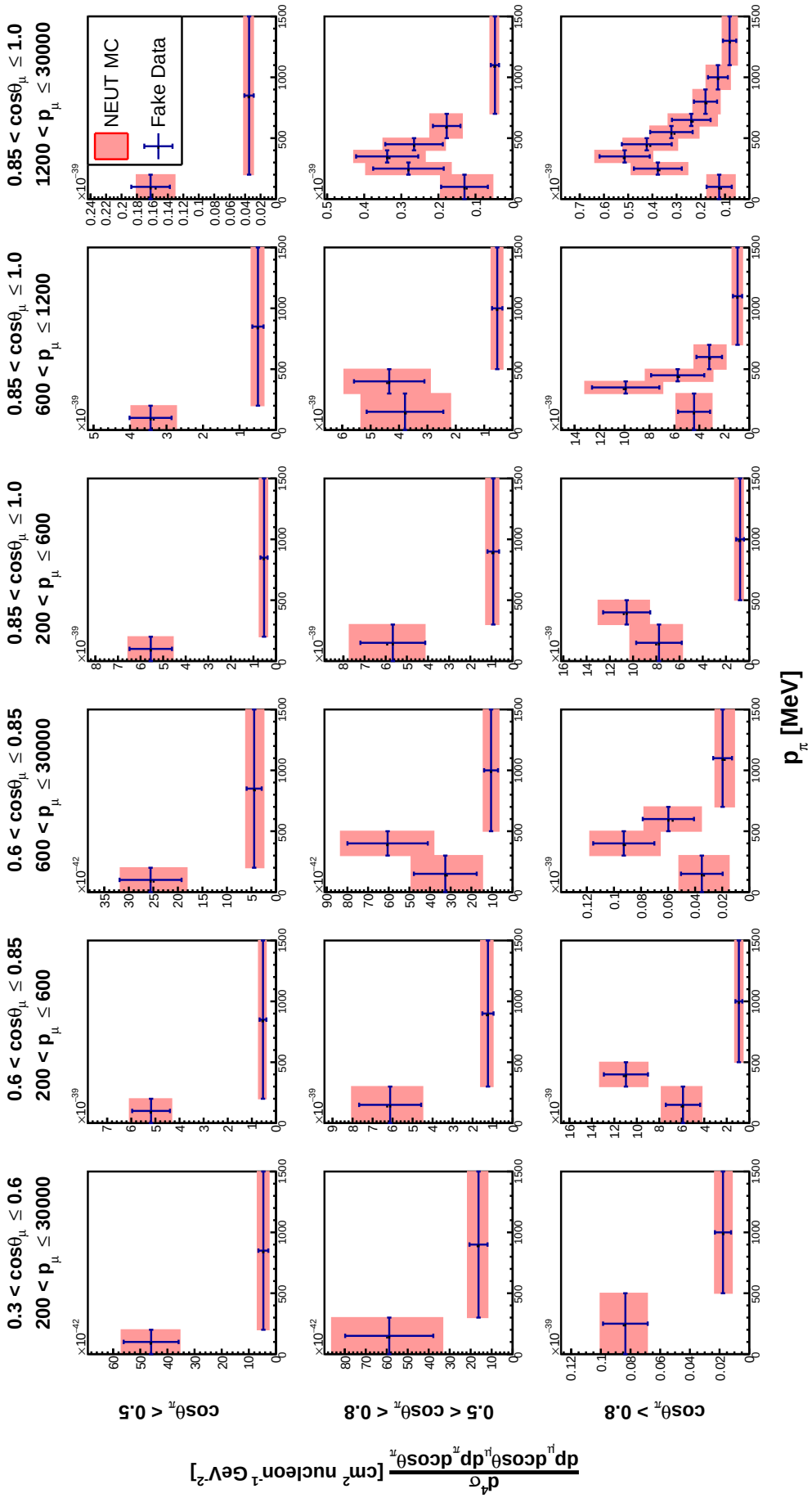


Figure 8.18: 4D differential cross-section result for interactions on hydrocarbon, for DIS enhanced fake data.

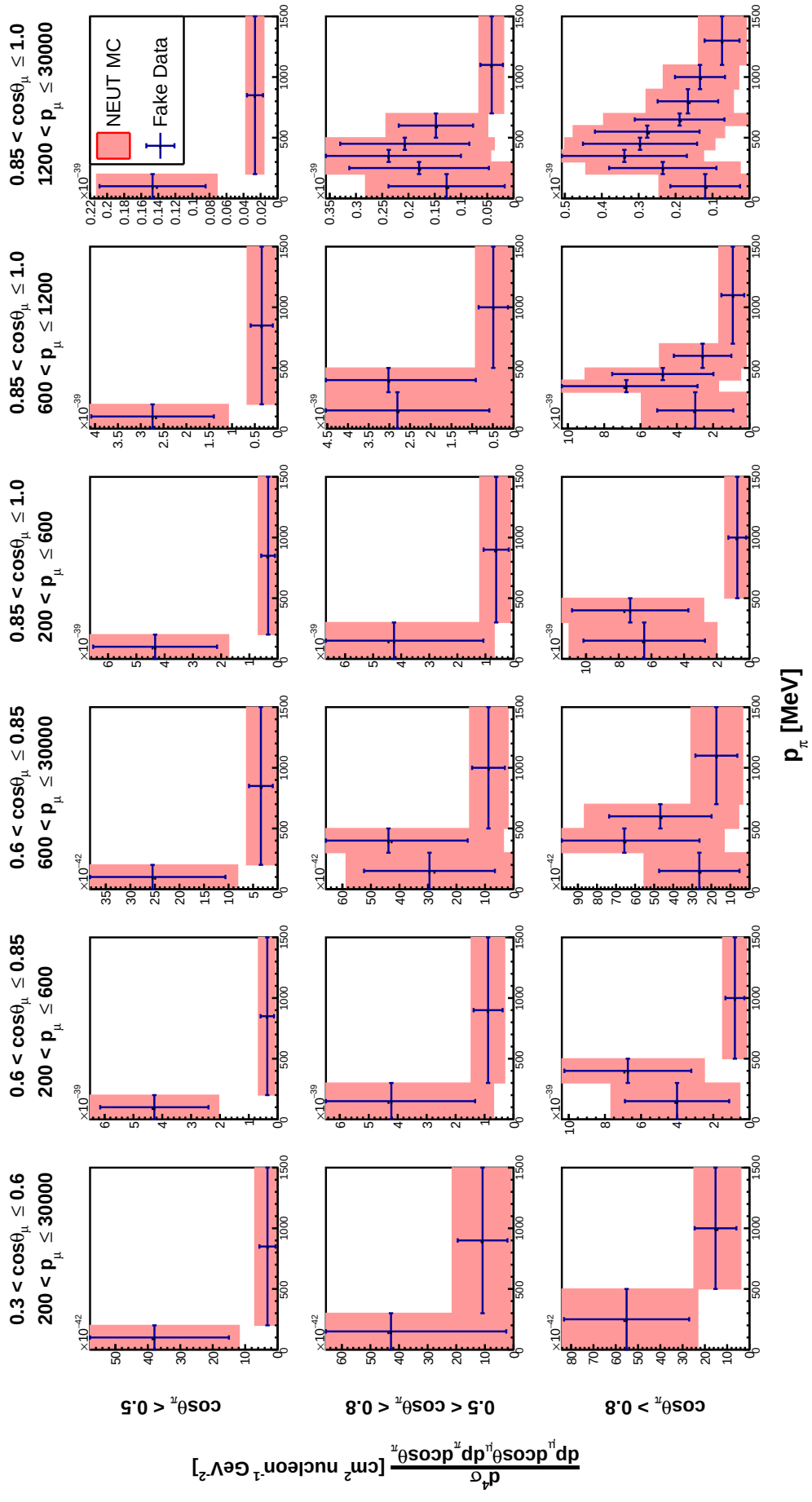
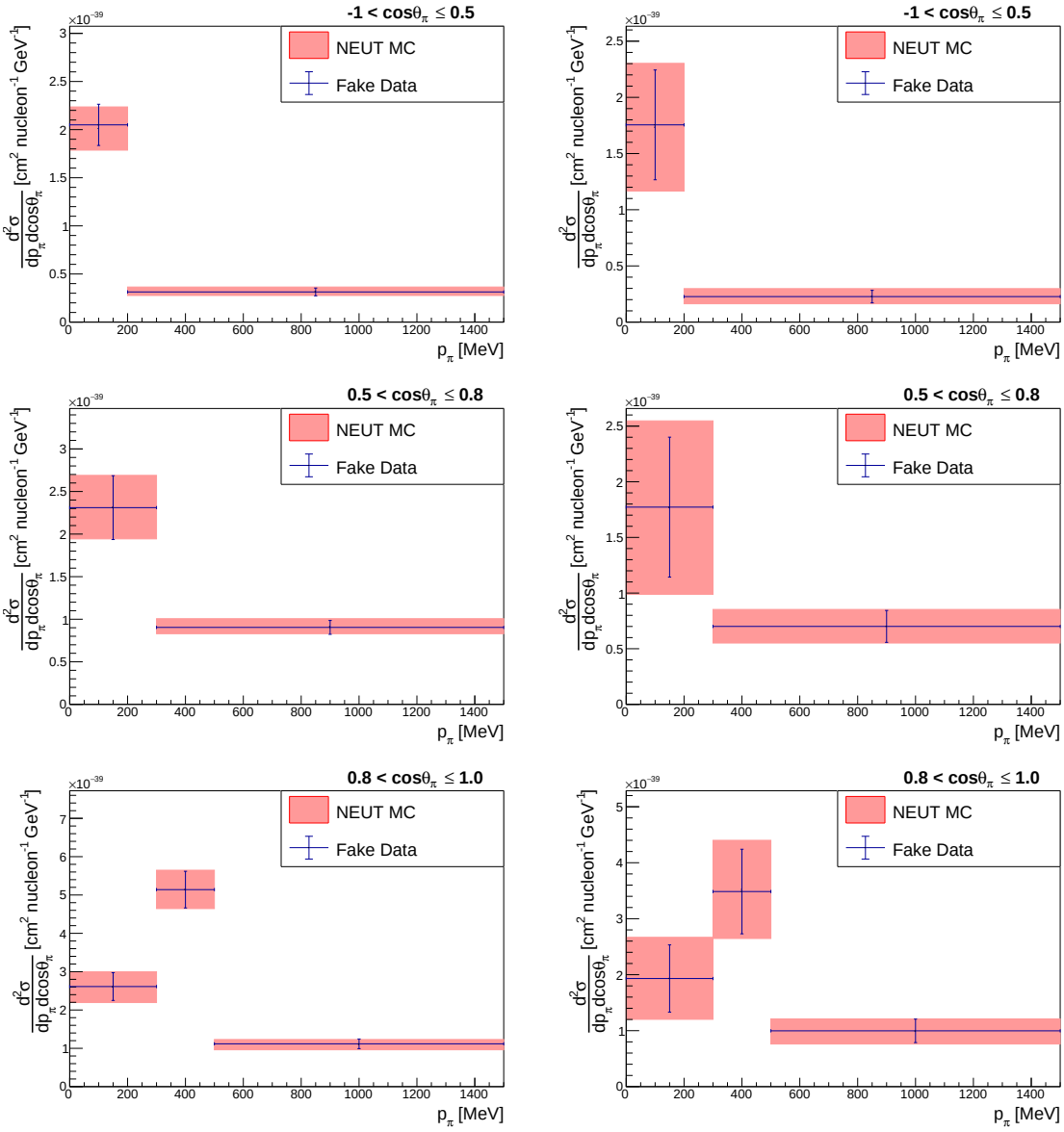


Figure 8.19: 4D differential cross-section result for interactions on water, for DIS enhanced fake data.



(a) CH (b) H_2O

Figure 8.20: 2D differential cross section in pion momentum and $\cos\theta$ using DIS enhanced fake data, for interactions on (a) hydrocarbon and (b) water. Phase space is limited to $p_\mu > 200$ MeV and $\cos\theta_\mu > 0.6$.

Chapter 9

Results and Conclusions

This chapter presents the final unfolded cross-section result, from a fit performed to all available ND280 FHC data.

9.1 Reconstructed Event Distributions

Comparisons of the reconstructed event rates in data, with the prior values in red and post-fit in blue, can be seen in Figures 9.1 to 9.3 for the FGD1, FGD2x and FGD2y samples, respectively. Across the distributions for all of these samples, it can be seen that the agreement with data is generally improved for the post-fit value obtained from the likelihood fit, compared to the prior. Certain samples, namely the isoFGD signal samples, still show some disagreement between the data and post-fit values. This is likely due to a combination of the low statistics and poor resolution in these samples, which means migration between neighbouring bins may cause more fluctuation than the fit is currently accounting for.

9.2 Fit Results

As with all previous Asimov and fake data tests, the nominal Monte Carlo distribution is fitted to ND280 data using the log-likelihood minimisation method described in Chapter 5. The post-fit parameter values from this fit are given in Figure 9.5. As we have come to expect from the statistically varied fits performed in Section 8.3.1, the post-fit template parameter values fluctuate around the prior values of one, with only a few taking negative values. While this will cause some issues in the full 4D cross-section extraction, the integration over muon bins to obtain the double-differential measurement in pion kinematics should resolve most of these issues. The flux parameters all show a systematic shift upwards, with most taking a post-fit value of 1.1, or slightly higher in the highest energy bins. This is then reflected in the detector parameter values; while the majority of these take post-fit values higher than that of the prior, the values associated with the higher energy multiple pion sidebands show the greatest increase. The majority of the cross-section dial values stay around their prior values, but in the dials that do move there is some considerable movement. This is most notable in the Bodek-Yang MultiPi and the high energy inelastic FSI dials, which increase and reduce by $\sim$50% respectively. There is also significant movement in the CCCOH dial,

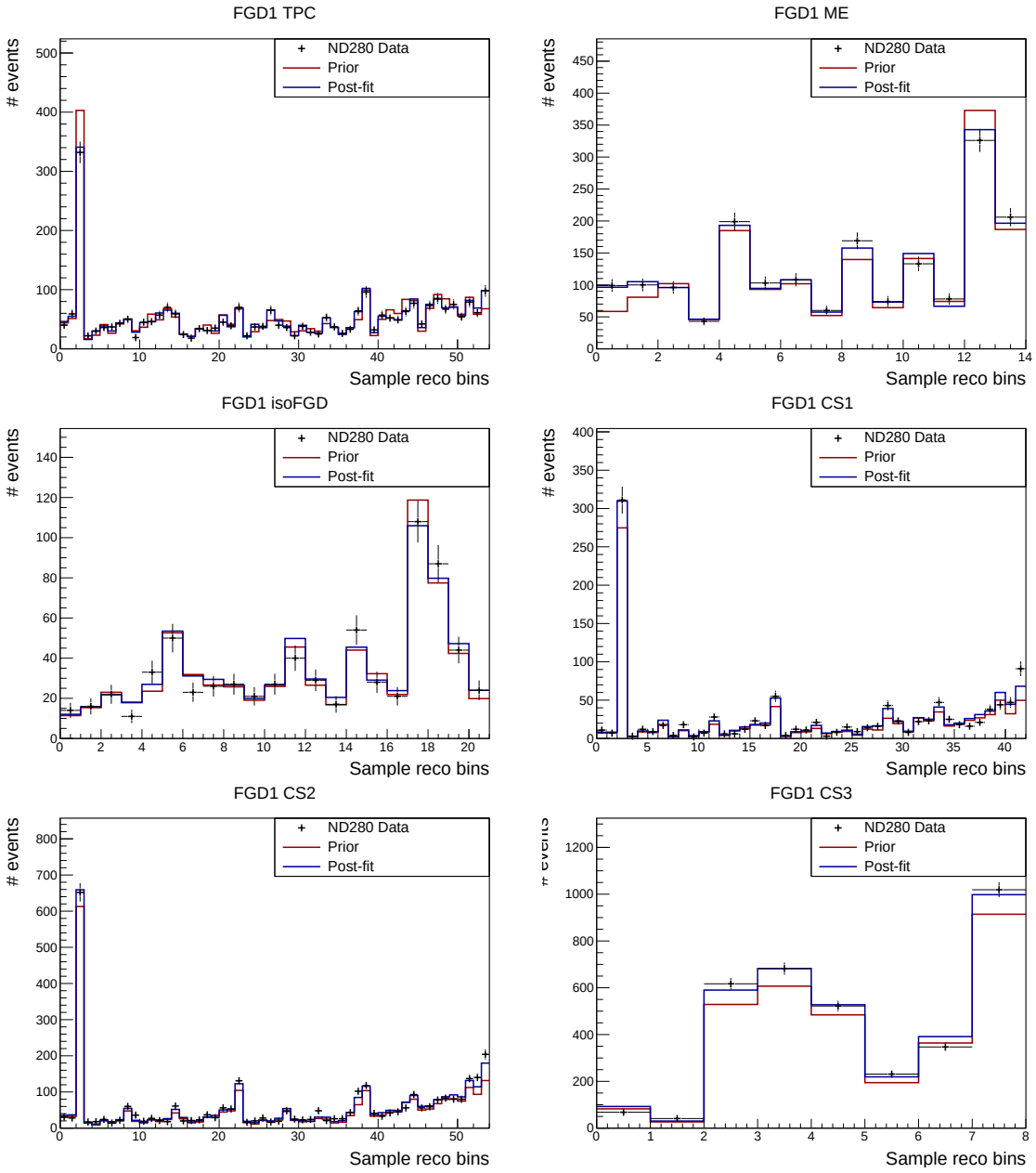


Figure 9.1: Prior (red), post-fit (blue) and fake data (black crosses) reconstructed event distributions for FGD1 samples.

which takes a post-fit value just over 100% greater than its prior, suggesting that coherent pion production is underestimated in the model used. The movement of the DIS, multiple pion and higher energy FSI dials is to be expected, as the higher energy flux parameters are raised more, which are where these interactions are more likely to occur.

In the fit to real data it is also interesting to examine the size of the nuisance parameter pulls, which are shown in Figure 9.4, again for the 20 parameters with the highest post-fit constraint. It is interesting to note that, out of 614 total nuisance parameters, 19 of those in the highest 20 are the same as for the Asimov fit, albeit with a different ordering. Examining the pull values, it can easily be seen that several of these parameters get pulled to their limits, with a couple completely pulled out of their prior constraints. This mostly happens

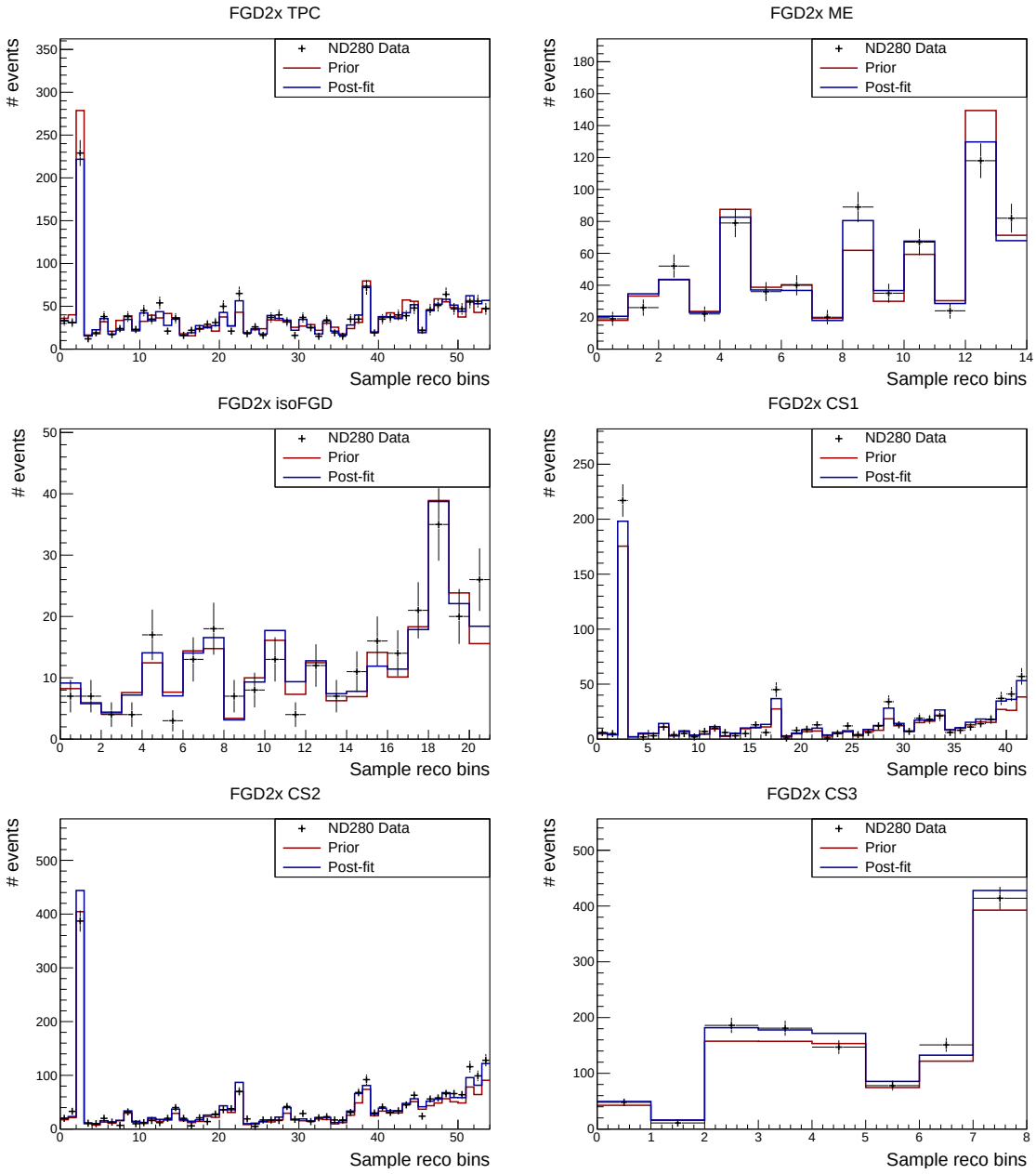


Figure 9.2: Prior (red), post-fit (blue) and fake data (black crosses) reconstructed event distributions for FGD2x samples.

for the cross-section parameters, and suggests that further freedoms are needed to properly explain the data observed.

Table 9.I gives the contribution to the total χ^2 value from each possible source. This shows that the majority of the disagreement between MC and data comes from the side-band samples, although the disagreement in the signal samples also increases as we move from FGD1 samples, through FGD2x and finally to the FGD2y samples. This is to be somewhat expected, as the FGD2y samples are the most statistically limited. Examining the contributions to the χ^2 total confirms that the majority of the tension between MC and data is statistical in nature. It can also be seen that the regularisation contribution is the smallest, which was required in an effort to prevent the fit being biased too much

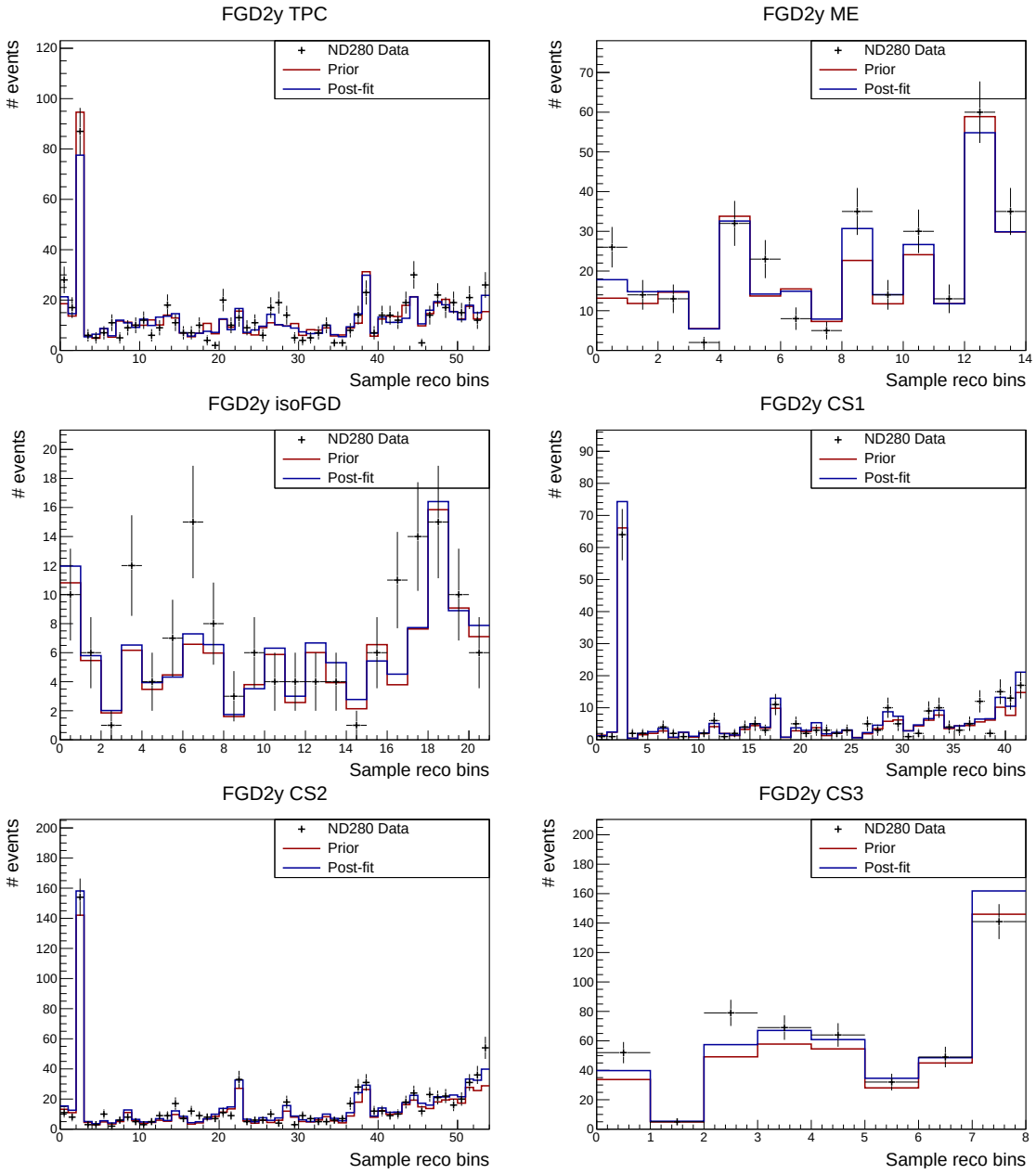


Figure 9.3: Prior (red), post-fit (blue) and fake data (black crosses) reconstructed event distributions for FGD2y samples.

towards the input MC. Using the effective number of degrees of freedom for the fit taken from Section 8.3.2, the χ^2/Ndf for the fit to data is found to be $\frac{681.30}{504.9} \approx 1.35$, which shows reasonable agreement.

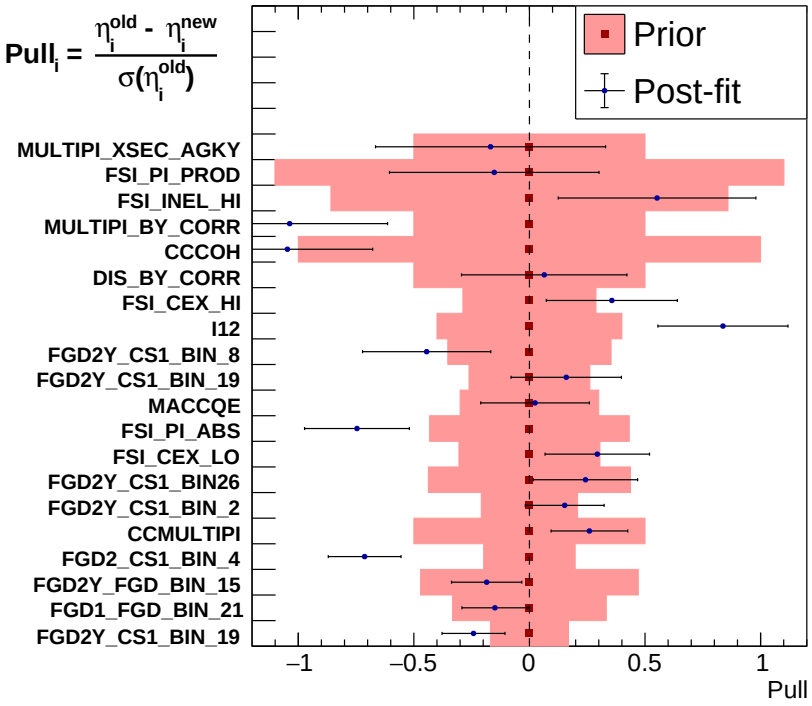


Figure 9.4: Pull values and associated errors for the 20 nuisance parameters with the largest post-fit error, given in order.

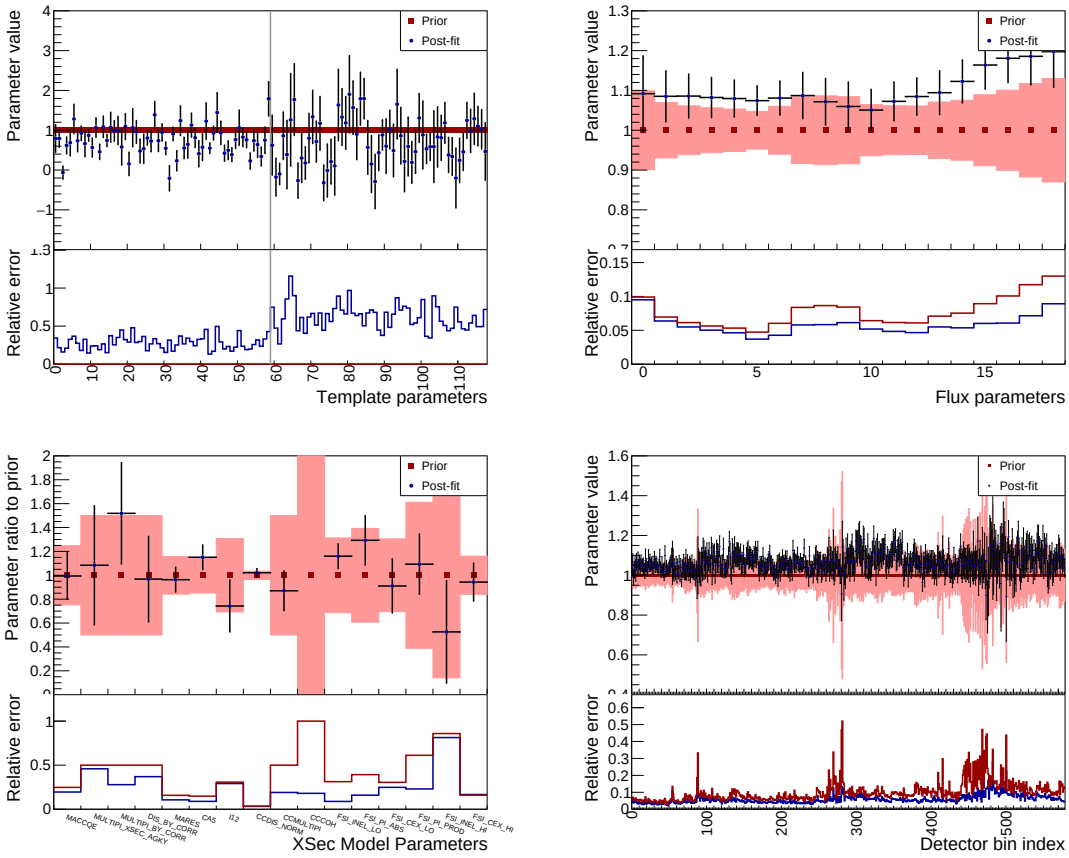


Figure 9.5: Template, flux, cross-section model and detector systematic parameter fit results, for the fit to ND280 data.

Source	χ^2	Ndf	χ^2/Ndf
FGD1 $1\pi^+$-TPC	19.90	54	0.37
FGD1 $1\pi^+$-ME	7.34	14	0.52
FGD1 $1\pi^+$-isoFGD	13.18	21	0.63
FGD1 CS1	49.96	42	1.19
FGD1 CS2	77.05	54	1.43
FGD1 CS3	18.09	8	2.26
FGD2x $1\pi^+$-TPC	24.42	54	0.45
FGD2x $1\pi^+$-ME	10.36	14	0.74
FGD2x $1\pi^+$-isoFGD	22.20	21	1.06
FGD2x CS1	45.96	42	1.09
FGD2x CS2	85.37	54	1.58
FGD2x CS3	9.50	8	1.19
FGD2y $1\pi^+$-TPC	71.78	54	1.33
FGD2y $1\pi^+$-ME	18.51	14	1.32
FGD2y $1\pi^+$-isoFGD	30.64	21	1.46
FGD2y CS1	40.39	42	0.96
FGD2y CS2	66.70	54	1.24
FGD2y CS3	13.91	8	1.74
Flux Parameters	5.08	19	0.27
Xsec Parameters	6.46	16	0.40
Detector Parameters	25.17	579	0.04
Statistical	625.27	–	–
Systematic	36.72	–	–
Regularisation	19.30	–	–
Total	681.30	–	–

Table 9.I: Contributions to the total χ^2 value for the fit to data from all sources. Using the calculated effective number of degrees of freedom, 504.9, the total χ^2/Ndf is found to be 1.35.

9.3 Cross-Section Data Result

Following the same procedure as used throughout Chapter 8, the fit parameters are unfolded to first obtain measurements of the 4-dimensional cross section,

$$\frac{d^4\sigma^\alpha}{dp_\mu d\cos\theta_\mu dp_\pi d\cos\theta_\pi}, \tag{9.1}$$

for interactions on both hydrocarbon and water. Results for these are shown in Figures 9.6 and 9.7, where it can be seen that there is a fair amount of disagreement in the data results with MC prediction. This is particularly apparent in the case of the water measurement, which was expected due to the lower event rates, making it much more susceptible to statistical fluctuations. For the hydrocarbon measurement however, the majority of measurements are seen to agree with MC, within the allowed error bars. To quantify this, the χ^2 value for the full 4D measurement is calculated as

$$\chi^2 = (\vec{\sigma}_{\text{MC}} - \vec{\sigma}_{\text{data}})\, \mathbf{V}_{\text{cov}}^{-\mathbf{1}}\, (\vec{\sigma}_{\text{MC}} - \vec{\sigma}_{\text{data}})^{\mathbf{T}}\,, \tag{9.2}$$

evaluating to 570.74. Normalising this by the effective number of degrees of freedom found in Section 8.3.2 (113.9), gives a χ^2/Ndf of 5.01, which shows a reasonable level of tension between data and MC. Figure 9.8 shows the correlation matrix for this cross-section measurement, where the first 56 bins are for the hydrocarbon measurements, and the latter 56 for water. It can be seen that the CH off-diagonal correlations are generally higher than for H_2O. This is due to the fact that the H_2O cross-section measurement is more statistically limited, which only affects individual bins. In the CH measurement, as statistics are higher, the systematic uncertainties play a larger role, which are expected to be correlated across bins.

After integrating over momentum bins and further restricting the muon phase space to $\cos\theta_\mu > 0.6$ (in addition to the regular phase space constraints), the double differential cross section in pion kinematics,

$$\frac{d^2\sigma^\alpha}{dp_\pi d\cos\theta_\pi}, \tag{9.3}$$

can be reported. This result is shown in Figure 9.9. It is clear that integrating over bins to reduce that effect of statistical fluctuations provides a more stable result; while there is still disagreement between data and MC, there are no bins that report negative values. Using again Equation (9.2), the value of the χ^2 for the ND280 data fit is found to be 55.29. Normalising by the effective number of degrees of freedom, calculated to be 12.83, we find a χ^2/Ndf of 4.31. This again shows reasonable tension between data and MC, but that there is more agreement when statistical fluctuations are integrated out. Finally, the cross-section correlation matrix for the double differential result is provided in Figure 9.10. This shows similar behaviour to that of the 4D result, with higher correlation between bins in the CH sample due to systematic effects.

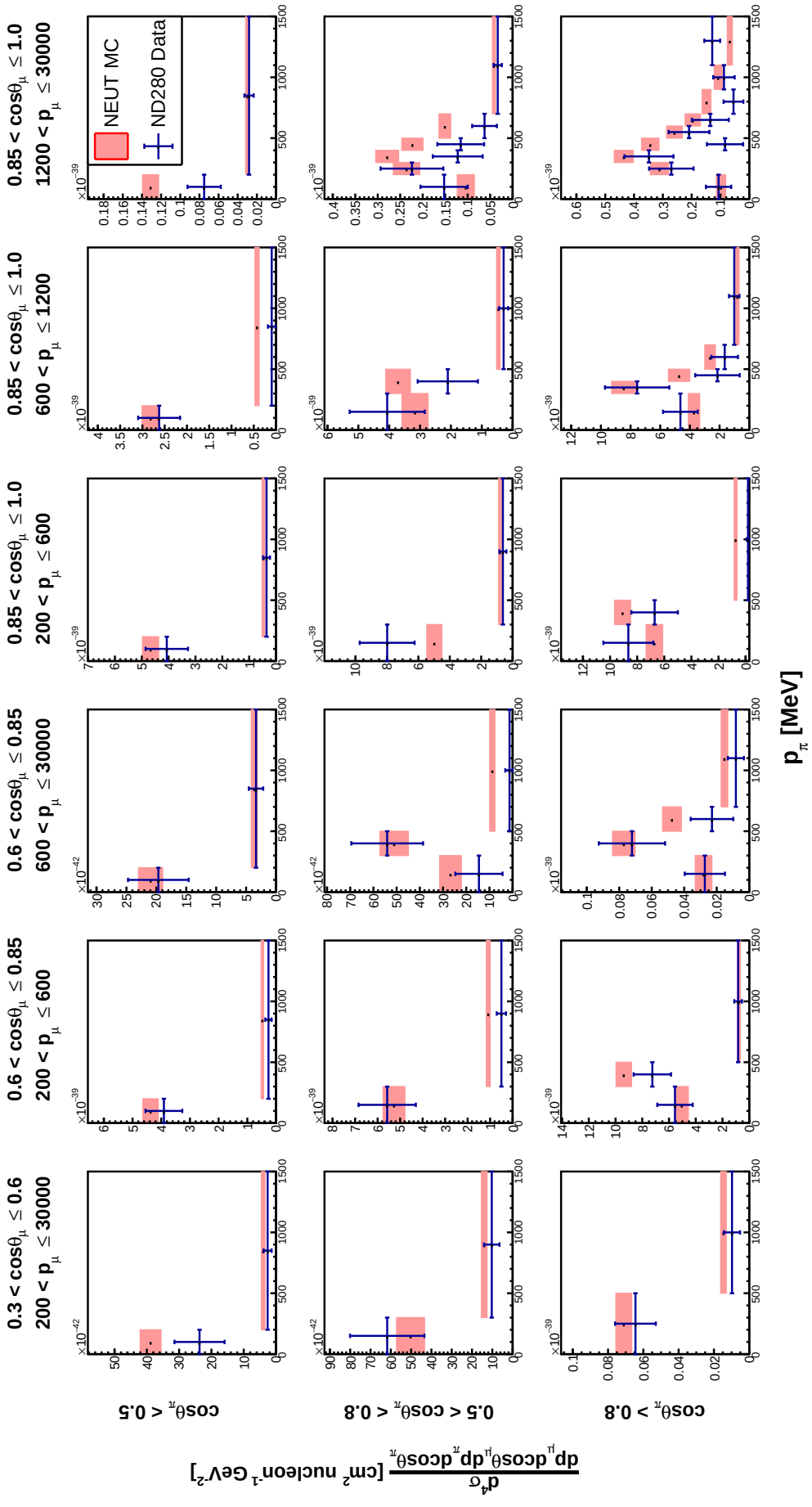


Figure 9.6: 4D differential cross-section result for interactions on hydrocarbon, for ND280 data.

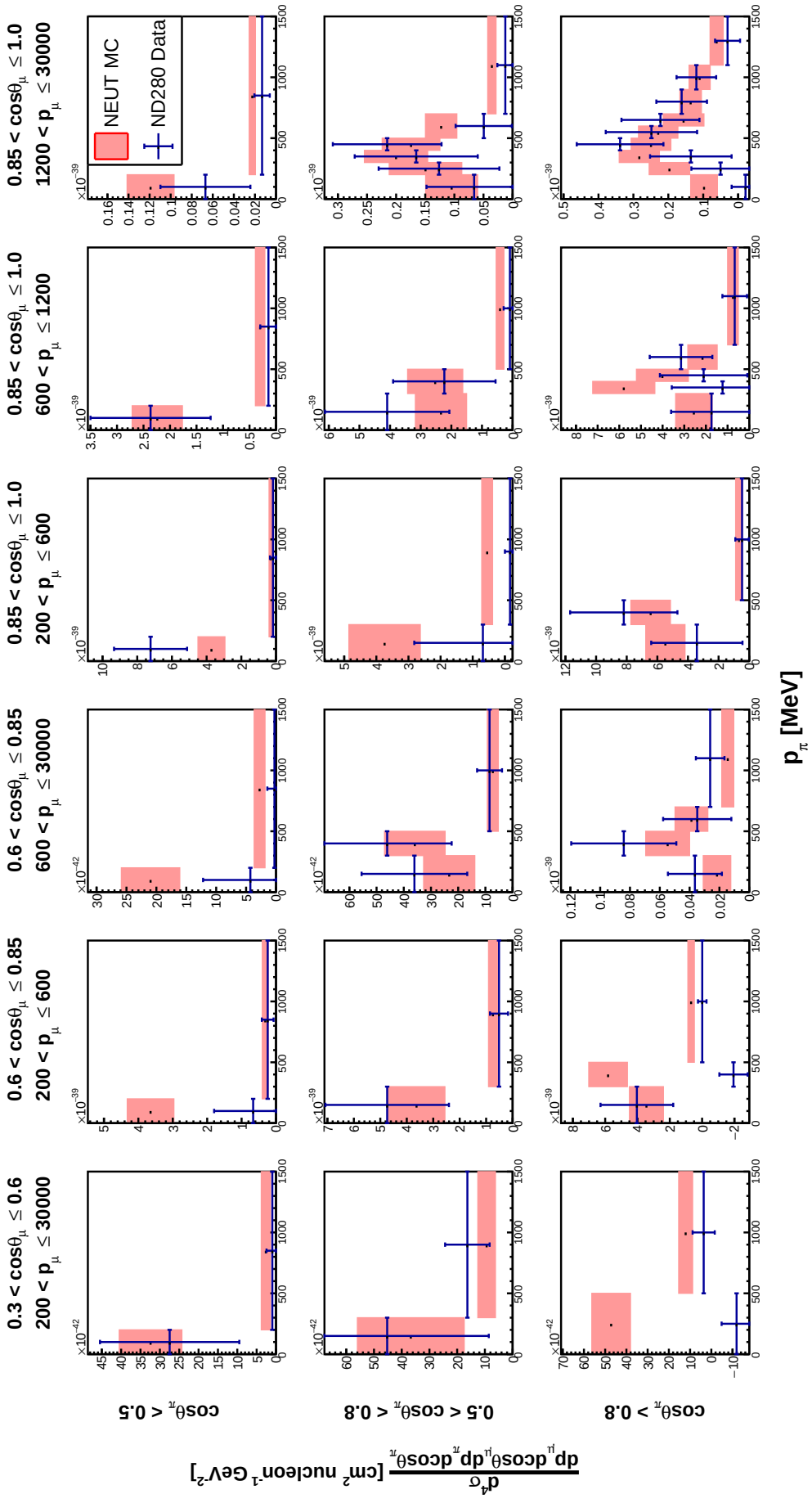


Figure 9.7: 4D differential cross-section result for interactions on water, for ND280 data.

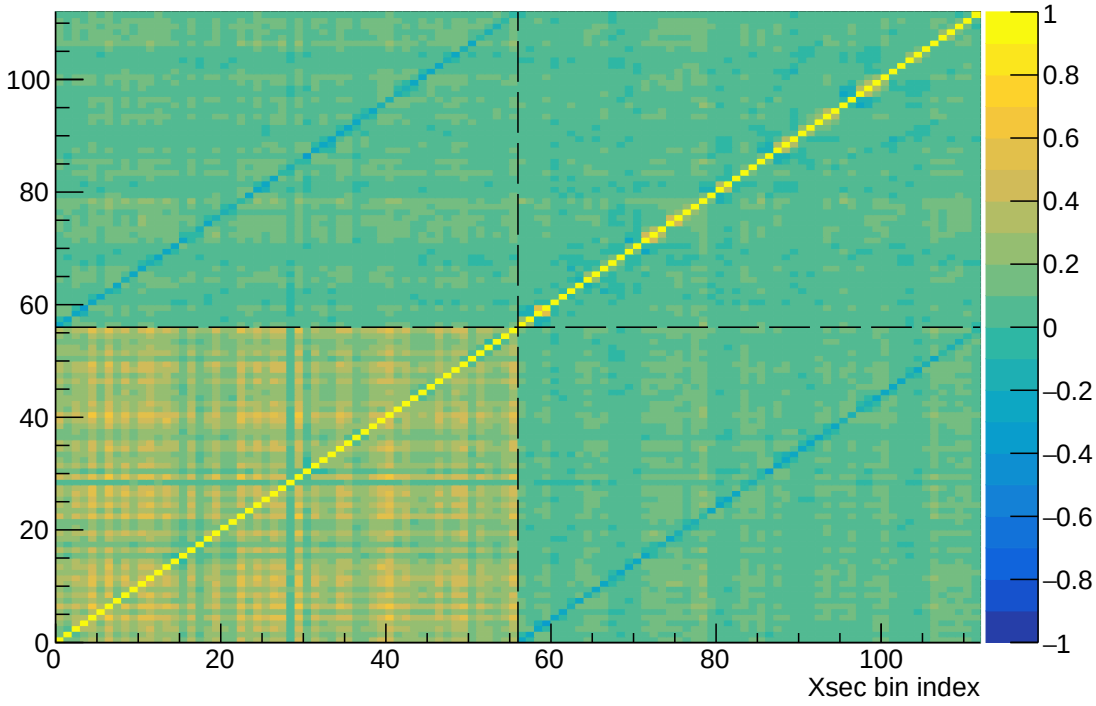
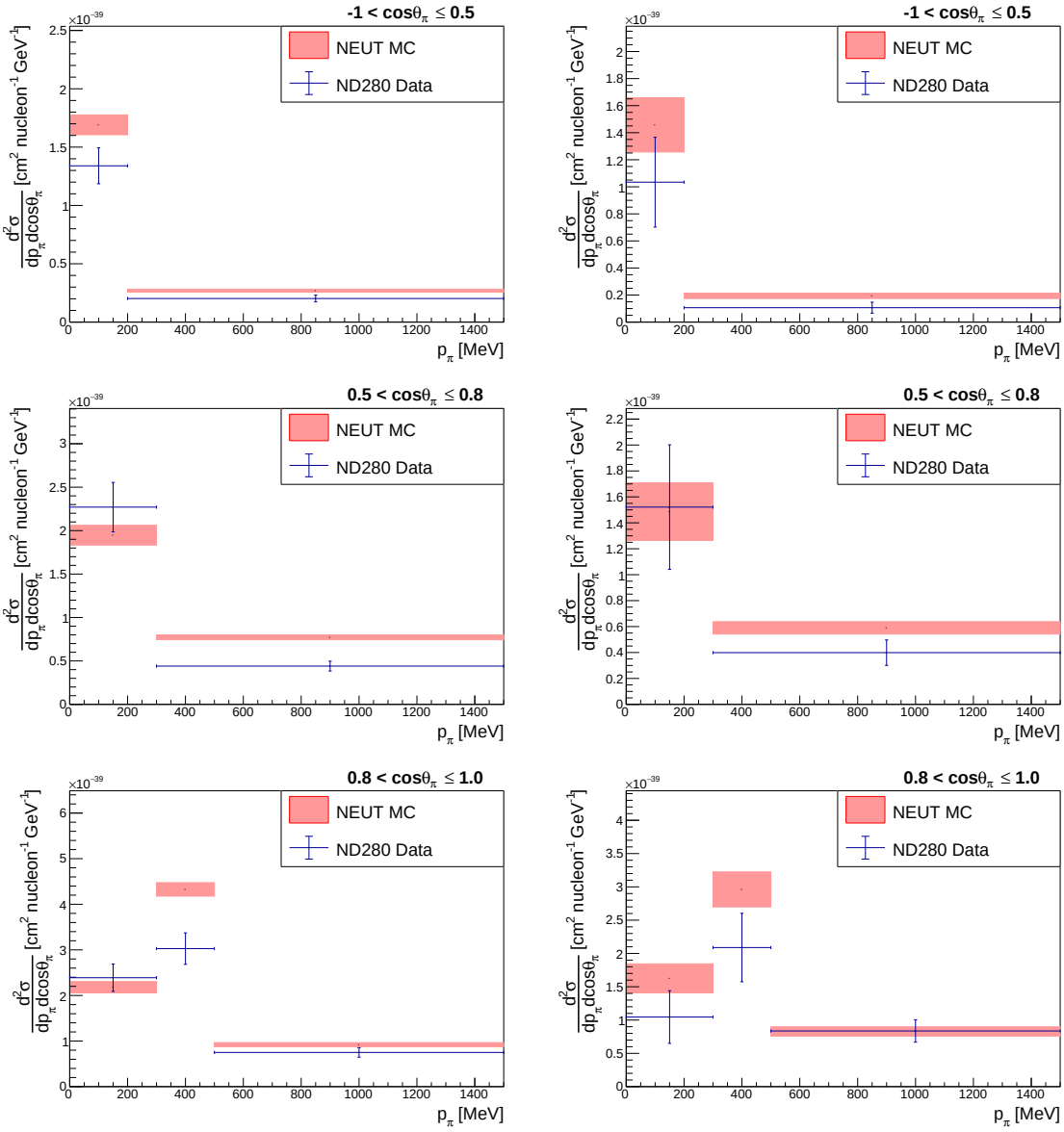


Figure 9.8: Cross-section correlation matrix for the 4D differential result, where the first 56 bins correspond to hydrocarbon, and the second 56 to water.



(a) CH (b) H_2O

Figure 9.9: 2D differential cross section in pion momentum and $\cos\theta$ using ND280 data, for interactions on (a) hydrocarbon and (b) water. Phase space is limited to $p_\mu > 200$ MeV and $\cos\theta_\mu > 0.6$.

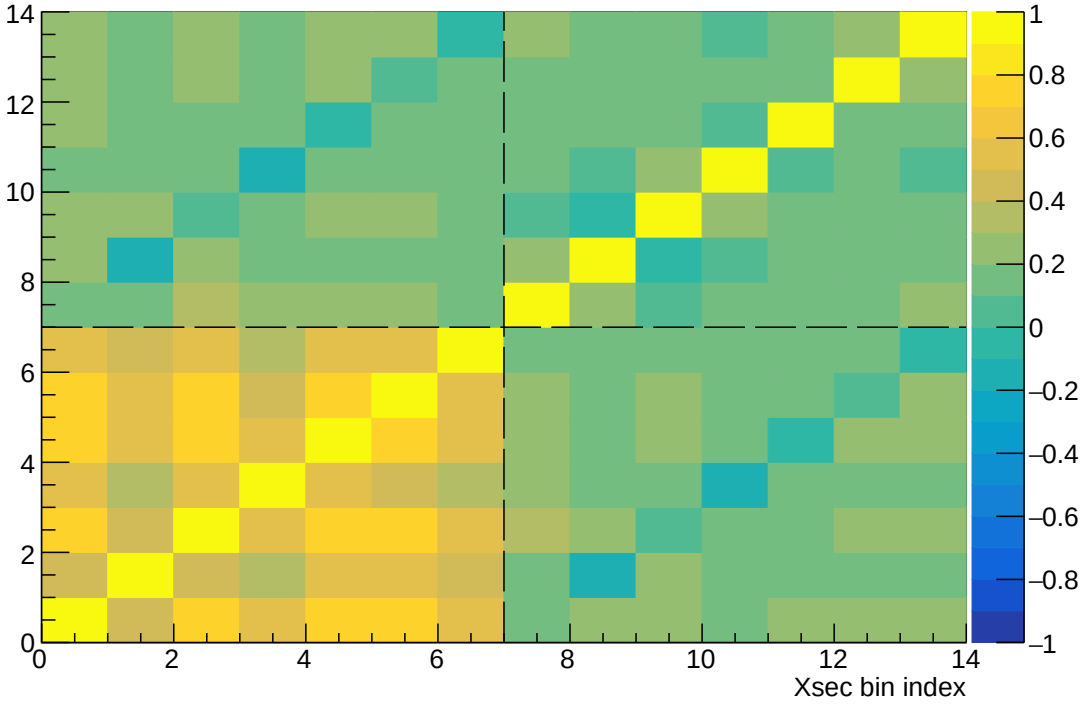


Figure 9.10: Cross-section correlation matrix for the 2D differential result in pion kinematics, where the first 7 bins correspond to hydrocarbon, and the second 7 to water.

9.4 Conclusions and Future Improvements

This work has presented a complete measurement of the ν_μCC1π^+ interaction cross section on both hydrocarbon and water, and is the first such measurement to be performed as both a 4-dimensional differential measurement in muon and pion kinematics, and a double differential measurement as a function of pion kinematics. This is achieved by performing a simultaneous log-likelihood fit to data, for both targets in 4-dimensions, using all available ND280 ν_μ data, and then integrating over momentum bins in order to extract the double differential results. The events input to the fit are selected through a refined procedure in comparison to previous measurements. The measurement also provides access to the low-momentum pion region, a first for T2K, through the use of the developed method for reconstructing the kinematics of pions which rapidly decay to muons and subsequently Michel electrons. This is a particularly important kinematic region to understand, as pions in this momentum regime also fall below the Cherenkov threshold of the far detector, Super-Kamiokande.

The unfolded cross-section results show clear tension between data and the input Monte Carlo model, with the 4D and 2D measurements giving χ^2/Ndf values of 5.01 and 4.31 respectively. While it is reasonable to assume that some of the observed disagreement in the 4D result is caused by statistical fluctuations due to how fine the binning scheme is, a clear disagreement is still observed when integrating down to the 2D result. This suggests that the current interaction model used in the `NEUT` generator is not fully representative of the effects observed in data. In particular, the fact that the fit to data sees the CCCOH dial parameter being pulled to a value ∼100% higher than the prior value suggests that the impact of coherent interactions is somewhat underestimated in the nominal MC.

Whilst this work represents a novel measurement for T2K, significant improvements to the analysis are already foreseen. In the time that this analysis has been finalised, the `T2KReWeight` framework has been largely overhauled, with multiple new cross-section dials now available. Of particular interest to this analysis are new single pion and proton production dials, which will allow the fit additional freedom with respect to the interaction modelling. Although performing the cross-section extraction in the fine 4D binning scheme is preferable for the purpose of efficiency corrections, it is clear that the reported 4D measurements are limited by statistical fluctuations, particularly for a water target, and this somewhat limits the power of the result to resolve model differences. Starting from this 4D form, multiple different results can be obtained, by integrating over different kinematic dimensions in the same manner as for the reported 2D result. It would be particularly interesting to obtain double differential results for both muon kinematics and pion kinematics separately, as it is expected that the hadronic side of the interaction will show more sensitivity to nuclear effects than the leptonic side. Comparing both of these measurements to different interaction models, such as the recent MK pion production model discussed in Section 2.3, would likely provide some interesting model discrimination. Finally, a simple extension of this analysis would be to additionally report a measurement of the cross-section

ratio, calculated as

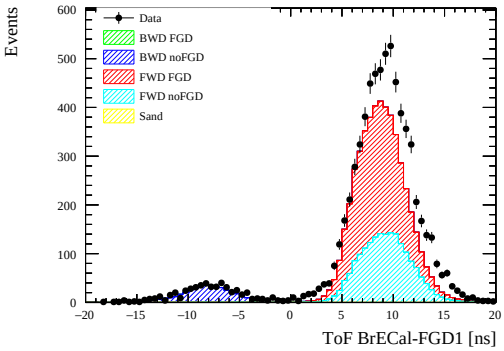
$$\left(\frac{d\sigma^{\mathrm{H_2O}}}{d\sigma^{\mathrm{CH}}}\right) = \frac{N^{\text{signal, H}_2\text{O}}}{N^{\text{signal, CH}}} \times \frac{\epsilon^{\mathrm{CH}} N_T^{\mathrm{CH}}}{\epsilon^{\mathrm{H_2O}} N_T^{\mathrm{H_2O}}}. \tag{9.4}$$

This has the advantage of removing the dependence on flux and the contribution of its uncertainty. Additionally, some of the systematic errors will cancel out due to correlations, further reducing the error on the measurement.

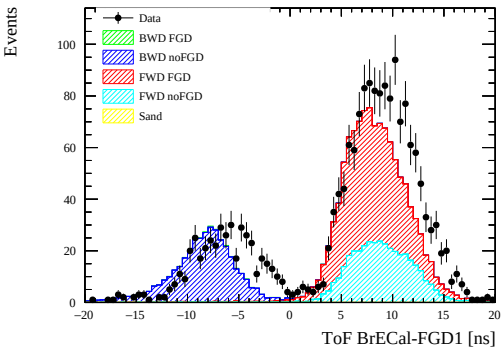
Looking to the future, the ND280 Upgrade planned for 2022 is expected to provide further improvements over the initial measurement presented here. Although the developed kinematic reconstruction of pions from Michel electrons provides a first opportunity to access the low momentum pion region as described above, it is clear that the reconstruction is not perfect. This is particularly apparent in the case of the angular reconstruction, which was found to have very low resolution due to the limited capacity of the FGDs to accurately reconstruct the Michel electron vertex. With the inclusion of the SuperFGD as part of the ND280 Upgrade, which will provide full 3-dimensional tracking through its use of hydrocarbon scintillating cubes, the resolution of the pion kinematic reconstruction from Michel electrons is expected to improve considerably. In addition to the improved reconstruction, the increased target mass of the upgraded ND280 detector will provide a greater rate of events to be used in such an analysis. Combining this with both the foreseen J-PARC beam upgrade necessary for the Tokai to Hyper-Kamiokande (T2HK) experiment, and the expected years of data taking as T2K transitions to T2HK, 4D cross-section measurements such as the ones presented in this analysis will be possible without the observed statistical limitations.

Appendix A

Run 3 ToF Uncorrected Data–MC Comparisons

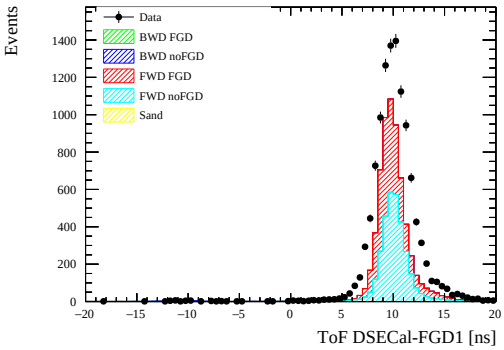


(a) Track-like

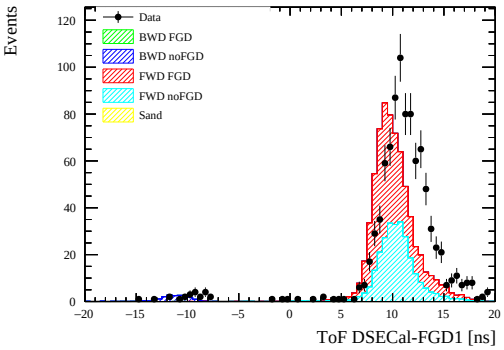


(b) Shower-like

Figure A.1: Run 3 ToF distribution in data and MC for forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like



(b) Shower-like

Figure A.2: Run 3 ToF distribution in data and MC for forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.

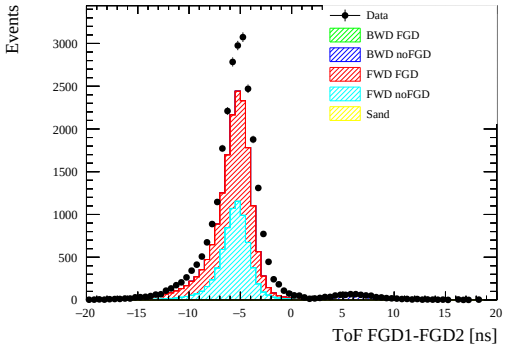
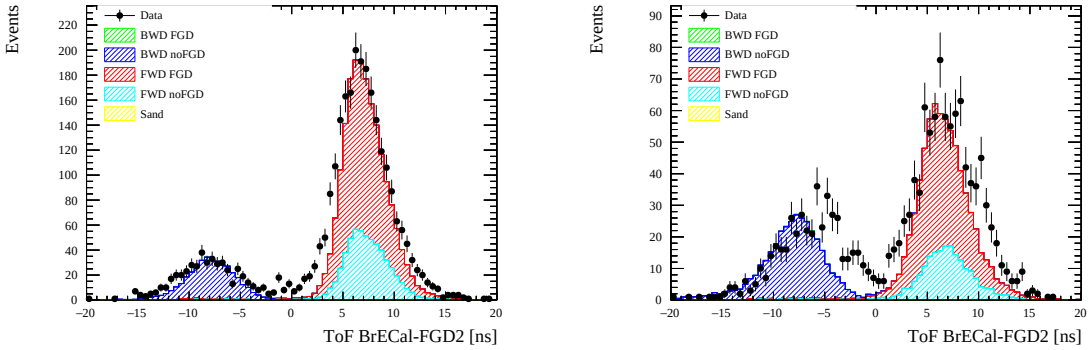
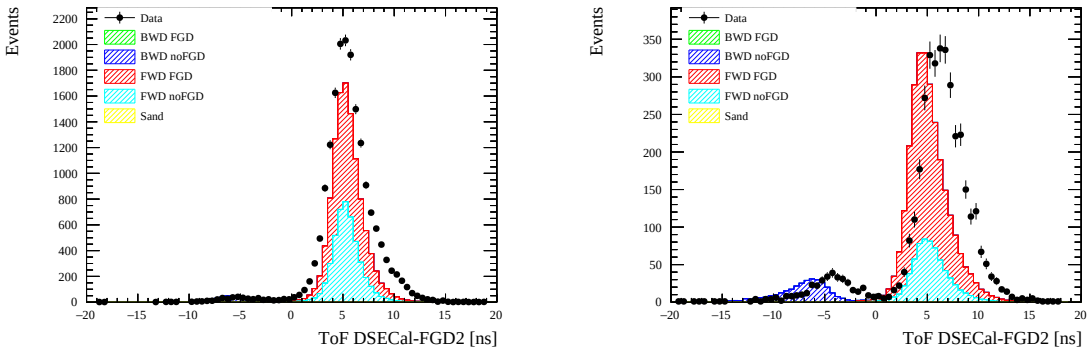


Figure A.3: Run 3 ToF distribution in data and MC for forward-going events between FGD1 and FGD2.



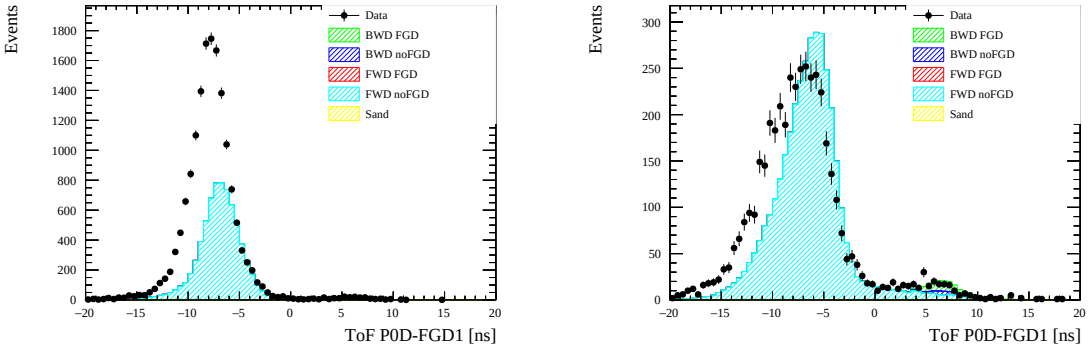
(a) Track-like (b) Shower-like

Figure A.4: Run 3 ToF distribution in data and MC for forward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.



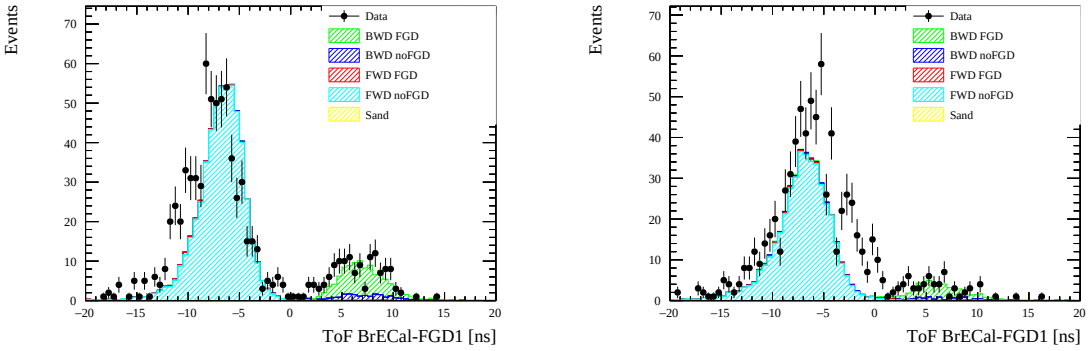
(a) Track-like (b) Shower-like

Figure A.5: Run 3 ToF distribution in data and MC for forward-going events between FGD2 and DsECal. Samples are split into (a) track-like and (b) shower-like.



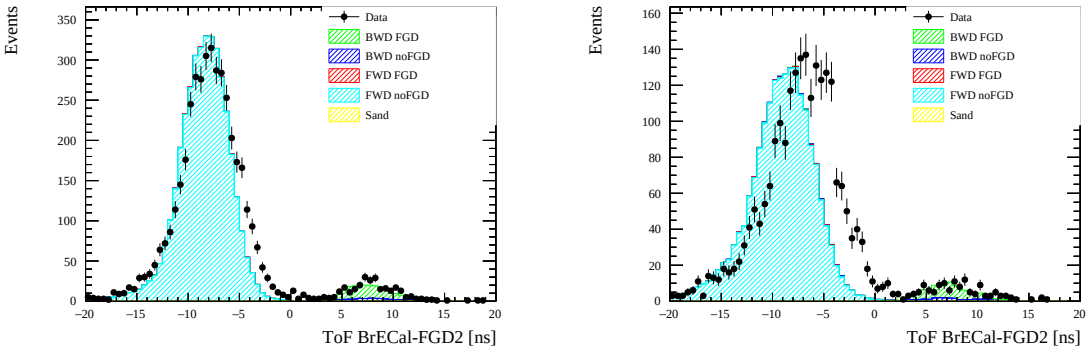
(a) Track-like (b) Shower-like

Figure A.6: Run 3 ToF distribution in data and MC for backward-going events between FGD1 and PØD. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure A.7: Run 3 ToF distribution in data and MC for backward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure A.8: Run 3 ToF distribution in data and MC for backward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.

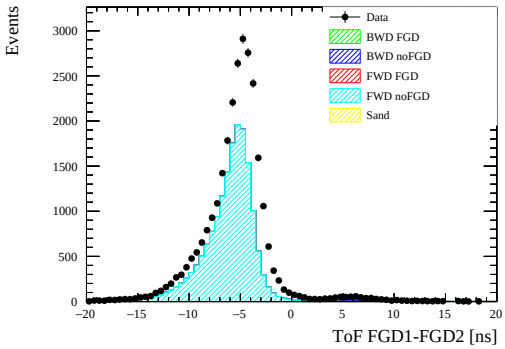
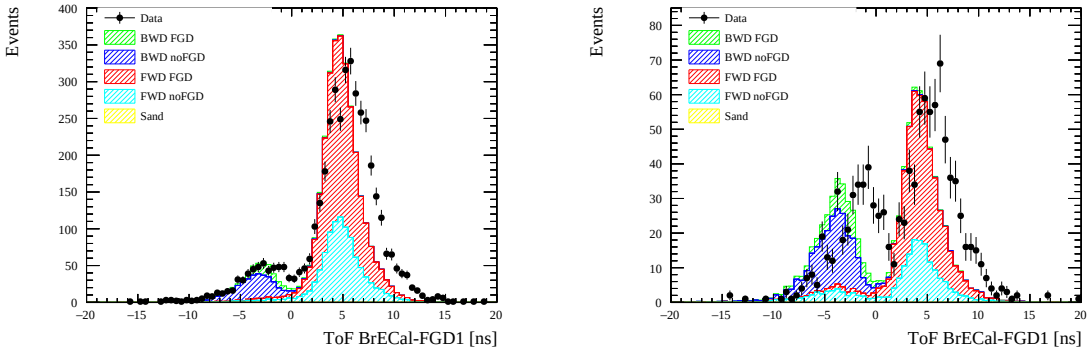
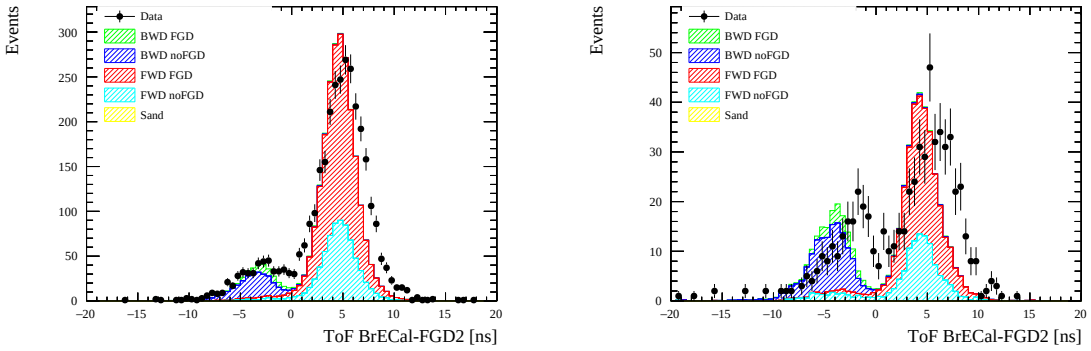


Figure A.9: Run 3 ToF distribution in data and MC for backward-going events between FGD1 and FGD2.



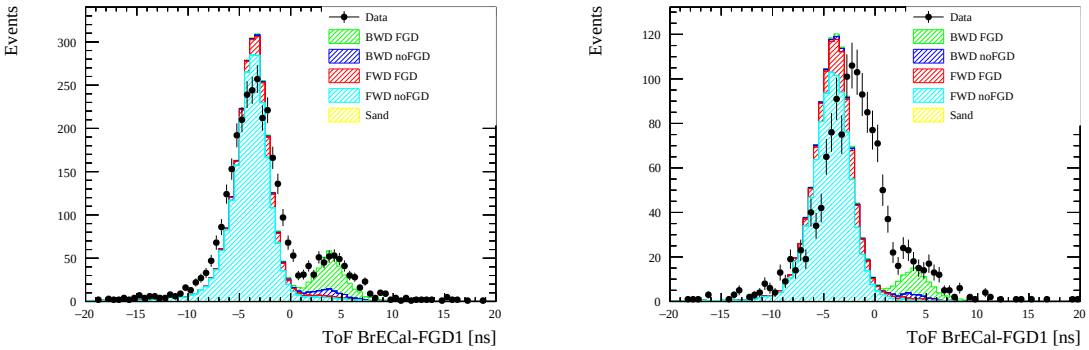
(a) Track-like (b) Shower-like

Figure A.10: Run 3 ToF distribution in data and MC for high-angle forward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



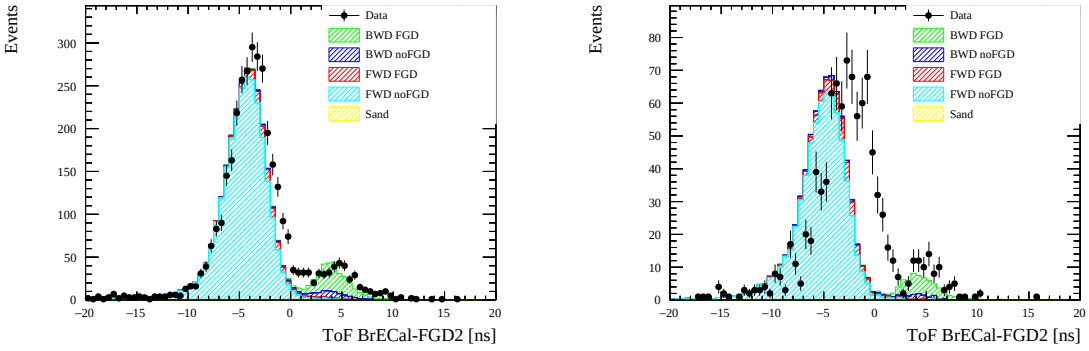
(a) Track-like (b) Shower-like

Figure A.11: Run 3 ToF distribution in data and MC for high-angle forward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure A.12: Run 3 ToF distribution in data and MC for high-angle backward-going events between FGD1 and BrECal. Samples are split into (a) track-like and (b) shower-like.



(a) Track-like (b) Shower-like

Figure A.13: Run 3 ToF distribution in data and MC for high-angle backward-going events between FGD2 and BrECal. Samples are split into (a) track-like and (b) shower-like.

Appendix B

ECal MIPEM Cut Optimisation

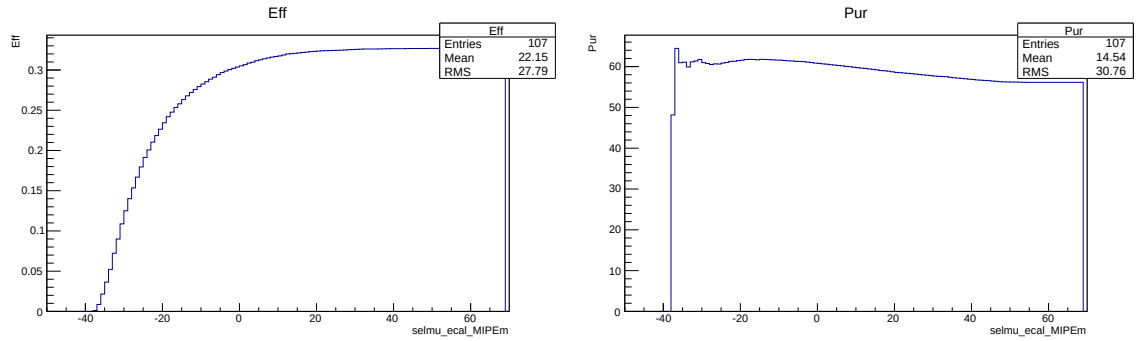
Before settling on a cut value of 0 on the ECal MIPEM variable in order to remove negative pion background from the muon candidate main track, a study was performed to determine the optimal value at which to cut. Using the FGD2 selection (which at the time was the only one implemented into the analysis), integer values of MIPEM were stepped through, cutting out events with a value greater than the cut value. For each integer value of the cut, the efficiency and purity post-cut were calculated. In doing so, two figures of merit were calculated, which were then used to choose the optimal cut value. The first figure of merit used is the standard one used throughout the development of the work given in Section 6.2, $\epsilon \times \rho$. The second is defined as

$$\text{FoM} = \frac{S}{\sqrt{S+B}}, \tag{B.1}$$

where S is the number of true signal events in the selection after the cut, and B is the number of true background events. This figure of merit is mostly used for low statistics selections, as it prioritises signal events over background, but is still of use to higher statistics analyses such as this one.

Figures B.1a and B.1b show the efficiency and purity values respectively, for each integer at which a cut is placed, accepting events with MIPEM values below the value in question. As expected, the efficiency rises rapidly from 0 at the most negative values of MIPEM, where almost all events would be removed by the cut. The efficiency distribution then plateaus at a cut value of slightly below 20, at which point very few events are being cut out. For the purity distribution, the maximal purity occurs at one of the most negative cut values, caused by the majority of events being cut out, and a high proportion of remaining events being true signal. From this it can be seen that the cut value at which the maximal efficiency occurs is generally very different to that at which the maximal purity occurs, and hence choosing an optimal cut value relies on balancing these two quantities.

Figure B.2 shows the distribution of $\epsilon \times \rho$ for the range of values over which the study is performed, with the inset plot showing the distribution along a smaller range around the maximal value. This study found that the optimal value at which to cut at is 12, which corresponds to an $\epsilon \times \rho$ of $\sim$19.05%. Although the second figure of merit, $\frac{S}{\sqrt{S+B}}$, is more suited to low statistics samples, it was also checked for the purposes of this cut, and the



(a) Efficiency (b) Purity

Figure B.1: Efficiency (a) and purity (b) values as a function of the value at which a cut on the muon candidate ECal MIPEM variable is placed. Cut removes all events with a MIPEM value above the cut value, and study is shown in for FGD2 selection only.

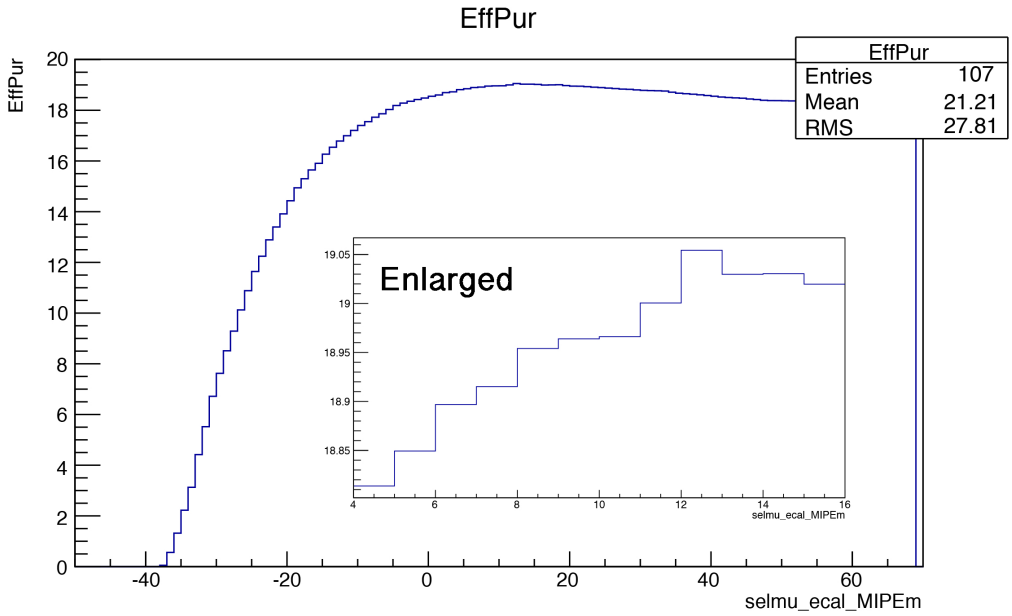
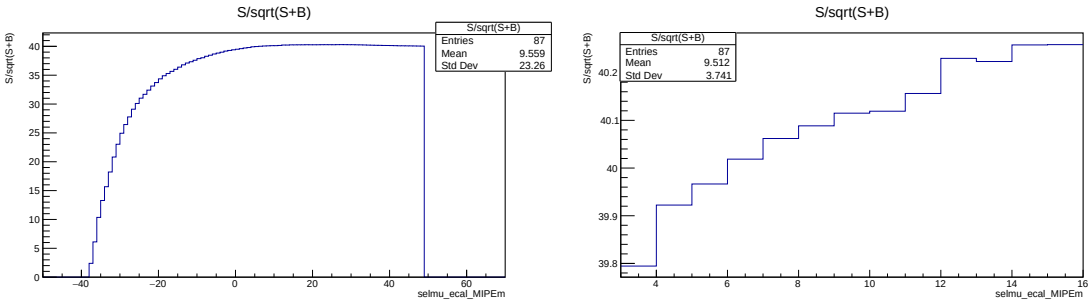


Figure B.2: Distribution of $\epsilon \times \rho$ for each value of MIPEM at which the cut was tested. (a) shows the distribution across the full range of MIPEM values, whereas (b) shows the range around the previously identified optimal cut value of 12.

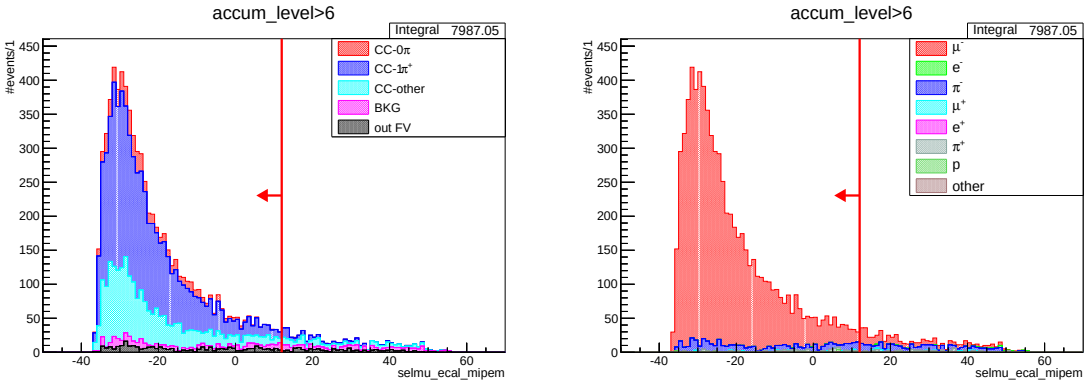
resultant distribution can be seen in Figure B.3. As would be expected from a figure of merit that prioritises event selection over background rejection, this study finds an optimal value at 16, marginally higher than that of the $\epsilon \times \rho$ optimisation. For this reason, 12 was chosen as the optimal value at which a cut should be placed.

Figures B.4a and B.4b show the ECal MIPEM distribution of events before the cut is applied, categorised by true event topology and muon candidate particle type respectively. The intended cut line at 12 is shown, with a very small amount of signal events being removed in comparison to the background.



(a) Full range (b) Close-up

Figure B.3: $\frac{S}{\sqrt{S+B}}$ values as a function of the value at which a cut on the muon candidate ECal MIPEM variable is placed. Cut removes all events with a MIPEM value above the cut value, and study shown is for FGD2 selection only.

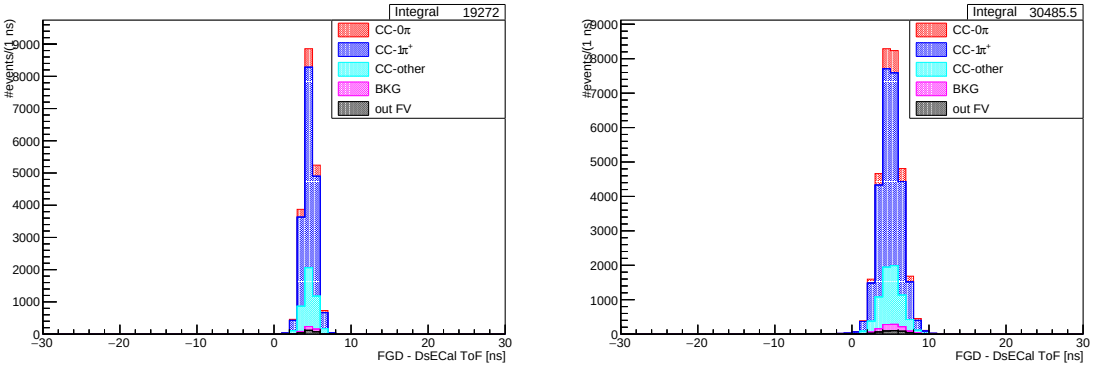


(a) Topology (b) Particle

Figure B.4: ECal MIPEM distribution for the selected muon candidate prior to applying the MIPEM cut, broken down as a function of (a) true topology or (b) true particle type. Cut line shows the optimal cut at 12, with all events to the right of the line rejected, which can be seen to be a very low amount of signal compared to background.

Appendix C

Supplementary Time of Flight Cut Plots

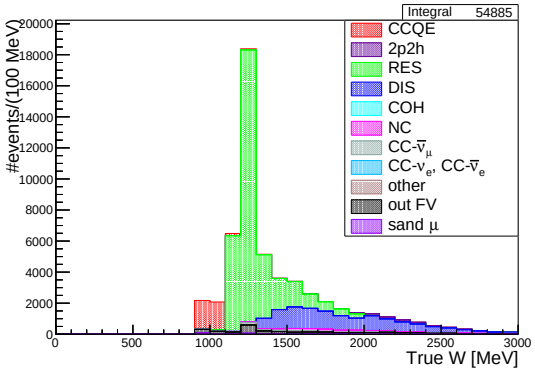


(a) FGD1 (b) FGD2

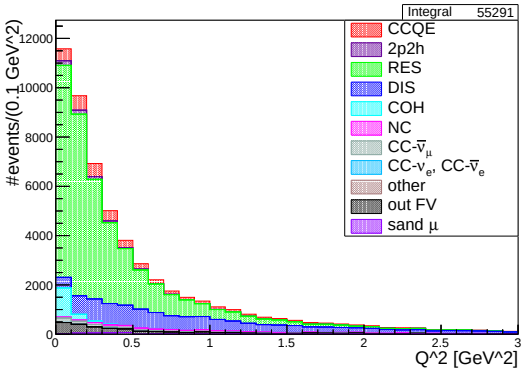
Figure C.1: Time of flight distribution for the μ^- candidate (main track), between (a) FGD1 and the Downstream ECal, and (b) FGD2 and Downstream ECal. Distribution is shown for total CC1π^+ sample. Little to no events are observed in the negative ToF region below zero.

Appendix D

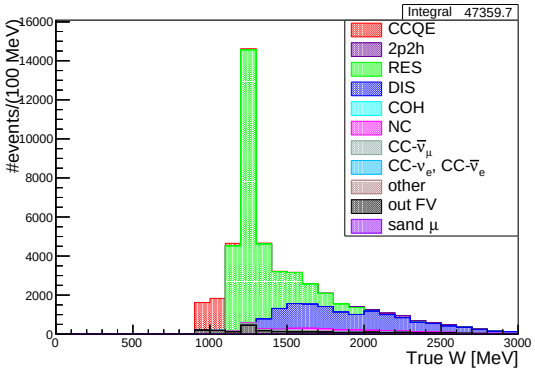
Selected Sample Distributions in True W and Q^2



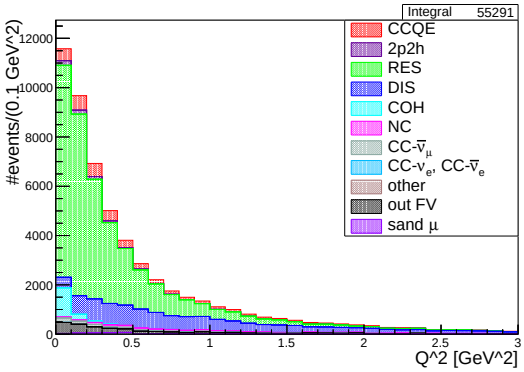
(a) True W - FGD1



(b) True Q^2 - FGD1

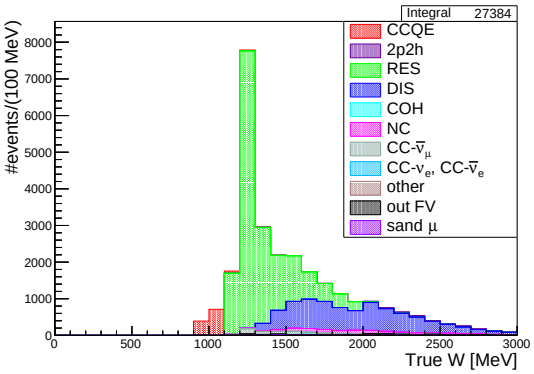


(c) True W - FGD2

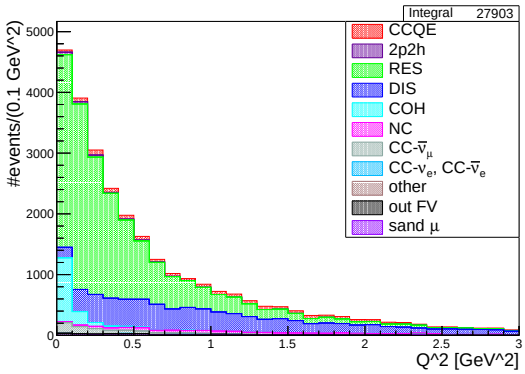


(d) True Q^2 - FGD2

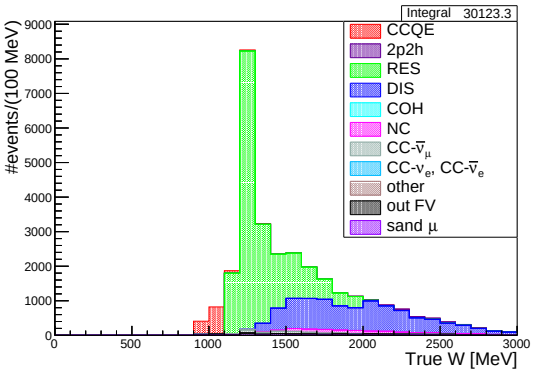
Figure D.1: True W (left) and Q^2 (right) distributions in FGD1 (top) and FGD2 (bottom), categorised by true reaction type, for the total CC1π^+ sample.



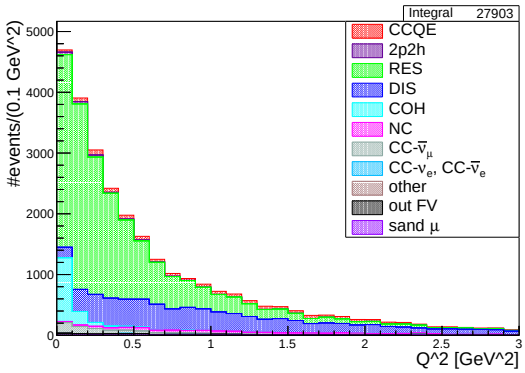
(a) True W - FGD1



(b) True Q^2 - FGD1

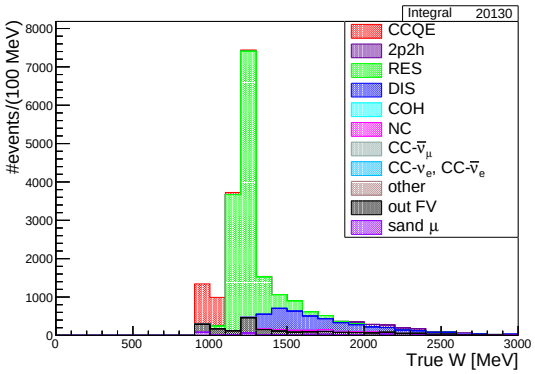


(c) True W - FGD2

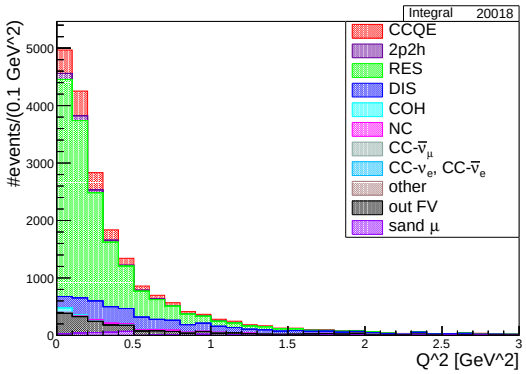


(d) True Q^2 - FGD2

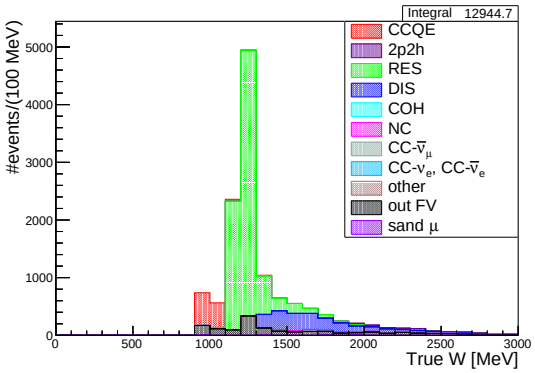
Figure D.2: True W (left) and Q^2 (right) distributions in FGD1 (top) and FGD2 (bottom), categorised by true reaction type, for the TPC CC1π^+ sample.



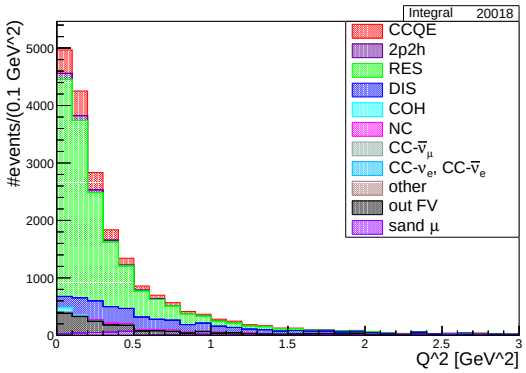
(a) True W - FGD1



(b) True Q^2 - FGD1

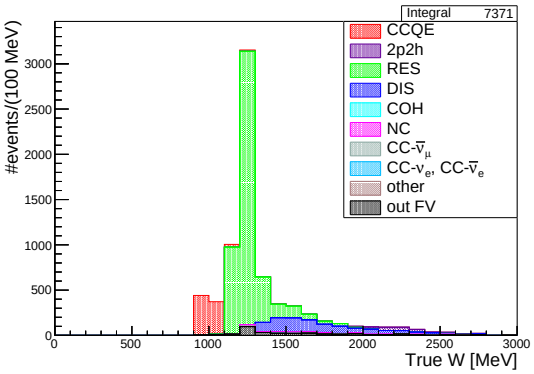


(c) True W - FGD2

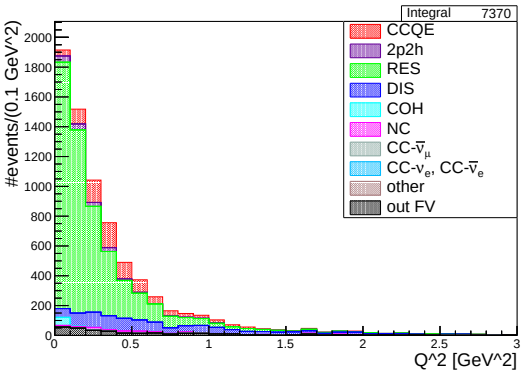


(d) True Q^2 - FGD2

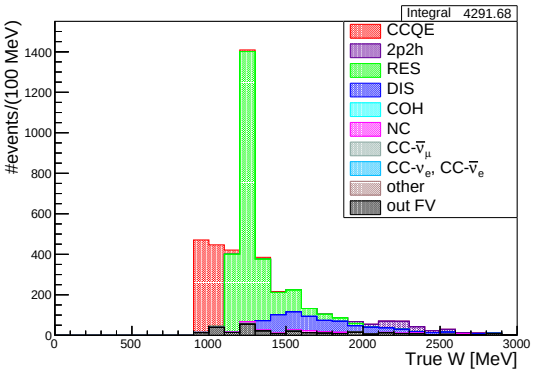
Figure D.3: True W (left) and Q^2 (right) distributions in FGD1 (top) and FGD2 (bottom), categorised by true reaction type, for the ME CC1π^+ sample.



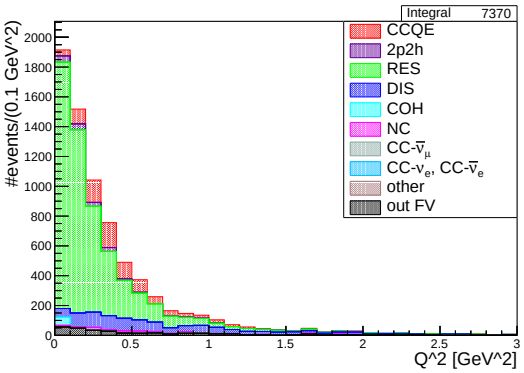
(a) True W - FGD1



(b) True Q^2 - FGD1

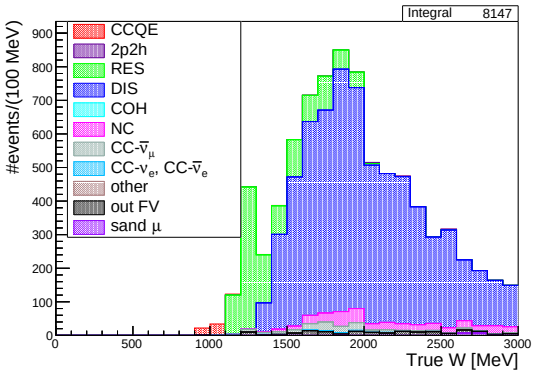


(c) True W - FGD2

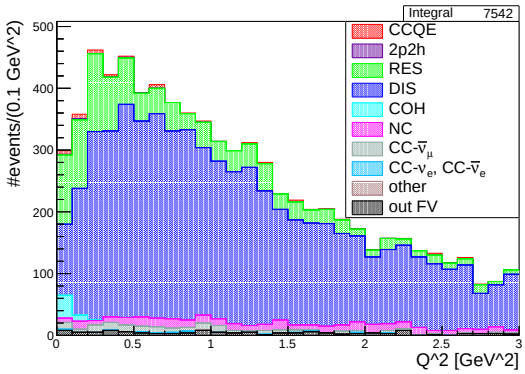


(d) True Q^2 - FGD2

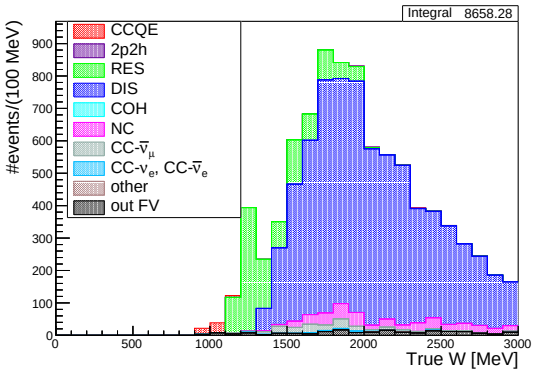
Figure D.4: True W (left) and Q^2 (right) distributions in FGD1 (top) and FGD2 (bottom), categorised by true reaction type, for the isolated FGD CC1π^+ sample.



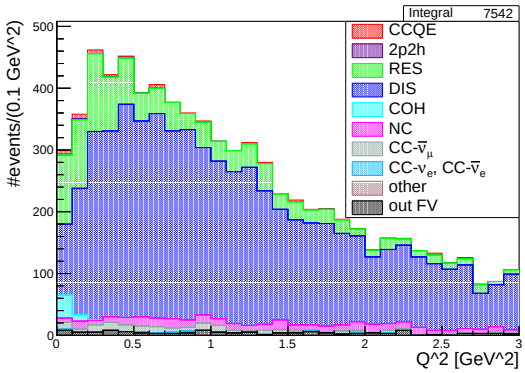
(a) True W - FGD1



(b) True Q^2 - FGD1

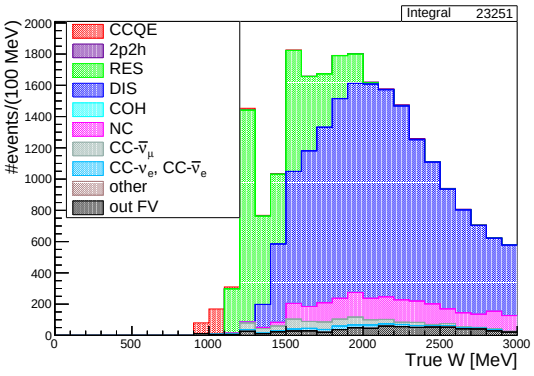


(c) True W - FGD2

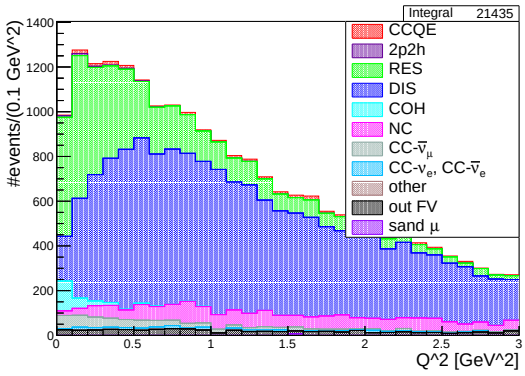


(d) True Q^2 - FGD2

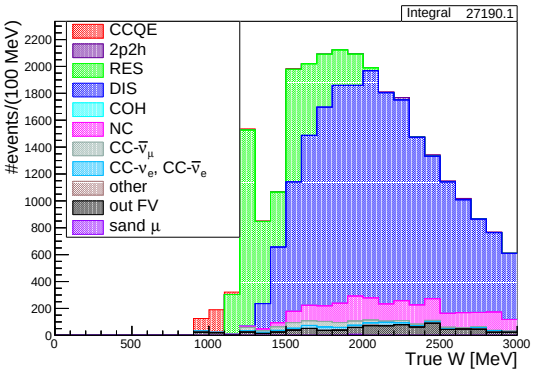
Figure D.5: True W (left) and Q^2 (right) distributions in FGD1 (top) and FGD2 (bottom), categorised by true reaction type, for the CC1π^+ + N$\pi^\pm$ control sample (CS1)



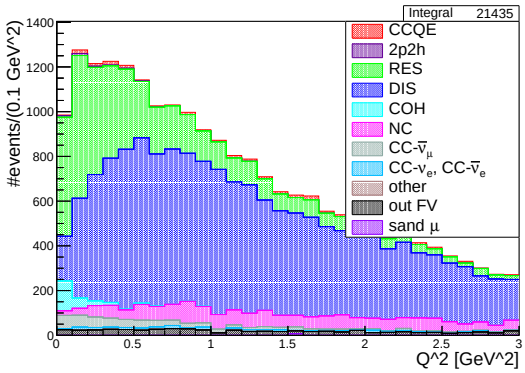
(a) True W - FGD1



(b) True Q^2 - FGD1

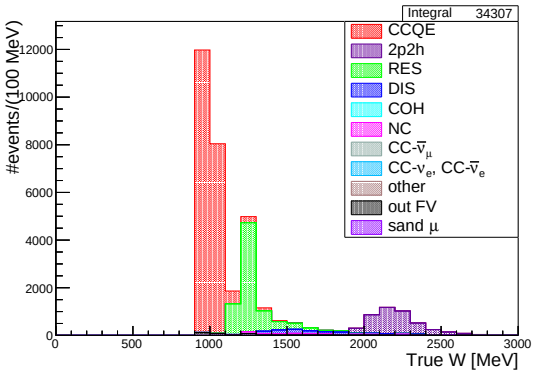


(c) True W - FGD2

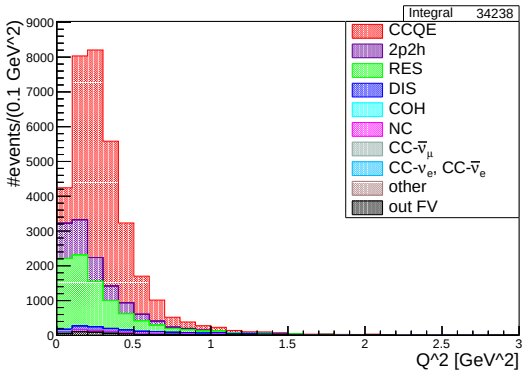


(d) True Q^2 - FGD2

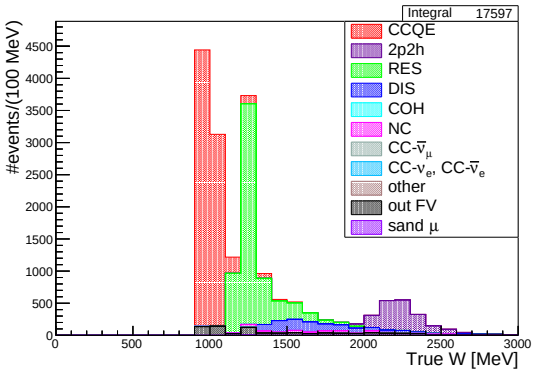
Figure D.6: True W (left) and Q^2 (right) distributions in FGD1 (top) and FGD2 (bottom), categorised by true reaction type, for the CC1π^+ + Nπ^0 control sample (CS2).



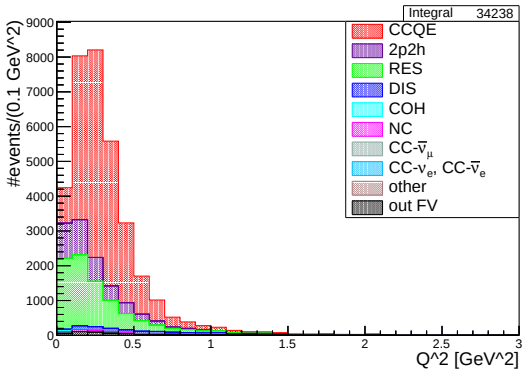
(a) True W - FGD1



(b) True Q^2 - FGD1



(c) True W - FGD2



(d) True Q^2 - FGD2

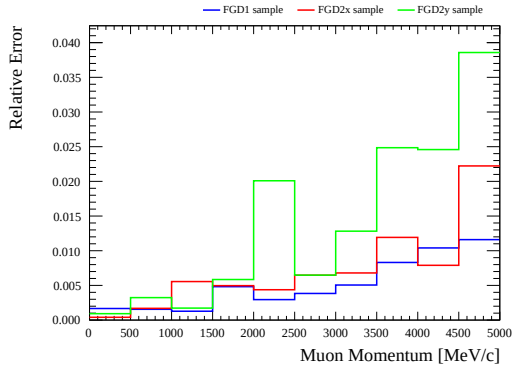
Figure D.7: True W (left) and Q^2 (right) distributions in FGD1 (top) and FGD2 (bottom), categorised by true reaction type, for the CC0π + Np control sample (CS3).

Appendix E

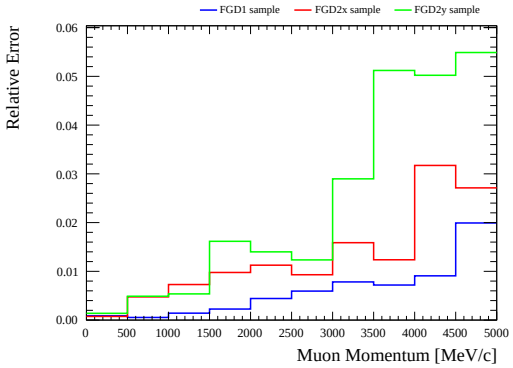
Detector Systematic Errors

Systematic	Figure
B-Field Distortions	Figure E.1
ECal PID	Figure E.2
ECal π^0 Veto Pileup	Figure E.3
FGD Hybrid Tracking Eff.	Figure E.4
FGD PID	Figure E.5
Michel Electron Eff.	Figure E.6
OOFV Background	Figure E.8
Pile-Up	Figure E.9
Pion Secondary Interactions	Figure E.10
Proton Secondary Interactions	Figure E.11
Sand Background	Figure E.12
TPC–ECal Matching	Figure E.13
TPC–FGD Matching	Figure E.14
TPC Charge ID Eff.	Figure E.15
TPC Cluster Eff.	Figure E.16
TPC Momentum Resolution	Figure E.17
TPC Momentum Scale	Figure E.18
TPC PID	Figure E.19
TPC Tracking Eff.	Figure E.20

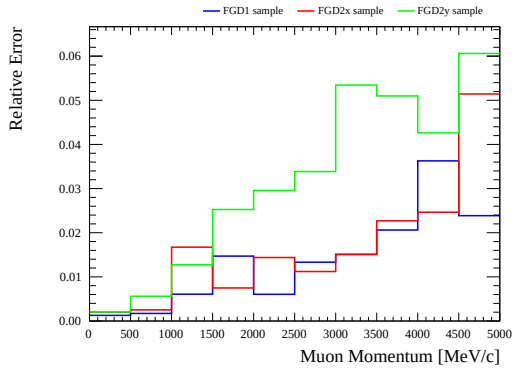
Table E.I: Relative error plots for detector systematic uncertainties.



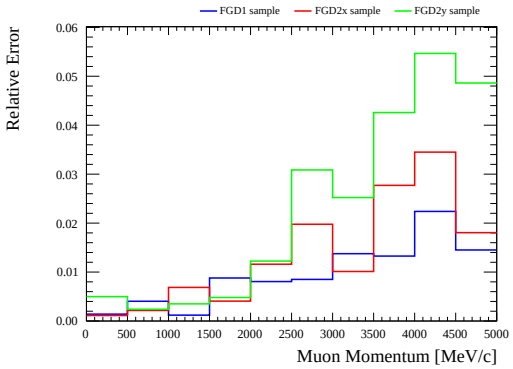
(a) CC1π^+ TPC



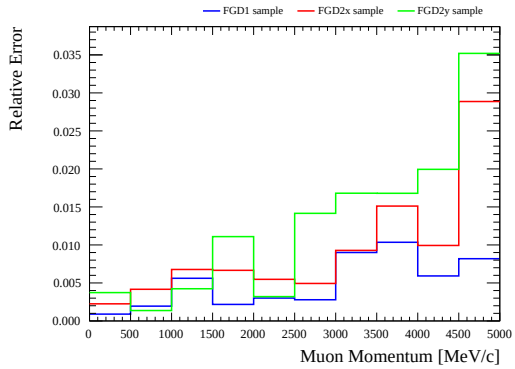
(b) CC1π^+ ME



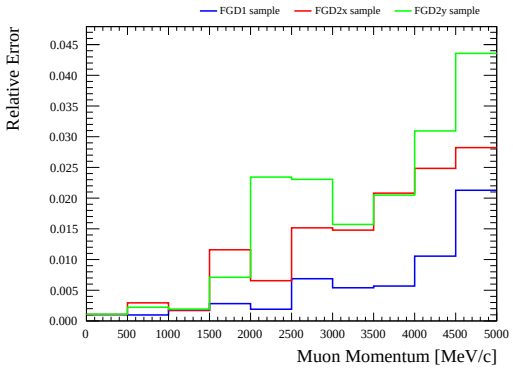
(c) CC1π^+ isoFGD



(d) CS1

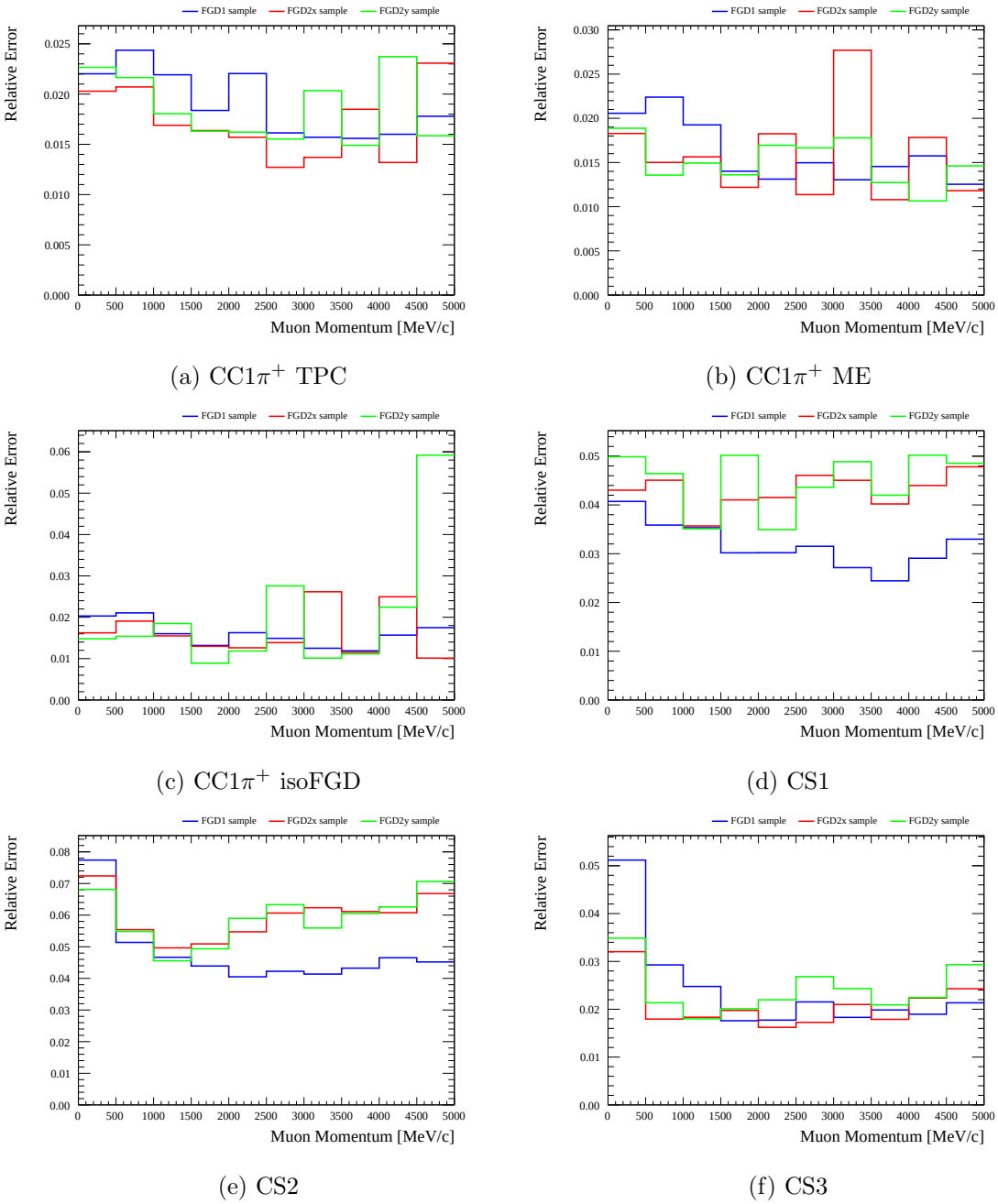


(e) CS2



(f) CS3

Figure E.1: Relative error values for the B field systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



(a) CC1π^+ TPC

(b) CC1π^+ ME

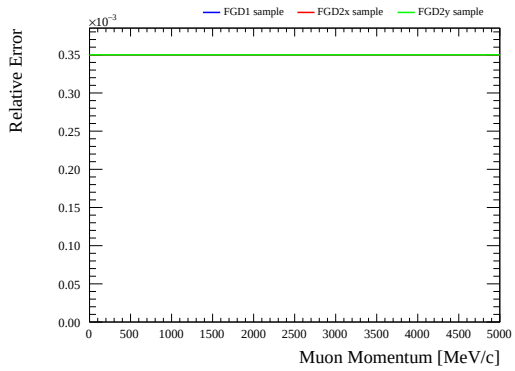
(c) CC1π^+ isoFGD

(d) CS1

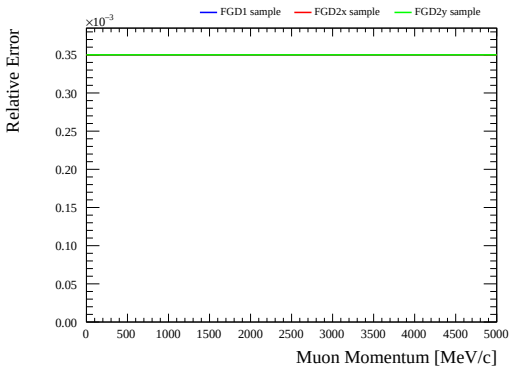
(e) CS2

(f) CS3

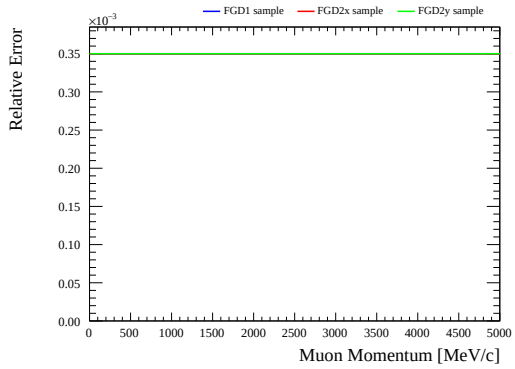
Figure E.2: Relative error values for the ECal PID systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



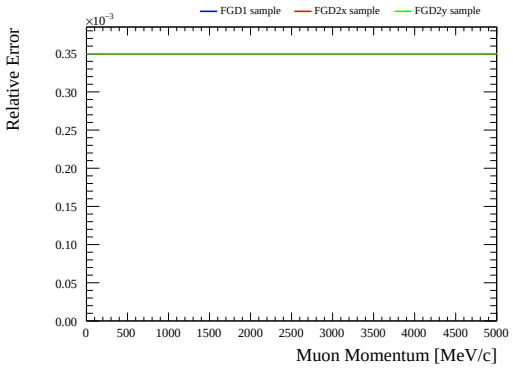
(a) CC1π^+ TPC



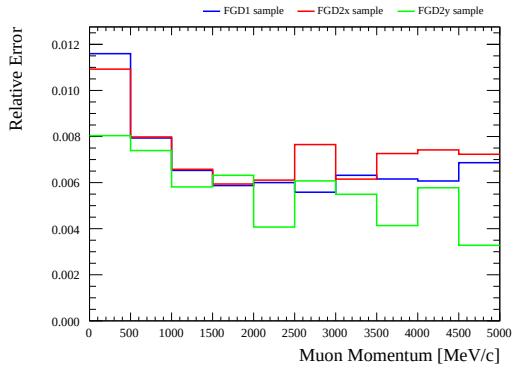
(b) CC1π^+ ME



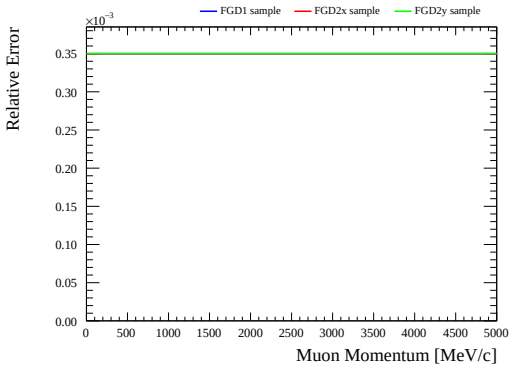
(c) CC1π^+ isoFGD



(d) CS1

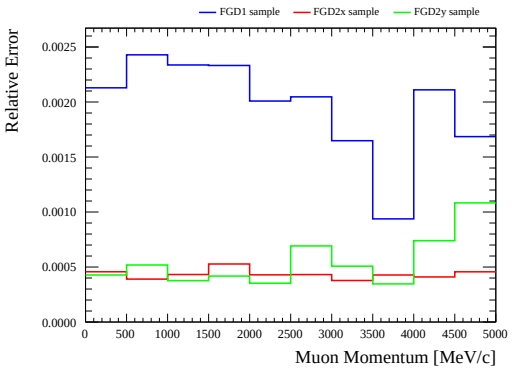


(e) CS2

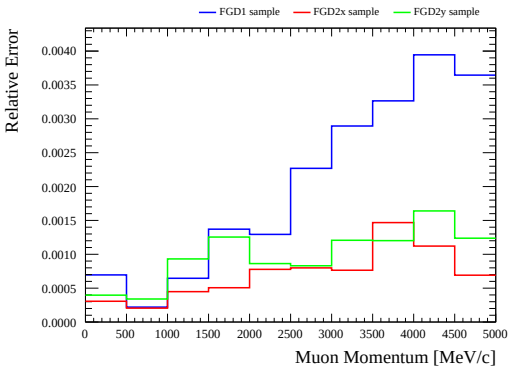


(f) CS3

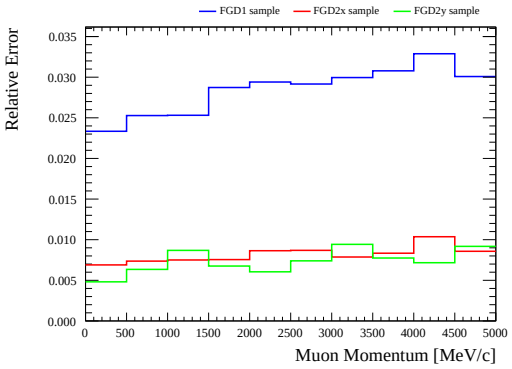
Figure E.3: Relative error values for the ECal π^0 veto systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



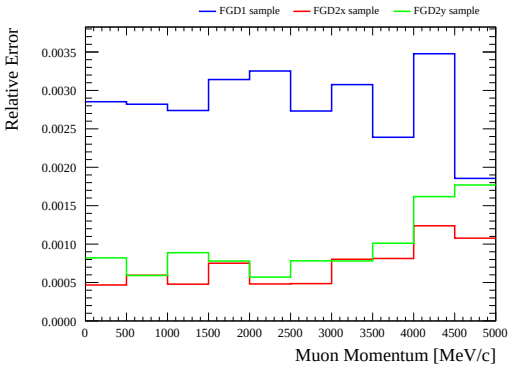
(a) CC1π^+ TPC



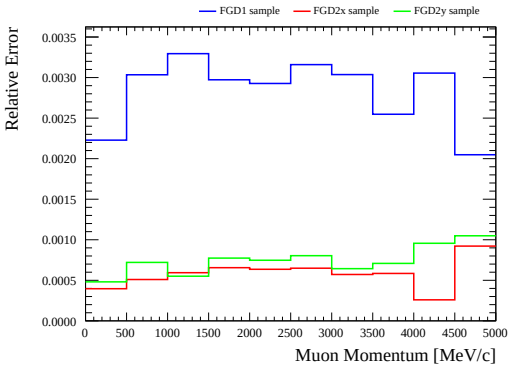
(b) CC1π^+ ME



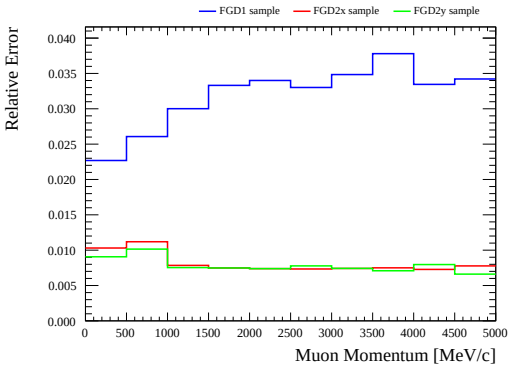
(c) CC1π^+ isoFGD



(d) CS1

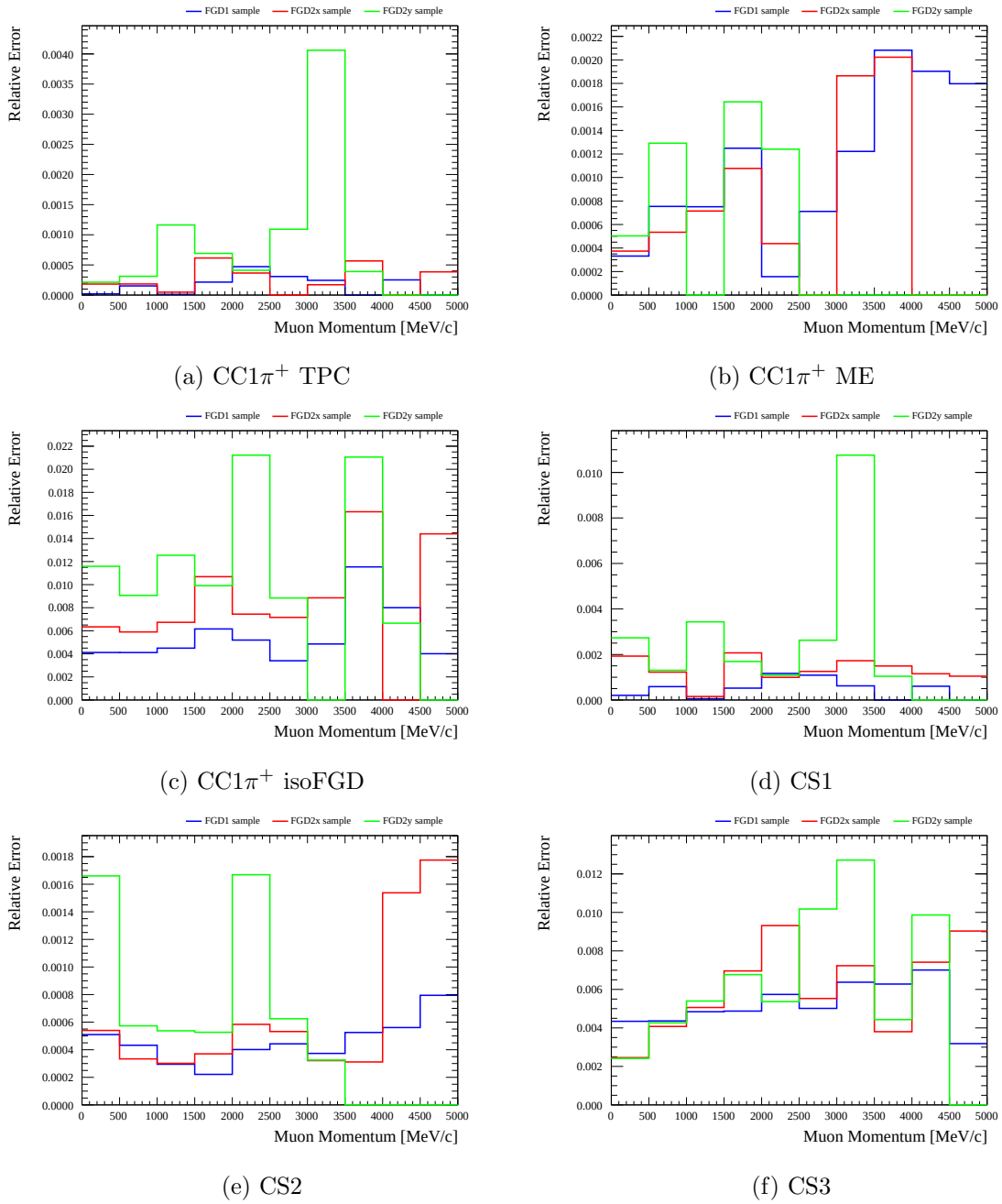


(e) CS2



(f) CS3

Figure E.4: Relative error values for the FGD hybrid track efficiency systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



(a) CC1π^+ TPC

(b) CC1π^+ ME

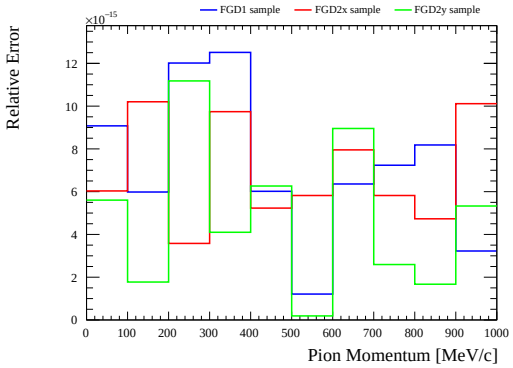
(c) CC1π^+ isoFGD

(d) CS1

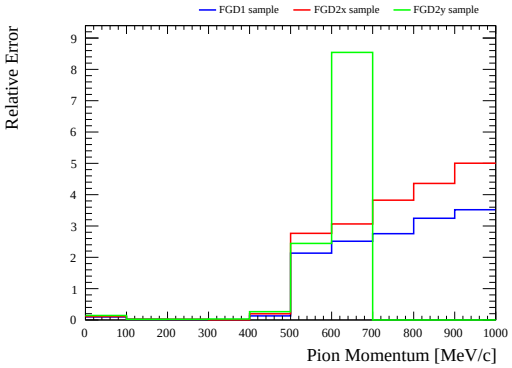
(e) CS2

(f) CS3

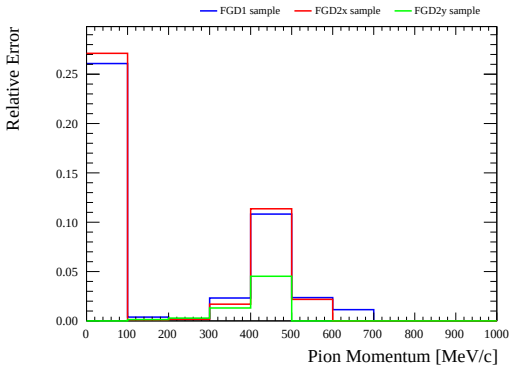
Figure E.5: Relative error values for the FGD PID systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



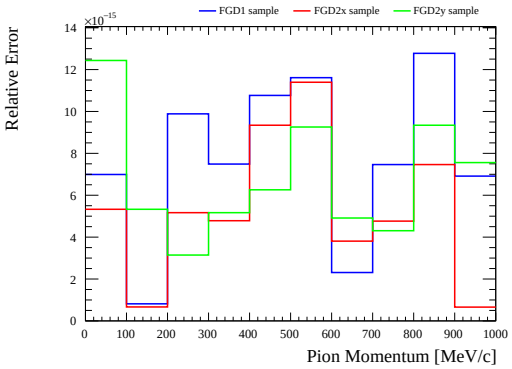
(a) CC1π^+ TPC



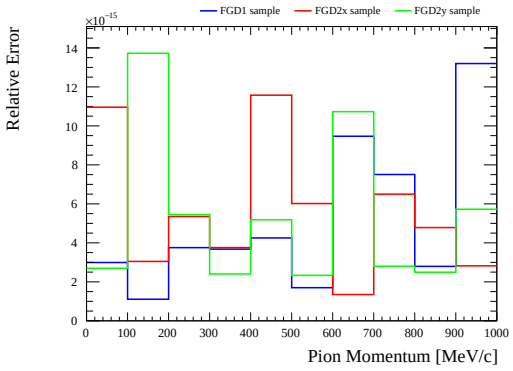
(b) CC1π^+ ME



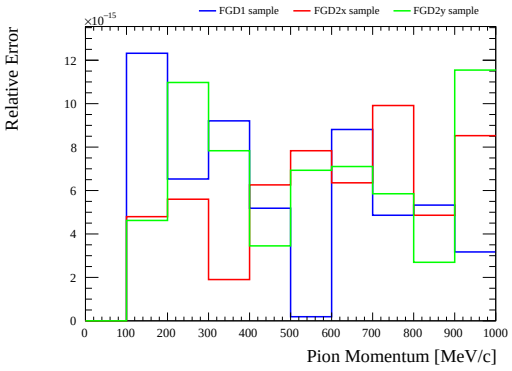
(c) CC1π^+ isoFGD



(d) CS1

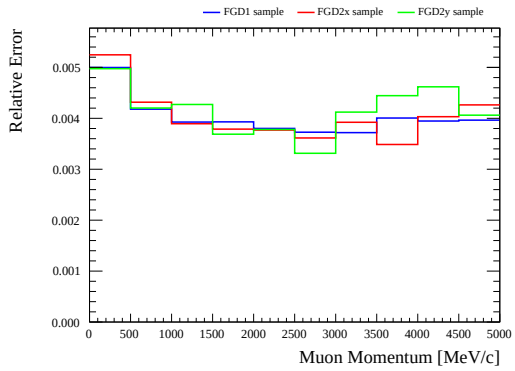


(e) CS2

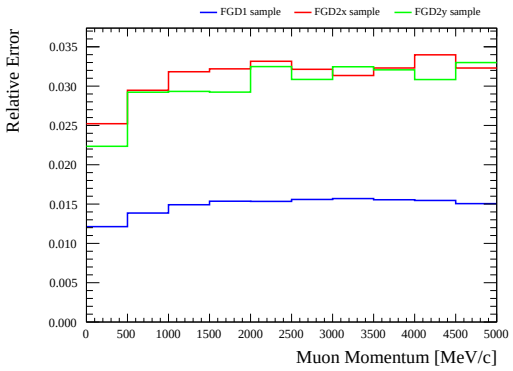


(f) CS3

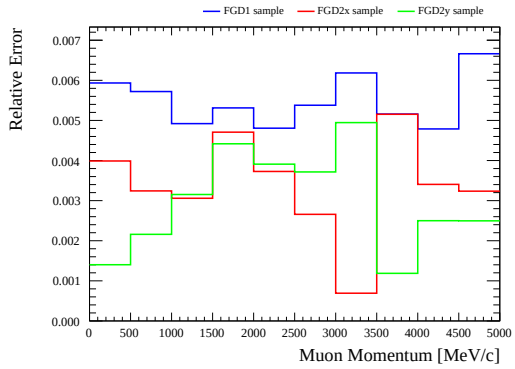
Figure E.6: Relative error values for the FGD pion momentum from Michel electron reconstruction systematic, as a function of reconstructed pion momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



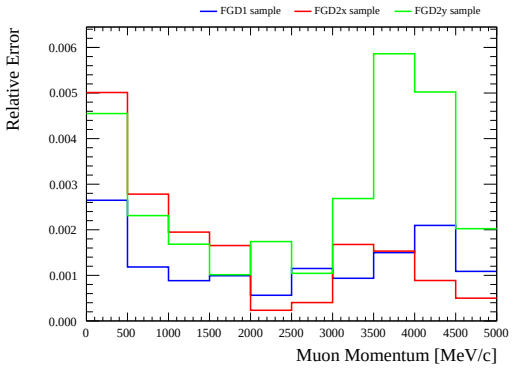
(a) CC1π^+ TPC



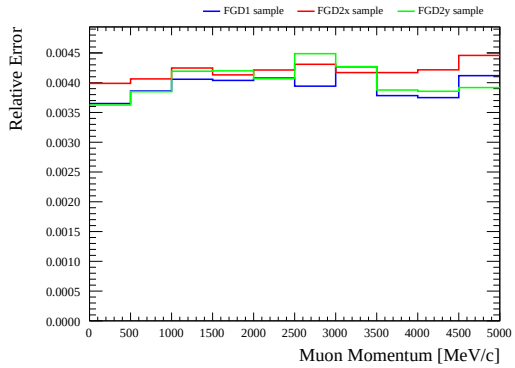
(b) CC1π^+ ME



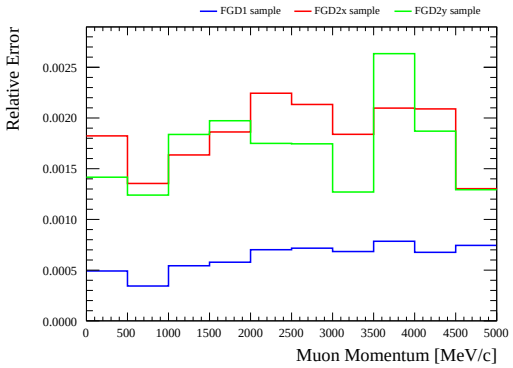
(c) CC1π^+ isoFGD



(d) CS1

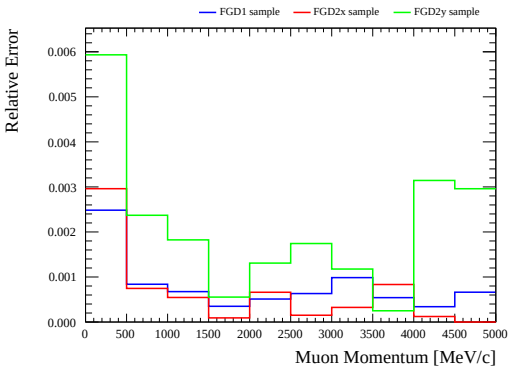


(e) CS2

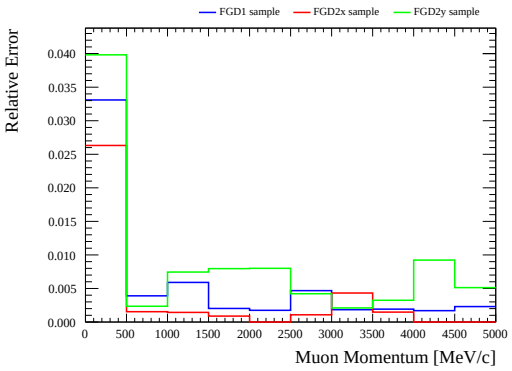


(f) CS3

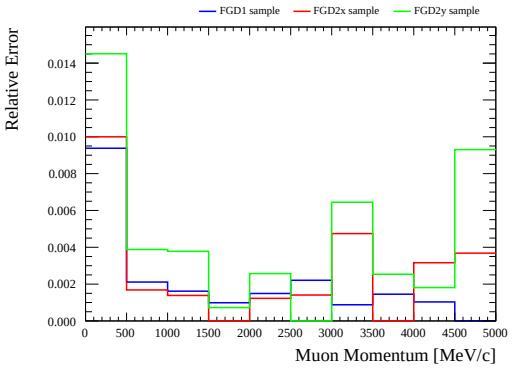
Figure E.7: Relative error values for the Michel electron efficiency systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



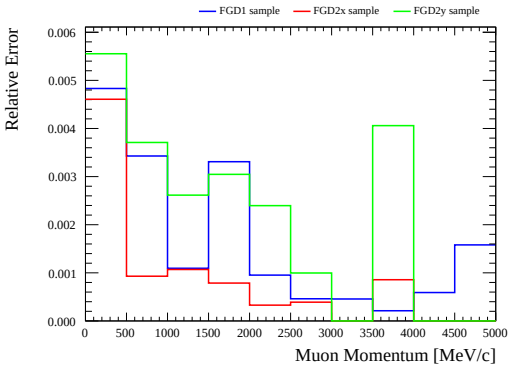
(a) CC1π^+ TPC



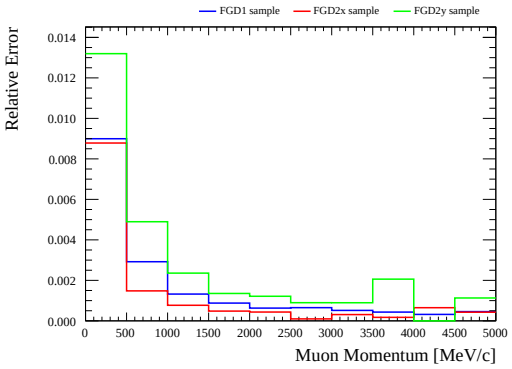
(b) CC1π^+ ME



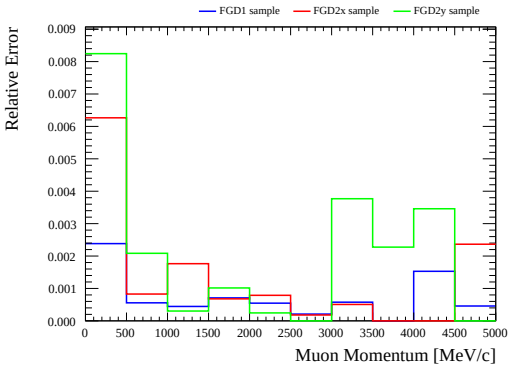
(c) CC1π^+ isoFGD



(d) CS1

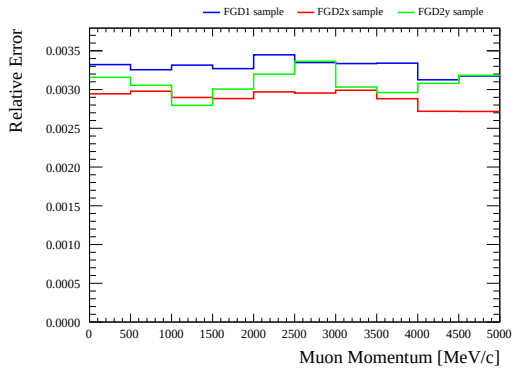


(e) CS2

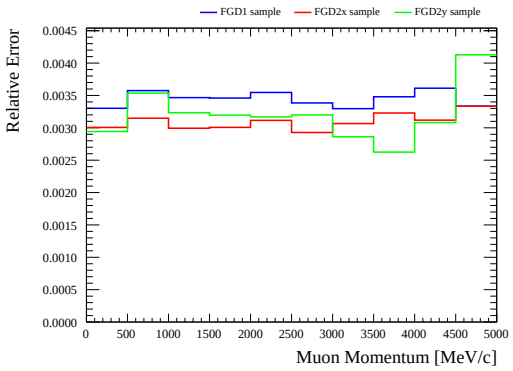


(f) CS3

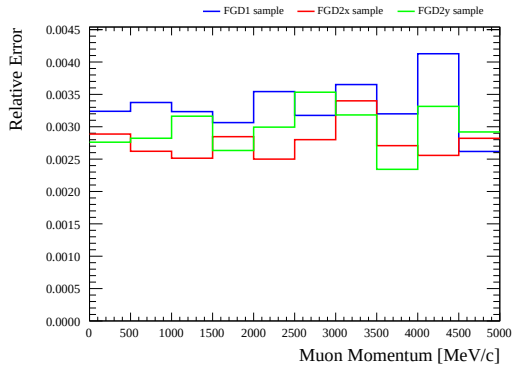
Figure E.8: Relative error values for the OOFV systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



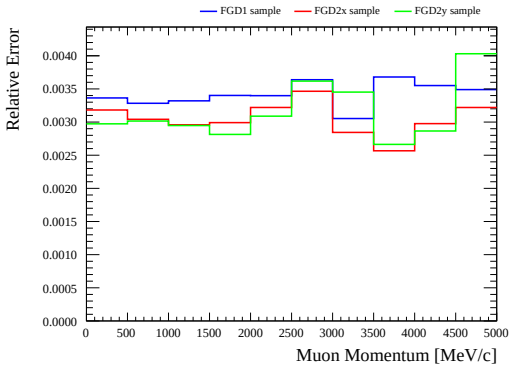
(a) CC1π^+ TPC



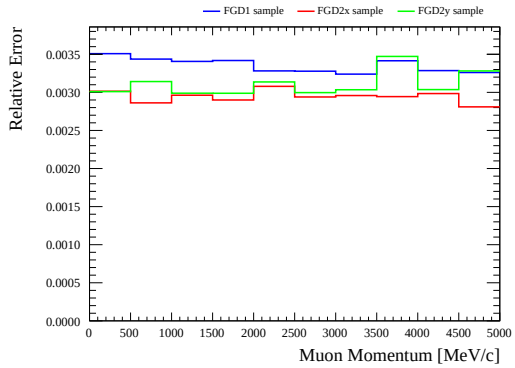
(b) CC1π^+ ME



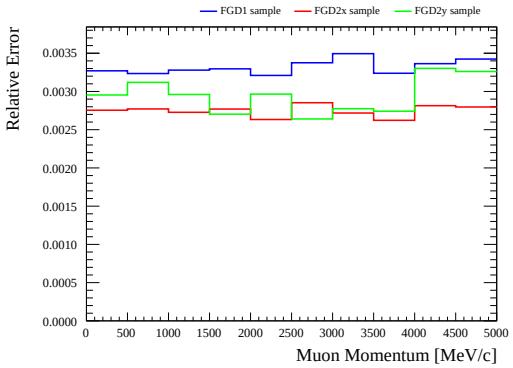
(c) CC1π^+ isoFGD



(d) CS1

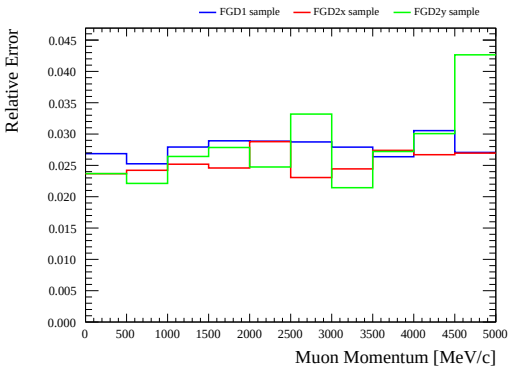


(e) CS2

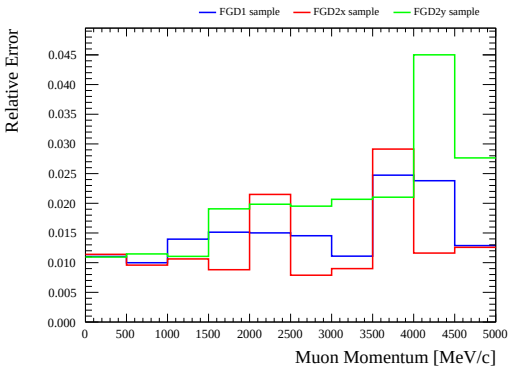


(f) CS3

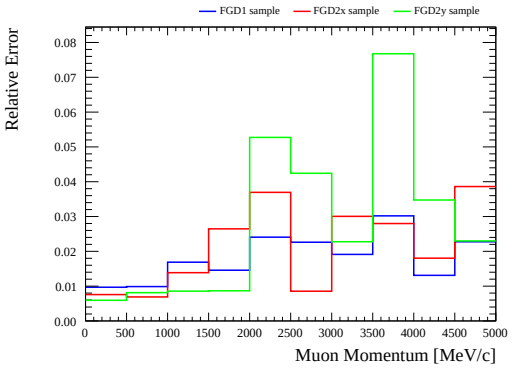
Figure E.9: Relative error values for the pileup systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



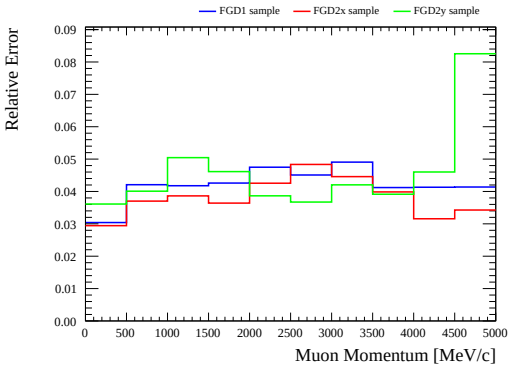
(a) CC1π^+ TPC



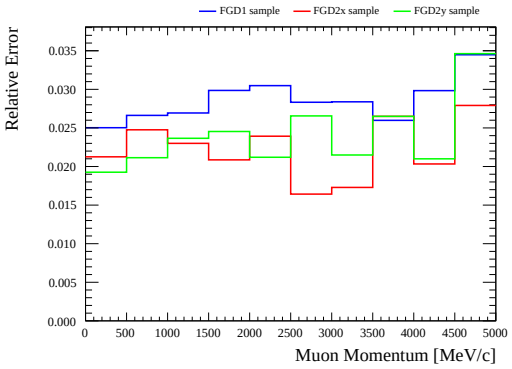
(b) CC1π^+ ME



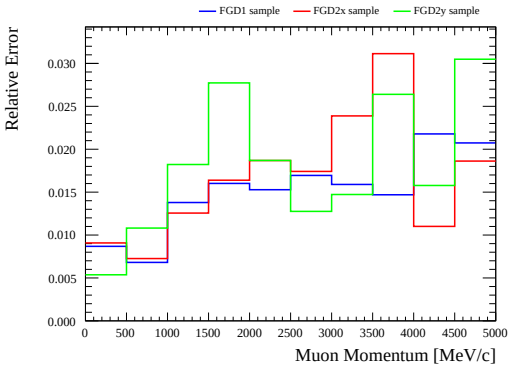
(c) CC1π^+ isoFGD



(d) CS1

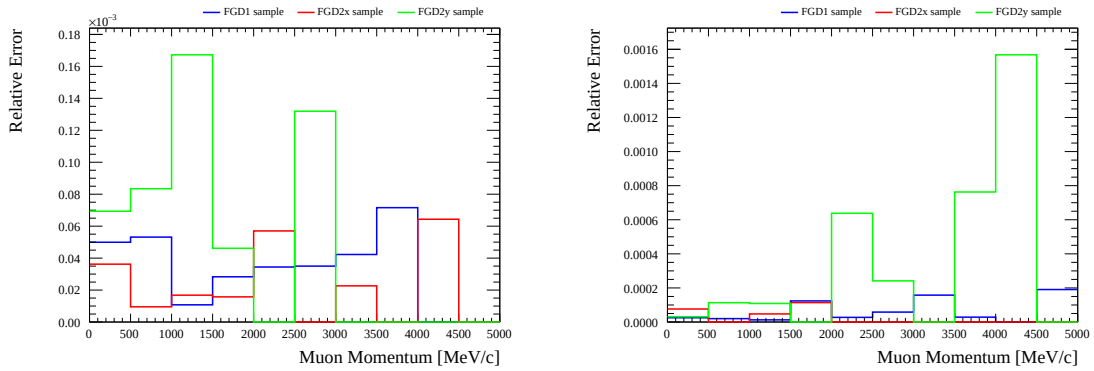


(e) CS2



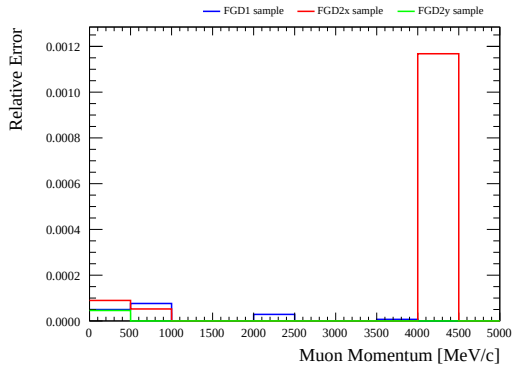
(f) CS3

Figure E.10: Relative error values for the pion SI systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).

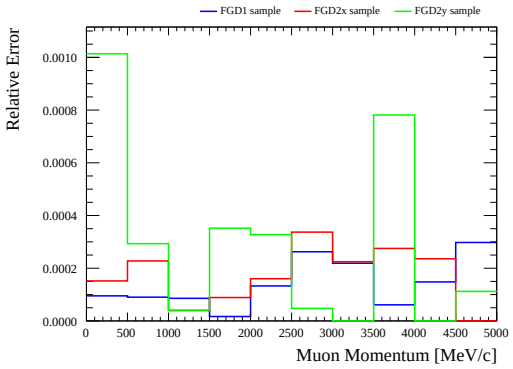


(a) CC1π^+ TPC

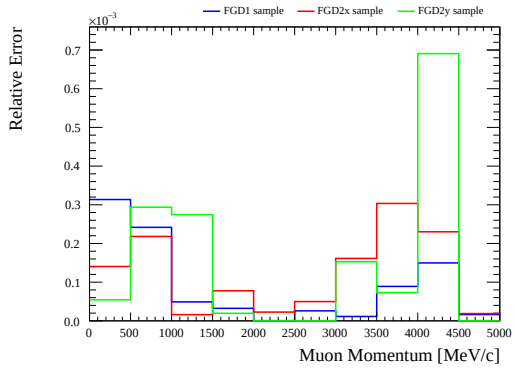
(b) CC1π^+ ME



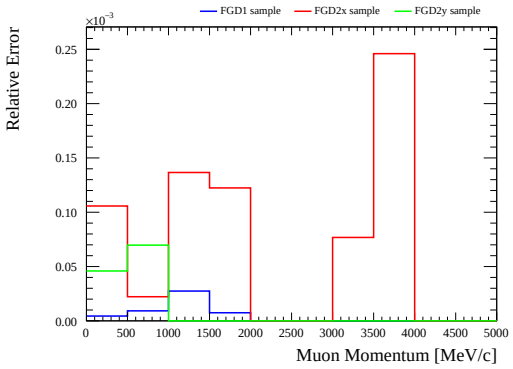
(c) CC1π^+ isoFGD



(d) CS1

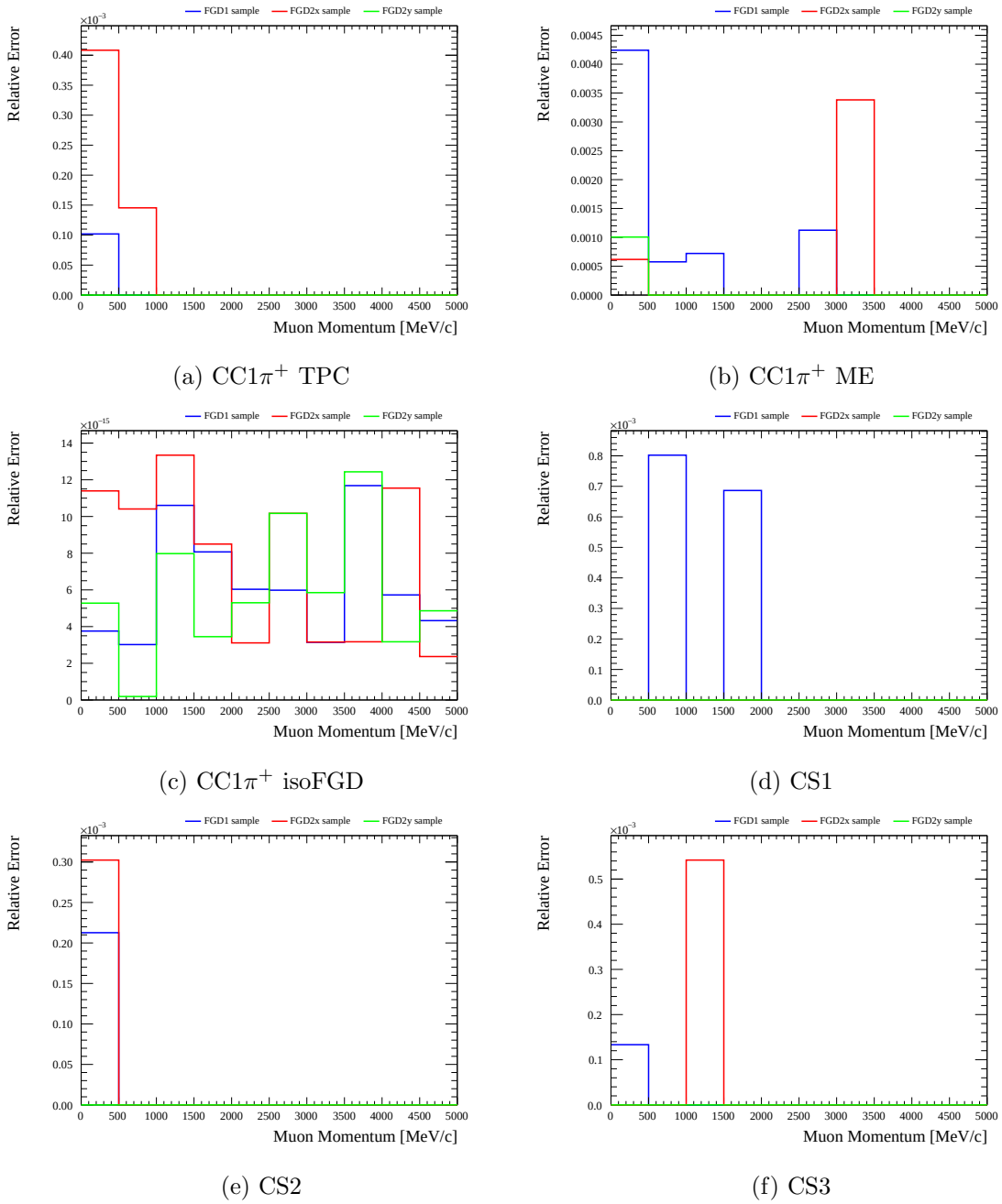


(e) CS2



(f) CS3

Figure E.11: Relative error values for the proton SI systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



(a) CC1π^+ TPC

(b) CC1π^+ ME

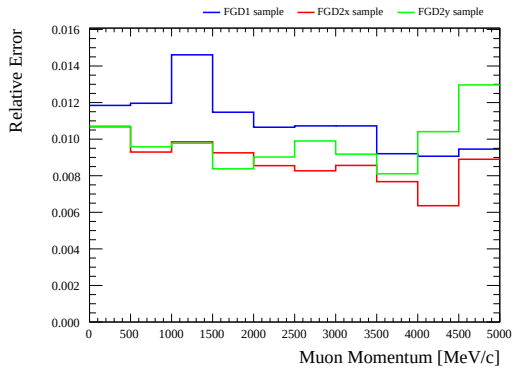
(c) CC1π^+ isoFGD

(d) CS1

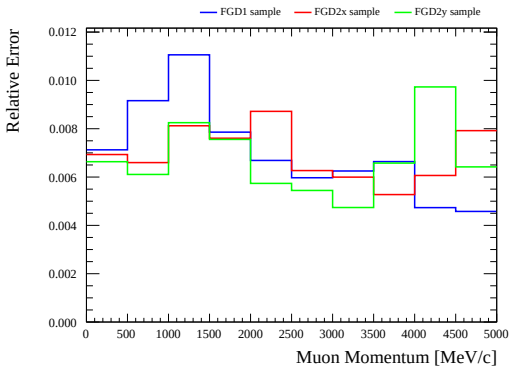
(e) CS2

(f) CS3

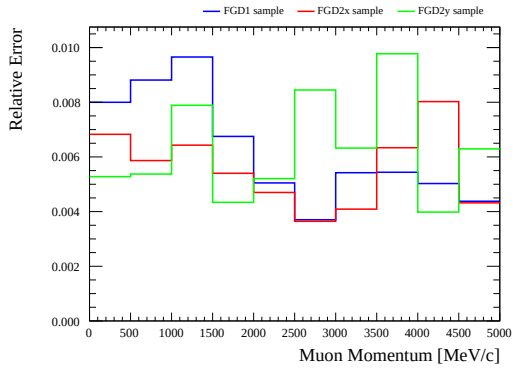
Figure E.12: Relative error values for the sand muon background systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



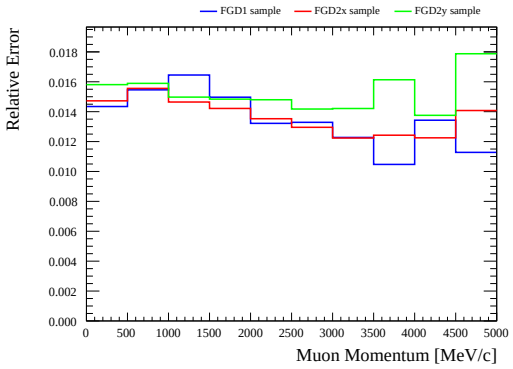
(a) CC1π^+ TPC



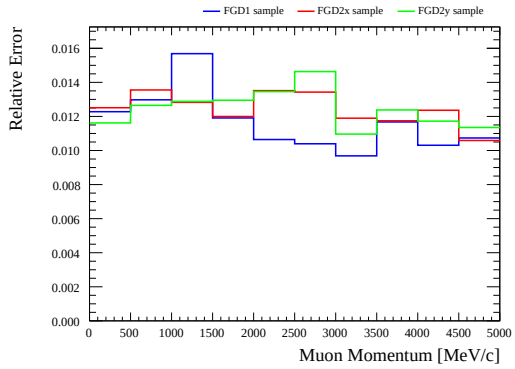
(b) CC1π^+ ME



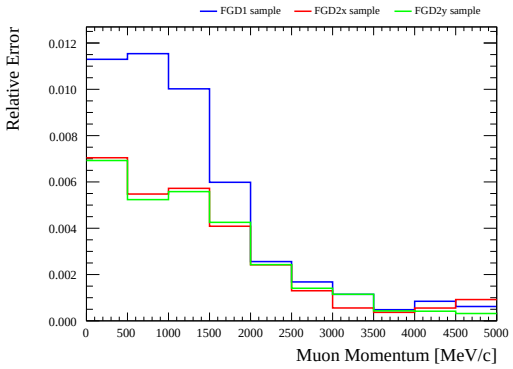
(c) CC1π^+ isoFGD



(d) CS1

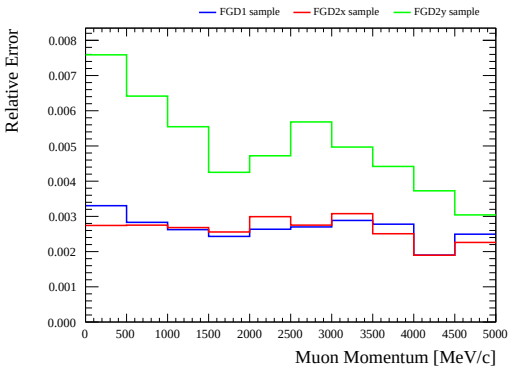


(e) CS2

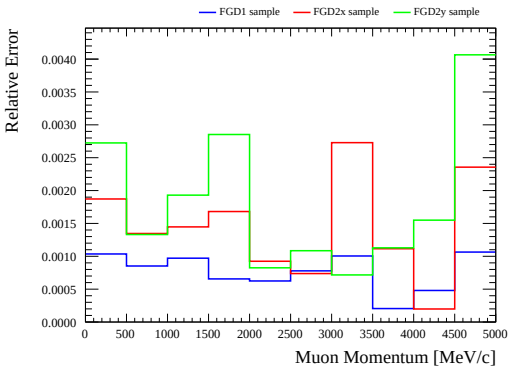


(f) CS3

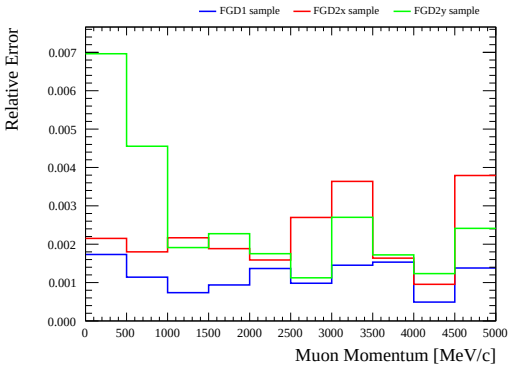
Figure E.13: Relative error values for the TPC-ECal matching efficiency systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



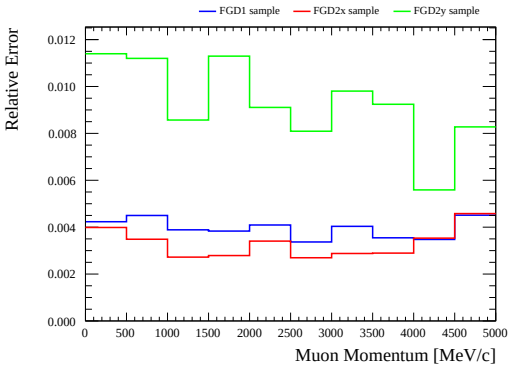
(a) CC1π^+ TPC



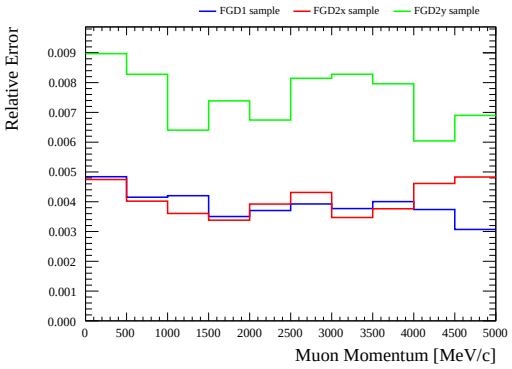
(b) CC1π^+ ME



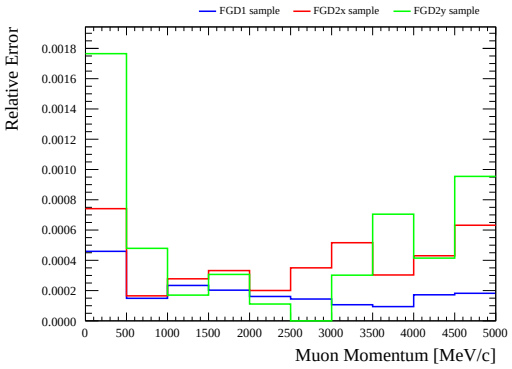
(c) CC1π^+ isoFGD



(d) CS1

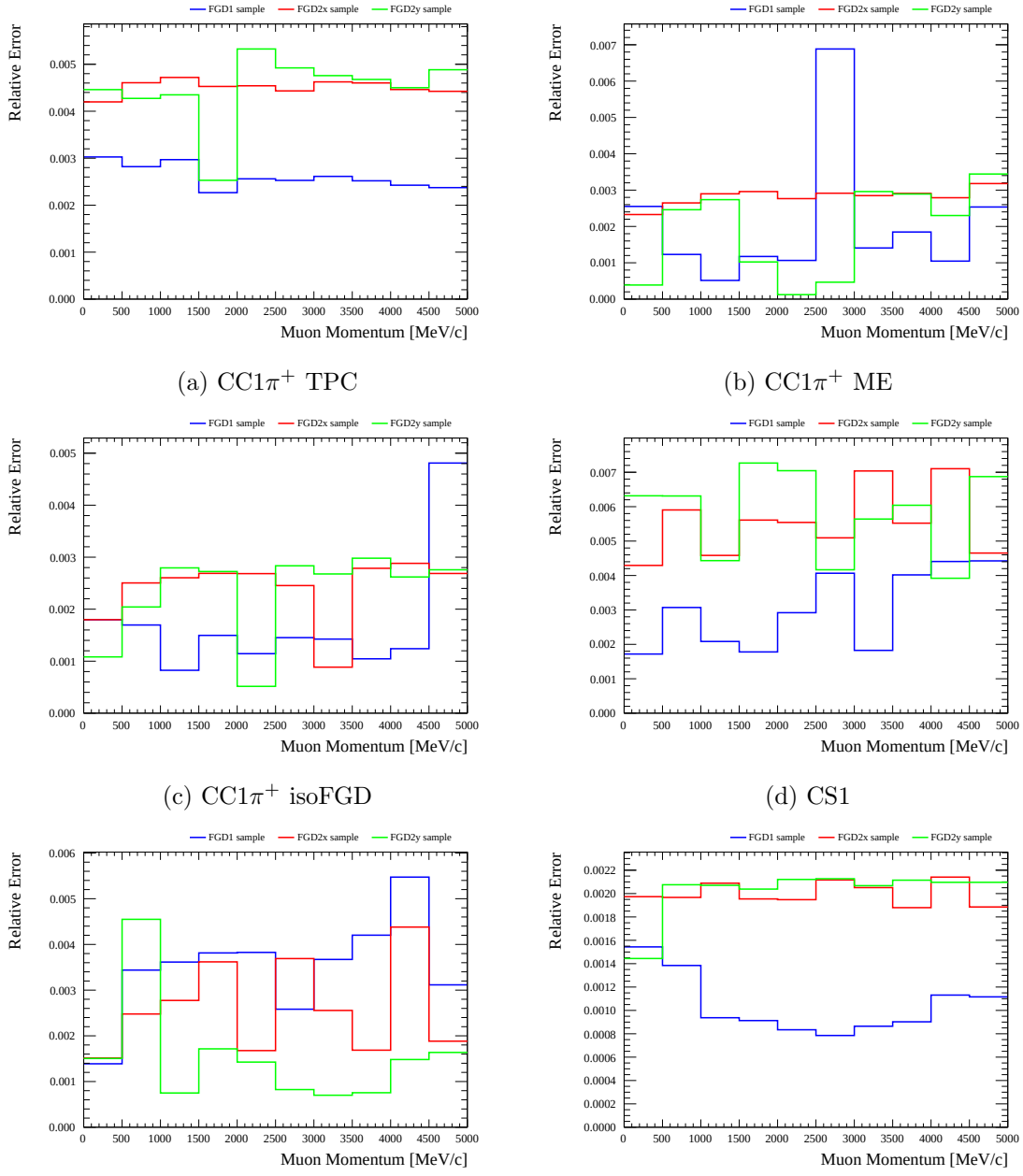


(e) CS2



(f) CS3

Figure E.14: Relative error values for the TPC-FGD matching efficiency systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



(a) CC1π^+ TPC

(b) CC1π^+ ME

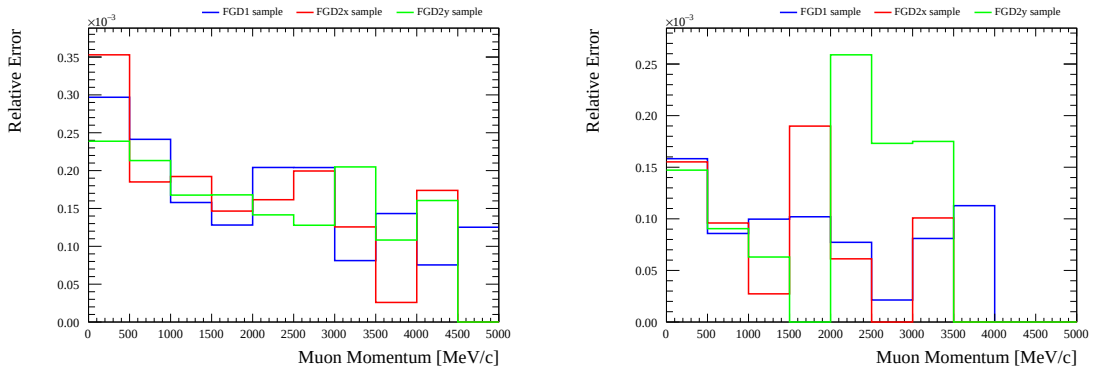
(c) CC1π^+ isoFGD

(d) CS1

(e) CS2

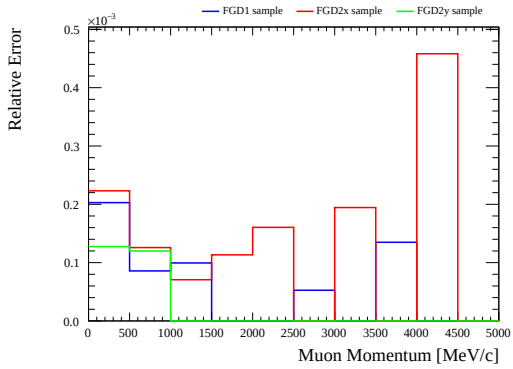
(f) CS3

Figure E.15: Relative error values for the TPC charge ID efficiency systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).

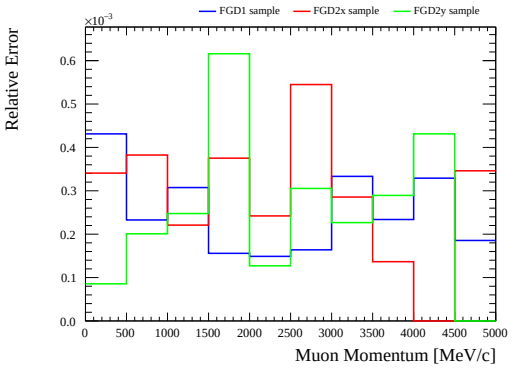


(a) CC1π^+ TPC

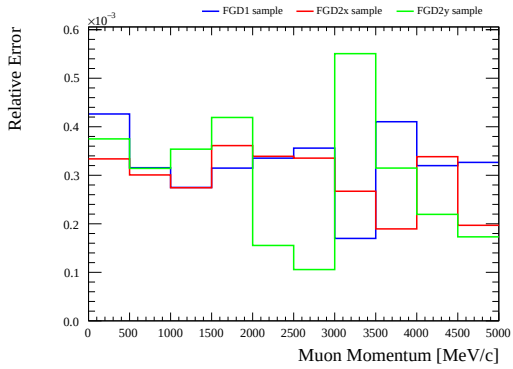
(b) CC1π^+ ME



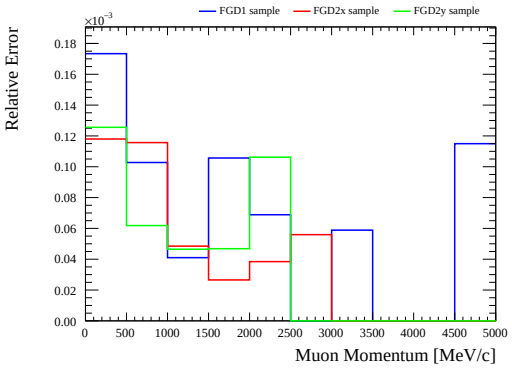
(c) CC1π^+ isoFGD



(d) CS1

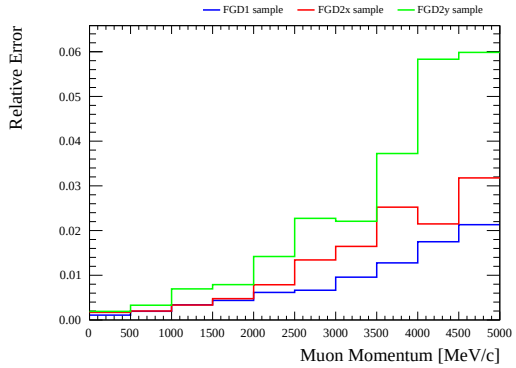


(e) CS2

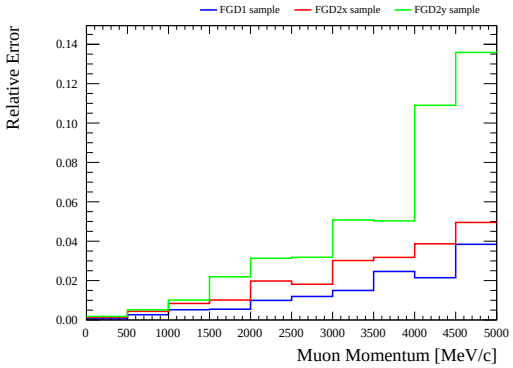


(f) CS3

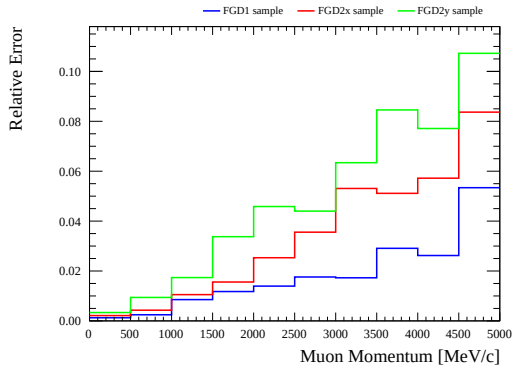
Figure E.16: Relative error values for the TPC clustering efficiency systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



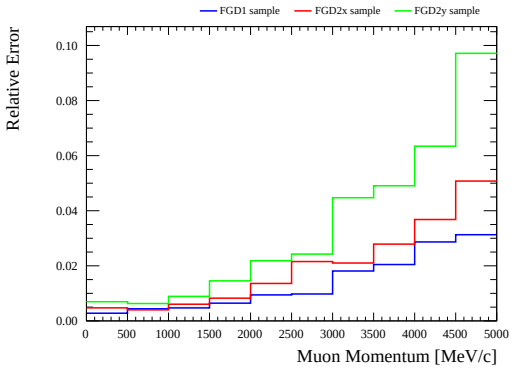
(a) CC1π^+ TPC



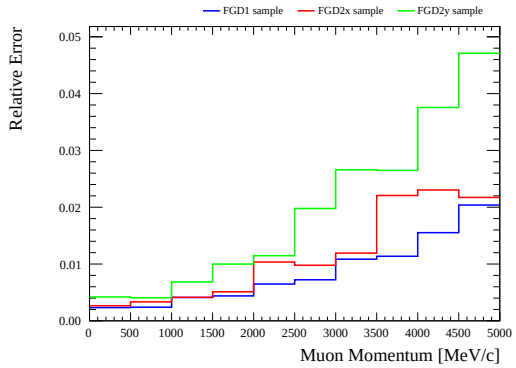
(b) CC1π^+ ME



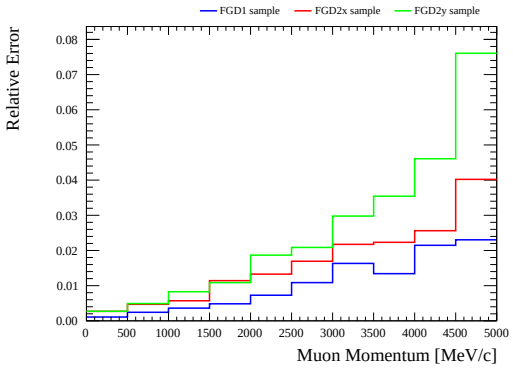
(c) CC1π^+ isoFGD



(d) CS1

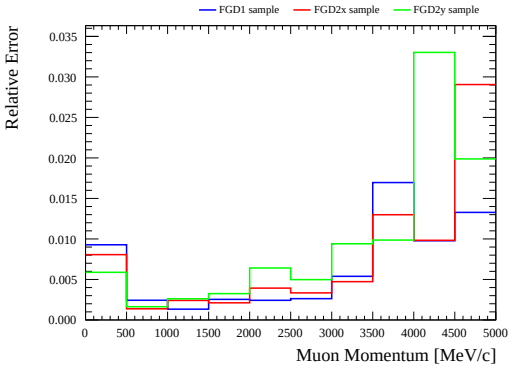


(e) CS2

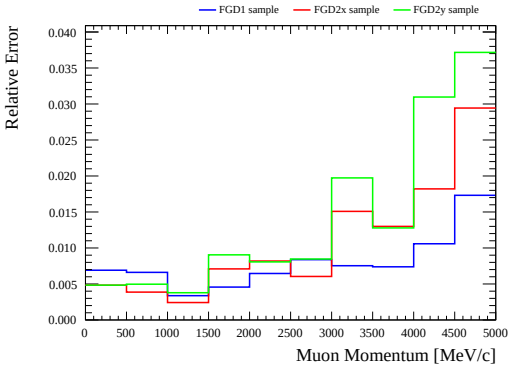


(f) CS3

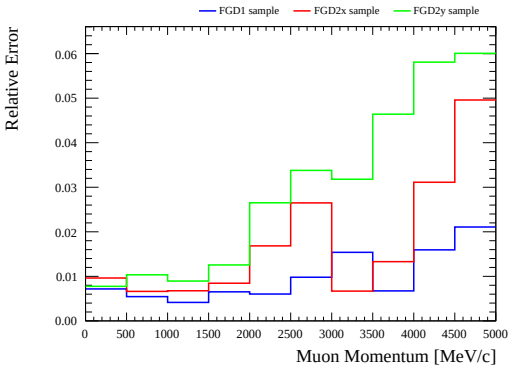
Figure E.17: Relative error values for the TPC momentum resolution systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



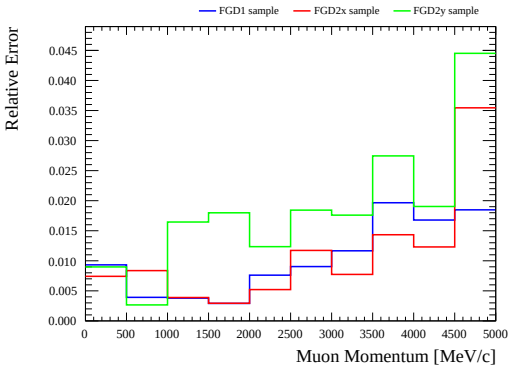
(a) CC1π^+ TPC



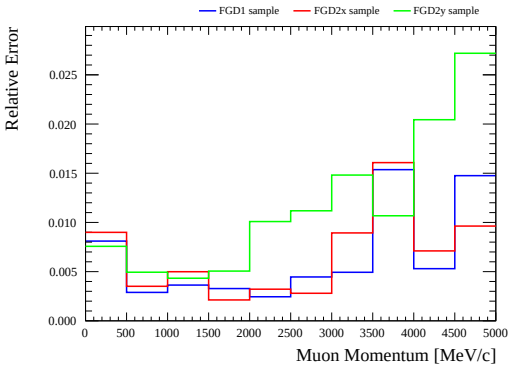
(b) CC1π^+ ME



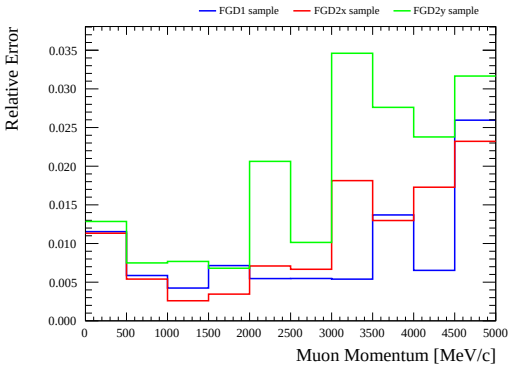
(c) CC1π^+ isoFGD



(d) CS1

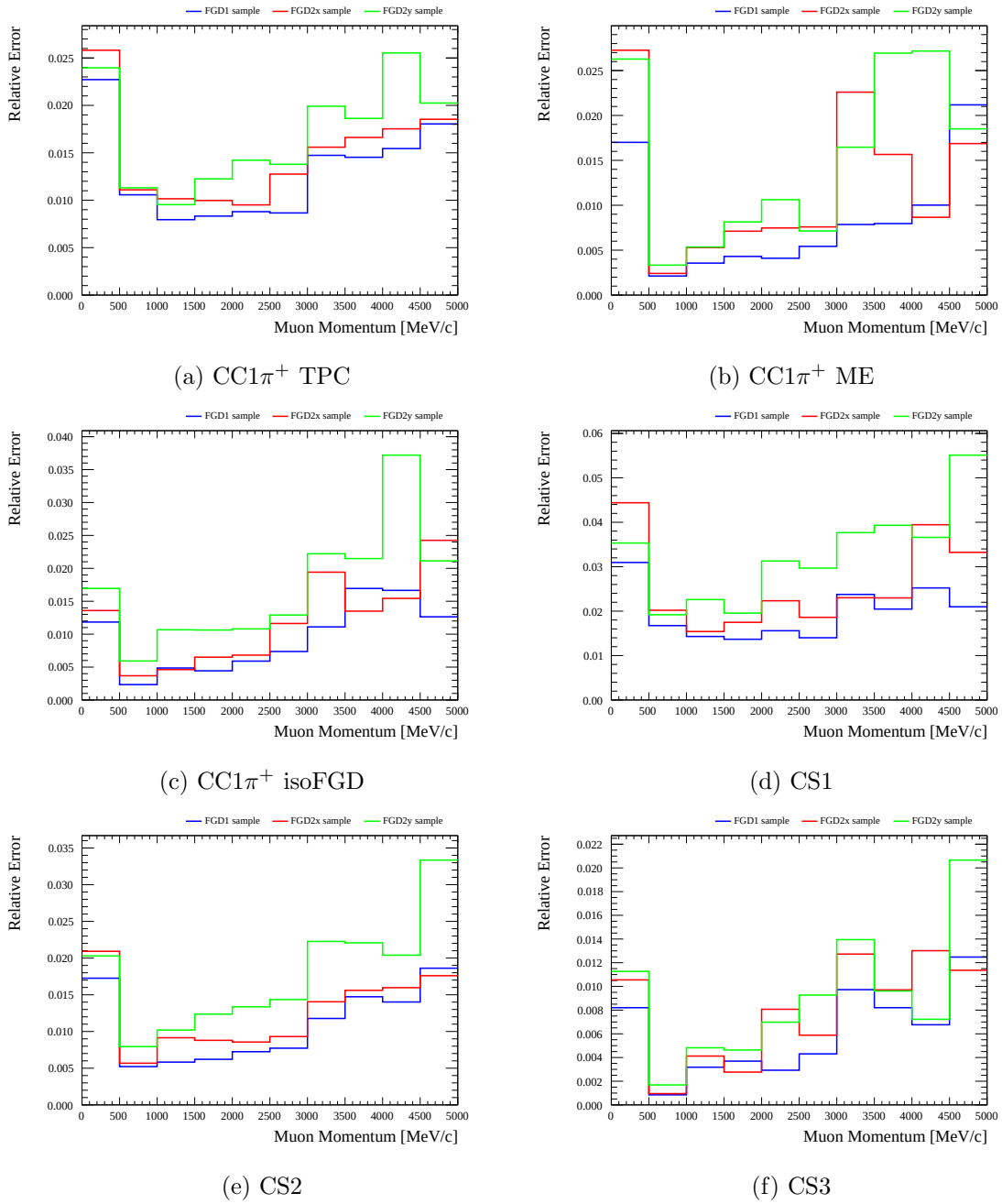


(e) CS2



(f) CS3

Figure E.18: Relative error values for the TPC momentum scale systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



(a) CC1π^+ TPC

(b) CC1π^+ ME

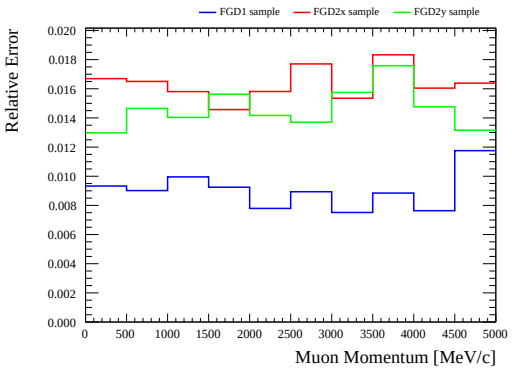
(c) CC1π^+ isoFGD

(d) CS1

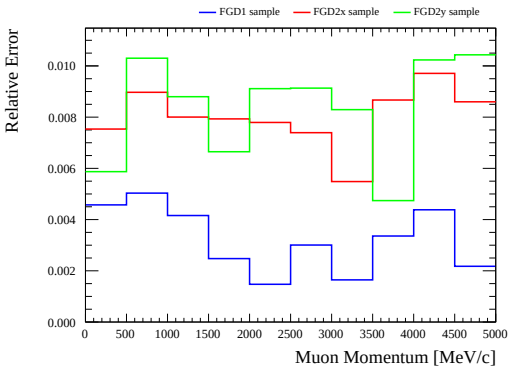
(e) CS2

(f) CS3

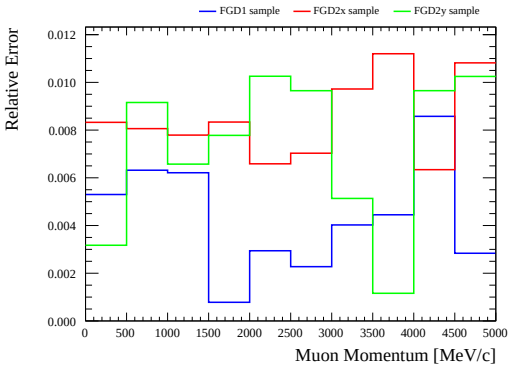
Figure E.19: Relative error values for the TPC PID systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).



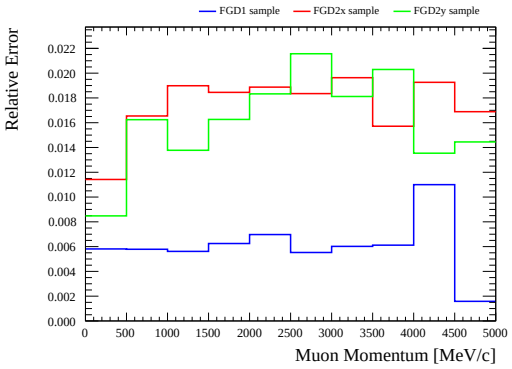
(a) CC1π^+ TPC



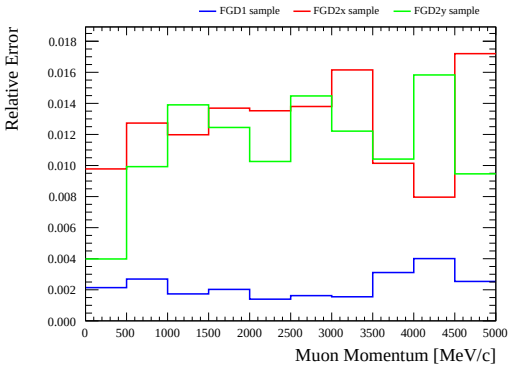
(b) CC1π^+ ME



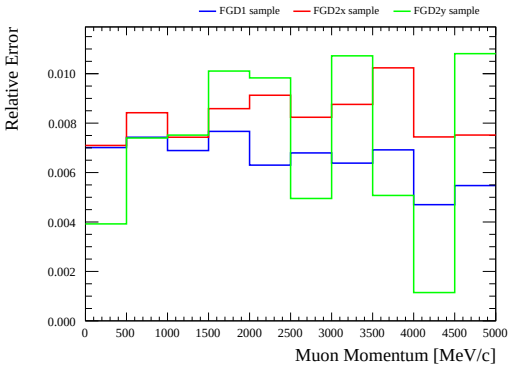
(c) CC1π^+ isoFGD



(d) CS1



(e) CS2



(f) CS3

Figure E.20: Relative error values for the TPC tracking efficiency systematic, as a function of reconstructed muon momentum. Distributions are shown for the three signal samples ((a) TPC, (b) ME, (c) isoFGD), and three control samples ((d) CS1, (e) CS2, (f) CS3). Each relative error distibution shows the values for FGD1 (blue), FGD2x (red) and FGD2y (green).

Appendix F

Control Sample Reconstructed Binning Schemes

Bin Number	$\cos\theta_\mu$	p_μ	$\cos\theta_\pi$	p_π
0	–1.0 – 0.3	0 – 30000	–1.0 – 1.0	0 – 30000
1	0.3 – 1.0	0 – 200	–1.0 – 1.0	0 – 30000
2	0.3 – 1.0	200 – 30000	–1.0 – 1.0	1500 – 30000
3	0.3 – 0.6	200 – 30000	–1.0 – 0.5	0 – 1500
4	0.3 – 0.6	200 – 30000	0.5 – 0.8	0 – 1500
5	0.3 – 0.6	200 – 30000	0.8 – 1.0	0 – 500
6	0.3 – 0.6	200 – 30000	0.8 – 1.0	500 – 1500
7	0.6 – 0.85	200 – 600	–1.0 – 0.5	0 – 1500
8	0.6 – 0.85	200 – 600	0.5 – 0.8	0 – 1500
9	0.6 – 0.85	200 – 600	0.8 – 1.0	0 – 300
10	0.6 – 0.85	200 – 600	0.8 – 1.0	300 – 500
11	0.6 – 0.85	200 – 600	0.8 – 1.0	500 – 1500
12	0.6 – 0.85	600 – 30000	–1.0 – 0.5	0 – 1500
13	0.6 – 0.85	600 – 30000	0.5 – 0.8	0 – 500
14	0.6 – 0.85	600 – 30000	0.5 – 0.8	500 – 1500
15	0.6 – 0.85	600 – 30000	0.8 – 1.0	0 – 500
16	0.6 – 0.85	600 – 30000	0.8 – 1.0	500 – 700
17	0.6 – 0.85	600 – 30000	0.8 – 1.0	700 – 1500
18	0.85 – 1.0	200 – 600	–1.0 – 0.5	0 – 1500
19	0.85 – 1.0	200 – 600	0.5 – 0.8	0 – 1500
20	0.85 – 1.0	200 – 600	0.8 – 1.0	0 – 500
21	0.85 – 1.0	200 – 600	0.8 – 1.0	500 – 1500
22	0.85 – 1.0	600 – 1200	–1.0 – 0.5	0 – 1500
23	0.85 – 1.0	600 – 1200	0.5 – 0.8	0 – 500
24	0.85 – 1.0	600 – 1200	0.5 – 0.8	500 – 1500
25	0.85 – 1.0	600 – 1200	0.8 – 1.0	0 – 300
26	0.85 – 1.0	600 – 1200	0.8 – 1.0	300 – 500
27	0.85 – 1.0	600 – 1200	0.8 – 1.0	500 – 700
28	0.85 – 1.0	600 – 1200	0.8 – 1.0	700 – 1500
29	0.85 – 1.0	1200 – 30000	–1.0 – 0.5	0 – 1500
30	0.85 – 1.0	1200 – 30000	0.5 – 0.8	0 – 300
31	0.85 – 1.0	1200 – 30000	0.5 – 0.8	300 – 500
32	0.85 – 1.0	1200 – 30000	0.5 – 0.8	500 – 700
33	0.85 – 1.0	1200 – 30000	0.5 – 0.8	700 – 1500
34	0.85 – 1.0	1200 – 30000	0.8 – 1.0	0 – 300
35	0.85 – 1.0	1200 – 30000	0.8 – 1.0	300 – 400
36	0.85 – 1.0	1200 – 30000	0.8 – 1.0	400 – 500
37	0.85 – 1.0	1200 – 30000	0.8 – 1.0	500 – 600
38	0.85 – 1.0	1200 – 30000	0.8 – 1.0	600 – 700
39	0.85 – 1.0	1200 – 30000	0.8 – 1.0	700 – 900
40	0.85 – 1.0	1200 – 30000	0.8 – 1.0	900 – 1100
41	0.85 – 1.0	1200 – 30000	0.8 – 1.0	1100 – 1500

Table F.I: CS1 sample binning scheme, for FGD1, FGD2x and FGD2y samples.

Bin Number	$\cos\theta_\mu$	p_μ	$\cos\theta_\pi$	p_π
0	–1.0 – 0.3	0 – 30000	–1.0 – 1.0	0 – 30000
1	0.3 – 1.0	0 – 200	–1.0 – 1.0	0 – 30000
2	0.3 – 1.0	200 – 30000	–1.0 – 1.0	1500 – 30000
3	0.3 – 0.6	200 – 30000	–1.0 – 0.5	0 – 1500
4	0.3 – 0.6	200 – 30000	0.5 – 0.8	0 – 300
5	0.3 – 0.6	200 – 30000	0.5 – 0.8	300 – 1500
6	0.3 – 0.6	200 – 30000	0.8 – 1.0	0 – 300
7	0.3 – 0.6	200 – 30000	0.8 – 1.0	300 – 500
8	0.3 – 0.6	200 – 30000	0.8 – 1.0	500 – 1500
9	0.6 – 0.85	200 – 600	–1.0 – 0.5	0 – 1500
10	0.6 – 0.85	200 – 600	0.5 – 0.8	0 – 300
11	0.6 – 0.85	200 – 600	0.5 – 0.8	300 – 1500
12	0.6 – 0.85	200 – 600	0.8 – 1.0	0 – 300
13	0.6 – 0.85	200 – 600	0.8 – 1.0	300 – 500
14	0.6 – 0.85	200 – 600	0.8 – 1.0	500 – 1500
15	0.6 – 0.85	600 – 30000	–1.0 – 0.5	0 – 1500
16	0.6 – 0.85	600 – 30000	0.6 – 0.8	0 – 300
17	0.6 – 0.85	600 – 30000	0.6 – 0.8	300 – 500
18	0.6 – 0.85	600 – 30000	0.6 – 0.8	500 – 1500
19	0.6 – 0.85	600 – 30000	0.8 – 1.0	0 – 300
20	0.6 – 0.85	600 – 30000	0.8 – 1.0	300 – 500
21	0.6 – 0.85	600 – 30000	0.8 – 1.0	500 – 700
22	0.6 – 0.85	600 – 30000	0.8 – 1.0	700 – 1500
23	0.85 – 1.0	200 – 600	–1.0 – 0.5	0 – 1500
24	0.85 – 1.0	200 – 600	0.5 – 0.8	0 – 300
25	0.85 – 1.0	200 – 600	0.5 – 0.8	300 – 1500
26	0.85 – 1.0	200 – 600	0.8 – 1.0	0 – 300
27	0.85 – 1.0	200 – 600	0.8 – 1.0	300 – 500
28	0.85 – 1.0	200 – 600	0.8 – 1.0	500 – 1500
29	0.85 – 1.0	600 – 1200	–1.0 – 0.5	0 – 1500
30	0.85 – 1.0	600 – 1200	0.5 – 0.8	0 – 300
31	0.85 – 1.0	600 – 1200	0.5 – 0.8	300 – 500
32	0.85 – 1.0	600 – 1200	0.5 – 0.8	500 – 1500
33	0.85 – 1.0	600 – 1200	0.8 – 1.0	0 – 300
34	0.85 – 1.0	600 – 1200	0.8 – 1.0	300 – 400
35	0.85 – 1.0	600 – 1200	0.8 – 1.0	400 – 500
36	0.85 – 1.0	600 – 1200	0.8 – 1.0	500 – 700
37	0.85 – 1.0	600 – 1200	0.8 – 1.0	700 – 1500
38	0.85 – 1.0	1200 – 30000	–1.0 – 0.5	0 – 1500
39	0.85 – 1.0	1200 – 30000	0.5 – 0.8	0 – 200
40	0.85 – 1.0	1200 – 30000	0.5 – 0.8	200 – 300
41	0.85 – 1.0	1200 – 30000	0.5 – 0.8	300 – 400
42	0.85 – 1.0	1200 – 30000	0.5 – 0.8	400 – 500
43	0.85 – 1.0	1200 – 30000	0.5 – 0.8	500 – 700
44	0.85 – 1.0	1200 – 30000	0.5 – 0.8	700 – 1500
45	0.85 – 1.0	1200 – 30000	0.8 – 1.0	0 – 200
46	0.85 – 1.0	1200 – 30000	0.8 – 1.0	200 – 300
47	0.85 – 1.0	1200 – 30000	0.8 – 1.0	300 – 400
48	0.85 – 1.0	1200 – 30000	0.8 – 1.0	400 – 500
49	0.85 – 1.0	1200 – 30000	0.8 – 1.0	500 – 600
50	0.85 – 1.0	1200 – 30000	0.8 – 1.0	600 – 700
51	0.85 – 1.0	1200 – 30000	0.8 – 1.0	700 – 900
52	0.85 – 1.0	1200 – 30000	0.8 – 1.0	900 – 1100
53	0.85 – 1.0	1200 – 30000	0.8 – 1.0	1100 – 1500

Table F.II: CS2 sample binning scheme, for FGD1, FGD2x and FGD2y samples.

Bin Number	$\cos\theta_\mu$	p_μ	$\cos\theta_\pi$	p_π
0	–1.0 – 0.3	0 – 30000	–1.0 – 1.0	0 – 30000
1	0.3 – 1.0	0 – 200	–1.0 – 1.0	0 – 30000
2	0.3 – 0.6	200 – 30000	–1.0 – 1.0	0 – 1500
3	0.6 – 0.85	600 – 30000	–1.0 – 1.0	0 – 1500
4	0.6 – 0.85	600 – 30000	–1.0 – 1.0	0 – 1500
5	0.85 – 1.0	200 – 600	–1.0 – 1.0	0 – 1500
6	0.85 – 1.0	600 – 1200	–1.0 – 1.0	0 – 1500
7	0.85 – 1.0	1200 – 30000	–1.0 – 1.0	0 – 1500

Table F.III: CS3 sample binning scheme, for FGD1, FGD2x and FGD2y samples.

Bibliography

[1] W. Pauli, "Dear Radioactive Ladies and Gentlemen", *Phys. Today* 31N9 (1978) 27, http://cds.cern.ch/record/83282

[2] C. Sutton "Spaceship Neutrino - Foreword by F. Reines", Cambridge University Press, 15 Oct 1992

[3] F. Close, "Neutrino", Oxford University Press, 2010

[4] W.J. Henderson, "The Upper Limits of the Continuous β-ray Spectra of Thorium C and C$''$", *Proc. R. Soc. Lond. A* **147**, 572–582, http://doi.org/10.1098/rspa.1934.0237

[5] H.R. Crane and J. Halpern, "New Experimental Evidence for the Existence of a Neutrino", *Phys. Rev. 53* **789**, 15 May 1938, https://doi.org/10.1103/PhysRev.53.789

[6] M.E. Nahmias, "An Attempt to Detect the Neutrino", *Mathematical Proceedings of the Cambridge Philosophical Society* **31**, Issue 1, January 1935, pp. 99 - 107, https://doi.org/10.1017/S0305004100012986

[7] F. Reines, "The neutrino: from poltergeist to particle", *Reviews of Modern Physics* **68**, 317-327 (1996), https://doi.org/10.1103/RevModPhys.68.317

[8] N.G. Cooper, "The Reines-Cowan Experiments - Detecting the Poltergeist", *Los Alamos Science* **Number 25**, 1997

[9] F. Reines and C.L. Cowan Jr, "Detection of the free neutrino", *Phys. Rev.* **92**, 830 (1953), https://doi.org/10.1103/PhysRev.92.830

[10] C.L. Cowan, F. Reines, F.B. Harrison, H.W. Kruse, A.D. McGuire, "Detection of the free neutrino: a confirmation", *Science* **124** 103-104 (1956), https://doi.org/10.1126/science.124.3212.103

[11] Nobel Prize Outreach, "The Nobel Prize in Physics 1995", accessible from https://www.nobelprize.org/prizes/physics/1995/summary/

[12] G. Danby, J-M. Gaillard, K. Goulianos, L.M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger, "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos", *Phys. Rev. Lett.* **9**, 36 (1962), https://doi.org/10.1103/PhysRevLett.9.36

[13] Nobel Prize Outreach, "The Nobel Prize in Physics 1988", accessible from https://www.nobelprize.org/prizes/physics/1988/summary/

[14] M.L. Perl *et al.*, "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation, *Phys. Rev. Lett.* **35**, 1489, 1 December 1975, https://doi.org/10.1103/PhysRevLett.35.1489

[15] J. Kirkby, "Review of e^+e^- Reactions in the Energy Range 3 GeV to 9 GeV", *9th International Symposium on Lepton and Photon Interactions at High Energy*, SLAC-PUB-2419, 1979

[16] C.A. Blocker *et al.*, "A Study of the Decay $\tau^- \to \pi^-\nu_\tau$" , *Phys. Lett.* **109B**, 119–123 (1982)

[17] V. Vuillemin, "$W^\pm$ and Z^0 Production in the UA1 Experiment at the CERN Proton-Antiproton Collider", *Annals N. Y. Acad. Sci.* **461**, 99 (1986), https://doi.org/10.1111/j.1749-6632.1986.tb52412.x

[18] K. Kodama *et al.* (DONUT Collaboration), "Observation of tau neutrino interactions", *Physics Letters B*, **504**, Issue 3, 218-224, (2001), https://doi.org/10.1016/S0370-2693(01)00307-0

[19] P.A. Zyla et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2020, 083C01 (2020), https://doi.org/10.1093/ptep/ptaa104

[20] F.J. Hasert *et al.*, (Gargamelle Collaboration), "Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment", *Phys. Lett. B* **46** (1973) 138-140, https://doi.org/10.1016/0370-2693(73)90499-1

[21] L. Di Lella and C. Rubbia, "The Discovery of the W and Z Particles", *Adv. Ser. Direct. High Energy Phys.* **23** (2015), https://doi.org/10.1142/9441

[22] A. Blondel, "The third family of neutrinos", *Proceedings of the International Conference on History of the Neutrino: 1930-2018 : Paris, France*, September 5-7, 2018, arXiv:1812.11362v2 [physics.hist-ph]

[23] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups, "Precision Electroweak Measurements on the Z Resonance", *Phys. Rept.* **427**, 257-454 (2006), https://doi.org/10.1016/j.physrep.2005.12.006

[24] The ALEPH Collaboration, "Determination of the number of light neutrino species", *Physics Letters B* **Vol 231**, Issue 4, 519-529, (1989), https://doi.org/10.1016/0370-2693(89)90704-1

[25] B. Pontecorvo, "Inverse β process", *National Research Council of Canada (NRC) Report No. PD 205*, 20 November 1946

[26] J.N. Bahcall, "Neutrinos from the Sun", *Scientific American* **221** (1), July 1969, pp. 28-37, https://www.jstor.org/stable/24926407

[27] R. Davis Jr., "Attempt to Detect the Antineutrinos from a Nuclear Reactor by the $Cl^{37}(\bar{\nu}, e^{-})Ar^{37}$ Reaction", *Phys. Rev.* **97**, 766, 1 February 1955, https://doi.org/10.1103/PhysRev.97.766

[28] B.T. Cleveland, T. Daily, R. Davis Jr., J.R. Distel, K. Lande, C.K. Lee, P.S. Wildenhain and J. Ullman, "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector", *The Astrophyiscal Journal* **496**, 505-526 (1998), https://doi.org/10.1086/305343

[29] J.N. Bahcall, P.I. Krastev, A. Yu. Smirnov, "Where do we stand with solar neutrino oscillations?", *Phys. Rev. D* **58**, 096016 (1998), https://doi.org/10.1103/PhysRevD.58.096016

[30] M. Cribier for the GALLEX Collaboration, "Results of the whole GALLEX experiment", *Nuclear Physics B (Proc. Suppl.)* **70** 284-29 (1999), https://doi.org/10.1016/S0920-5632(98)00438-1

[31] D. Vignaud, "The GALLEX solar neutrino experiment", *Nuclear Physics B (Proc. Suppl.)* **60**, 20-29 (1998), https://doi.org/10.1016/S0920-5632(97)00498-2

[32] J.N. Abdurashitov *et al.*, "Results from SAGE (The Russian-American Gallium solar neutrino Experiment)", *Physics Letters B* **328**, 234-248 (1994), https://doi.org/10.1016/0370-2693(94)90454-5

[33] Y. Suzuki, "Kamiokande Solar Neutrino Results", *Nuclear Physics B (Proc. Suppl.)* **38**, 54-59 (1995), https://doi.org/10.1016/0920-5632(94)00733-C

[34] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), "Measurement of the Solar Neutrino Energy Spectrum Using Neutrino-Electron Scattering", *Phys. Rev. Lett.* **82**, 2430, 22 March 1999, https://doi.org/10.1103/PhysRevLett.82.2430

[35] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), "Evidence for Oscillation of Atmospheric Neutrinos", *Phys. Rev. Lett.* **81**, 1562, 24 August 1998, https://doi.org/10.1103/PhysRevLett.81.1562

[36] A.B. McDonald for the SNO Collaboration, "First neutrino observations from the Sudbury Neutrino Observatory", *Nucl.Phys.B Proc. Suppl.* **91** 21-28 (2001), https://doi.org/10.1016/S0920-5632(00)00918-X

[37] Nobel Prize Outreach, "The Nobel Prize in Physics 2015", accessible from https://www.nobelprize.org/prizes/physics/2015/press-release/

[38] M. Gell-Mann and A. Pais, "Behavior of Neutral Particles under Charge Conjugation", *Phys. Rev.* **97**, 1387 (1955), https://doi.org/10.1103/PhysRev.97.1387

[39] B. Pontecorvo, "Neutrino Experiments and the Problem of Conservation of Leptonic Charge", *Soviet Physics JETP* **26** (5) May, 1968

[40] Z. Maki, M. Nakagawa and S. Sakata, "Remarks on the Unified Model of Elementary Particles", *Progress of Theoretical Physics*, **Vol. 28**, No. 5, November 1962, https://doi.org/10.1143/PTP.28.870

[41] G.C. Branco and M.N. Rebelo, "Building the Full PMNS Matrix From Six Independent Majorana-Type Phases", *Phys. Rev. D* **79**, 013001 (2009), https://doi.org/10.1103/PhysRevD.79.013001

[42] M.C. Gonzalez-Garcia, M. Yokoyama, "Neutrino Masses, Mixing and Oscillations", *Prog. Theor. Exp. Phys. 2020*, 083C01 (2020)

[43] H.M. O'Keeffe, "Low Energy Background in the NCD Phase of the Sudbury Neutrino Observatory", University of Oxford Doctoral Thesis, 2008, https://ora.ox.ac.uk/objects/uuid:b53c04b1-5dba-4cfb-98ec-c0d8b87fb58b

[44] A.Y. Smirnov, "Solar neutrinos: Oscillations or no-oscillations?", arXiv:1609.02386v2 [hep-ph] 19 Sep 2017

[45] L. Wolfenstein, "Neutrino oscillations in matter", *Phys. Rev. D* **17**, 2369, 1 May 1978, https://doi.org/10.1103/PhysRevD.17.2369

[46] S.P. Mikheyev and A.Y. Smirnov, "Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos", *Sov.J.Nucl.Phys.* **42** (1985) 913-917, *Yad.Fiz.* **42** (1985) 1441-1448

[47] I. Esteban, M.C. Gonzalez-Garcia, M. Maltoni, T. Schwetz and A. Zhou, "The fate of hints: updated global analysis of three-flavor neutrino oscillations", *JHEP* **2020**, 178 (2020), https://doi.org/10.1007/JHEP09(2020)178, NuFIT 5.0 (2020), www.nu-fit.org

[48] S. Vagnozzi, "Weigh them all! - Cosmological searches for the neutrino mass scale and mass ordering", Stockholm University Doctoral Thesis, 2019, https://doi.org/10.1007/978-3-030-53502-5

[49] X. Qian and P. Vogel, "Neutrino Mass Hierarchy", *PPNP* **83**, 1-30 (2015), https://doi.org/10.1016/j.ppnp.2015.05.002

[50] A. Himmel for the NOνA Collaboration, "New Oscillation Results from the NOνA Experiment", *Neutrino 2020 Presentation*, 2 July 2020, https://indico.fnal.gov/event/43209/contributions/187840/attachments/130740/159597/NOvA-Oscilations-NEUTRINO2020.pdf

[51] P. Dunne for the T2K Collaboration, "Latest Neutrino Oscillation Results from T2K", *Neutrino 2020 Presentation*, https://doi.org/10.5281/zenodo.3959558

[52] A. Timmons, "The Results of MINOS and the Future with MINOS+", *Adv. High Energy Phys.* **2016** 7064960 (2016), https://doi.org/10.1155/2016/7064960

[53] T. Nakaya, "K2K Results", *Nuclear Physics B - Proceedings Supplements* **143** (1):96-103, (2005), https://doi.org/10.1016/j.nuclphysbps.2005.01.093

[54] P.F. de Salas, D.V. Forero, S. Gariazzo *et al.*, "2020 global reassessment of the neutrino oscillation picture",*J. High Energ. Phys.* **2021**, 71 (2021), https://doi.org/10.1007/JHEP02(2021)071

[55] The T2K Collaboration, "Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations", *Nature* **580**, 339–344 (2020), https://doi.org/10.1038/s41586-020-2177-0

[56] Y. Kudenko for the Hyper-Kamiokande Proto Collaboration, "Hyper-Kamiokande", *JINST* **15** (2020) 07, C07029, https://doi.org/10.1088/1748-0221/15/07/C07029

[57] B. Abi *et al.* (DUNE Collaboration), "Long-baseline neutrino oscillation physics potential of the DUNE experiment", *Eur. Phys. J. C* **80**, 978 (2020), https://doi.org/10.1140/epjc/s10052-020-08456-z

[58] B. Roskovec for the Daya Bay Collaboration, "Latest Reactor Neutrino Oscillation Results from the Daya Bay Experiment", *PoS* **ICHEP2020** (2021) 170, https://doi.org/10.22323/1.390.0170

[59] H. Seo for the RENO Collaboration "Recent Result from RENO", *J.Phys. Conf. Ser.* **1216** (2019) 1, 012003, https://doi.org/10.1088/1742-6596/1216/1/012003

[60] The Double Chooz Collaboration, "Double Chooz θ_{13} measurement via total neutron capture detection", *Nature Physics* **16**, pages 558–564 (2020), https://doi.org/10.1038/s41567-020-0831-y

[61] A. Gando *et al.*, "Constraints on θ_{13} from a three-flavor oscillation analysis of reactor antineutrinos at KamLAND", *Phys. Rev. D* **83**, 052002 (2011), https://doi.org/10.1103/PhysRevD.83.052002

[62] D. Adey *et al.* (The Daya Bay Collaboration), "Measurement of the Electron Antineutrino Oscillation with 1958 Days of Operation at Daya Bay", *Phys. Rev. Lett.* **121**, 241805 (2018), https://doi.org/10.1103/PhysRevLett.121.241805

[63] J. Yoo for the RENO Collaboration, "Recent Results from the RENO Experiment", *Neutrino 2020 Presentation*, https://doi.org/10.5281/zenodo.4123573

[64] F.P. An *et al.*, "New Measurement of Antineutrino Oscillation with the Full Detector Configuration at Daya Bay", *Phys. Rev. Lett.* **115**, 111802, 11 September 2015, https://doi.org/10.1103/PhysRevLett.115.111802

[65] H. Lu for the JUNO Collaboration, "The physics potentials of JUNO", *Phys.Scripta* **96** (2021) 9, 094013, https://doi.org/10.1088/1402-4896/ac0a29

[66] K.K. Joo, "New results from RENO & prospects with RENO-50", *J. Phys.: Conf. Ser.* **888** 012012, https://doi.org/10.1088/1742-6596/888/1/012012

[67] M.V. Diwan, V. Galymov, X. Qian, and A. Rubbia, "Long-Baseline Neutrino Experiments", *Annu. Rev. Nucl. Part. Sci.* **66:47–71**, 1 June 2016, https://doi.org/10.1146/annurev-nucl-102014-021939

[68] M. Altmann *et al.* (GNO Collaboration), "Complete results for five years of GNO solar neutrino observations", *Physics Letters B* **616** (3-4):174-190 June 2005, https://doi.org/10.1016/j.physletb.2005.04.068

[69] J.N. Abdurashitov *et al.* (SAGE Collaboration), "Measurement of the solar neutrino capture rate with gallium metal. III. Results for the 2002–2007 data-taking period", *Phys. Rev. C* **80**, 015807 30 July 2009, https://doi.org/10.1103/PhysRevC.80.015807

[70] B.T. Cleveland *et al.* "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector", *ApJ* **496** 505 1998, https://doi.org/10.1086/305343

[71] G. Bellini *et al.* (Borexino Collaboration), "Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy", *Phys. Rev. D* **89**, 112007, 25 June 2014, https://doi.org/10.1103/PhysRevD.89.112007

[72] Y. Nakajima for the Super-Kamiokande Collaboration, "Recent results and future prospects from Super-Kamiokande", *Neutrino 2020 Presentation*, https://doi.org/10.5281/zenodo.4134680

[73] B. Aharmim *et al.*, "Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory", *Phys. Rev. C* **88**, 025501, 1 August 2013, https://doi.org/10.1103/PhysRevC.88.025501

[74] D. Williams for the IceCube Collaboration, "Recent Results from IceCube", *21st Particles and Nuclei International Conference (PANIC 2017) International Journal of Modern Physics: Conference Series* **46** (2018), http://dx.doi.org/10.1142/S2010194518600480

[75] A. Albert *et al.* (The ANTARES Collaboration), "Measuring the atmospheric neutrino oscillation parameters and constraining the 3+1 neutrino model with ten years of ANTARES data", *J. High Energ. Phys.* **2019**, 113 (2019), https://doi.org/10.1007/JHEP06(2019)113

[76] M.G. Aartsen *et al.*, "Measurement of atmospheric tau neutrino appearance with IceCube DeepCore", *Phys. Rev. D* **99** 032007, 15 February 2019, https://doi.org/10.1103/PhysRevD.99.032007

[77] M.G. Aartsen *et al.*, "PINGU: a vision for neutrino and particle physics at the South Pole", *J. Phys. G: Nucl. Part. Phys.* **44** 054006 (2017), https://doi.org/10.1088/1361-6471/44/5/054006

[78] S. Aiello *et al.* (The KM3NeT Collaboration), "Determining the Neutrino Mass Ordering and Oscillation Parameters with KM3NeT/ORCA", arXiv:2103.09885 [hep-ex]

[79] J.A. Formaggio, G.P. Zeller, "From eV to EeV: Neutrino Cross-Sections Across Energy Scales", *Rev. Mod. Phys.* **84**, 1307 (2012), https://doi.org/10.1103/RevModPhys.84.1307

[80] C.H. Llewellyn Smith, "Neutrino reactions at accelerator energies", *Physics Reports* **Vol 3**, Issue 5, June 1972, 261-379, https://doi.org/10.1016/0370-1573(72)90010-5

[81] R. Bradford, A. Bodek, H. Budd, J. Arrington, "A New Parameterization of the Nucleon Elastic Form Factors", *Nuclear Physics B (Proc. Suppl.)* **159** (2006) 127–132, https://doi.org/10.1016/j.nuclphysbps.2006.08.028

[82] R. Gran *et al.* (K2K Collaboration), "Measurement of the quasi-elastic axial vector mass in neutrino-oxygen interactions", *Phys.Rev.D 74* (2006) 052002, https://doi.org/10.1103/PhysRevD.74.052002

[83] M.L. Goldberger and S.B. Treiman, "Form Factors in β Decay and μ Capture", *Phys. Rev.* **111**, 354, 1 July 1958, https://doi.org/10.1103/PhysRev.111.354

[84] J.T. Suhonen, "Value of the Axial-Vector Coupling Strength in β and $\beta\beta$ Decays: A Review", *Front. Phys.* **5**:55 (2017), DOI:10.3389/fphy.2017.00055

[85] V. Bernard, L. Elouadrhiri, U.G. Meißner, "Axial structure of the nucleon", *J. Phys. G: Nucl. Part. Phys.* **28** R1 (2001), https://doi.org/10.1088/0954-3899/28/1/201

[86] A. Bodek, H.S. Budd, M.E. Christy, "Resolving the Axial Mass Anomaly in Neutrino Scattering", *AIP Conference Proceedings* **1441**, 420 (2012), https://doi.org/10.1063/1.3700572

[87] J. Nieves, I. Ruiz Simo, M.J. Vicente Vacas, *The nucleon axial mass and the MiniBooNE Quasielastic Neutrino–Nucleus Scattering problem*, *Phys. Lett. B* **707**:1 (2012) https://doi.org/10.1016/j.physletb.2011.11.061

[88] A.A. Aguilar-Arevalo et al., *First Measurement of the Muon Neutrino Charged Current Quasielastic Double Differential Cross Section*, *Phys. Rev. D* **81**, 092005 (2010), https://doi.org/10.1103/PhysRevD.81.092005

[89] B. Bhattacharya, R.J. Hill, G. Paz, "Model-independent determination of the axial mass parameter in quasielastic neutrino-nucleon scattering", *Phys. Rev. D* **84**, 073006 (2011), https://doi.org/10.1103/PhysRevD.84.073006

[90] M. Kabirnezhad, "Single pion production in neutrino-nucleon interactions", *Phys. Rev. D* **97**, 013002 (2018), https://doi.org/10.1103/PhysRevD.97.013002

[91] D. Rein, L.M. Sehgal, "Neutrino-excitation of baryon resonances and single pion production", *Annals of Physics* **133**, 79-153 (1981), https://doi.org/10.1016/0003-4916(81)90242-6

[92] R.P. Feynman, M.Kislinger, and F.Ravndal, "Current Matrix Elements from a Relativistic Quark Model", *Phys. Rev. D* **3**, 2706, 1 June 1971, https://doi.org/10.1103/PhysRevD.3.2706

[93] K.M. Graczyk, D. Kiełczewska, P. Przewłocki, J.T. Sobczyk, "C_5^A axial form factor from bubble chamber experiments", *Phys. Rev. D* **80**, 093001 (2009) https://doi.org/10.1103/PhysRevD.80.093001

[94] Y. Hayato, "A neutrino interaction simulation program library NEUT", *Acta Phys.Polon.B* **40** (2009) 2477-2489

[95] C. Berger, L.M. Sehgal, "Lepton Mass Effects in Single Pion Production by Neutrinos", *Phys. Rev. D* **76**, 113004 (2007), https://doi.org/10.1103/PhysRevD.76.113004

[96] M.S. Athar and J.G. Morfin, "Neutrino(antineutrino)–nucleus interactions in the shallow- and deep-inelastic scattering regions", *J. Phys. G: Nucl. Part. Phys.* **48** 034001 (2021), https://doi.org/10.1088/1361-6471/abbb11

[97] J.M. Conrad, M.H. Shaevitz, T. Bolton, "Precision measurements with high-energy neutrino beams", *Rev. Mod. Phys.* **70**, 1341, 1 October 1998, https://doi.org/10.1103/RevModPhys.70.1341

[98] O. Benhar, N. Farina, H. Nakamura, M. Sakuda, R. Seki, "Electron- and neutrino-nucleus scattering in the impulse approximation regime", *Phys.Rev. D* **72** (2005) 053005, https://doi.org/10.1103/PhysRevD.72.053005

[99] R.A. Smith, E.J. Moniz, "Neutrino reactions on nuclear targets", *Nuclear Physics B* **43** (1972) 605-622, https://doi.org/10.1016/0550-3213(72)90040-5

[100] L. Alvarez-Ruso, Y. Hayato, J. Nieves, "Progress and open questions in the physics of neutrino cross sections at intermediate energies", *New J. Phys.* **16**, 075015 (2014), https://doi.org/10.1088/1367-2630/16/7/075015

[101] A. Bodek, M.E. Christy, B. Coopersmith, "Effective Spectral Function for Quasielastic Scattering on Nuclei", *Eur. Phys. J. C* **74** (2014) 3091, https://doi.org/10.1140/epjc/s10052-014-3091-0

[102] O. Benhar, A. Fabrocini, S. Fantoni, I. Sick, "Spectral function of finite nuclei and scattering of GeV electrons", *Nuclear Physics A* **579** (1994) 493-517, https://doi.org/10.1016/0375-9474(94)90920-2

[103] C. Wilkinson, "Constraining neutrino interaction uncertainties for oscillation experiments", PhD thesis, University of Sheffield (2015), https://etheses.whiterose.ac.uk/9198/

[104] J.E. Amaro, M.B. Barbaro, J.A. Caballero, T.W. Donnelly, C.F. Williamson, "Meson-exchange currents and quasielastic neutrino cross sections in the superscaling approximation model", *Physics Letters B* **696** (2011) 151-155, https://doi.org/10.1016/j.physletb.2010.12.007

[105] O. Benhar, P. Huber, C. Mariani, D. Meloni, "Neutrino–nucleus interactions and the determination of oscillation parameters", *Physics Reports* **Vol 700** 1-47 (2017), https://doi.org/10.1016/j.physrep.2017.07.004

[106] T. Golan, "What is inside MC generators... and why it is wrong", "NuSTEC School in Neutrino-Nucleus Scattering Physics 2017", http://www.ift.uni.wroc.pl/~tgolan/talks/nustec_mc_02.pdf

[107] B.J. Kim *et al.*, "Tracking performance of the scintillating fiber detector in the K2K experiment", *Nucl. Instrum. Meth. A* **497** (2003) 450-466, https://doi.org/10.1016/S0168-9002(02)01980-0

[108] R. Gran, E.J. Jeon *et al.* (The K2K Collaboration), "Measurement of the quasi-elastic axial vector mass in neutrino-oxygen interactions", *Phys. Rev. D* **74**, 052002 (2006), https://doi.org/10.1103/PhysRevD.74.052002

[109] J.L. Alcaraz-Aunion, J. Walding, "Measurement of the muon-neutrino charged-current quasi-elastic cross-section in the SciBooNE experiment", *AIP Conference Proceedings* **1189**, 145 (2009), https://doi.org/10.1063/1.3274145

[110] M. Dorman for the MINOS Collaboration, "Preliminary Results for CCQE Scattering with the MINOS Near Detector", *AIP Conference Proceedings 1189*, 133 (2009), https://doi.org/10.1063/1.3274143

[111] K. Abe *et al.*, "Measurement of the ν_μ CCQE cross section on carbon with the ND280 detector at T2K", *Phys. Rev. D* **92**, 112003 (2015), https://doi.org/10.1103/PhysRevD.92.112003

[112] L. Aliaga *et al.* (MINERνA Collaboration), "Design, calibration, and performance of the MINERνA detector", *Nuclear Instruments and Methods in Physics Research A* **743** 130-159 (2014), https://doi.org/10.1016/j.nima.2013.12.053

[113] G.A. Fiorentini *et al.*, "Measurement of Muon Neutrino Quasielastic Scattering on a Hydrocarbon Target at $E_\nu \sim 3.5$ GeV", *Phys. Rev. Lett.* **111**, 022502 (2013), https://doi.org/10.1103/PhysRevLett.111.022502

[114] J. Nieves, I. Ruiz Simo, M.J. Vicente Vacas, "The nucleon axial mass and the MiniBooNE Quasielastic Neutrino-Nucleus Scattering problem", *Physics Letters B* **707** 72-25 (2012), https://doi.org/10.1016/j.physletb.2011.11.061

[115] C. Wilkinson *et al.*, "Testing charged current quasi-elastic and multinucleon interaction models in the NEUT neutrino interaction generator with published datasets

from the MiniBooNE and MINERνA experiments", *Phys. Rev. D* **93**, 072010 (2016), https://doi.org/10.1103/PhysRevD.93.072010

[116] C.E. Patrick *et al.* (MINERνA Collaboration), "Measurement of the muon antineutrino double-differential cross section for quasielastic-like scattering on hydrocarbon at $E_\nu\sim$ 3.5 GeV", *Phys. Rev. D* **97**, 052002 (2018), https://doi.org/10.1103/PhysRevD.97.052002

[117] D. Ruterbories *et al.* (MINERνA Collaboration), "Measurement of quasielastic-like neutrino scattering at $\langle E_\nu\rangle \sim$3.5 GeV on a hydrocarbon target", *Phys. Rev. D* **99**, 012004 (2019), https://doi.org/10.1103/PhysRevD.99.012004

[118] K. Abe *et al.* (The T2K Collaboration), "Measurement of double-differential muon neutrino charged-current interactions on C_8H_8 without pions in the final state using the T2K off-axis beam", *Phys. Rev.* **D** 93, 112012 (2016), https://doi.org/10.1103/PhysRevD.93.112012

[119] K. Abe *et al.* (The T2K Collaboration), "First measurement of the ν_μ charged-current cross section on a water target without pions in the final state", *Phys. Rev. D* **97**, 012001 (2018), https://doi.org/10.1103/PhysRevD.97.012001

[120] K. Abe *et al.* (The T2K Collaboration), "First measurement of the charged current $\bar{\nu}_\mu$ double differential cross section on a water target without pions in the final state", *Phys. Rev. D* **102**, 012007 (2020), https://doi.org/10.1103/PhysRevD.102.012007

[121] C.L. McGivern *et al.* (MINERνA Collaboration), "Cross sections for ν_μ and $\bar{\nu}_\mu$ induced pion production on hydrocarbon in the few-GeV region using MINERvA", *Phys. Rev. D* **94**, 052005 (2016), https://doi.org/10.1103/PhysRevD.94.052005

[122] O. Altinok *et al.* (MINERνA Collaboration), "Measurement of ν_μ charged-current single π^0 production on hydrocarbon in the few-GeV region using MINERvA", *Phys. Rev. D* **96**, 072003(2017), https://doi.org/10.1103/PhysRevD.96.072003

[123] T. Le *et al.* (MINERνA Collaboration), "Measurement of $\bar{\nu}_\mu$ charged-current single π^- production on hydrocarbon in the few-GeV region using MINERvA", *Phys. Rev. D* **100**, 052008 (2019), https://doi.org/10.1103/PhysRevD.100.052008

[124] A.A. Aguilar-Arevalo *et al.* (The MiniBooNE Collaboration), "Measurement of ν_μ and $\bar{\nu}_\mu$ induced neutral current single π^0 production cross sections on mineral oil at $E_\nu \sim \mathcal{O}(1$ GeV)", *Phys. Rev. D* **81**, 013005 (2010), https://doi.org/10.1103/PhysRevD.81.013005

[125] A.A. Aguilar-Arevalo *et al.* (The MiniBooNE Collaboration), "Measurement of neutrino-induced charged-current charged pion production cross sections on mineral oil at $E_\nu \sim 1$ GeV", *Phys. Rev. D* **83**, 052007 (2011), https://doi.org/10.1103/PhysRevD.83.052007

[126] A.A. Aguilar-Arevalo *et al.* (The MiniBooNE Collaboration), "Measurement of ν_μ-induced charged-current neutral pion production cross sections on mineral oil at $E_\nu \in 0.5 - 2.0$ GeV", *Phys. Rev. D* **83**, 052009 (2011), https://doi.org/10.1103/PhysRevD.83.052009

[127] K. Abe *et al.* (The T2K Collaboration), "First measurement of the muon neutrino charged current single pion production cross section on water with the T2K near detector", *Phys. Rev. D* **95**, 012010 (2017), https://doi.org/10.1103/PhysRevD.95.012010

[128] K. Abe *et al.* (The T2K Collaboration), "Measurement of the muon neutrino charged-current single π^+ production on hydrocarbon using the T2K off-axis near detector ND280", *Phys. Rev. D* **101**, 012007 (2020), https://doi.org/10.1103/PhysRevD.101.012007

[129] X.-G. Lu *et al.*, "Measurement of nuclear effects in neutrino interactions with minimal dependence on neutrino energy", *Phys. Rev. C* **94**, 015503 (2016), https://doi.org/10.1103/PhysRevC.94.015503

[130] X.-G. Lu *et al.* (MINERνA Collaboration), "Measurement of Final-State Correlations in Neutrino Muon-Proton Mesonless Production on Hydrocarbon at $\langle E_\nu \rangle = 3$ GeV", *Phys. Rev. Lett.* **121**, 022504 (2018), https://doi.org/10.1103/PhysRevLett.121.022504

[131] K. Abe *et al.* (The T2K Collaboration), "Characterization of nuclear effects in muon-neutrino scattering on hydrocarbon with a measurement of final-state kinematics and correlations in charged-current pionless interactions at T2K", *Phys. Rev. D* **98**, 032003 (2018), https://doi.org/10.1103/PhysRevD.98.032003

[132] D. Coplowe *et al.* (MINERνA Collaboration), "Probing nuclear effects with neutrino-induced charged-current neutral pion production", *Phys. Rev. D* **102**, 072007 (2020), https://doi.org/10.1103/PhysRevD.102.072007

[133] K. Abe *et al.* (The T2K Collaboration), First T2K measurement of transverse kinematic imbalance in the muon-neutrino charged-current single-π^+ production channel containing at least one proton", *Phys. Rev. D* **103**, 112009 (2021), https://doi.org/10.1103/PhysRevD.103.112009

[134] The T2K Collaboration, "T2K Experiment Public Webpage", accessed from https://t2k-experiment.org/t2k/

[135] Japan Proton Accelerator Research Complex, "Accelerators", accessed from https://j-parc.jp/public/Acc/en/index.html

[136] Japan Atomic Energy Research Institute, "Accelerator technical design report for high-intensity proton accelerator facility project, J-PARC", *JAERI-TECH* **2003-044**, http://dx.doi.org/10.11484/JAERI-Tech-2003-044

[137] K. Abe *et al.*, (The T2K Collaboration), "The T2K Experiment", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **Volume 659**, Issue 1, 2011, Pages 106-135, ISSN 0168-9002, https://doi.org/10.1016/j.nima.2011.06.067.

[138] T. Sekiguchi, "Neutrino Beamline for T2K", *AIP Conference Proceedings* **981**, 275 (2008), https://doi.org/10.1063/1.2898959

[139] Nakamura, T. *et al* "The OPERA film: New nuclear emulsion for large-scale, high-precision experiments", *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **556** (1), 80-86. https://doi.org/10.1016/j.nima.2005.08.109

[140] K. Abe *et al.*, (The T2K Collaboration), "Measurement of the muon neutrino inclusive charged-current cross section in the energy range of 1-3 GeV with the T2K INGRID detector", *Physical Review D* **93**, 072002 (2016), https://doi.org/10.1103/PhysRevD.93.072002

[141] T. Vladisavljevic, "The T2K Flux Predictions", University of Oxford/Kavli IPMU, *NuInt2018 Conference Presentation*, Gran Sasso Science Institute

[142] L. Berns *et al.*, "Flux Prediction and Uncertainty with NA61/SHINE 2010 Replica Target Data", *T2K-TN-401*, 5th January 2021

[143] A. Ferrari, P.R. Sala, A. Fasso, J. Ranft, "FLUKA: A Multi-Particle Transport Code (Program version 2005)", October 2005, https://doi.org/10.2172/877507

[144] T.T. Böhlen, F. Cerutti, M.P.W. Chin, A. Fassò, A. Ferrari, P.G. Ortega, A. Mairani, P.R. Sala, G. Smirnov, V. Vlachoudis, "The FLUKA Code: Developments and Challenges for High Energy and Medical Applications", *Nuclear Data Sheets* **120** (2014) 211-214, https://doi.org/10.1016/j.nds.2014.07.049

[145] R. Brun, F. Bruyant, F. Carminati, S. Giani, M. Maire, A. McPherson, G. Patrick, L. Urban, "GEANT Detector Description and Simulation Tool", *CERN Program Library Long Writeup* **W5013**, https://doi.org/10.17181/CERN.MUHF.DMJ1

[146] C. Zeitnitz, T.A. Gabriel, "The GEANT-CALOR interface and benchmark calculations of ZEUS test calorimeters", *Nuclear Instruments and Methods in Physics Research A*, **349** (1994) 106-111, https://doi.org/10.1016/0168-9002(94)90613-0

[147] N. Abgrall *et al.*, "NA61/SHINE facility at the CERN SPS: beams and detector system", *JINST* **9** (2014) P06005, https://doi.org/10.1088/1748-0221/9/06/P06005

[148] K. Abe *et al.*, (The T2K Collaboration), "Measurements of the T2K neutrino beam propertied using the INGRID on-axis near detector", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **Vol 694**, 211-223 (2012), https://doi.org/10.1016/j.nima.2012.03.023

[149] M. Yokoyama *et al*, "Application of Hamamatsu MPPC to T2K Neutrino Detectors", *Nuclear Instruments and Methods in Physics Research A* **610** (2009) 128-130, https://doi.org/10.1016/j.nima.2009.05.077

[150] M. Otani *et al.*, "Design and construction of INGRID neutrino beam monitor for T2K neutrino experiment", *Nuclear Instruments and Methods in Physics Research A* **623** (2010) 368–370, https://doi.org/10.1016/j.nima.2010.02.251

[151] T. Kikawa for the T2K Collaboration, "Development of the new T2K on-axis neutrino detector, INGRID proton module", *Nuclear Physics B (Proc. Suppl.)* **229-232** (2012) 451, https://doi.org/10.1016/j.nuclphysbps.2012.09.088

[152] A. Ferrero for the T2K Collaboration, "The ND280 Near Detector of the T2K Experiment", *AIP Conference Proceedings* **1189**, 77 (2009), https://doi.org/10.1063/1.3274192

[153] E. Frank *et al.*, "A dedicated device for measuring the magnetic field of the ND280 magnet in the T2K experiment", *Journal of Instrumentation* **7**, P01018, January 2012, https://doi.org/10.1088/1748-0221/7/01/P01018

[154] S. Assylbekov, G. Barr, B.E. Berger, H. Berns, D. Beznosko *et al.*, "The T2K ND280 off-axis pi-zero detector", *Nuclear Instruments and Methods in Physics Research A* **686** (2012) 48–63, https://doi.org/10.1016/j.nima.2012.05.028

[155] L. Anthony, "A model independent determination of the π^0 background in Super-Kamiokande in the T2K ν_e appearance measurement", Doctor of Philosophy thesis, University of Liverpool, 2020, http://doi.org/10.17638/03076633

[156] N. Abgrall, B. Andrieu, P. Baron, P. Bene, V. Berardi *et al.*, "Time projection chambers for the T2K near detectors", *Nuclear Instruments and Methods in Physics Research A* **637** (2011) 25–46, https://doi.org/10.1016/j.nima.2011.02.036

[157] I. Giomataris, R. De Oliveira, S. Andriamonje, S. Aune G. Charpak *et al.*, "Micromegas in a bulk", *Nuclear Instruments and Methods in Physics Research A* **560** (2006) 405–408, https://doi.org/10.1016/j.nima.2005.12.222

[158] P.-A. Amaudruz, M. Barbi, D. Bishop, N. Braam, D.G. Brook-Roberge, *et al.*, "The T2K Fine-Grained Detectors", *Nuclear Instruments and Methods in Physics Research A* **696** (2012) 1–31, https://doi.org/10.1016/j.nima.2012.08.020

[159] M. Antonova for the T2K Collaboration, "Ageing of the scintillator detectors of the T2K off-axis and on-axis detectors, ND280 and INGRID", *XIX International Workshop on Neutrino Telescopes*, February 2021, https://doi.org/10.5281/zenodo.4899051

[160] D. Allan *et al.*, "The electromagnetic calorimeter for the T2K near detector ND280", *JINST* **8** P10019 (2013), http://dx.doi.org/10.1088/1748-0221/8/10/P10019

[161] A. Vacheret, M. Noy, M. Raymond, A. Weber, "First results of the Trip-t based T2K front end electronics performance with GM-APD", *PoS* **PD07** (2006) 027, https://doi.org/10.22323/1.051.0027

[162] S. Aoki, G. Barr, M. Batkiewicz, J. Błocki, J.D. Brinson *et al.*, "The T2K Side Muon Range Detector (SMRD)", *Nuclear Instruments and Methods in Physics Research A* **698** (2013) 135–146, https://doi.org/10.1016/j.nima.2012.10.001

[163] M.D. Lawe, "Study of Charged Current Neutral Pion Production in the T2K Near Detector", PhD Thesis, University of Sheffield, September 2014, https://etheses.whiterose.ac.uk/8562/

[164] F. Dufour for the T2K Collaboration, "T2K experiment: Status and first results", *J. Phys.: Conf. Ser.* **335** 012053, https://doi.org/10.1088/1742-6596/335/1/012053

[165] T. Honjo,"BSD for MR Run 86", *T2K internal presentation*, 3rd May 2021

[166] T. Lindner, "ND280 Computing Tutorial", *T2K internal presentation*, 11th February 2017

[167] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo *et al.*, "Geant4 - A simulation toolkit", *Nuclear Instruments and Methods in Physics Research A* **506** (2003) 250–303, https://doi.org/10.1016/S0168-9002(03)01368-8

[168] T. Ovsiannikova *et al.*, "The new experiment WAGASCI for water to hydrocarbon neutrino cross section measurement using the J-PARC beam", *J. Phys.: Conf. Ser.* **675** 012030 (2016), https://doi.org/10.1088/1742-6596/675/1/012030

[169] N. Chikuma *et al.*, "A new experiment at J-PARC to measure the neutrino cross section ratio between water and plastic", *PoS* **248**, (FPCP2015) 069, https://doi.org/10.22323/1.248.0069

[170] M. Antonova *et al.*, "Baby MIND: A magnetised spectrometer for the WAGASCI experiment", NuPhys2016-Hallsjö, 26 Apr 2017, arXiv:1704.08079v1 [physics.ins-det]

[171] G. Pintaudi, *T2K-WAGASCI: First physics run of the WAGASCI-BabyMIND detector with full setup*, *PoS* (LeptonPhoton2019) 142 (2019), https://doi.org/10.22323/1.367.0142

[172] K. Abe *et al.*, "Measurements of $\bar{\nu}_\mu$ and $\bar{\nu}_\mu + \nu_\mu$ charged-current cross-sections without detected pions nor protons on water and hydrocarbon at mean antineutrino energy of 0.86 GeV", *Progress of Theoretical and Experimental Physics* **Volume 2021**, Issue 4, April 2021, 043C01, https://doi.org/10.1093/ptep/ptab014

[173] S. Fukuda *et al.*, "The Super-Kamiokande detector", *Nuclear Instruments and Methods in Physics Research A* **501** (2003) 418–462, https://doi.org/10.1016/S0168-9002(03)00425-X

[174] K. Abe *et al.*, "Calibration of the Super-Kamiokande detector", *Nuclear Instruments and Methods in Physics Research A* **737** (2014) 253–272, https://doi.org/10.1016/j.nima.2013.11.081

[175] Hamamatsu Photonics, "Large photocathode area photomultiplier tubes", accessible from http://hep.ucsb.edu/people/hnn/n/pmt_info/R3600-02_20inch_superKmain.pdf

[176] A. Bernstein *et al.*, " Report on the Depth Requirements for a Massive Detector at Homestake", https://doi.org/10.2172/946806

[177] C. Bronner, "Details of T2K Oscillation Analysis", *PoS* **369** (NuFact2019) 037, *https://doi.org/10.22323/1.369.0037*

[178] K. Abe *et al.* (T2K Collaboration), "Measurement of Muon Antineutrino Oscillations with an Accelerator-Produced Off-Axis Beam", *Phys. Rev. Lett.* **116**, 181801 (2016), https://doi.org/10.1103/PhysRevLett.116.181801

[179] A. Izmaylov, S. Jenkins, L. Molina Bueno, C. Riccio, C, Schloesser, "Time of flight correction and systematic", *T2K Internal Presentation*, 25th March 2019

[180] Xilinx, "RocketIO™ Transceiver User Guide", UG024 (v3.0) February 22, 2007, accessible from https://www.xilinx.com/support/documentation/user_guides/ug024.pdf

[181] CERN ROOT, "TFractionFitter Class Reference", https://root.cern.ch/doc/v606/classTFractionFitter.html

[182] CERN ROOT, "TRandom3 Class Reference", https://root.cern.ch/doc/master/classTRandom3.html

[183] S. Schmitt, "Data Unfolding Methods in High Energy Physics", *EPJ Web of Conferences* **137**, 11008 (2017), https://doi.org/10.1051/epjconf/201713711008

[184] S.S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", *Ann. Math. Statist.* **9** (1): 60-62 (March, 1938), https://doi.org/10.1214/aoms/1177732360

[185] S. Baker, R.D. Cousins, "Clarification of the Use of Chi-Square and Likelihood Functions in Fits to Histograms", *Nuclear Instruments and Methods in Physics Research*, **221** (1984), https://doi.org/10.1016/0167-5087(84)90016-4

[186] J.Dobson and C. Andreopoulos, "Propagating ν-Interaction Uncertainties Via Event Reweighting", *Acta Phys. Pol. B* **40**, 2613 (2009), https://www.actaphys.uj.edu.pl/R/40/9/2613/pdf

[187] C. Andreopoulos, J. Dobson, S. Dytman, "Handling Neutrino Interaction Uncertainties using Event Reweighing", *T2K-TN-007 v1.0*, 1st April 2009

[188] A. Cudd, "Super-xsllhFitter", accessible from `https://cuddandr.gitlab.io/xsLLhFitter/`

[189] F. James, M. Winkler, "Minuit 2 Documentation", accessible from `https://root.cern.ch/root/htmldoc/guides/minuit2/Minuit2.html`

[190] A.N. Tikhonovm *Dokl. Akad. Nauk. SSSR* **151** (1963), pp. 501–504; English translation: "Solution of incorrectly formulated problems and the regularization method", *Soviet Math. Dokl.* 4 (1963), pp. 1035–1038

[191] D. Dereniowski, M. Kubale, "Cholesky Factorization of Matrices in Parallel and Ranking of Graphs", *PPAM 2003: Parallel Processing and Applied Mathematics* pp 985-992 (2004), `https://doi.org/10.1007/978-3-540-24669-5_127`

[192] CERN ROOT, "TDecompChol Class Reference", `https://root.cern.ch/doc/master/classTDecompChol.html`

[193] S. King *et al.*, "FHC muon neutrino charged current multiple pion samples in the ND280 tracker for the 2020 oscillation analysis inputs: Run 2+3+4+8 with P6T NEUT_D MC", *T2K-TN-407 v1.2*, 18th June 2020

[194] R. Castillo, "Measurement of the ν_μCC1π^+ cross-section on carbon using the near detector", *T2K TN-199 Version 3.1*, 3rd January 2016

[195] C. Giganti, M. Zito, "Particle Identification with the T2K TPC", *T2K TN-001 Version 2*, 9th October 2009

[196] J. Kim, C. Nielsen, M. Wilking, "Michel Electron Tagging in the FGDs", *T2K TN-104*, 22nd January 2015

[197] A.P. Kolaceke, "FGD-only PID",*ND280 Reconstruction Group Analysis Meeting*, 14th April 2014

[198] F. Di Lodovico *et al.*, "CC-multiple-pion ν_μ event selections in the ND280 tracker using Run 1+2+3+4 data", *T2K TN-152-V4.2*, 7th November 2013

[199] C. Licciardi, M. Barbi, "Particle Identification with the Fine Grained Detectors", *T2K TN-103*, 16th July 2011

[200] D. Hadley *et al.*, "Implementation of the Second Generation PID for the ND280 Tracker ECals", *T2K TN-111*, 8th August 2012

[201] S.J. Jenkins, "Reinvestigation of the ECal π^0 Veto", *T2K TN-392*, 20th November 2019

[202] G. Christodoulou, S. King, "Measurement of electron (anti-)neutrino cross-sections in the ND280 tracker using (anti-)neutrino beam data up to Run 8", *T2K-TN-277-v3.2*, 7th January 2020

[203] AKIMOTO Yuki, "Higgstan Elementary Particle Illustrations", accessible from `https://higgstan.com`

[204] L. Cremonesi, "Measurement of the muon neutrino charged current pion production cross-section on water using the T2K near detector", PhD thesis, April 2015

[205] A. Blondel *et al.*, "The SuperFGD Prototype Charged Particle Beam Tests", *JINST* **15** (2020) 12, P12003, 20 August 2020, `https://doi.org/10.1088/1748-0221/15/12/P12003`

[206] A.L. Stancik, E.B. Brauns, "A simple asymmetric lineshape for fitting infrared absorption spectra", *Vibrational Spectroscopy* **47** (2008) 66–69, `https://doi.org/10.1016/j.vibspec.2008.02.009`

[207] C. Bojechko *et al.*, "Measurement and Correction of Magnetic Field Distortions in the Time Projection Chambers", *T2K-TN-061 v2*, 21st June 2013

[208] G. Christodoulou *et al.*, "Study of the tracker ECal systematic uncertainties", *T2K-TN-279 v2*, 27th June 2017

[209] P. Martins, "ECal π^0 Veto Studies", *T2K-TN-270 v0.1*, 4th April 2016

[210] W. Oryszczak, W. Warzycha, "FGD Systematics: PID and IsoRecon Hybrid Efficiency", *T2K-TN-223*, 20th April 2015

[211] L. Maret *et al.*, "Backward Migration Uncertainty in FGD2", *T2K-TN-368*, 26th November 2018

[212] F. Di Lodovico *et al.*, "CCQE-like and CC-non-QE-like ν_μ event selections in the ND280 tracker using Run 1+2+3 data", *T2K-TN-128-V4.0*, 7th March 2013

[213] J. Lagoda *et al.*, "ν_μ CC event selections in the ND280 tracker using Run 2+3+4 data", *T2K-TN-212-V2.1*, 13th October 2015

[214] A. Cervera *et al.*, "CC-0Pi Multi-Topology Selection and Systematics in FGD1", *T2K-TN-216*, 6th January 2015

[215] L. Haegel *et al.*, "Systematics on Out-of-Fiducial Volume Backgrounds in the ND280 Tracker Muon Neutrino Analysis", *T2K-TN-098 v2.1*, 14th June 2015

[216] J. Myslik,"Determination of Pion Secondary Interaction Systematics for the ND280 numu Analysis", *T2K-TN-125 v2*, 22nd October 2013

[217] K. Kowalik, J. Lagoda, "The simulation of beam neutrinos interactions outside the ND280 near detector", *T2K-TN-077 v2.1*, 12th February 2012

[218] A. Hillairet, T. Lindner, J. Myslik, P. Stamoulis, "ND280 tracker tracking efficiency in the 2011 analysis", *T2K-TN-075 v1.1*, 15th May 2012

[219] F. Sanchez, J. Vo Medina, "ND280 global charge identification systematic error", *T2K-TN-229 v3*, 4th May 2016

[220] A. Fiorentini, "ND280 TPC Cluster Efficiency", *T2K TN-234*, 3rd March 2015

[221] A. Cervera, L. Escudero, "Study of momentum resolution and scale using tracks that cross multiple TPCs", *T2K-TN-222 v1*, 4th November 2014

[222] E. Frank, A. Marchionni, M. Messina, "B-field calibration and systematic errors", *T2K-TN-081 v1*, 21st November 2011

[223] S. Bordoni, C. Giganti, A. Hillairet, F. Sanchez, "The TPC Particle Identification algorithm with production 6B", *T2K-TN-221*, 2nd April 2015

[224] A. Hillairet, Y. Petrov, "ND280 TPC Track Reconstruction Efficiency", *T2K-TN-163 v2.1*, 22nd February 2016

[225] P. Bartet *et al.*, "ν_μ 4pi CC event selections in the ND280 tracker using Run 2+3+4", *T2K-TN-245*, 20th June 2017

[226] E.T. Atkin *et al.*, "NIWG model and uncertainties for 2019-2020 oscillation analysis", *T2K-TN-344-v4*, October 2019

[227] L. Berns (T2K Beam Group), "T2K Flux Release 13av7.1", 10th March 2020

[228] S. Oser, "Elemental Composition of the FGD XY Modules", *T2K-TN-091 v1.2*, 25th February 2010

[229] D. Roberge, "Elemental Composition of FGD Passive Water Modules", *T2K-TN-198 v1.0*, 22nd January 2010