Measurement of the Electron Neutrino Charged-Current Inclusive Cross-section on Carbon Using the T2K Near Detector

Benjamin Mark Smith

Imperial College London

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Physics

Abstract

T2K is a long baseline neutrino oscillation experiment which uses a beam of muon neutrinos to study muon neutrino disappearance and electron neutrino appearance. The measurement of ν_e appearance is sensitive to CP violation, and reducing the systematic uncertainties is critical to enabling experimental determination of whether there is CP violation in the lepton sector.

This thesis describes the first measurement of the ν_e CC inclusive cross-section on carbon at energies relevant to T2K and other long baseline neutrino oscillation experiments. The T2K near detector, ND280, is used to select a sample of ν_e events, and a Bayesian unfolding technique is used to extract differential cross-sections as a function of electron momentum, electron angle and Q^2. The total flux-averaged cross-section is also extracted, and is found to be $1.11 \pm 0.20 \times 10^{-38}$ cm^2/nucleon, which agrees well with both the NEUT neutrino interaction generator prediction of 1.23×10^{-38} cm^2/nucleon and the GENIE prediction of 1.08×10^{-38} cm^2/nucleon.

A restricted phase-space analysis is also performed, in which only events where the ejected electron has momentum greater than 550 MeV/c and $\cos(\theta_e) > 0.72$ are selected. In this case, the flux-averaged cross-section is found to be $6.54 \pm 1.22 \times 10^{-39}$ cm^2/nucleon, again in good agreement with both the NEUT prediction of 7.38×10^{-39} cm^2/nucleon and the GENIE prediction of 6.41×10^{-39} cm^2/nucleon.

Declaration

As in all high energy physics experiments, the work in this thesis would not be possible without the work of many collaborators. Where appropriate, citations are provided to public references; where no public reference is available the text attempts to describe who performed the work.

My most significant personal contributions to the results in this thesis include developing the selection of ν_e interactions, evaluating ECal-related systematics, being one of the two lead developers of the "high-level analysis at the near detector" (highland) package, contributing to the "BANFF interface" for propagating the effect of ND280 systematics to the analysis, and writing the code for converting the selection to a cross-section result. The highland package is now used by most T2K collaborators who analyse ND280 data, and the BANFF interface was used to propagate detector systematic uncertainties for the 2013 T2K oscillation analyses. I also spent a significant amount of time improving the ECal reconstruction software, although the improved algorithms are not used in this analysis, and are therefore not described in the text.

As is explained in the text, this is the first measurement of the ν_e CC inclusive cross-section since 1978, and is the first ν_e differential cross-section measurement. The aim is to publish the results in a prestigious peer-reviewed journal soon after the submission of this thesis.

Acknowledgements

Firstly, I would like to acknowledge the enormous help I received from Dr. Claudio Giganti at LPNHE and Dr. Asher Kaboth at Imperial College London. They both provided excellent guidance and probing, thoughtful comments during the development of this analysis, and are both thoroughly kind and patient.

I would also like to thank the rest of the ND280 ν_e group for their help, in particular Javi Caravaca at IFAE and Dr. Neil McCauley at Liverpool University. Many thanks also go to the ND280 ν_μ group, in particular for their work evaluating many of the ND280 detector systematics. Special thanks go to Dr. Anselmo Cervera and Dr. Federico Sanchez for their work developing highland and the BANFF interface.

Thank you, of course, to my supervisor Dr. Yoshi Uchida. You were always there with sage advice whenever it was needed.

During my PhD I lived in Tokai, Japan for 11 months, and visited numerous times whilst not living there. Thank you to Dr. Sarah Foley, Dr. Jevon Keane-Brennan, Dr. Megan Friend, Matthew Lawe, Andrew Furmanski and Linda Cremonesi for making my time in Japan so enjoyable. ありがとうございます！ Thank you also to the residents of Tokai for being so welcoming, in particular the staff at Masago International Lodging, Ayako Kiuchi and the other secretaries at the J-PARC users office, and Rika Nagai at PonoPono. 乾杯！

Thank you also to my fellow T2K students at Imperial—Peter Sinclair, Dominic Brailsford and Pip Hamilton—for making the office such a friendly and enjoyable place to work.

Lastly, I would like to thank my family for being so supportive of my decision to pursue a PhD. Mum, Dad, Chris and Helen: thank you for your generosity and support. Up the clarets!

Contents

List of figures

List of tables

Chapter 1

Neutrino theory

1.1 Overview of neutrino physics

1.1.1 Neutrino oscillations

Neutrinos were first postulated by Pauli in 1930 as an explanation for why the electron emitted in beta decay has a continuous energy spectrum [1]. Experimental observation followed in 1956, when Reines and Cowan observed $\overline{\nu}_e + p \rightarrow e^+ + n$ at the Savannah River reactor [2]. A problem emerged in 1968 when Ray Davis detected only one-third of the flux of solar ν_e as predicted by the Standard Solar Model [3]. This deficit was seen in further experiments, including SAGE [4] and GALLEX [5], and became known as the Solar Neutrino Problem. It was found that Maki, Nagakawa and Sakata's extension [6] to Pontecorvo's [7] theory of neutrino oscillations could explain the discrepancy when the Mikheyev-Smirnov-Wolfenstein (MSW) effect—which explains the effect matter has on neutrino oscillations—was also included. Compelling support for neutrino oscillations was provided by the Super-Kamiokande (SK) experiment in 1998 [8], when it saw a deficit of muon neutrinos produced by pion decay in the upper atmosphere. The deficit was found to be strongly dependent on the distance the neutrinos travel between creation and detection, and the shape of this dependence was best explained by neutrino oscillations.

The fundamental concept of neutrino oscillations is that the flavour eigenstates and mass eigenstates of neutrinos are not the same, but are related through the unitary

transformation

$$|\nu_\alpha\rangle = \sum_{j=1}^{3} U^*_{\alpha j} |\nu_j\rangle, \tag{1.1}$$

where $\alpha \in \{e, \mu, \tau\}$, ν_α are the flavour eigenstates, ν_j are the mass eigenstates and U is the PMNS mixing matrix [6]. The notation and treatment of Giunti and Kim [9] has been adopted, and for the rest of this section the summation over index j is implied. The mixing matrix is a 3×3 complex unitary matrix which can be parameterised as

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{-i\alpha_1} & 0 & 0 \\ 0 & e^{-i\alpha_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.2}$$

where $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$, θ_{ij} are known as mixing angles, δ is a CP-violating phase, and α_i are Majorana phases that have no effect on neutrino oscillations [10].

Supposing that at time $t = 0$ a neutrino is produced in the state $|\nu_\alpha\rangle$, the time evolution of the quantum state will be

$$|\nu(t)\rangle = U^*_{\alpha j} e^{-iE_j t} |\nu_j\rangle, \tag{1.3}$$

with a probability of being found in the different state $|\nu_\beta\rangle$ (where $\beta \in \{e, \mu, \tau\}$) of

$$P(\nu_\alpha \to \nu_\beta) = |\langle \nu_\beta | \nu(t)\rangle|^2 = \left|U^*_{\alpha j} e^{-iE_j t} \langle \nu_\beta | \nu_j\rangle\right|^2 = \left|U_{\beta j} e^{-iE_j t} U^*_{\alpha j}\right|^2, \tag{1.4}$$

and as neutrinos are highly relativistic the approximation

$$E_j \approx p + \frac{m_j^2}{2E} \tag{1.5}$$

is made. The T2K neutrino oscillation experiment is the subject of this thesis, and this uses a beam of muon neutrinos. The probabilities of ν_μ survival and $\nu_\mu \to \nu_e$ oscillation are

$$P(\nu_\mu \to \nu_\mu) \simeq 1 - \cos^4\theta_{13} \sin^2 2\theta_{23} \sin^2\left(1.27 \frac{\Delta m^2_{32}}{(\mathrm{eV}^2)} \frac{L}{(\mathrm{km})} \frac{(\mathrm{GeV})}{E}\right) \tag{1.6}$$

$$P(\nu_\mu \to \nu_e) \simeq \sin^2 2\theta_{13} \sin^2\theta_{23} \sin^2\left(1.27 \frac{\Delta m^2_{32}}{(\mathrm{eV}^2)} \frac{L}{(\mathrm{km})} \frac{(\mathrm{GeV})}{E}\right), \tag{1.7}$$

where L is the distance the neutrino has travelled, E is the energy of the neutrino, and $\Delta m^2_{ij} \equiv m^2_i - m^2_j$. Equations (1.6) and (1.7) show that the magnitude of $\nu_\mu \to \nu_e$ oscillations is governed by the mixing angles θ_{13} and θ_{23}, and the frequency of the oscillations depends on Δm^2_{32}.

A more thorough derivation which accounts for the effect of matter—which contains electrons but not muons or taus—yields equation (1.8), in which the CP-violating phase δ appears in second-order terms [11].

$$P(\nu_\mu \to \nu_e) \simeq \sin^2 2\theta_{13} T_1 - \alpha \sin^2 2\theta_{13}(T_2 - T_3) + \alpha^2 T_4, \tag{1.8}$$

where

$$T_1 = \sin^2\theta_{23} \frac{\sin^2[(1-x_\nu)\Delta]}{(1-x_\nu)^2} \tag{1.9}$$

$$T_2 = \sin\delta \sin^2 2\theta_{12} \sin^2 2\theta_{23} \sin\Delta \frac{\sin(x_\nu\Delta)}{x_\nu} \frac{\sin[(1-x_\nu)\Delta]}{1-x_\nu} \tag{1.10}$$

$$T_3 = \cos\delta \sin^2 2\theta_{12} \sin^2 2\theta_{23} \cos\Delta \frac{\sin(x_\nu\Delta)}{x_\nu} \frac{\sin[(1-x_\nu)\Delta]}{1-x_\nu} \tag{1.11}$$

$$T_4 = \cos^2\theta_{23} \sin^2 2\theta_{12} \frac{\sin^2(x_\nu\Delta)}{x^2_\nu}, \tag{1.12}$$

and the extra definitions $\Delta \equiv \Delta m^2_{31} L/4E$ and $\alpha \equiv \Delta m^2_{21}/\Delta m^2_{31} \simeq 1/30$ have been made. x_ν is a term describing the electron density of the matter the neutrinos are passing through, and is defined as

$$x_\nu = \frac{VL}{2\Delta}, \tag{1.13}$$

where V is proportional to the matter density and the number of electrons per nucleon, and describes the effective charged-current potential felt by ν_e.

1.1.2 Current status

Evidence from a large number of neutrino oscillation experiments shows excellent agreement with the 3-flavour model of neutrino oscillations described in the previous section. There are four main categories of oscillation experiments, and each is sensitive to a different set of oscillation parameters.

Solar experiments detect neutrinos produced in nuclear reactions in the centre of the Sun and are mainly sensitive to θ_{12} and Δm^2_{21}. These neutrinos are created as very low energy MeV-scale ν_e in the core of the sun, but the MSW effect causes the neutrinos to be almost pure ν_2 at the surface of the Sun. Ray Davis' Homestake experiment [3], SAGE [4], GALLEX [5], Super-Kamiokande [8] and SNO [12] are all examples of solar experiments.

Atmospheric experiments detect the neutrinos produced when cosmic rays strike the upper atmosphere and create a cascade of pions, kaons, muons and electrons. These neutrinos are a mixture of ν_μ, $\overline{\nu}_\mu$, ν_e and $\overline{\nu}_e$ and span a range of neutrino energies from MeV to TeV. Atmospheric experiments are mainly sensitive to θ_{23} and Δm^2_{31}, and examples include Super-Kamiokande, Soudan 2 [13] and MINOS [14].

Reactor experiments look for the disappearance of MeV-scale $\overline{\nu}_e$ produced in nuclear reactors, over a baseline of several to over a hundred km. These experiments are able to make extremely precise measurements of θ_{12} (KamLAND [15]) and θ_{13} (Daya Bay [16], RENO [17] and Double CHOOZ [18]).

Long baseline accelerator experiments such as MINOS, T2K and NOνA [19] use high energy GeV-scale neutrinos and baselines of hundreds of kilometres. As was shown in equations (1.6) and (1.7), these experiments are primarily sensitive to θ_{13}, θ_{23} and Δm^2_{32}. However, precision measurements give the possibility of constraining the CP-violating phase δ, as shown in equation (1.8).

A global analysis by the Particle Data Group of all neutrino oscillation data leads to the constraints shown in Table 1.1.

1.1.3 Open questions

The major open questions in neutrino physics are outlined below.

The value of δ. Now that it has been experimentally determined that all the mixing angles in the PMNS matrix are non-zero, a non-zero value of δ would produce CP violation in the leptonic sector. CP violation has so far only been observed

Parameter	Current Limit
$\sin^2(2\theta_{12})$	0.857 ± 0.024
$\sin^2(2\theta_{23})$	> 0.95
$\sin^2(2\theta_{13})$	0.095 ± 0.010
Δm^2_{21}	$(7.50 \pm 0.20) \times 10^{-5}$ eV^2
$\lvert\Delta m^2_{32}\rvert$	$\left(2.32^{+0.12}_{-0.09}\right) \times 10^{-3}$ eV^2
δ_{CP}	$0 - 2\pi$ (Unknown)

Table 1.1: Experimentally determined values of neutrino oscillation parameters, as reported by the Particle Data Group in 2012 [20].

in the quark sector, and that is not sufficient to explain the matter/anti-matter asymmetry in the present Universe [21, 22]. Discovering leptonic CP violation would have wide-ranging consequences in particle physics, astrophysics and cosmology. As mentioned previously, long baseline oscillation experiments such as T2K, NOνA and the proposed LBNE have the capacity to constrain the value of δ, especially when combined with precision measurements of θ_{13} from nuclear reactor experiments [23].

The mass hierarchy. Although the absolute value of the mass-squared splittings have been determined with 5% precision (see Table 1.1), it is not known which mass state is heaviest. Matter effects in solar neutrino oscillation experiments have shown that m_2 is larger than m_1, but is it is not known whether $m_3 \gg m_2 > m_1$ or $m_2 > m_1 \gg m_3$. Due to the presence of matter effects, long baseline experiments again have the possibility to determine the correct hierarchy.

The absolute mass of neutrinos. Oscillation experiments are able to probe the mass-squared splittings of the mass states, but say nothing about the absolute mass scale of neutrinos. The most direct measure of neutrino mass comes from observing the endpoint of the Tritium decay $^3\mathrm{H} \rightarrow {}^3\mathrm{He} + e^- + \overline{\nu}_e$. The KATRIN experiment will soon start taking data and has sensitivity to $m_\nu > 0.2\,\mathrm{eV}$ [24]. Constraints from cosmology—including the cosmic microwave background and galaxy surveys—indicate that the summed mass of the three neutrinos flavours should be less than $0.3\,\mathrm{eV}$ [25].

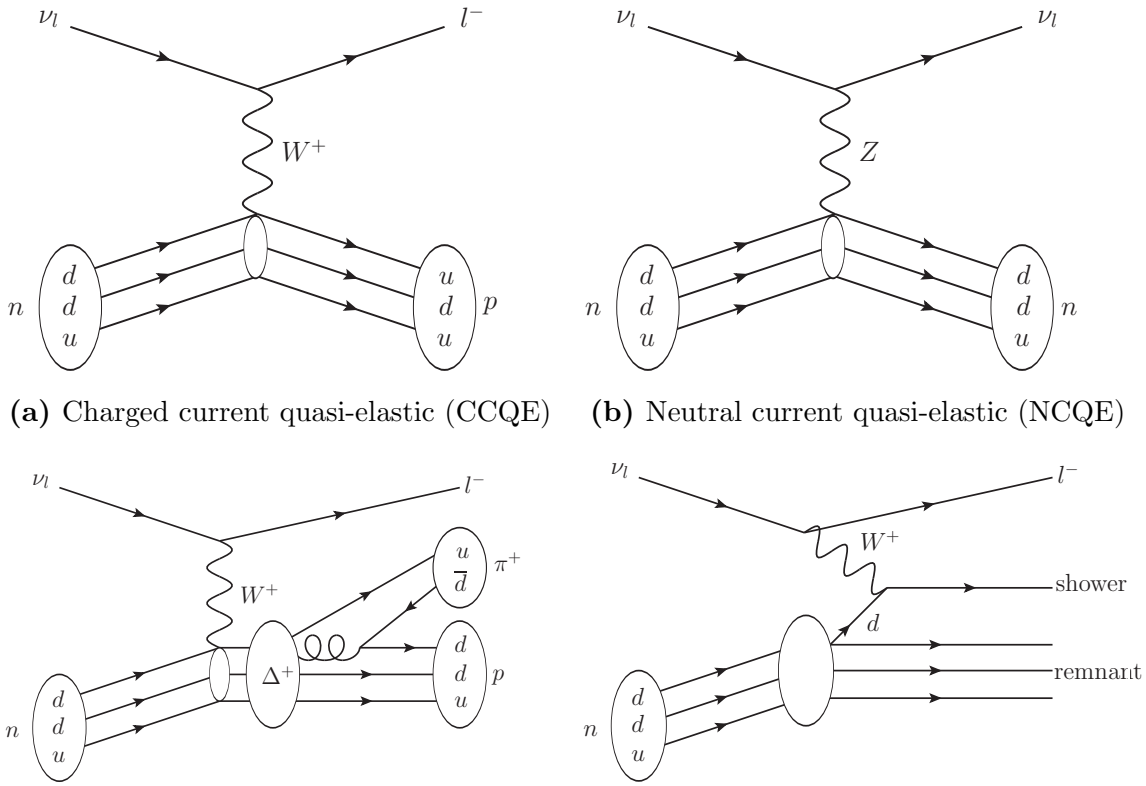
The nature of neutrino mass. Neutrinos are massless in the Standard Model, and the two main formalisms for adding mass is as Dirac masses or Majorana masses. A Dirac neutrino mass is generated using the standard Higgs mechanism, and requires the addition of right-handed components of the neutrino field. Such right-handed

neutrinos are referred to as sterile neutrinos, as they do not interact through weak, strong or electromagnetic interactions. A Majorana particle is described by a real wave equation (rather than a complex one), and so is its own anti-particle. If neutrinos are Majorana particles they can mediate neutrinoless double β decay, and this is the most promising way to determine whether neutrinos are Dirac particles or Majorana particles. SuperNEMO [26] and SNO+ [27] are examples of the many experiments searching for neutrinoless double β decay.

Short baseline neutrino anomalies. There are several anomalies from short baseline neutrino oscillation experiments that do not yet have a clear explanation. The LSND experiment observed $\overline{\nu}_e$ appearance in a $\overline{\nu}_\mu$ beam, consistent with neutrino oscillations with a mass-squared splitting of 0.2–10 eV^2 [28]. This large mass-squared splitting is inconsistent with the three-neutrino mixing model described in Section 1.1.1, and motivates the possible existence of a fourth generation of neutrinos. Experiments at the LEP collider showed that there are only 2.9840 ± 0.0082 neutrino species that couple to the Z boson [29], and so this fourth generation must not interact through the weak interaction, and is known as a sterile neutrino. The MiniBooNE experiment also observed an excess of $\overline{\nu}_e$ in a $\overline{\nu}_\mu$ beam, consistent with neutrino oscillations with $0.01 < \Delta m^2 < 1.0$ eV^2 [30], and further evidence for sterile neutrinos is provided by the reactor antineutrino anomaly [31] and gallium anomaly [32]. However, there is significant tension [33] with null results from ν_μ and $\overline{\nu}_\mu$ disappearance searches such as CDHS [34], Super-Kamiokande [35] and MINOS [36], and further tension from cosmological bounds also complicates the question of whether sterile neutrinos exist [37].

1.2 Interactions of GeV-scale neutrinos

Neutrinos are neutral particles that interact through the weak force. It is not possible to observe them directly, so experiments search for the particles produced when a neutrino interacts with a nucleus. In charged-current (CC) interactions, mediated by the W boson, a charged lepton is produced. If this lepton can be identified, the flavour of the incoming neutrino can be determined. In neutral-current (NC) interactions, mediated by the Z boson, the neutrino appears in the final state, so its flavour cannot be determined. A wide range of neutrino interactions are possible, and diagrams of the dominant modes for GeV-scale neutrinos are shown in Figure 1.1.



(a) Charged current quasi-elastic (CCQE) **(b)** Neutral current quasi-elastic (NCQE)

(c) CC resonant pion production (CCRES) **(d)** CC deep inelastic scattering (CCDIS)

Figure 1.1: Diagrams of several neutrino-nucleon interaction modes. The smaller ellipses indicate that the boson is effectively interacting with the nucleon as a whole.

Neutrino experiments detect the number of neutrino interactions in the detector, which is the product of the flux, cross-section, number of target nucleons and the detector efficiency. In long-baseline oscillation experiments such as T2K, the flux is measured and constrained using detectors near the neutrino production point. The size and performance of the far detector are also well studied, so knowledge of the neutrino cross-sections is crucial to determining the expected number of events for a given set of oscillation parameters. If neutrino cross-sections are not well understood, large systematic uncertainties have to be applied, and the sensitivity of the experiment is reduced.

Quasi-elastic scattering—CCQE shown in Figure 1.1a and NCQE shown in Figure 1.1b—are dominant for neutrino energies below 1 GeV. In this energy region, the neutrino is effectively interacting with the nucleon as a whole. There is insufficient energy to break the nucleon apart, so although a d quark is converted to a u quark (in the CCQE case), kinematically the interaction appears to be affecting the entire nucleon. At neutrino energies between 1 GeV and 5 GeV, resonant pion production (CCRES shown in Figure 1.1c) is the dominant mode. Again, there is insufficient energy to completely break the nucleon apart, so the initial interaction is effectively with the nucleon as a whole. The nucleon is elevated to a more energetic state, such as a Δ^+, and then decays to produce a pion and a neutron. Above 5 GeV, deep inelastic scattering (CCDIS shown in Figure 1.1d) dominates. In this mode, the neutrino can resolve an individual quark in the nucleon, and hadronic showers are produced in the final state.

Neutrino experiments do not always have the luxury of using free nucleons as their target, and these nucleons are instead bound in nuclei. Common target elements include carbon (in plastic scintillator), iron (in sampling calorimeters) and water (in Cherenkov detectors), and the neutrino-nucleon cross-section predictions are significantly complicated by the nuclear environment. The main effects that must be modelled are the momentum of nucleons within the nucleus, interactions between the products of the interaction and other nucleons (final state interactions, FSI), and interactions between correlated pairs of nucleons (meson exchange currents, MEC [38]).

Although the nuclear environment is different for different elements, experiments can be compared using the measured cross-section per nucleon. Figure 1.2 summarises recent ν_μ cross-section measurements for neutrino energies between 100 MeV and 300 GeV. The T2K neutrino beam peaks at approximately 700 MeV (in the transition region between CCQE and CCRES) and has a high-energy tail extending to over 10 GeV (including the transition between CCRES and CCDIS).

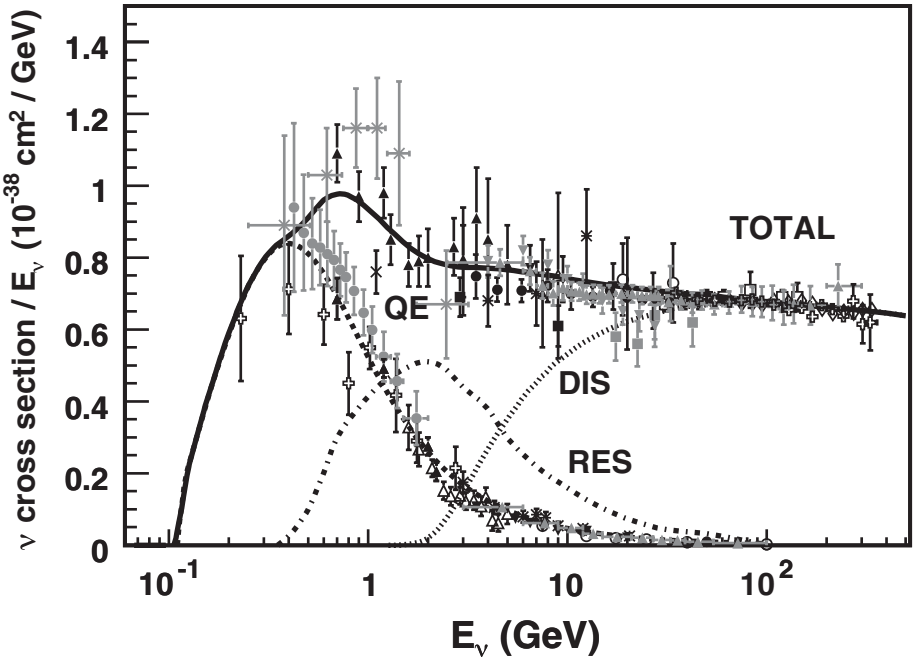


Figure 1.2: ν_μ-nucleon cross-sections at intermediate energies showing quasi-elastic (QE), resonant pion (RES) and deep inelastic scattering (DIS) contributions. The T2K beam is peaked at 700 MeV. Figure taken from Reference [39]: see the reference for details of the 27 experimental results shown.

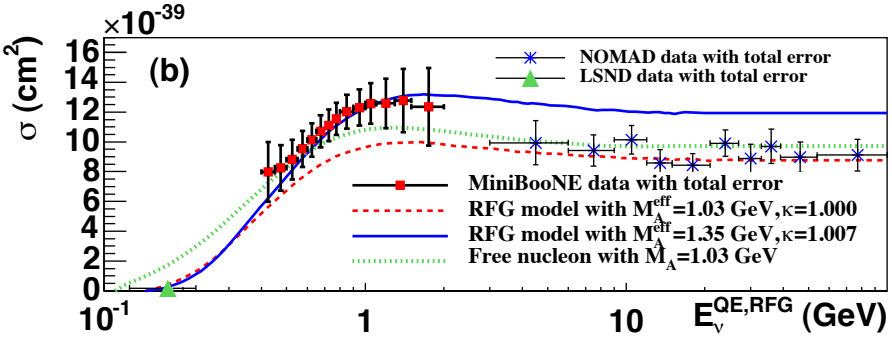


Figure 1.3: Comparison of MiniBooNE and NOMAD ν_μ CCQE cross-section data, and model predictions with M_A^{QE} values that best describe each dataset [40].

CCQE interactions are the most experimentally useful interactions to study. The two-body kinematics allow the energy of the incoming neutrino to be determined, and the interaction models are relatively simple. CCQE interaction models parameterise the nucleon structure with a priori unknown form factors. The leading terms are those associated with the vector form factor $F_V^1(Q^2)$—which is well-measured from electron scattering data—and the axial-vector form factor $F_A(Q^2)$. $F_A(Q^2)$ is modelled as

$$F_A(Q^2) = \frac{F_A(0)}{\left(1 + Q^2/\left(M_A^{\mathrm{QE}}\right)^2\right)^2}, \tag{1.14}$$

where Q^2 is the four momentum transferred from the leptonic system to the hadronic system, $F_A(0)$ is the form factor at $Q^2 = 0$, and M_A^{QE} is a parameter that will shortly be discussed in more detail. $F_A(0)$ has been determined from neutron beta decay, but much of the knowledge at higher Q^2 comes from ν_μ CCQE scattering measurements. M_A^{QE} affects the Q^2 distribution of CCQE interactions, and has been measured by the MiniBooNE and NOMAD Collaborations, among others (although the experiments really measure an *effective* M_A^{QE}, due to the nuclear environments in which the interactions take place). NOMAD used 3–100 GeV neutrinos and measured $M_A^{\mathrm{QE}} = 1.05 \pm 0.02(\mathrm{stat}) \pm 0.06(\mathrm{syst})$ GeV. MiniBooNE used neutrinos with a mean energy of 800 MeV, and measured $M_A^{\mathrm{QE}} = 1.35 \pm 0.17$ GeV. The MiniBooNE and NOMAD data are shown in Figure 1.3, where the tension between the two M_A^{QE} values is clearly shown. More data is required to improve the neutrino interaction and nuclear models, and this will be provided in the immediate future by T2K, NOνA, and the dedicated neutrino cross-section experiment MINERνA.

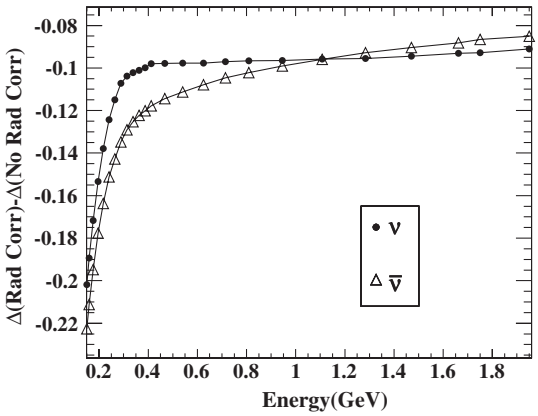


Figure 1.4: Change in Δ, the fractional difference between ν_e and ν_μ CCQE cross-sections, when accounting for radiative corrections [41].

ν_e interactions

Although the theory of lepton universality suggests that the ν_e and ν_μ cross-sections should be the same, differences are expected due to the different charged lepton masses, radiative corrections and uncertainties in the nucleon form factors [41]. The different final state lepton masses affect the kinematic limits of interactions, and cause large effects near the threshold for ν_μ CCQE interactions. These effects are calculable, and are accounted for in all modern neutrino interaction generators.

Radiative corrections are generally not accounted for in neutrino interaction generators, but can distort elastic scattering kinematics. Of particular importance are diagrams where the lepton radiates a photon. An estimate in Reference [41], and reproduced in Figure 1.4, shows a 10% effect on the difference between ν_μ and ν_e cross-sections, although they point out that a more thorough study must be undertaken, using a full neutrino generator and simulating the realistic neutrino flux and detector performance of a given experiment. In the figure, Δ is the fractional difference between the ν_μ and ν_e cross-sections, and is defined as

$$\Delta(E_\nu) = \frac{\sigma_\mu - \sigma_e}{\sigma_\mu}, \tag{1.15}$$

where σ_μ is the ν_μ cross-section and σ_e the ν_e cross-section.

As mentioned previously, there is large uncertainty on the M_A^{QE} parameter of the axial vector form factor. Allowing M_A^{QE} to vary within the experimentally allowed values, as shown in Figure 1.5, gives a 1% effect on the difference between ν_μ and ν_e

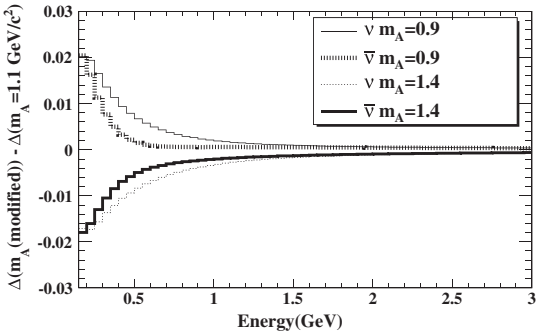


Figure 1.5: Change in Δ, the fractional difference between ν_e and ν_μ CCQE cross-sections, when varying M_A^{QE} from the nominal value of $1.1\,\mathrm{GeV}/c^2$ [41].

CCQE cross-sections at a few hundred MeV. Variations of other form factors, including second-class currents which violate charge or time symmetry, can also be significant, even when varying these factors within the small experimentally-allowed violations.

All the data shown in the previous section related to ν_μ interactions. There are currently no measurements of ν_e cross-sections on carbon at energies relevant to T2K. KARMEN, LSND and E225 all measured the ν_e cross-section on carbon, but the ν_e were produced by muons decaying at rest, so have a maximum energy of approximately 50 MeV [42, 43, 44]. Figure 1.6 shows the expected T2K ν_e flux and the predicted CC ν_e cross-section on carbon; the existing low-energy data is seen to be of very little use to constrain the ν_e cross-sections for T2K. The cross-section predictions in Figure 1.6 are made using the NEUT neutrino interaction generator [45], which is the main neutrino interaction simulator used by T2K. Other generators are available, and each provides implementations of different neutrino interaction models. In particular, GENIE [46] is used as an alternative generator by T2K.

The Gargamelle bubble chamber experiment measured the total CC ν_e cross-section on CF_3Br in 1978 [47]. They observed a linear increase of the cross-section with neutrino energy, which is best parameterised as $\sigma = (0.7 \pm 0.2)E_\nu \times 10^{-38}$ cm^2/nucleon. The data and best fit are shown in Figure 1.7. The ratio of the ν_e cross-section to the ν_μ cross-section was also measured, and found to be $\sigma_{\nu_e}/\sigma_{\nu_\mu} = 0.95 \pm 0.30$, in good agreement with lepton universality.

Measuring and constraining the differences between ν_e and ν_μ cross-sections is critical as the neutrino community searches for CP violation in the lepton sector. Measuring $\nu_\mu \to \nu_e$ oscillations in long baseline oscillation experiments is the most promising way to measure δ, and CC ν_e interactions are both the signal and the dominant background

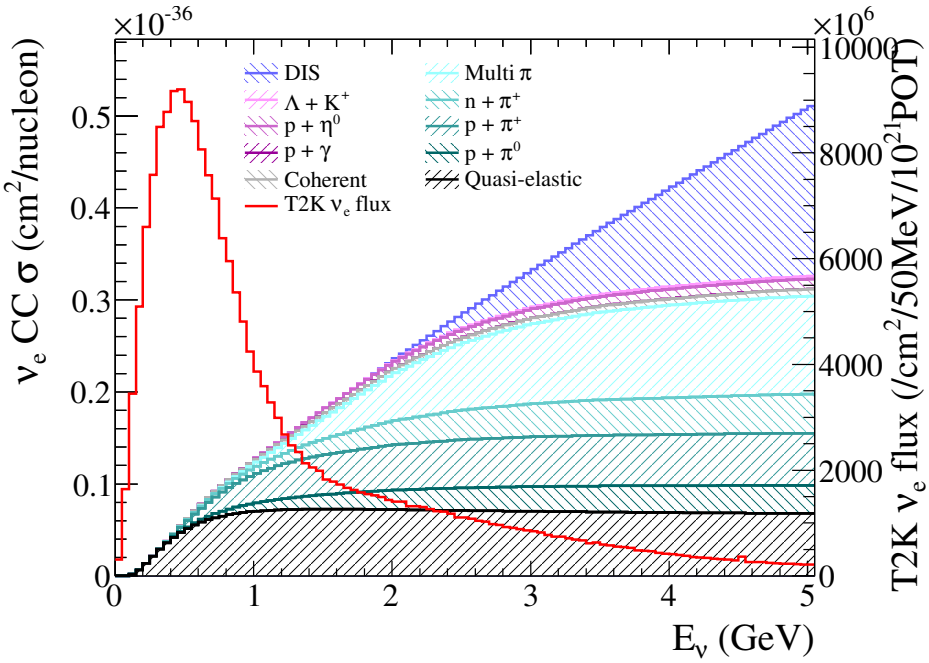


Figure 1.6: CC ν_e cross-section predictions from the NEUT generator and T2K ν_e flux prediction. The cross-section prediction is split into the 10 distinct modes NEUT simulates.

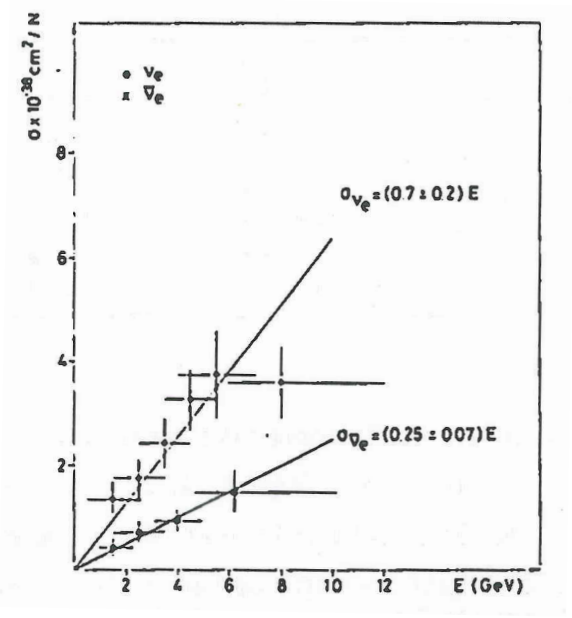


Figure 1.7: CC ν_e and $\overline{\nu}_e$ inclusive cross-section results from the Gargamelle experiment [47].

in these experiments, as will be shown in Section 2.5. Without tighter data-driven constraints on the cross-section models, it will be impossible for future long baseline experiments to reduce the systematic uncertainties to the level required for precision measurements of δ. This thesis is the first measurement of the ν_e CC cross-section on carbon at energies relevant to T2K and other long baseline experiments. As discussed, the only other ν_e cross-section measurement at such energies is from the Gargamelle experiment in 1978. The results in this thesis will therefore provide valuable input to the neutrino community, as models of ν_e interactions with nuclei are developed.

Chapter 2

T2K and the ND280 near detector

The Tokai to Kamioka long baseline neutrino oscillation experiment is designed to improve our knowledge of the parameters governing the PMNS matrix introduced in Chapter 1. The 30 GeV J-PARC proton accelerator in Tokai-mura on the east coast of Japan is used to generate a neutrino beam that is mostly ν_μ, and a suite of near detectors measure the properties of the neutrino beam soon after production. Super-Kamiokande (SK) is located 295 km away and detects the neutrinos after they have oscillated. Constraints on the oscillation parameters are found by comparing the near and far detector measurements of the neutrino beam.

The main goals of T2K as outlined in the original experiment proposal [48] are

1. measure $\sin^2 2\theta_{13}$ by searching for $\nu_\mu \to \nu_e$ appearance
2. make precise measurements of $\sin^2 2\theta_{23}$ and Δm^2_{23} using ν_μ disappearance
3. search for a $\nu_\mu \to \nu_\tau$ component of the oscillation, to distinguish from oscillations to a sterile neutrino.

T2K was the first experiment to show direct evidence for $\nu_\mu \to \nu_e$ appearance [49], and has since excluded $\theta_{13} = 0$ at 7.3σ significance [23]. Combining T2K results with precise measurements of θ_{13} from reactor experiments allows regions of δ to be excluded at the 90% confidence level. T2K's ν_μ disappearance results provide independent measurements of $\sin^2 \theta_{23}$ and Δm^2_{23} [50]. The T2K near detectors, INGRID and ND280, are also measuring a variety of neutrino cross-sections [51, 52], which as well as reducing systematic uncertainties on the T2K oscillation analyses, are important for the neutrino community as a whole.

This chapter details each component of the T2K experiment. The proton accelerator and neutrino beamline are described in Section 2.1, the on-axis near detector INGRID in Section 2.2, the off-axis near detector ND280 in Section 2.3 and the far detector Super-Kamiokande in Section 2.4. Finally, the relevance of this analysis to the T2K experiment is outlined in Section 2.5.

2.1 Accelerator and neutrino beam

T2K generates its neutrino beam by firing 30 GeV protons at a graphite target, producing charged pions and kaons. These hadrons are then focussed using magnetic horns, and decay into products that include neutrinos.

2.1.1 Proton accelerators

A schematic of the J-PARC accelerator complex is shown in Figure 2.1. The three main accelerator phases are the linear accelerator (LINAC), the rapid-cycling synchrotron (RCS) and the main ring synchrotron (MR). The 300 m LINAC accelerates H^- anions to 181 MeV, before charge-stripping foils remove the electrons to produce H^+ ions. The RCS accelerates the protons to 3 GeV, and they are then extracted to be injected into the main ring. Eight bunches of protons (six before June 2010) are injected into the MR every 2–3 seconds, and are accelerated to 30 GeV. The 8 bunches are then extracted in a single turn by a set of kicker magnets, and are directed down the neutrino beamline (NU) to the target. Each spill lasts less than 5 µs and consists of 8 bunches, each containing approximately 3×10^{14} protons and lasting 58 ns. The short duration of the spill is critical to rejecting background events at the near and far detectors.

2.1.2 Neutrino beamline

The neutrino beamline is conceptually broken down into two segments: the primary beamline which transports protons from the MR to the target, and the secondary beamline which handles products of the collisions.

The primary beamline consists of a preparation section, an arc section and a final focussing section, as shown in Figure 2.2. The preparation section aligns the beam

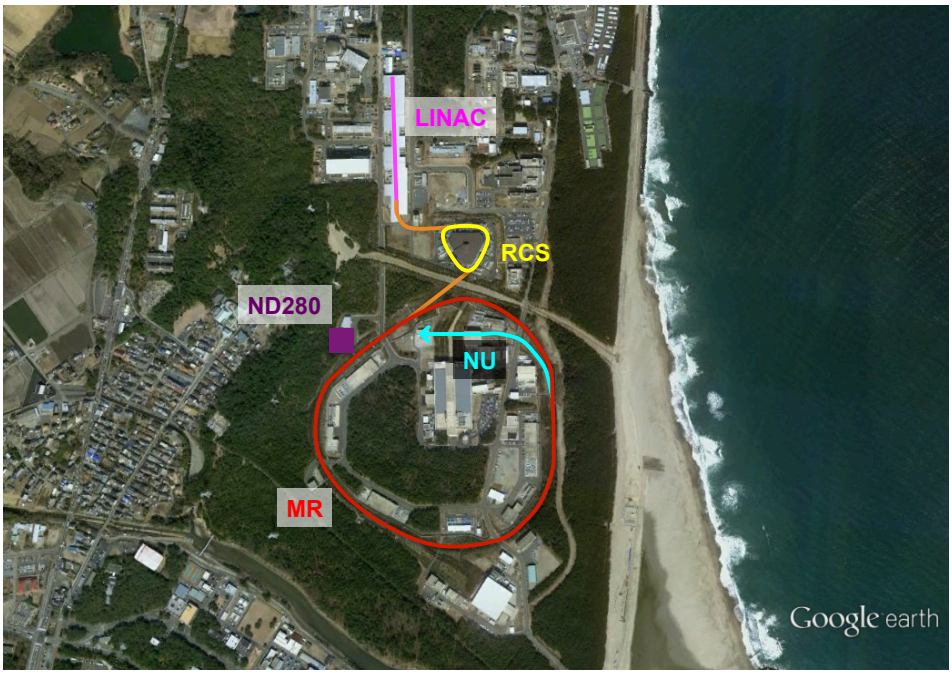


Figure 2.1: Schematic of the J-PARC accelerator complex.

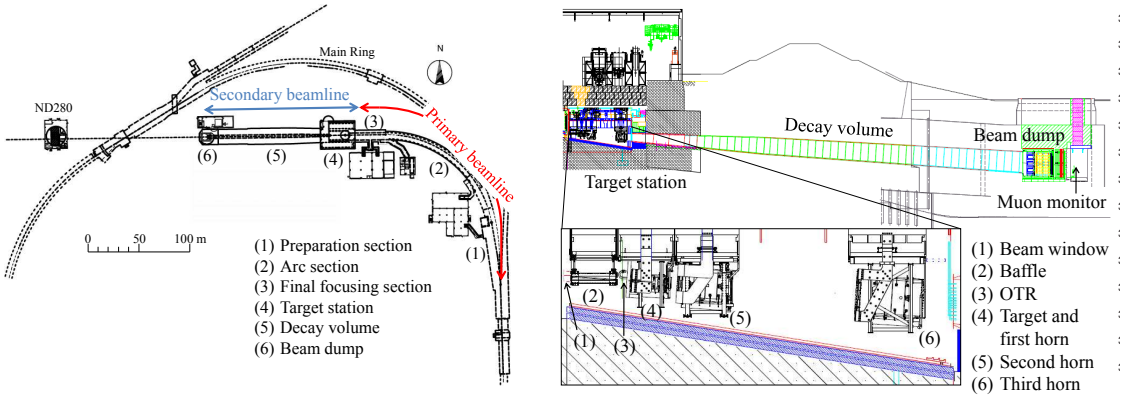


Figure 2.2: Schematics of the whole neutrino beamline (left, viewed from above) and the secondary neutrino beamline (right, viewed from the side and looking south) [53].

for entry to the arc section, where superconducting magnets direct the protons in the desired direction. As will be discussed in the next section, this is 2.5° away from Super-Kamiokande. The final focussing section then aligns the beam for entry to the secondary beamline.

Preventing beam loss and understanding the characteristics of the proton beam is vital for creating a stable and intense neutrino beam, and the primary beamline is equipped with 96 separate instruments for measuring the position, intensity, profile and loss of

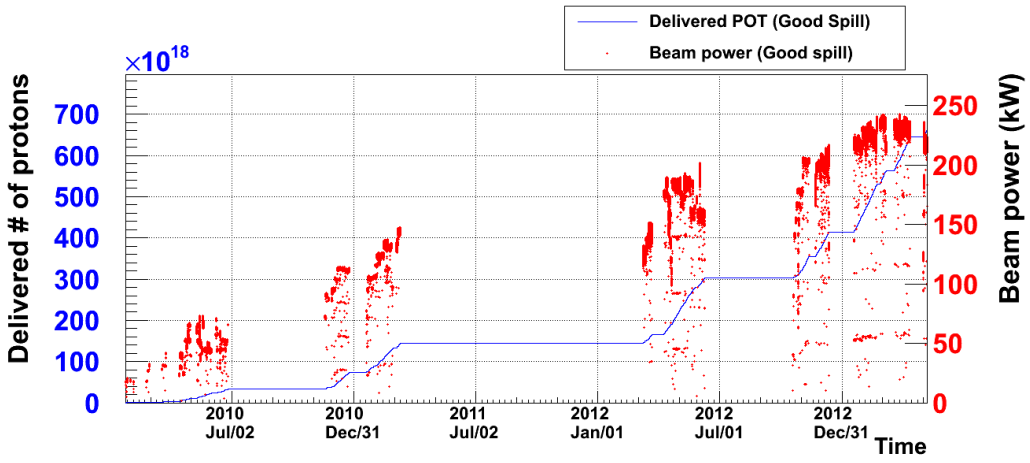


Figure 2.3: Number of protons delivered to the target by the neutrino beamline.

the beam. Among these devices are the current transformers (CT) which measure the intensity of the beam. Each CT consists of a 50-turn toroid around the beam pipe, and the induced current is measured when each bunch passes through the coil. This current is then converted into the number of protons in the spill. The final CT, CT5, is located at the end of the final focussing section just before the protons reach the target. It is the measurement of this CT that is used to determine the number of protons on target (POT), which is the figure of merit used to show how much data T2K has collected.

Figure 2.3 shows the total POT delivered by the neutrino beamline since T2K started taking data. The large gap between March 2011 and February 2012 when no data was taken is due to damage caused by the Great East Japan Earthquake. The disaster caused significant damage to the LINAC and MR, with many components becoming misaligned by several cm. It is a testament to the hard work of the J-PARC staff that beam was resumed so swiftly. From March 2010 to June 2013 a total of 6.40×10^{20} good POT has been delivered to the T2K target.

The secondary beamline contains a graphite target for the protons to strike, magnetic horns to focus the resulting charged hadrons, a decay volume for the hadrons to decay into neutrinos and other particles, and a beam dump to stop the non-neutrino by-products. A schematic of the secondary beamline is shown in Figure 2.2.

The interaction target is a graphite rod which is 2.6 cm in diameter and 91.4 cm long (1.9 interaction lengths). The central rod is surrounded by a graphite tube 2 mm thick, and a 0.3 mm titanium case. Helium gas is pumped through the assembly to cool the target. Interactions between the protons and graphite produce charged pions and kaons,

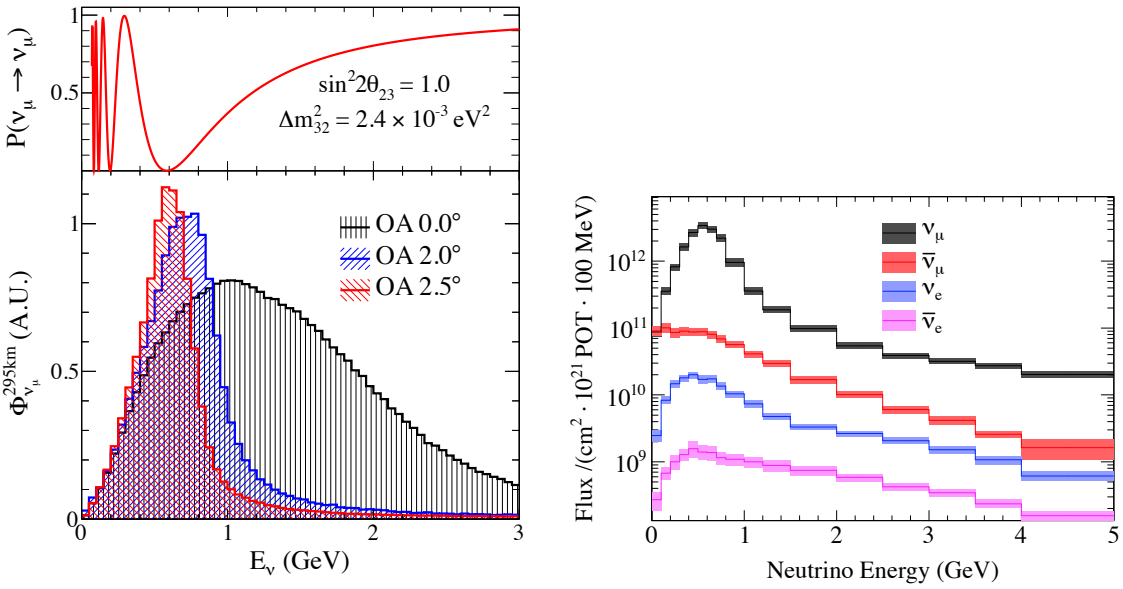


Figure 2.4: Predicted T2K flux, showing that a 2.5° off-axis beam has a much narrower energy spectrum (left) and the neutrino type composition (right) [54, 51].

and these are focussed by a series of three magnetic horns. Each horn consists of two coaxial conductors which produce a toroidal magnetic field which is inversely proportional to the distance from the beam axis. The charged particles are either focussed or deflected depending on the charge of the particle and the direction of the current. So far, T2K has chosen to focus positive hadrons and defocus negative hadrons.

After being focussed by the horns, the hadrons travel through a 96 m long decay volume, which is filled with Helium rather than air to reduce pion absorption and prevent the build-up of tritium and other unwanted contaminants. The hadrons decay to produce neutrinos, and Figure 2.4 shows the predicted neutrino flux as a function of the angle from the beam axis. Due to the kinematics of pion decay, an off-axis beam has a much narrower energy spectrum, and this means that there is a much lower background of high-energy unoscillated ν_μ at the far detector, improving the experiment's sensitivity to ν_μ disappearance and ν_e appearance. It is for this reason that the beam is directed such that ND280 and Super-Kamiokande are 2.5° off-axis.

Figure 2.4 also shows the predicted flux at ND280 broken down by neutrino type. Approximately 1% of the beam is composed of ν_e, but as will be shown, this becomes the dominant background in the ν_e appearance oscillation analysis. The majority of pions decay through $\pi^+ \to \mu^+\nu_\mu$, and decays of these muons through $\mu^+ \to e^+\overline{\nu}_\mu\nu_e$ are the dominant source of ν_e background in the primarily ν_μ beam. A small fraction of

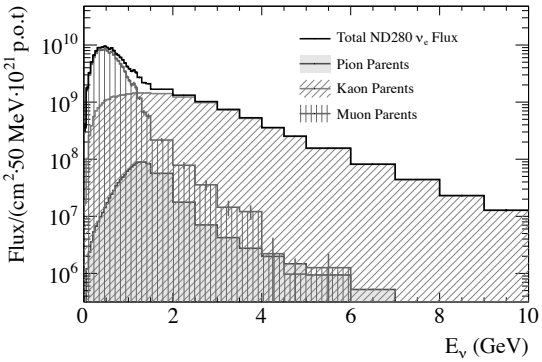


Figure 2.5: Predicted ν_e flux, broken down by the particle that decays to produce the neutrino [54].

pions produce ν_e directly, decaying through $\pi^+ \to e^+\nu_e$. Finally, 5.1% of charged kaons decay through $K^+ \to \pi^0 e^+ \nu_e$, and 40.5% of neutral kaons decay through $K^0_L \to \pi^- e^+ \nu_e$. Figure 2.5 shows the predicted ν_e flux broken down by the particle that decayed to produce the neutrino.

After the neutrinos have been produced, the hadrons, charged leptons and other by-products must be stopped. This is done using a 75 ton graphite beam dump at the end of the decay volume. This stops all hadrons, but only stops muons with energy below 5 GeV. A muon monitor (MUMON) is therefore placed behind the beam dump, and is used to help monitor the direction of the beam by detecting these muons [55].

2.2 INGRID on-axis near detector

INGRID is another detector used to monitor the direction of the beam. It is located 280 m from the target, on the axis of the beam. Whereas MUMON detects muons from pion decay, INGRID instead detects neutrino interactions. The detector consists of 14 identical modules arranged in a plus-sign configuration, with 2 extra off-diagonal modules, as shown in Figure 2.6. Each INGRID module consists of iron plates interleaved with plastic scintillator layers. The plastic scintillator layers use the same technology as the scintillator detectors in ND280, and will be described in detail in the next section. In brief, however, the muons created through ν_μ CC interactions cause the scintillator to produce light, which is then guided to the end of each bar through a wavelength-shifting (WLS) fibre. The light is then detected using a Multi Pixel Photon Counter (MPPC). By having

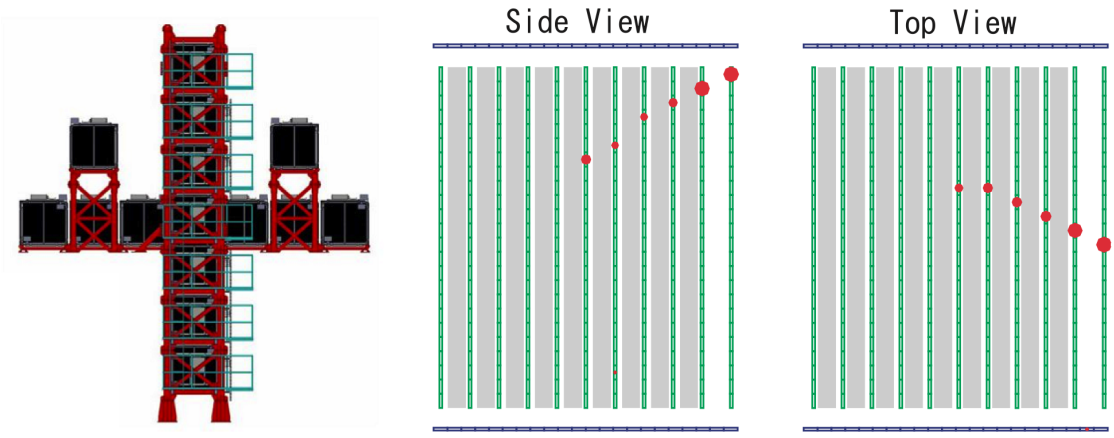


Figure 2.6: Schematic of the INGRID on-axis near detector (left) and an example reconstructed muon track in one of the modules (right).

multiple layers of scintillator in alternating orientations, it is possible to reconstruct a 3-dimensional track of the path the muon took, as is also shown in Figure 2.6.

2.3 ND280 off-axis near detector

ND280 is the off-axis near detector, and is the detector used for this analysis. Like INGRID, it is located 280 m from the target, but is 2.5° off-axis. Both INGRID and ND280 are located in a cylindrical pit dug into the ground. The walls of the pit are lined with concrete, and the ground surrounding it is mostly sand. Figure 2.7 shows an exploded view of ND280, highlighting the many subdetectors that comprise it. The central Tracker region contains three gas-based time projection chambers (TPCs [56]) and two scintillator-based fine-grained detectors (FGDs [57]). The FGDs and TPCs are complementary detectors, with the relatively dense FGDs serving as the target for neutrino interactions, and the light TPCs then tracking any charged particles that are produced. Upstream of the Tracker is the π^0 detector (PØD [58])—optimised for studying interactions with a π^0 in the final state—which consists of scintillator, water and brass layers. Surrounding the Tracker and PØD are electromagnetic calorimeters (ECals [59]) containing layers of scintillator and lead. All these subdetectors are in a horizontal magnetic field perpendicular to the beam direction, which is approximately 0.2 T in strength. Surrounding the entire detector is a magnet return yoke to make the magnetic field more uniform and limit its extent outside the detector. The yoke is instrumented with more plastic scintillator to form the side muon range detectors (SMRDs [60]).

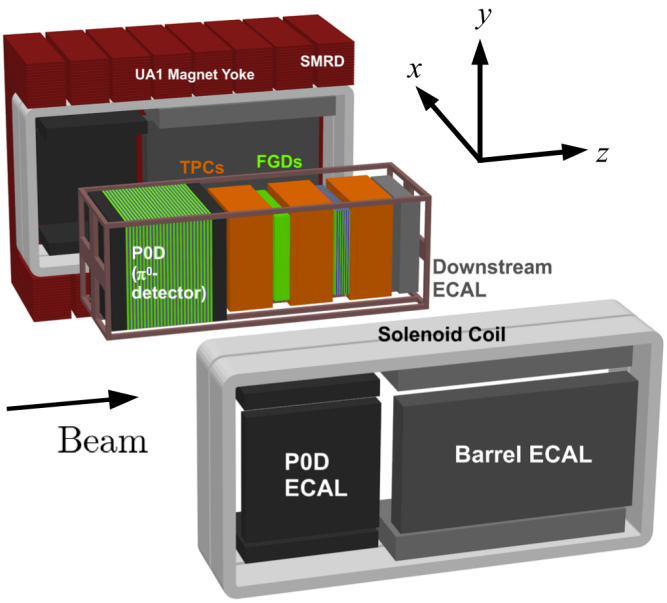


Figure 2.7: Exploded view of the ND280 off-axis near detector.

The ND280 co-ordinate system is defined such that horizontal in the direction of the beam is $+z$, vertical is $+y$, and $+x$ completes the right-hand rule, as shown in Figure 2.7.

The analysis in this thesis primarily uses the TPCs, FGDs and ECals, so the PØD and SMRD will not be described in detail.

2.3.1 Fine-grained detectors: FGDs

ND280 contains two FGDs: the upstream one, FGD1, consists solely of plastic scintillator layers, while the downstream one, FGD2, also contains water layers. The analysis in this thesis uses FGD1 as the active interaction target. This means that in a signal event the neutrino interacts with an atom in the plastic scintillator, and the charged lepton resulting from a CC interaction then causes scintillation. The plastic scintillator is polystyrene ($(C_8H_8)_n$) doped with 1% PPO ($C_{15}H_{11}NO$) and 0.03% POPOP ($C_{24}H_{16}N_2O$).

Each FGD is composed of layers of scintillator bars, with each bar having dimension 9.61 mm × 9.61 mm × 1864.3 mm. The bars are orientated perpendicular to the beam, with subsequent layers alternating between x and y orientations. FGD1 consists of 30 layers of 192 bars each, and FGD2 has 14 layers arranged in 7 xy modules, with 6 water layers interleaved between the modules. This design is chosen such that neutrino cross-sections on water can be calculated by comparing interaction rates in FGD1 and

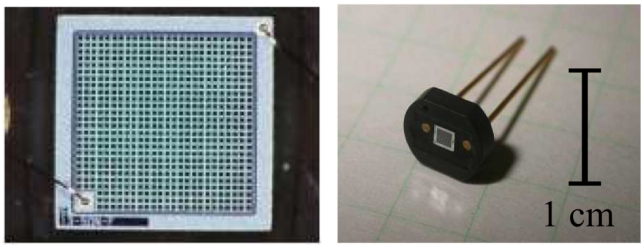


Figure 2.8: A $1.3 \times 1.3\,\mathrm{mm}^2$ MPPC showing the pixel grid (left) and the packaging in which it is mounted (right).

FGD2. As SK is a water Cherenkov detector, it is the cross-section on water that is most important to be understood.

Each scintillator bar has a wavelength-shifting fibre running through its length, in a hole in the middle of the bar. The plastic scintillator emits light at a peak wavelength of 420 nm, and the WLS fibre is well-matched with an absorption spectrum centred at 430 nm. The emission spectrum of the fibre is centred at 476 nm however, and as there is little overlap between the emission and absorption spectra, there is little self-absorption as the light travels along the fibre. One end of the fibre is coated with aluminium to form a mirror, while the other end is attached to a Multi Pixel Photon Counter (MPPC).

An MPPC measures and digitises the light signal and, unlike a photo-multiplier tube (PMT), is able to operate in the 0.2 T magnetic field. An example of an MPPC is shown in Figure 2.8. The MPPC consists of 667 individual pixels covering an area of $1.3 \times 1.3\,\mathrm{mm}^2$. When a photon from the fibre strikes a pixel it creates a photoelectron which then generates an avalanche. Each pixel operates as a binary device—either a photon was detected or not—and the signal of the MPPC as a whole is the sum of the number of pixels fired. The size of the signal depends on the overvoltage applied to the MPPC, and careful calibration is required to normalise the response of each bar to the same level. Individual pixels can also generate a signal even when there is no incident photon. This "dark noise" means that if an MPPC only has a few pixels fire, it was likely due to noise rather than incident photons, and should be ignored when trying to reconstruct particle tracks.

The granularity with which particles can be tracked is limited by the $1 \times 1\,\mathrm{cm}^2$ cross-sectional area of the scintillator bars. Particles that only travel a short distance may not create enough "hits" (bars which had their MPPC triggered) for the track to be reconstructable, and two co-linear tracks may appear to be a single track with higher charge.

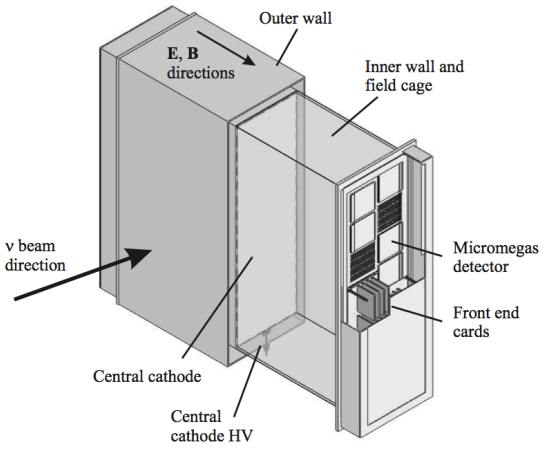


Figure 2.9: Cutaway view of a TPC module, showing the outer and inner walls and the MicroMEGAS detectors and their electronics.

2.3.2 Time projection chambers: TPCs

Each of the three ND280 TPCs are filled with 3000 litres of Argon-based gas. An electric field is applied through the gas, parallel to the 0.2 T magnetic field, from a central cathode to anodes at each side of the TPC. Charged particles ionise the gas, and the ionisation electrons drift towards the electrical anodes. Here, MicroMEGAS detectors amplify the signal and the analogue signal is recorded. A diagram of a TPC is shown in Figure 2.9. Each side of each TPC is instrumented with 12 MicroMEGAS modules, with each module split into 1728 pads in 48 rows and 36 columns. Each pad is 7.0 mm $\times$ 9.8 mm, so is similar in size to the polystyrene bars in the FGD. All the pads in the TPC are in the same (yz) plane, and the x co-ordinate is determined by the relative time difference between hits.

The TPC is able to provide much better tracking of particles than the FGDs due to its larger volume, and the fact that all the pads are in the same orientation and provide three-dimensional hits, rather than the alternating xz and yz layers of two-dimensional hits of the FGD. The TPCs are also able to provide excellent momentum resolution by measuring the curvature of a particle in the magnetic field, and are also able to identify the particle type based on the energy loss along the track. These capabilities will be explained in more detail in Section 3.3.

2.3.3 Electromagnetic calorimeters: ECals

The ECals use the same technology as the FGDs, but are coarser in granularity. Each bar is 1 cm × 4 cm, and there are layers of 1.75 mm-think lead between each layer. The lead helps to contain electromagnetic showers, and so the ECal is able to distinguish between electrons (which create a shower) and muons (which do not). The lead layers also increase the probability for a photon to shower, making their detection much more efficient. There are 13 separate ECal modules, as was shown in Figure 2.7. The Downstream ECal (DsECal) consists of 34 layers, for a total thickness of 10.6 radiation lengths (10.6 X_0), with each bar read out by two MPPCs—one at each end of the fibre. The six Barrel ECal modules surround the Tracker region, and consist of 32 layers each, for a total thickness of 9.7 X_0. The long bars running in the z direction are read out by two MPPCs. The bars perpendicular to these are much shorter, so are read out by one MPPC, and have the other end of the fibre mirrored. The six PØD ECal modules surround the PØD, and are not used in this analysis.

2.3.4 Data acquisition and processing

The ND280 data acquisition system (DAQ) is responsible for triggering the readout of ND280 data and storing it for future retrieval.

The three main triggers for reading physics data are

- **Beam trigger:** when beam is extracted from the main ring and through the neutrino beamline, a signal is sent from the neutrino beamline to the ND280 DAQ to instruct it to record data
- **FGD cosmic trigger:** if there is no beam trigger, but coincident hits are seen in both FGDs, the data is recorded as the hits are likely to have been caused by a cosmic ray muon, and these events are useful for calibration and calculating systematic uncertainties
- **TRIP-t cosmic trigger:** if there is no beam trigger, but hits are seen on opposite sides of the detector (top and bottom SMRDs, left and right SMRDs, PØD and DsECal), then data is recorded as these are again likely to have been caused by cosmic ray muons. A pre-scaling is applied based on the approximate direction of the track, so that the data is not dominated by vertical muons. This is referred to

as the TRIP-t cosmic trigger as the TRIP-t ASIC chip is used in the electronics of the PØD, ECal and SMRD subdetectors.

The data is first stored at the KEK computing centre in Japan, before being replicated to TRIUMF in Canada and RAL in the UK. Processing of the data through the software chain detailed in Chapter 3 currently takes place at TRIUMF, allowing for a very rapid turnaround as new data is taken. Monte Carlo production is more distributed, with processing taking place at many institutions in North America and Europe.

2.4 Super-Kamiokande far detector

Super-Kamiokande (SK) is a cylindrical water Cherenkov detector containing 50 ktons of pure water. The inner detector consists of 11,146 20″ PMTs and surrounds the central 35,000 tons of water. The outer detector consists of 1,885 8″ PMTs facing outwards, and acts as a veto for particles originating outside of the inner volume of the detector.

A charged particle that is travelling faster than the speed of light in the material it is passing though produces Cherenkov radiation, and it is this radiation that the PMTs detect. The radiation is emitted at an angle of $\cos\theta = 1/n\beta$, where n is the refractive index of the material, $\beta = v/c$, and v is the speed of the particle. For relativistic particles in water, the light is emitted at an angle of 42°. As the particle loses energy and slows down it stops emitting Cherenkov radiation, so a ring of light is observed on the side of the detector. Electrons and muons can be distinguished by how sharp and well-defined this ring is, and example rings are shown in Figure 2.10. This is because electrons create an electromagnetic cascade producing multiple Cherenkov rings in slightly different directions, whilst a muon generates a single ring with well-defined edges.

The primary cause of electron mis-identification is due to a π^0 decaying to two photons. Photons and electrons cannot be distinguished, as they both produce electromagnetic cascades. If one of the photons is not reconstructed, either because it is very low energy or the two photons are very co-linear, then a single electron-like ring is observed, mimicking the ν_e signal. Improving knowledge of the NC1π^0 cross-section on water is therefore one of the most important goals of ND280.

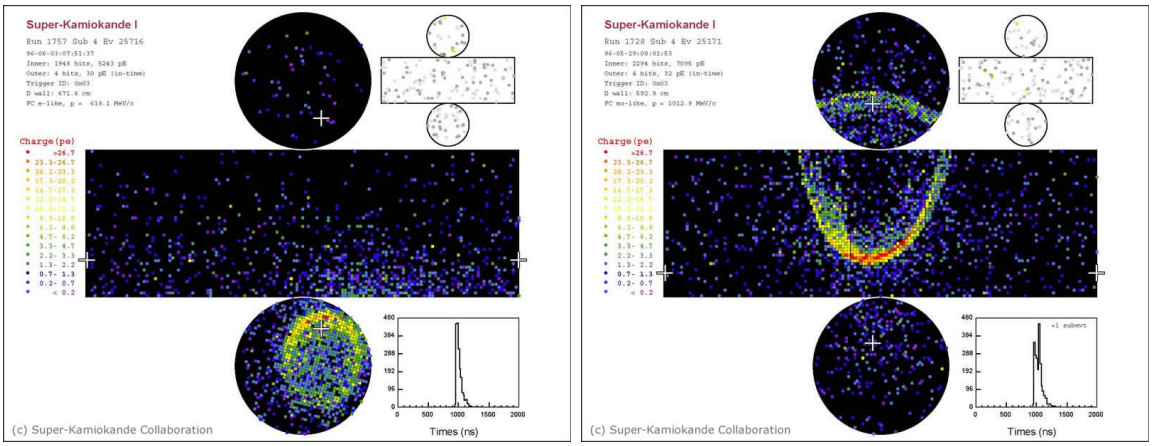


Figure 2.10: SK event display for an electron-like data event (left) and a muon-like data event (right).

2.5 Importance of understanding ν_e characteristics

Knowledge of the ν_e component of the T2K beam is critically important when studying $\nu_\mu \rightarrow \nu_e$ appearance. Although ν_e account for only 1% of the produced neutrinos (see Section 2.1.2), they are the dominant background at Super-Kamiokande due to ν_μ disappearance and cuts being applied to select electron-like events. For $\sin^2 2\theta_{13} = 0.1$, 21.6 events are expected in the current analysis, of which 17.3 are true ν_e from ν_μ oscillations, 3.2 are from intrinsic ν_e in the beam, and 1.1 are from other backgrounds [23]. So, intrinsic ν_e are three times as large a background as all other backgrounds combined.

Improved knowledge of ν_e cross-sections and the ν_e flux is therefore critical to not only reduce systematic errors on the expected number of signal events, but also on the dominant background. This thesis describes the first measurement of the ν_e cross-section on carbon at energies relevant to T2K.

Chapter 3

ND280 software

ND280 uses a custom software suite based on ROOT [61] and Geant4 [62] to simulate and analyse data. The majority of the software uses the oaEvent library, which was developed specifically for ND280 and provides a common framework for reading, manipulating and saving information about an event. The "oa" prefix of oaEvent stands for "off-axis", and the same prefix is used for naming many other pieces of the ND280 software.

As ND280 is a complicated detector, the software is split into multiple packages that each perform a specific role. The ND280 Monte Carlo simulation proceeds in the following stages.

1. Simulate the flux of neutrinos arriving at ND280.
2. Simulate neutrino interactions—using either the NEUT or GENIE generators—determining where in the detector the interaction occurs, and the products of the interaction.
3. Combine multiple neutrino interactions to simulate a spill, and use Geant4 to propagate the products of the interactions (this package is called nd280mc).
4. Simulate the response of the detector and electronics (elecSim).
5. Apply calibration to the detector output (oaCalib).
6. Reconstruct particle tracks and showers from the hits (oaRecon).
7. Summarise the reconstruction and truth information into a format ready for analysis (oaAnalysis).
8. Analyse the results (highland).

When processing data, the first four stages are replaced by unpacking the data and converting it into a format ready for oaCalib. The rest of this chapter details the most important aspects of the software chain.

3.1 Monte Carlo event simulation

3.1.1 Neutrino flux prediction

The neutrino flux is predicted using a simulation that starts by tracking protons in the primary beamline and ends when hadrons decay to produce neutrinos in the decay volume. FLUKA2008 [63] is used to simulate 30 GeV protons impinging on the target and baffle, and GEANT3 [64] is used to propagate the resulting particles through the secondary beamline. Particles are tracked through the decay volume until they either interact or decay, with the particle decays calculated using the latest PDG branching fractions and decay rates. The properties of any resulting neutrinos are saved, including the full interaction chain that produced the neutrino. This allows events to be re-weighted based on external hadron interaction data, the measured beam profile, or other external data.

The initial flux simulation is tuned using both external data and beamline measurements. NA61/SHINE is a dedicated hadron interaction experiment at CERN, and was designed to measure the hadrons produced by 31 GeV/c protons colliding with a graphite target [65]. The experiment has collected data using both a thin (2 cm) graphite target and a replica of the T2K target, and the experiment can detect pions in 90% of the phase-space relevant to T2K, and kaons in 60% of the relevant phase-space. Differences between the NA61 data and FLUKA2008 simulation are used to re-weight the neutrino spectra expected at ND280 and Super-Kamiokande, and data from other experiments are used in the regions of phase-space that NA61 does not cover [66, 67].

As well as NA61 data, differences between the measured and simulated beam profile in each T2K run are used to re-weight the flux. The T2K beam group provides re-weighting histograms as a function of true neutrino energy for each T2K run. As an example, the re-weighting applied to the ν_e flux for one T2K run is shown in Figure 3.1.

The FLUKA2008-based simulation described above is used in this analysis, and is tuned using NA61 thin target data and beamline measurements. In future T2K analyses a new flux prediction will be used, which uses FLUKA2011 instead of FLUKA2008, and

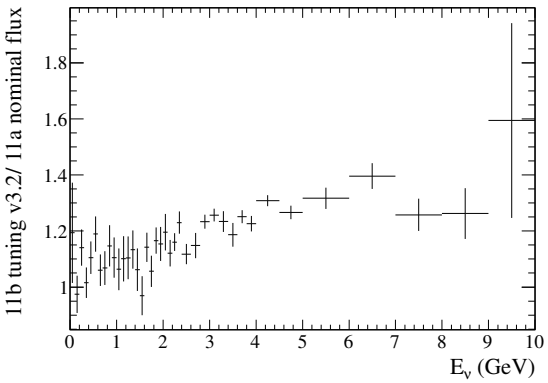


Figure 3.1: ND280 ν_e flux re-weighting applied to simulated data, based on NA61 data and the measured beam profile for one T2K run (Run 3c).

includes the NA61 replica target data. The effect of flux re-weighting in future analyses will therefore be much less than the 10–20% effect shown in Figure 3.1.

3.1.2 Neutrino interaction simulation: NEUT

The NEUT [45] event generator is used to simulate neutrino interactions in ND280. Neutrino energies from tens of MeV to hundreds of TeV can be simulated, and interactions on all the elements in ND280 are handled. Signal and background interactions are simulated in the entirety of ND280, up to and including the magnet return yoke. In reality, neutrinos can also interact outside the magnet volume, in the concrete walls of the pit or sand surrounding it, and a special "sand" Monte Carlo is used to estimate the impact of these events on an analysis. This will be described in more detail in later sections.

The properties of neutrinos simulated by the beam group are used as the input to NEUT, and dictate the energy, position, direction and flavour of incoming neutrinos. Combining this with a model of the ND280 geometry (described in Section 3.1.3), NEUT tracks the neutrino and calculates the interaction probability on all the matter that it crosses. A pseudo-random number generator then determines whether an interaction should be simulated, along with its position and mode.

As mentioned in Section 1.2, the dominant interaction modes at ND280 are quasi-elastic scattering (CCQE) at low energies, single pion production (CC1π) above the pion production threshold, and deep inelastic scattering (DIS) at high energies. This

section will briefly outline the interaction models used in NEUT for each process. A more complete description can be found in References [51] and [53].

NEUT uses the Llewellyn Smith model [68] to simulate CCQE interactions, and uses Smith and Moniz's relativistic Fermi gas model to describe the nucleons within the nucleus [69, 70]. For carbon, the Fermi momentum is set to 217 MeV/c and the binding energy to 25 MeV.

Pion production is simulated using Rein and Sehgal's model [71]. Eighteen resonances and their interference terms are simulated, and 20% of the Delta resonances undergo pion-less Delta decay, in which the Delta de-excites without emitting a pion.

Multi-pion and DIS processes are simulated using two distinct processes, depending on the energy of the hadronic system. PYTHIA/JETSET [72] is used for energies above 2 GeV, and internal NEUT code is used below this limit, as PYTHIA/JETSET does not reproduce experimental data well at lower energies.

After simulating the initial interaction, final state interactions must be simulated by propagating the resulting particles to the edge of the nucleus. Each particle is stepped through the nucleus, with the probability of an interaction (such as charge exchange, absorption or scattering) calculated at each step. If an interaction occurs, the resulting particles are then used for stepping through the rest of the nucleus. This cascade continues until all particles have reached the edge of the nucleus.

3.1.3 ND280 detector simulation: nd280mc

NEUT only simulates individual neutrino interactions, so the ND280 detector simulation first groups them into spills. The number of interactions in a spill is defined by the beam intensity that is being simulated, with Poisson variations around the expected number. The interactions are then distributed in time according to the beam bunch structure.

nd280mc then uses ROOT and Geant4 to simulate the passage of the outgoing particles from each interaction through the ND280 detector. The geometry of the detector is simulated in ROOT, with the composition and size of each element taken from design specifications. If the "as-built" geometry differs from the design, components can be shifted using alignment constants. The simulated geometry is very detailed, with each scintillator bar modelling the central hole through which the wavelength-shifting fibre

runs, the rounded corners of the bar and the epoxy used to glue bars together, for example.

Geant4 uses the ROOT geometry to propagate particles through the detector, simulating particle interactions, decay and energy deposition. The output of nd280mc includes the time, position and amount of energy deposited by each particle—in a list of Geant4 "hits"—and this information is then used in elecSim to simulate the detector response to particles passing through the detector.

3.1.4 Detector and electronics simulation: elecSim

elecSim takes the list of Geant4 hits from nd280mc and simulates the response of the detector. In the scintillator detectors, the light produced by the scintillator, the propagation of light along the fibre and the response of the MPPCs are all simulated. In the TPC, the drift of ionised electrons and the response of the MicroMEGAS is simulated. In both cases the effect of the readout electronics is then simulated, to produce an output in the same format as the real data.

3.2 Detector calibration: oaCalib

The oaCalib package is responsible for calibrating the ND280 data. The calibration group produce calibration constants that are valid for a specific period of data, and these are stored in a database, which is then queried by oaCalib.

The scintillator-based detectors all use the same technology and use similar techniques to calibrate the light yield and time of hits. The two main datasets used for calibration are "pedestal" triggers and "cosmic" triggers. The pedestal triggers are taken at random times and generally are empty events with no particles travelling through them. This allows the "dark noise" of the MPPCs to be measured, which is the signal produced when no light is incident on the sensor. The cosmic triggers select high energy cosmic ray muons. The high energy means that they travel in straight lines and are not significantly deflected by the magnetic field. Simple reconstruction algorithms can then be used to extrapolate the path the particle took and calculate the expected energy deposit in each bar. The response of different bars to these minimally-ionising particles allows bar-to-bar variations to be calibrated. Timing calibration accounts for both bar-to-bar variation

and delays introduced by the readout electronics. Bar-to-bar variations, for example, are calibrated by comparing the difference in recorded hit times between bars to the difference expected for a particle travelling at the speed of light. There is no precise inter-detector timing calibration in the data used in this analysis, and deviations of up to 10 ns are observed between the TPCs and the ECal modules. Both the reconstruction and analysis are designed so that these deviations do not significantly affect the analysis. Full inter-detector timing calibration at the 1 ns level is expected in future T2K analyses, and this will allow the direction of tracks—either forwards-going or backwards-going—to be more reliably determined.

Unlike the scintillator detectors, the TPCs must also calibrate for the reconstructed position of particles due to distortions of the electric and magnetic fields. A dedicated laser system is used for this, and will be discussed more in Section 5.1.1 in the detector systematics section. The energy deposited in the gas by charged particles, the drift velocity of ionised electrons in the gas, and the gain of the electric field are also calibrated.

3.3 Event reconstruction: oaRecon

The ND280 reconstruction is broken down into two phases: local reconstruction and global reconstruction. In local reconstruction, each subdetector groups hits (MPPCs or MicroMEGAS pads with charge above a noise threshold) together to form tracks and showers. Global reconstruction then combines these tracks and showers to form a complete picture of an event. Figure 3.2 shows an example of the local and global reconstruction results for the same event. Although an attempt is made to identify the type of particle that created each global track, it is generally more powerful for each analyser to combine the particle identification (PID) capabilities of the local reconstruction algorithms according to their own needs. As will be described in Chapter 4, this analysis uses the TPC and ECal modules for PID purposes. The momentum of each global track is reconstructed assuming various particle hypotheses, and the analyser is then free to choose whichever is most appropriate for their analysis.

3.3.1 TPC reconstruction: tpcRecon

The TPCs are key to reconstructing particles in the Tracker region, due to the excellent tracking capabilities they provide. As will be shown in Chapter 4, this analysis requires

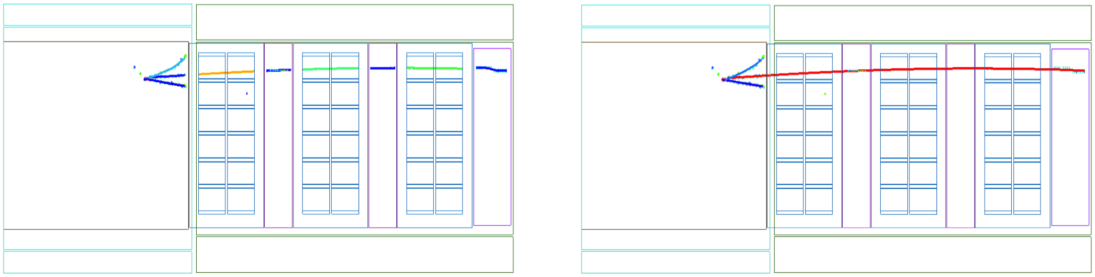


Figure 3.2: Results of the local (left) and global (right) reconstruction algorithms for the same event. The long global track consists of segments in the PØD, TPC1, FGD1, TPC2, FGD2, TPC3 and the DsECal. The colour of each track is related to the energy deposited by the object.

tracks to pass through the TPC for them to be considered in the analysis, so that good momentum resolution and PID can be attained.

The output of each MicroMEGAS pad is a waveform of the charge readout as a function of time, and the first stage of tpcRecon is to scan this waveform to determine the periods in which the charge was above a noise threshold. Adjacent pads in the same horizontal row that have overlapping charge peaks are combined into clusters, and these clusters are then combined using a cellular automaton to create the two-dimensional set of pads corresponding to a track. Reconstruction occurs separately in each of the three TPCs, and long tracks that span multiple TPCs are merged later in trackerRecon and globalRecon.

Although the time of each peak in the waveform is known, there is an ambiguity when trying to determine the x co-ordinate of a particle and the time it passed through the detector (remembering that the x co-ordinate is along the drift direction of the TPCs). A particle that arrives early and far from the anode will produce peaks in the waveform at the same time as a particle that arrives late but close to the anode. tpcRecon therefore looks at hits in the neighbouring FGDs to break the degeneracy. The TPC track is extrapolated in the yz plane, and the time of the closest FGD hit to the extrapolated track is taken as T_0, the start time of the TPC track. If there are no suitable hits in the FGDs, then hits in the PØD and ECal are also examined, for TPC1 and TPC3 tracks, respectively.

Particle identification in the TPCs is based on the energy the particle loses as it travels. Energy loss is governed by the Bethe equation, and is dependent on $\beta \equiv v/c$. If the momentum of the particle is known, one can compare the measured energy loss with that expected for particles with different masses. The "pull" from particle hypothesis α,

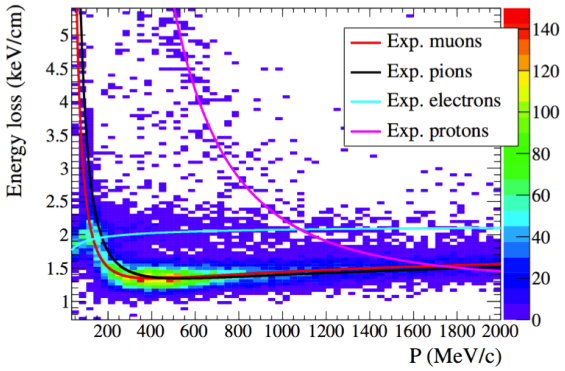


Figure 3.3: Measured and expected energy loss in the TPC, demonstrating the excellent particle identification capabilities.

δ_α, is then defined as

$$\delta_\alpha = \frac{dE/dx_{\text{meas}} - dE/dx^{\alpha}_{\text{exp}}}{\sigma^{\alpha}_{\text{meas}}}, \tag{3.1}$$

where dE/dx_{meas} is the measured energy loss, $dE/dx^{\alpha}_{\text{exp}}$ is the expected energy loss for particle type α and $\sigma^{\alpha}_{\text{meas}}$ is the uncertainty on the measured energy loss. Figure 3.3 shows the measured dE/dx and the energy loss expected for different true particle types.

3.3.2 FGD reconstruction: fgdRecon

FGD reconstruction is performed in two stages: TPC-FGD matching and FGD-only reconstruction. In TPC-FGD matching, the TPC tracks reconstructed by tpcRecon are extrapolated into the FGD volumes. A Kalman filter is used to determine which hits in the FGD should be added to the track, and the incremental addition of hits to the track proceeds from the layer nearest the TPC track to the layer furthest away. If no hits are added in two consecutive layers, then propagation stops and no more hits are added to the track.

In FGD-only reconstruction, a simple clustering algorithm is used to create tracks using only the FGD hits that were not used in the TPC-FGD matching stage. FGD hits are first clustered into time-ordered groups, with a gap of at least 100 ns between adjacent groups. In each time group, hits in adjacent layers are clustered into segments, segments are connected if they form a line that is nearly straight, and the connected segments are then combined into tracks. Due to the requirement of having hits in multiple layers, this algorithm breaks down for tracks perpendicular to the beam. Although specific exceptions

have been coded to handle these cases, the performance of FGD-only reconstruction is very poor for tracks with $\cos(\theta) < 0.3$, where θ is measured from the $+z$ axis. Further, FGD-only reconstruction generally fails to find tracks that are less than 10 cm in length.

FGD hits are also used to tag the presence of a Michel electron (a delayed electron produced when a muon decays). Any time groups that contain a significant amount of charge, but are more than 100 ns from the beam time, are tagged as Michel electron candidates.

3.3.3 ECal reconstruction: ecalRecon

The TPCs and FGDs only attempt to reconstruct particle tracks. The ECal is different as the high density of lead can cause electrons and photons to shower, so both tracks and showers must be reconstructed. One of the features of a particle shower is that a photon may traverse several layers before converting into an e^+e^- pair, and so hits in the same shower may not be immediately adjacent. ecalRecon starts by using a simple nearest-neighbour algorithm to create basic 2-dimensional clusters of hits that are within two layers and one bar of each other. These clusters are then combined if they are close together in time and are pointing in the same direction. Any unclustered hits that are close to the combined cluster are then added. Finally, the 2D clusters from the two different views—for example the xz and yz views for the DsECal—are combined to form 3D clusters.

Particle identification in the ECal uses a multi-variate analysis (MVA) based on the shape and charge distribution of the 3D object. Tracks from muons and other minimally-ionising particles are expected to be long and thin, with uniform energy deposition along the track; electromagnetic showers are expected to appear more spherical, with highly non-uniform energy deposition. The specific variables used by the MVA are

- circularity: the ratio of the major and minor axes from a principal component analysis of the position of hits in the cluster
- QRMS: the RMS of the charge distribution of hits
- FrontBackRatio: the ratio between the charges in the first and last quarter of the track
- TruncatedMaxRatio: the ratio between the charges in the layers with highest and lowest charge, after removing the higher and lower 20% of hits.

Log-likelihood ratios are formed to distinguish between how well the cluster matches different particle category hypotheses. The variables available are

- $R_{\mathrm{MIP/EM}}$ which distinguishes between minimally-ionising particles—such as muons—and electromagnetic showers
- $R_{\mathrm{EM/HIP}}$ which distinguishes between electromagnetic showers and highly-ionising particles such as protons
- $R_{\mathrm{MIP/Pion}}$ which distinguishes between minimally-ionising particles and showering pions.

An estimate is also made of the electromagnetic energy deposited in the ECal by the object, and this is referred to as E_{EM}.

3.3.4 Global reconstruction: globalRecon

The final stage of the reconstruction chain is to combine the results of all the local reconstruction algorithms. The RecPack toolkit [73] is used for this, which provides tools for extrapolating tracks beyond the hits they use, Kalman filters for determining how well other reconstructed tracks match the extrapolated track, and Kalman filters for fitting the resulting merged track. RecPack accounts for processes such as bremsstrahlung radiation and multiple scattering, and utilises a simplified geometry of the detector with the correct average density for each major component.

As the TPCs provide the best spatial and momentum resolution, they are used as the basis for extrapolation. A reduced version of globalRecon, called trackerRecon, first attempts to merge TPC and FGD tracks into longer Tracker tracks. globalRecon then takes over and attempts to match Tracker tracks with PØD and ECal objects, and to hits in the SMRD. PØD ECal objects are not considered for matching by globalRecon in the software version used in this analysis. This does not impact this analysis, however, as it is based in the Tracker region, downstream of the PØD ECal.

Once all the global tracks have been formed, they are each refit assuming different particle hypotheses. The main use of these alternate fits is to determine most accurately the momentum of the track, once the analyser has decided the particle type. Fits are currently performed assuming the electron, muon and proton hypotheses. Tracks are generally assumed to be travelling forwards (towards $+z$), unless the track passes through

both FGD1 and FGD2. In that situation, the track is reversed if the time of the FGD2 segment is before the time of the FGD1 segment.

Chapter 4

Event selections

The aim of this analysis is to measure the charged-current ν_e cross-section on carbon. A selection of ν_e CC interactions, characterised by having an electron in the final state, is therefore required. A significant background in the selection comes from photons converting into electron-positron pairs, so a second sample is used to constrain this background. In this chapter, Section 4.1 details the ν_e event selection and Section 4.2 details the γ event selection.

The analysis is performed with all ND280 data taken in 2010–2013 that passes the standard beam quality and ND280 data quality checks. This corresponds to a total of 5.90×10^{20} POT. T2K data is assigned a run number based on the calendar year in which it was taken, with letter suffixes used for logically distinct subsets of each run. Because of the different beam and detector conditions, the Monte Carlo for each run is weighted to the POT for the equivalent data run separately, as shown in Table 4.1. The main configuration differences between each run are:

- In Run 1 (in 2010), the Barrel ECal modules were not installed, and only the DsECal was present. In all subsequent runs, the DsECal and Barrel ECal are both present.
- In Run 1, each beam spill consisted of 6 bunches. There are 8 bunches for all subsequent runs.
- The PØD contains water bags that can be filled with either air or water. The details of whether the bags are filled are shown in Table 4.1.
- The beam power has increased over time, and three separate beam powers are simulated (beama, beamb and beamc). Differences between the simulated and true

T2K run	Data POT ($\times 10^{19}$)	MC POT ($\times 10^{20}$)	PØD water status	MC beam power
Run 1	1.7	2.0	water	beama
Run 2 (water)	4.3	5.0	water	beamb
Run 2 (air)	3.6	5.0	air	beamb
Run 3b	2.1	2.3	air	beamc
Run 3c	13.5	13.5	air	beamc
Run 4 (water)	16.2	25.0	water	beamc
Run 4 (air)	17.6	14.0	air	beamc
Total	59.0	66.8		

Table 4.1: Definition of T2K runs and the amount of data and MC POT used in the analysis.

beam powers do not affect the number of neutrino interactions (as that is based on the amount of POT analysed), but does affect the piling up of events. This is discussed in the detector systematics section.

- In Run 3a, the magnetic horns that focus the beam were switched off, so data from this period is not used.
- In Run 3b, the magnetic horns operated at 205 kA, rather than the nominal 250 kA.

Overall, the Monte Carlo simulates 11.3 times the data POT. The NEUT version 5.1.4.2_nd280 prediction is used for the MC plots in this chapter, with the NA61 flux tuning applied, as explained in Section 3.1.1. The Monte Carlo prediction is normalised to the amount of data POT.

4.1 ν_e interaction selection

The selection of ν_e interactions in ND280 is difficult due to the small fraction of ν_e in the T2K beam. This component is expected to be of the order of 1% of the total neutrino flux, and to select it the large number of ν_μ interactions have to be rejected. For this reason a fundamental tool for this analysis is the particle identification (PID). As will be shown, this analysis combines the TPC and the ECal PID capabilities to reject more than 99.8% of the muons from ν_μ interactions. However, ν_μ interactions—either inside or outside the FGD—can also produce photons that then convert into electron-positron

pairs in the FGD. These γ conversions are the main source of background for the analysis, and are a much larger component than the few muons that are misidentified as electrons.

As the goal of this analysis is to measure the ν_e cross-section on carbon, FGD1 is used as the interaction target. FGD2 contains water layers, and using that as an interaction target would yield a combined carbon/oxygen cross-section, which is less useful for the community. The most directly relevant measurement for T2K would be to measure the cross-section on water, by combining the FGD1 and FGD2 information. Unfortunately, such a statistical subtraction is infeasible using the statistics expected with the current dataset, although such a measurement should be pursued in the future.

The first stage of selecting CC ν_e interactions is to select events with lepton tracks starting in FGD1. These criteria are shared with the ND280 ν_μ analyses, and aim to

- reject the large background from "sand muons", produced by neutrino interactions in the sand or the concrete walls of the ND280 pit
- reject the background from neutrino interactions in the magnet, which produce tracks that enter the Tracker or produce neutral particles that may interact in the FGD
- retain a high efficiency.

The ν_e-specific selection then has the additional goals of

- rejecting the muons produced in the large number of ν_μ interactions
- rejecting neutrino interactions inside or outside the FGD producing photons that convert into e^+e^- in the FGD.

Taking into account these points, the following criteria are used to select ν_e candidates.

1. **Event quality:** use only events selected with the beam trigger and compatible with one of the 8 bunches (6 bunches for Run 1).
2. **Track selection:** select the highest-momentum negative track of the event that contains a reconstructed segment in the TPC, and require it to have a momentum greater than 200 MeV/c. This track is referred to as the lepton candidate.
3. **FGD FV:** require that the start position of the lepton candidate is in the FGD1 Fiducial Volume (FV).

4. **TPC quality:** require that the track passes through enough of the TPC that good particle identification and momentum resolution is attained.
5. **PID:** use TPC2 and the ECal to select electrons.
6. **2nd TPC PID:** if the track also uses TPC3, apply further particle identification.
7. **TPC veto:** require there are no high-momentum tracks in TPC1.
8. **Pair veto:** to reject $\gamma \rightarrow e^+e^-$ conversions, require that there is not an electron-positron pair with a low invariant mass.
9. **PØD veto:** require there are no reconstructed objects in the PØD.
10. **ECal veto:** require there to be no reconstructed objects in the ECal upstream of the lepton candidate.

After the ECal veto cut, the selection is split into two branches: one to increase the purity of CCQE-like events, and one to increase the purity of CCnonQE-like events. Although the two branches are mutually exclusive (so no event passes both sets of cuts), some events will fail both branches. In this way, the overall purity of the selection is improved.

The following cuts are applied to enhance the purity of CCQE-like interactions in the CCQE branch:

11. **No Michel:** require that there are no Michel electron candidates in the spill.
12. **1 FGD track:** the lepton candidate is the only reconstructed track in FGD1.

The following criteria are applied to enhance the purity of CCnonQE-like interactions in the CCnonQE branch of the analysis:

11. **CCnonQE:** there is *either*:
    - a Michel electron candidate *and/or*
    - there is at least one other track starting near the lepton candidate.

Example event displays of events entering the CCQE and CCnonQE branches are shown in Figure 4.1. The final selection used for the analysis is the combination of the CCQE and CCnonQE branches.

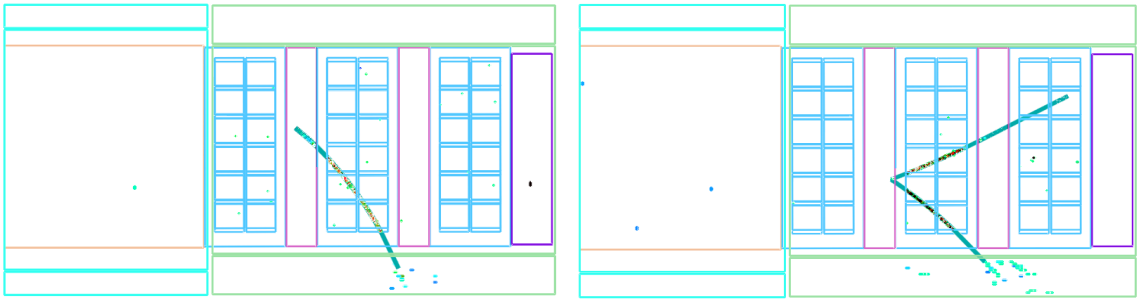


Figure 4.1: Event displays of data events that pass the CCQE (left) and CCnonQE (right) selections.

The selection criteria were developed by studying the NEUT Monte Carlo simulation. After the selection criteria had been chosen, the real data was first analysed to evaluate detector systematic uncertainties, and these are described in more detail in Section 5.1. Once the detector systematic studies had shown that there were no significant issues affecting the selection, the ν_e selection cuts were applied to the data. The rest of this section describes each of the selection cuts in detail, and shows plots comparing the full dataset to the NEUT Monte Carlo prediction.

4.1.1 Selecting CC ν_e interactions

Lepton selection

The first cut applied is to require good data quality. This requires that both the beam and all the ND280 subdetectors were operating correctly. Of a total 6.40×10^{20} POT delivered by the beam, 9.13×10^{17} POT was lost due to bad beam conditions and 4.92×10^{19} POT lost due to faults with ND280, leaving a total of 5.90×10^{20} POT for use in this analysis.

The first cut also requires that only events taken with the beam trigger are used, and tracks must be compatible with one of the 8 bunches (6 bunches for Run 1). Tracks must be within a 60 ns window either side of the mean bunch time for each run.

The second cut selects the highest-momentum negative track that enters the TPC, and requires it to have a momentum greater than 200 MeV/c. The reason for the 200 MeV/c requirement is that there is very little ν_e signal visible below this threshold, and the selection becomes dominated by the $\gamma \to e^+e^-$ background.

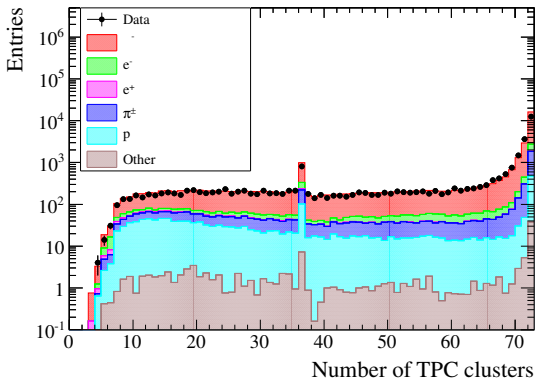


Figure 4.2: Number of TPC clusters for the lepton candidate tracks, after selecting the highest momentum negative track in the bunch and requiring it to start in the FGD FV.

The third cut requires that the lepton candidate starts in the FGD1 Fiducial Volume (FV). The FV is defined such that tracks starting in the first two layers (in z) or the outer five bars (in x and y) are excluded. This requirement removes tracks that truly start outside the FGD but either do not create a reconstructed track upstream of FGD1, or where the reconstructed track is not merged with the lepton candidate.

The fourth cut requires that the TPC2 segment of the track contains at least 36 clusters. A cluster is made of hits in one vertical column of MicroMEGAS pads, and there are 72 columns in each TPC. This cut therefore requires that the track passes through at least half of TPC2. This requirement is imposed as the energy resolution and PID performance of the TPC worsens for shorter tracks, and the TPC PID is critical to the analysis. The distribution of the number of clusters after the first three cuts is shown in Figure 4.2. Most tracks have 72 clusters, indicating that the particle crossed the whole TPC. The slight spike at 36 clusters is caused by tracks that cross an entire MicroMEGAS module (of 36 columns), but are not matched to hits in the module next to it.

The momentum of the lepton candidate after the first four cuts is shown in Figure 4.3. The figure shows a disagreement between the data and the Monte Carlo prediction, even though the plots are normalised to the amount of data POT analysed. This is a consistent feature of current ND280 analyses, with the data showing a deficit with respect to the NEUT Monte Carlo. In the right plot of Figure 4.3, the Monte Carlo is broken down by the main signal and background categories of the analysis. The categories are mutually exclusive, and defined sequentially as

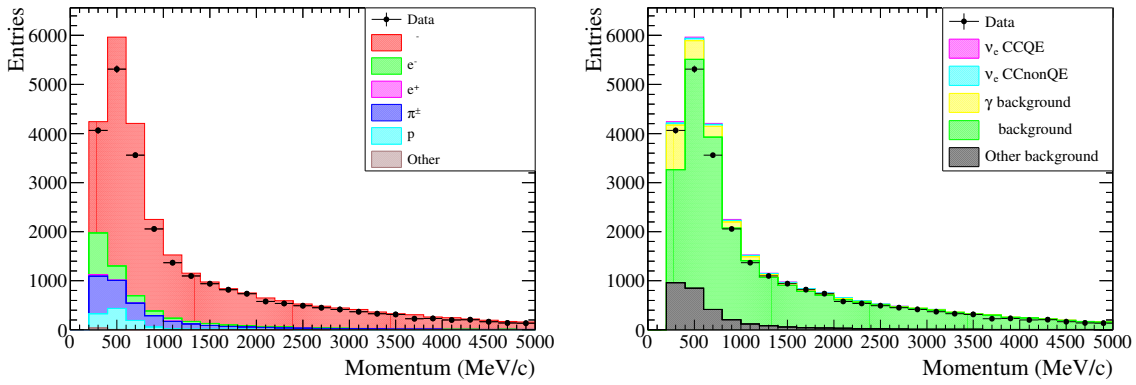


Figure 4.3: Momentum of the lepton candidate tracks, after requiring them to have at least 36 TPC clusters. The left plot shows the Monte Carlo broken down by the true particle type, and the right plot shows it broken down by the interaction type that caused the track.

1. tracks from true ν_e CCQE interactions occurring in the FGD1 FV (defined by the initial interaction, and ignoring any final state interaction effects)
2. tracks from other true CC ν_e interactions occurring in the FGD1 FV
3. true electrons or positrons that start in the FGD1 FV and have a photon as their parent
4. true muons
5. any other track.

Events in the first two categories are the signal for this selection.

Particle identification

To select electron-like tracks, a combination of TPC and ECal PID information is used. Although all the lepton candidates use the FGD and TPC, not all of them may enter the ECal. In these cases, only the TPC PID is used. Also, as the ability of the ECal to discriminate between muons and electrons degrades for low momentum particles, the TPC-only PID criteria are used if the momentum of the track as it enters the ECal is less than 300 MeV/c. Two separate ECal PID variables are used ($R_{\mathrm{MIP/EM}}$ and E_{EM}) depending on the momentum of the particle.

Figure 4.4 shows the momentum distribution of tracks entering the PID cut, broken down by the PID paths that are taken. Most high momentum tracks enter the DsECal,

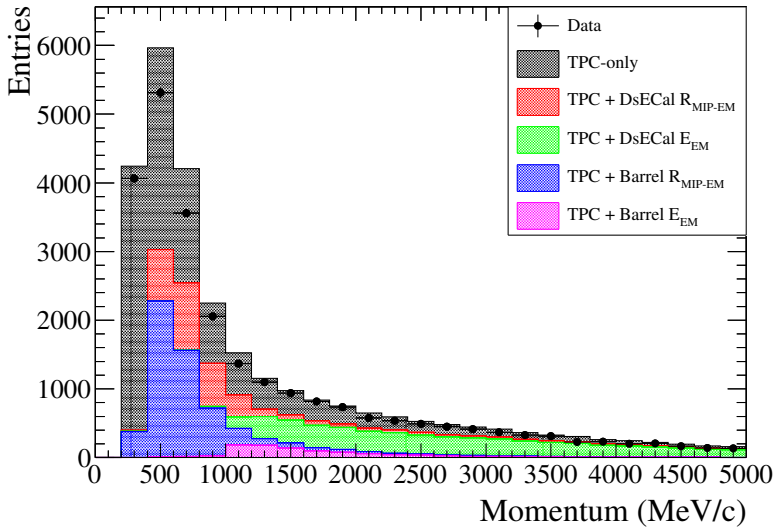


Figure 4.4: The momentum of tracks that reach the PID stage, showing the PID path that each takes.

PID path	Fraction (%)	Efficiency (%)	Purity (%)
TPC-only	45.5	56.6	92.6
TPC + DS $R_{\mathrm{MIP/EM}}$	12.0	87.6	97.6
TPC + DS E_{EM}	20.0	72.1	98.5
TPC + Barrel $R_{\mathrm{MIP/EM}}$	19.0	89.7	90.0
TPC + Barrel E_{EM}	3.6	55.1	86.9

Table 4.2: The overall fraction of tracks that take each PID path, and the efficiency and purity of selecting electrons.

whilst a greater variety of criteria are used for low momentum tracks. Table 4.2 shows the overall fraction of tracks that take each PID path after integrating over all momenta.

Using only the TPCs As mentioned in Section 3.3.1, the TPC PID is based on how the measured dE/dx for a track compares to the energy loss expected for a given true particle hypothesis. For tracks where only TPC information is used for PID, tracks are required to be electron-like, not muon-like, and not pion-like. Specifically, the following cuts are imposed, where δ_α is the pull in TPC2 for particle hypothesis α (as defined in equation (3.1)):

- $-1 < \delta_e < 2$

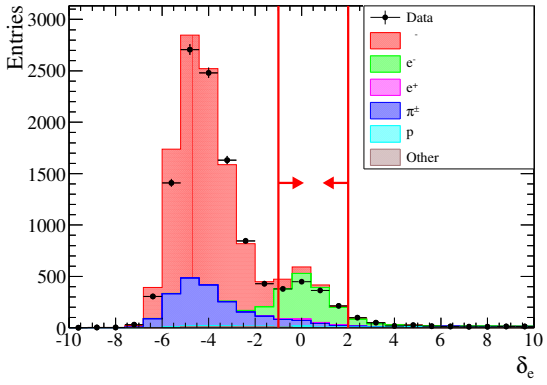


Figure 4.5: Electron pull of tracks reaching the PID cut and where only TPC information is used for particle identification. The red lines and arrows indicate the region that passes the corresponding cut.

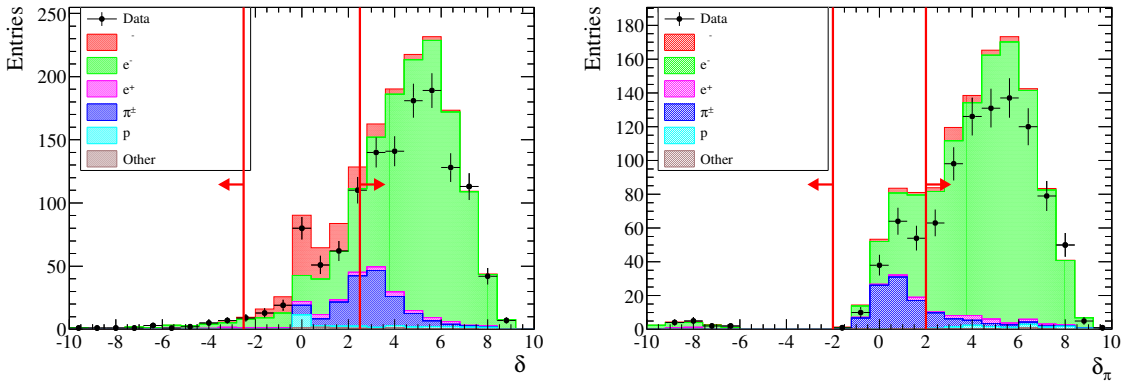


Figure 4.6: TPC pulls for tracks where only TPC information is used for particle identification. The left plot shows the muon pull of tracks that pass the electron pull cut, and the right plot the pion pull of tracks that pass the electron and muon pull cuts.

- $|\delta_\mu| > 2.5$
- $|\delta_\pi| > 2$

These cuts are shown in Figure 4.5 and Figure 4.6.

Using the Downstream ECal For tracks that enter the Downstream ECal, a combination of TPC and ECal PID information is used. In the TPC, the track is required to be electron-like, but a looser cut is used than for tracks that only use TPC information. Two variables are used for the ECal PID: $R_{\mathrm{MIP/EM}}$, a log-likelihood ratio variable for discriminating between MIPs and electromagnetic showers, and the amount of energy deposited in the ECal (E_{EM}). For electromagnetic showers, the amount of energy de-

posited in the ECal should equal the momentum of the particle as it enters the ECal, whilst for MIPs, much less energy will be deposited. E_{EM} was found to perform better for high-momentum tracks, and $R_{\mathrm{MIP/EM}}$ performed better for lower-momentum tracks.

However, E_{EM} can only be used if the shower is well-contained within the ECal. If not, energy will leak out of the side of the ECal, causing the track to appear more MIP-like. To judge whether the object is contained or not, the reconstructed position of the shower is required to be sufficiently far away from the edge of the ECal. If it is not, $R_{\mathrm{MIP/EM}}$ is used rather than E_{EM}. Specifically, the shower is said to be contained if the reconstructed x and y positions are in the range $-900\,\mathrm{mm} < x, y < 900\,\mathrm{mm}$, which excludes the outer 250 mm at each side. There is no requirement on the longitudinal containment of the shower, so E_{EM} is still used for showers that pass through the whole of the ECal, as these still deposit a large amount of energy.

The specific cuts applied are

- $-2 < \delta_e < 2.5$
- $R_{\mathrm{MIP/EM}} > 0$ if the momentum of the track as it enters the ECal is less than $1000\,\mathrm{MeV}/c$, or the shower is not contained
- $E_{\mathrm{EM}} > 1100\,\mathrm{MeV}$ if the momentum of the track is greater than $1000\,\mathrm{MeV}/c$ and the shower is contained. Note that the cut does not require E_{EM} to be similar to the momentum of the track (as would be expected for an electromagnetic shower) as there is no requirement on the longitudinal containment of the shower.

The electron pull distribution for tracks with a DsECal object is shown in Figure 4.7. The $R_{\mathrm{MIP/EM}}$ distribution is also shown in Figure 4.7, while the E_{EM} distribution is shown in Figure 4.8.

Using the Barrel ECal For tracks that enter the Barrel ECal, the same selection criteria are used as for tracks that enter the Downstream ECal. The only difference is the definition of what constitutes a contained shower. For the Barrel ECals, the reconstructed shower position must have $z < 2900\,\mathrm{mm}$, $-900\,\mathrm{mm} < x < 900\,\mathrm{mm}$ (for the top and bottom modules only), and $-900\,\mathrm{mm} < y < 900\,\mathrm{mm}$ (for the left and right side modules only). It should be noted that far fewer tracks enter the Barrel ECal than the Downstream ECal, as the lepton must either have a very low momentum (so it will bend in the magnetic field to reach the Barrel ECal), or be ejected from the nucleus at a high angle, which is kinematically less likely.

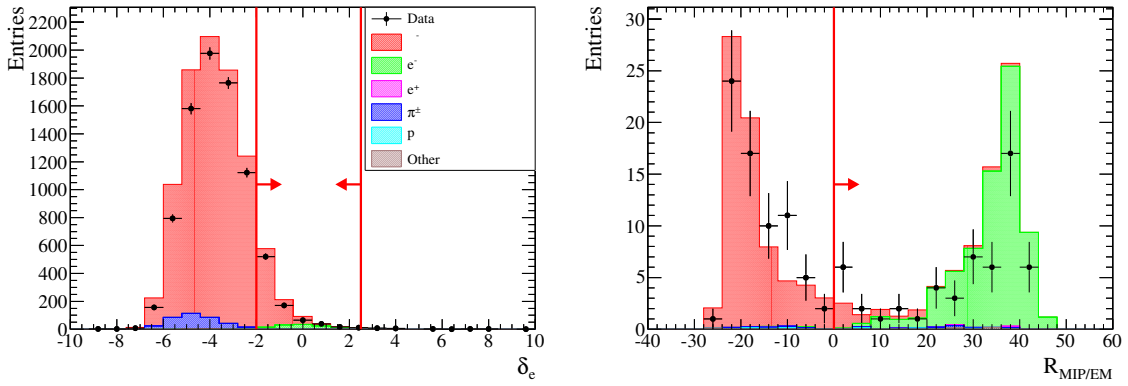


Figure 4.7: The left plot shows the electron pull for tracks where the DsECal is used, and the right plots shows $R_{\mathrm{MIP/EM}}$ for tracks that pass the electron pull cut.

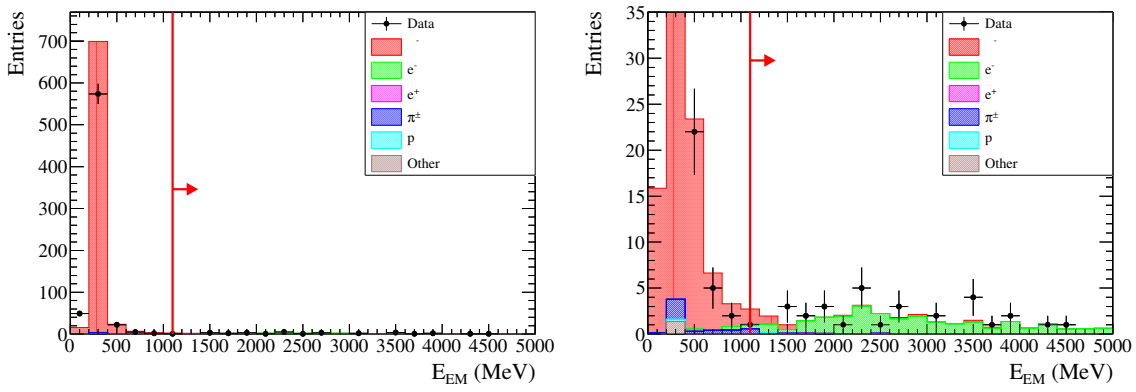


Figure 4.8: E_{EM} for tracks that pass the electron pull cut, and where the DsECal energy estimation is used. The right plot is a zoomed-in version of the left plot.

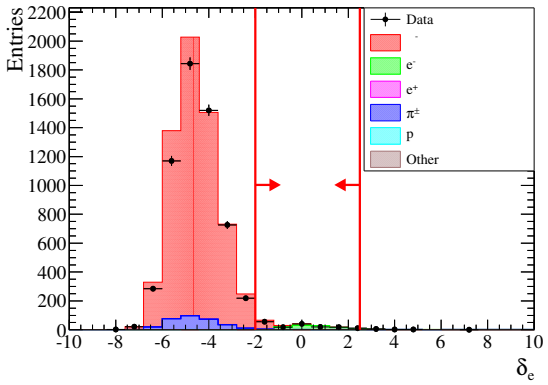


Figure 4.9: Electron pull for tracks where the Barrel ECal is used.

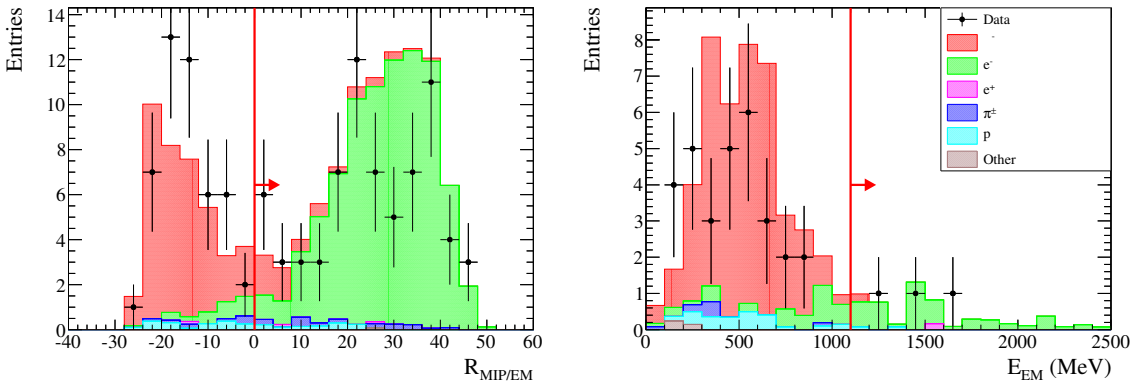


Figure 4.10: PID distributions for tracks that pass the electron pull cut and where the Barrel ECal is used. The left plot shows $R_{\mathrm{MIP/EM}}$ for tracks where the Barrel ECal PID is used, and the right plot shows E_{EM} for tracks where the Barrel ECal energy estimation is used.

The electron pull distribution for tracks with a Barrel ECal object is shown in Figure 4.9 and the Barrel ECal cuts are shown in Figure 4.10.

Overall particle identification Figure 4.11 shows the events that are rejected by the particle identification cuts, and Figure 4.12 shows the events that pass the selection. Over 99.8% of muon events are rejected by the selection, and the resulting sample is 91.5% pure electron events. However, the majority of these electrons come from γ conversions, rather than ν_e interactions, as shown in Table 4.3.

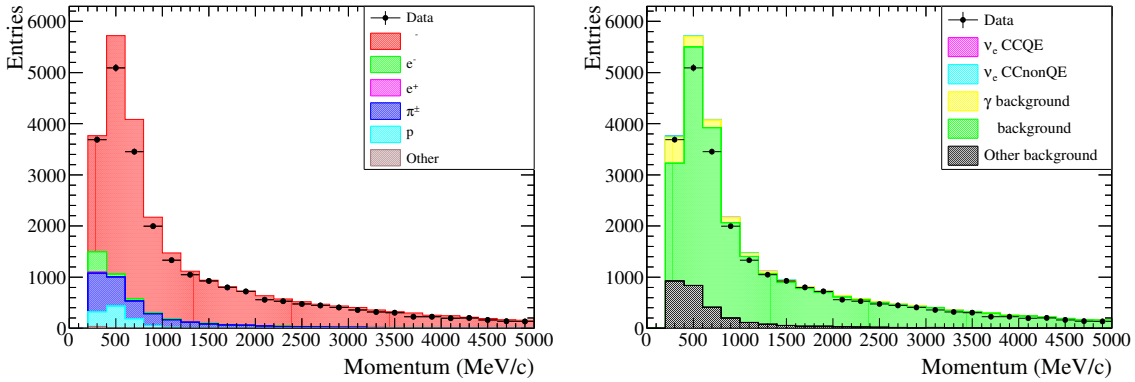


Figure 4.11: Events failing the full PID selection.

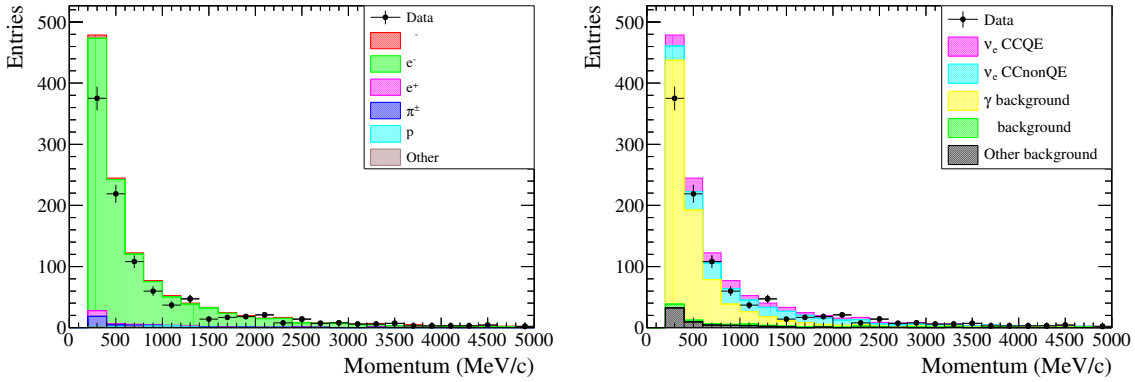


Figure 4.12: Events passing the full PID selection.

Category	Fraction of events (%)	MC expected events
ν_e CCQE	10.3	124.4
ν_e CCnonQE	18.6	225.3
γ background	62.4	754.4
μ background	3.4	41.6
Other	5.2	63.4

Table 4.3: Composition of the ν_e selection after the PID cuts.

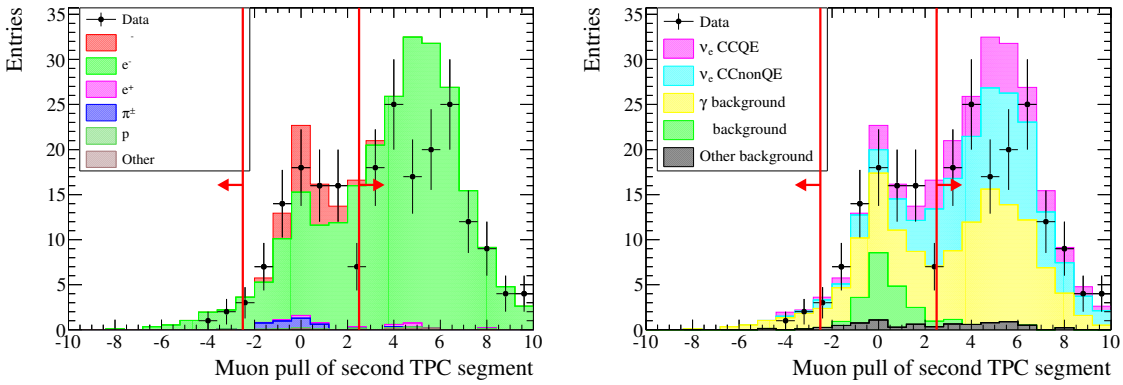


Figure 4.13: Muon pull in TPC3 for tracks that cross two TPCs.

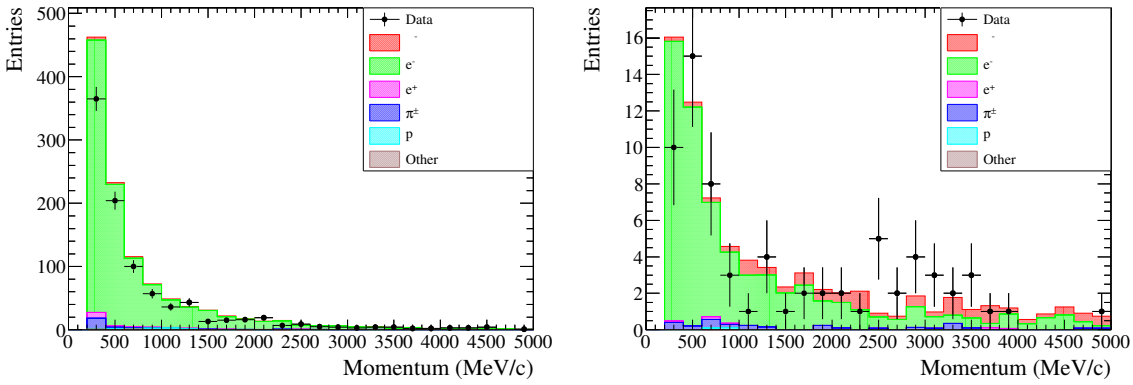


Figure 4.14: Events passing (left) and failing (right) the muon pull cut in TPC3.

Particle identification for tracks using two TPCs

In the previous cut, only information from TPC2 and the ECal modules was used. If the track crosses both TPC2 and TPC3, an extra cut is imposed which requires that it is not muon-like in TPC3. Figure 4.13 shows the δ_μ distribution, and a cut of $|\delta_\mu| > 2.5$ is applied. Figure 4.14 shows the events that pass and fail this cut.

TPC veto

The analysis includes several cuts that attempt to reject events that occurred outside of FGD1, but that produced an electron starting in FGD1. The first of these veto cuts looks for activity in TPC1. The simplest cut would simply require that there are no reconstructed tracks in TPC1, but this would also remove true signal events that have backwards-going particles. This cut therefore looks at the highest-momentum track that

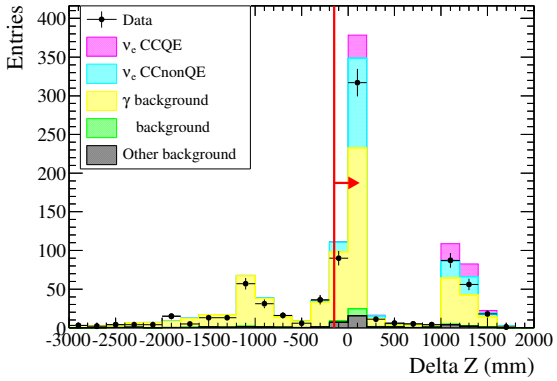


Figure 4.15: The distance in z between the start of the electron candidate track and the start of the highest-momentum track in the events that is not the electron candidate.

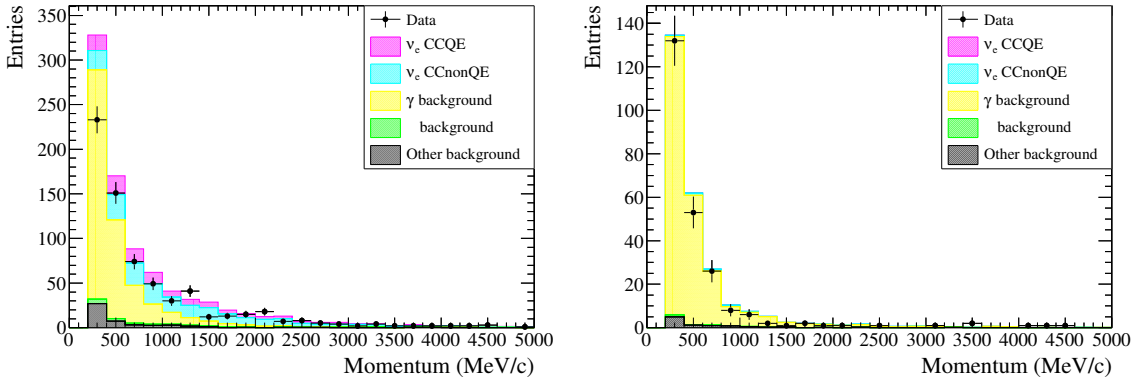


Figure 4.16: Events passing (left) and failing (right) the TPC veto cut.

is not the lepton candidate, and calculates the z distance between the starting position of this track and of the lepton candidate. If the two tracks come from the same vertex, the z distance will be a few centimetres, while if the other track originated in the PØD or in the Barrel ECal this distance will be much larger.

The z distance for the selected events is shown in Figure 4.15, and the events passing and failing the cut of $\Delta z > -150\,\text{mm}$ are shown in Figure 4.16.

Rejection of e^+e^- pairs

The next cut is devoted to the rejection of the electromagnetic background in which both the e^+ and the e^- produced in a γ conversion reach the TPC. Specifically, events are rejected in which there is a second track which

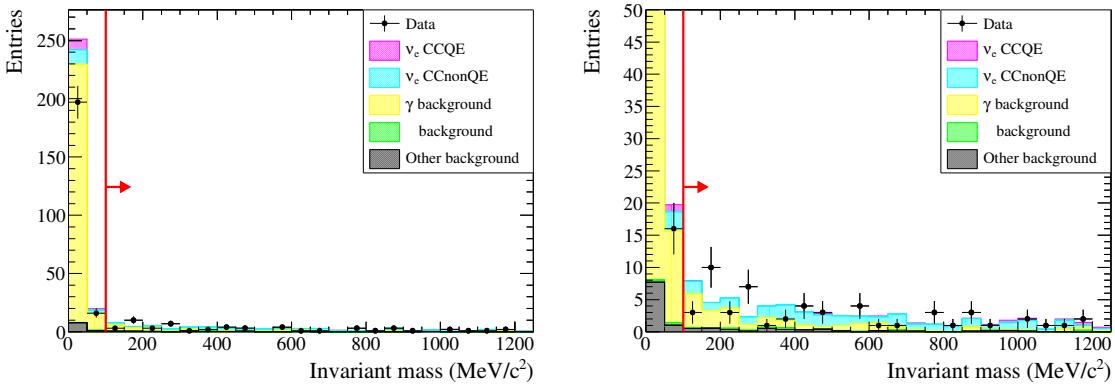


Figure 4.17: Invariant mass distribution for the gamma candidates. The plots are identical except for the vertical axis ranges.

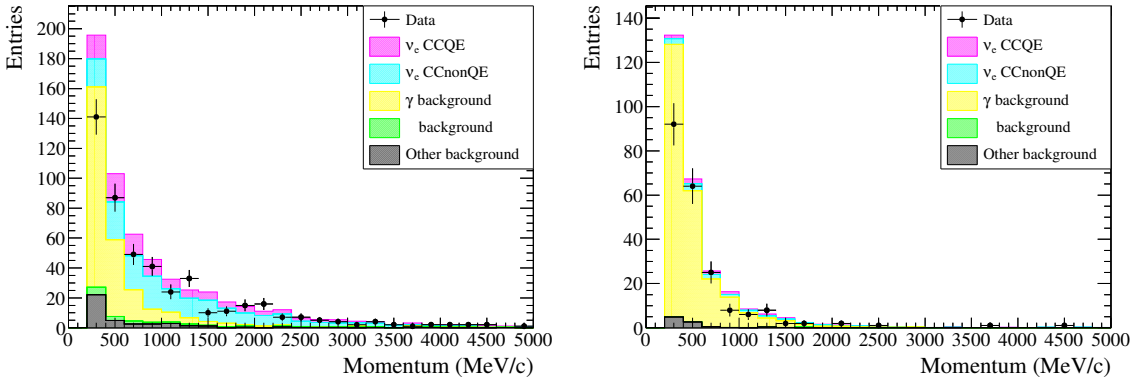


Figure 4.18: Events passing (left) and failing (right) the pair rejection cut.

- starts in the FGD1 FV
- starts within 10 cm of the lepton candidate
- is reconstructed as being positively-charged
- energy loss is compatible with an electron, having $|\delta_e| < 3$.

For events that meet these criteria, the invariant mass of the negative and positive tracks, m_{inv}, is computed assuming both particles have the mass of an electron. The invariant mass distribution is shown in Figure 4.17. Events are rejected if the invariant mass is less than 100 MeV/c^2, and the events passing and failing this cut are shown in Figure 4.18.

It is interesting to note that the pair rejection cut also removes ν_e events as well as the γ background the cut is designed to remove. Examining the positive particle in these

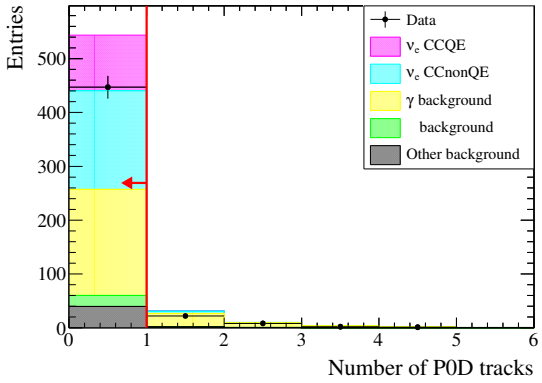


Figure 4.19: The number of reconstructed objects in the PØD in the same bunch as the electron candidate.

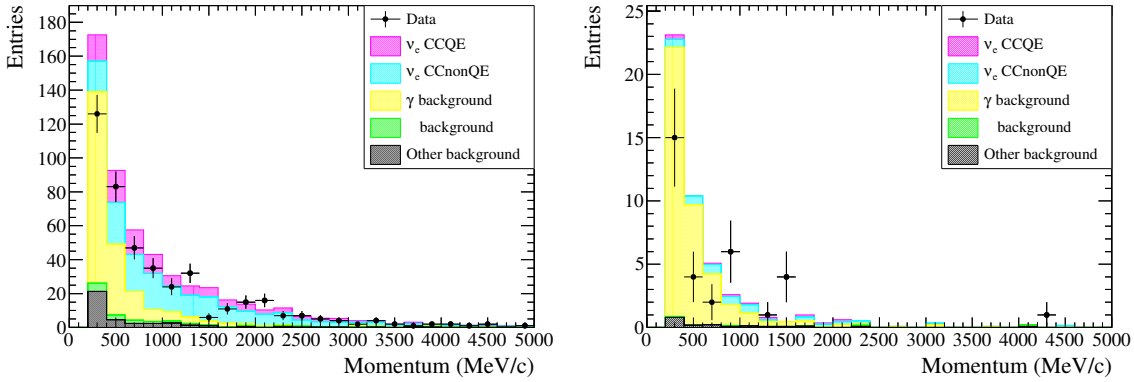


Figure 4.20: Events passing (left) and failing (right) the PØD veto cut.

events shows that they are generally positrons, formed in the chain $e^- \rightarrow \gamma \rightarrow e^+e^-$, where the first e^- is the particle ejected from the nucleus. The conclusion is that these removed events are those in which a bremsstrahlung photon is emitted and converts into an e^+e^- pair in FGD1.

PØD veto

To further reject γ conversions coming from a neutrino interaction in the PØD, there must be no other reconstructed objects in the PØD in the same bunch as the lepton candidate. The distribution of the number of PØD objects is shown in Figure 4.19, and the events passing and failing the cut are shown in Figure 4.20.

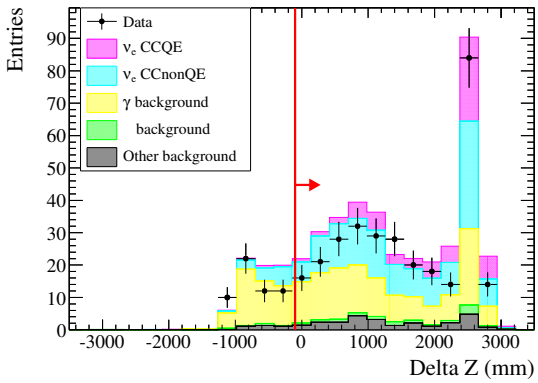


Figure 4.21: The z distance between the start of the electron candidate track and the most downstream ECal object.

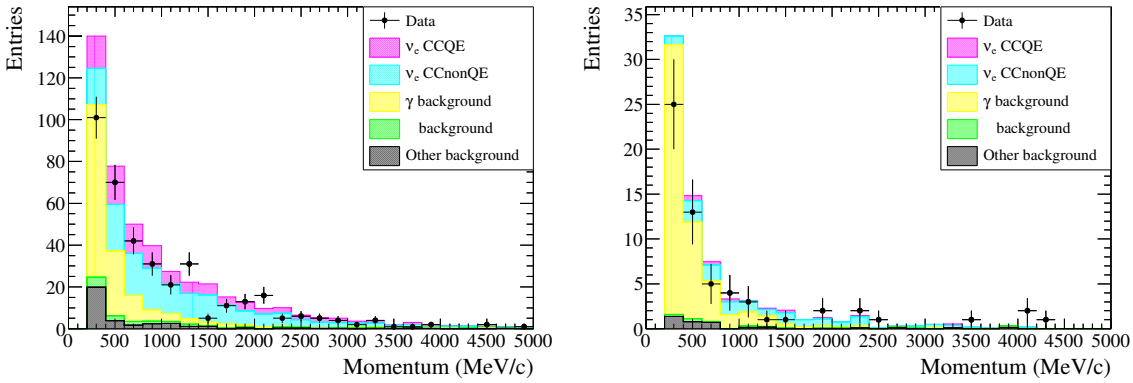


Figure 4.22: Events passing (left) and failing (right) the upstream ECal veto cut.

Upstream ECal veto

To further reject γ conversions coming from a neutrino interaction in the ECal, there must be no reconstructed ECal objects which are in the same bunch and more than 150 mm upstream of the electron candidate. ECal activity downstream of the electron candidate is permitted, as this may come from other particles ejected in the interaction, or as a result of the electron showering. The z distance between the start of the electron candidate and the most upstream ECal object in the same bunch is shown in Figure 4.21. The events passing and failing the cut are shown in Figure 4.22.

The upstream ECal veto is the last cut to be applied before the selection splits into separate CCQE and CCnonQE branches. The purity of the selection at this stage is shown in Table 4.4, and the overall ν_e purity is 57.3%. Photon conversions are still the

Category	Fraction of events (%)	MC expected events
ν_e CCQE	21.4	101.1
ν_e CCnonQE	35.9	168.8
γ background	31.0	145.7
μ background	4.0	18.9
Other	7.7	36.1

Table 4.4: Composition of the ν_e selections after the upstream ECal veto cut (before splitting into CCQE-enhanced and CCnonQE-enhanced branches).

dominant background, but the contamination has dropped significantly, from the 62.4% after the electron PID cut to 31.0% now.

4.1.2 Improving the purity of ν_e CCQE interactions

Electron neutrino CCQE events are characterised by there being no other particles ejected from the nucleus except the electron and a proton. The proton is often low-momentum and is not always reconstructed, so requiring a single reconstructed track is a simple way to select a clean sample of CCQE interactions. For this selection it is required that there are no other tracks in the Tracker, and no Michel electron candidates in the FGD (which would indicate the presence of a charged pion decaying, even if the pion itself is not reconstructed).

The first cut requires that there are no Michel electron candidates in the event, and the number of Michel electron candidates is shown in Figure 4.23. Very few events have a Michel electron, and the majority of those that do are CCnonQE events. The events passing and failing this cut are shown in Figure 4.24.

The second cut requires there are no other Tracker tracks that use FGD1. For FGD-only tracks, only those with a reconstructed $\cos(\theta) > 0.3$ are used, as the reconstruction performance of the FGD worsens at high angles, and systematic uncertainties were only calculated for tracks with $\cos(\theta) > 0.3$. TPC-FGD tracks at any angle are considered. The number of other tracks in the bunch is shown in Figure 4.25, and the events passing and failing this cut are shown in Figure 4.26.

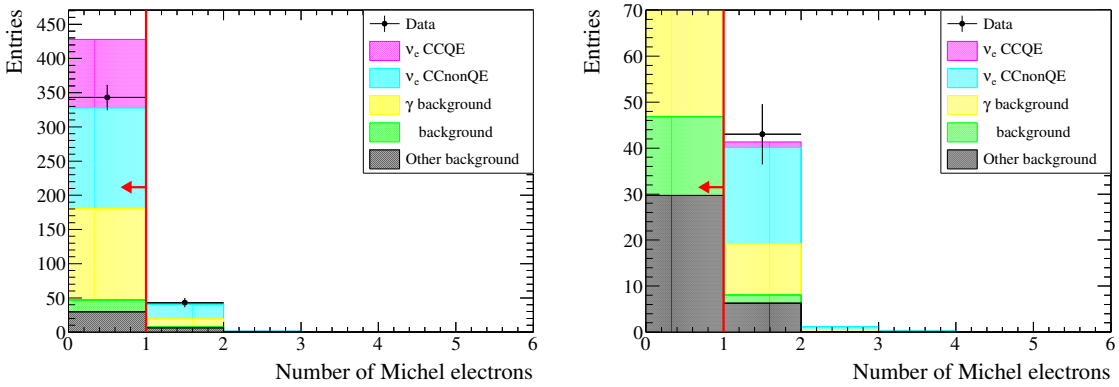


Figure 4.23: The number of Michel electron candidates. The right plot is a zoom of the plot on the left.

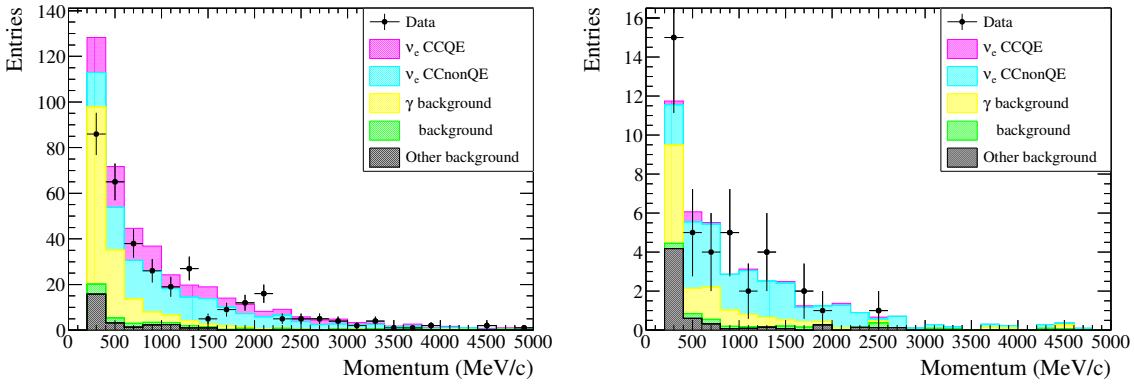


Figure 4.24: Events passing (left) and failing (right) the CCQE cut requiring there to be no Michel electron candidates.

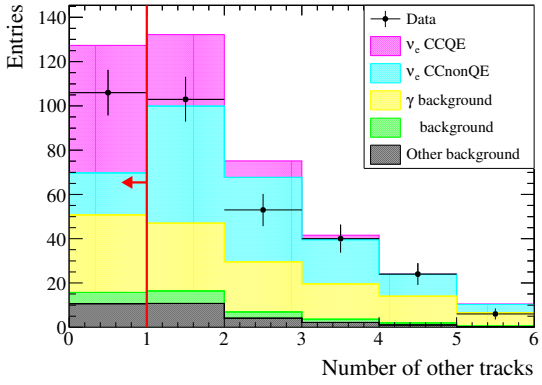


Figure 4.25: The number of other reconstructed tracks that use the same FGD as the electron candidate.

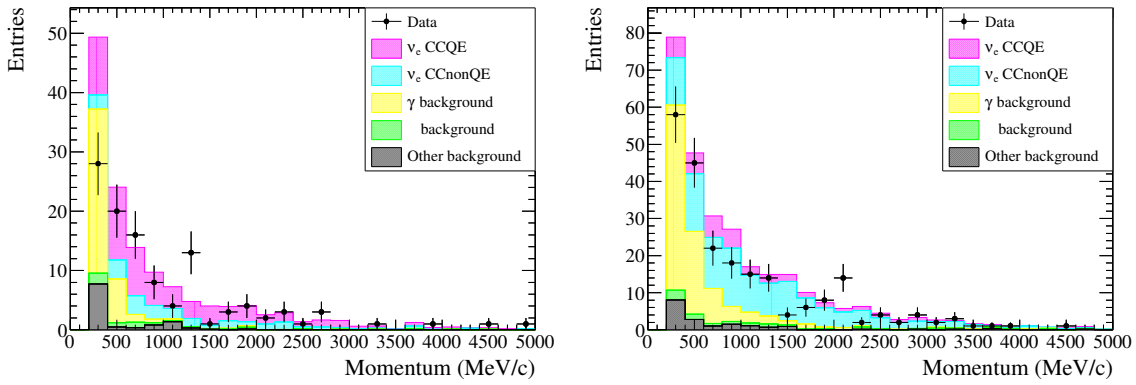


Figure 4.26: Events passing (left) and failing (right) the CCQE cut requiring there to be no other tracks that use the same FGD that the electron candidate started in.

4.1.3 Improving the purity of ν_e CCnonQE interactions

ν_e CCnonQE events are characterised by there being more particles ejected from the nucleus than just one electron and one proton. The selection requires there to either be a Michel electron candidate in the event, or for there to be another track starting near the electron candidate.

The number of Michel electron candidates in the events was shown in the previous section, where it was seen that the majority of events with a Michel electron candidate are ν_e CCnonQE events.

When looking for other tracks starting near the electron candidate, again for the FGD-only tracks, only those with $\cos(\theta) > 0.3$ are considered. TPC-FGD tracks of any angle are considered. The distance between the start of the electron candidate track and the start of the nearest other track is shown in Figure 4.27. Events are selected if there is a track starting within 50 mm of the electron candidate, and the events passing and failing the CCnonQE cuts are shown in Figure 4.28.

4.1.4 Properties of selected events

The CCQE and CCnonQE branches are combined into a CC inclusive sample for this analysis. The momentum and $\cos(\theta)$ distributions of the selected electron candidates are shown in Figure 4.29, and the reconstructed position of the start of the electron track is shown in Figure 4.30 (x and y position) and Figure 4.31 (z position).

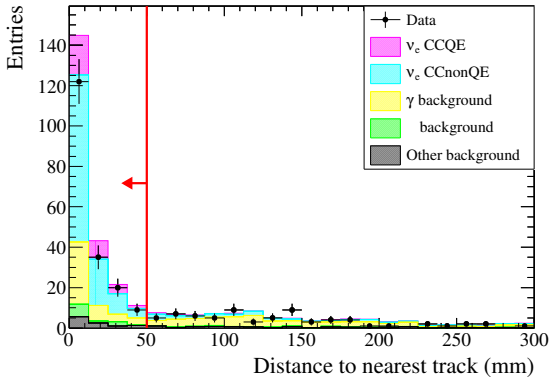


Figure 4.27: The distance between the start of the electron candidate track and the nearest other track.

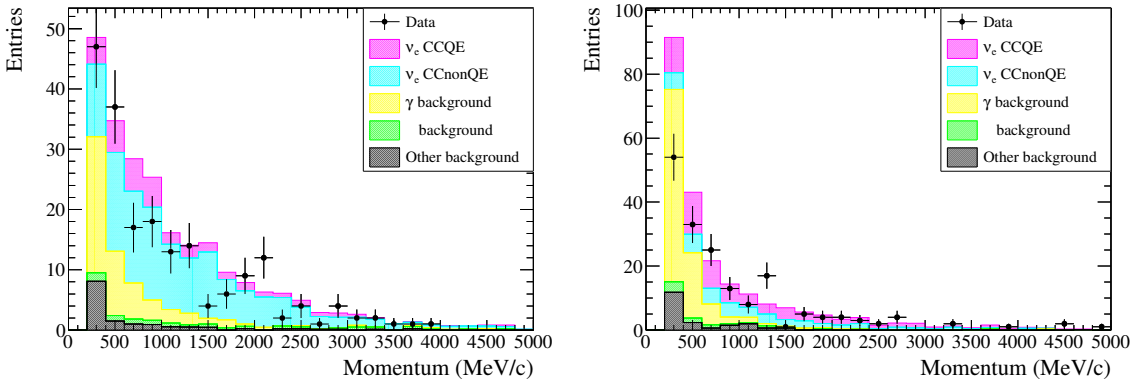


Figure 4.28: Events passing (left) and failing (right) the CCnonQE cuts requiring there to either be a Michel electron candidate or another track starting within 50 mm of the electron candidate.

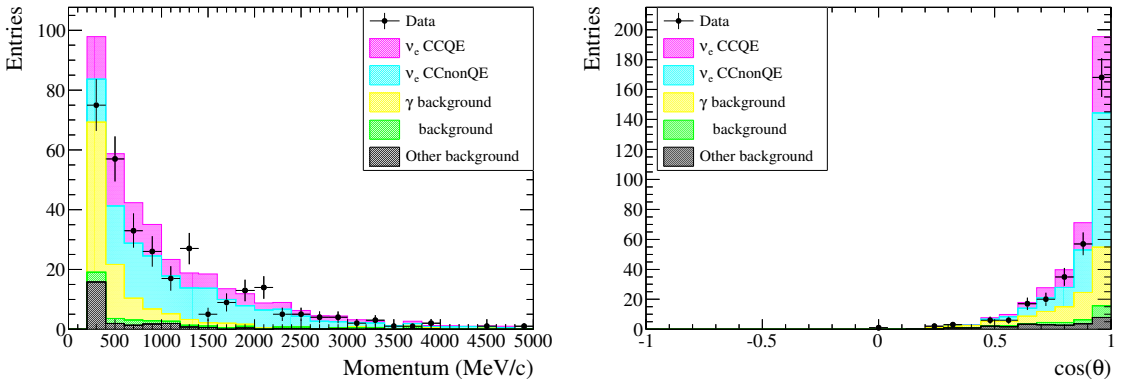


Figure 4.29: Momentum (left) and $\cos(\theta)$ (right) distributions of the final ν_e sample. In these plots, the ν_e component is broken down by the interaction type.

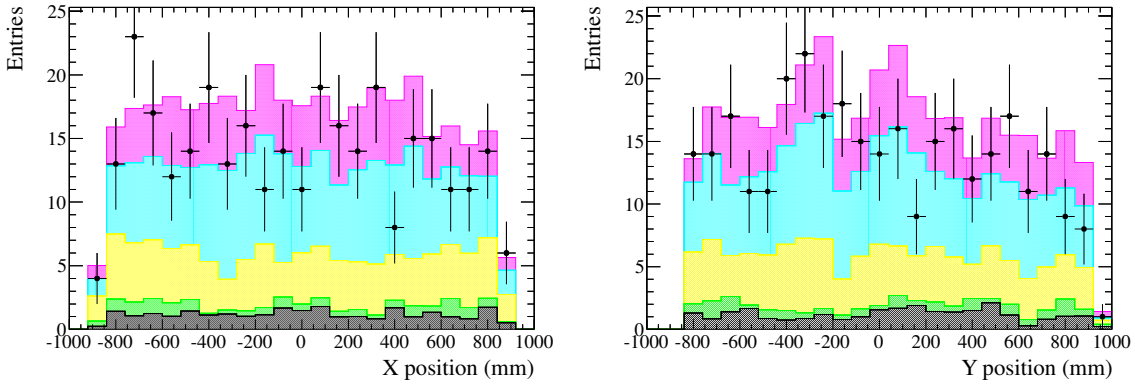


Figure 4.30: Reconstructed x (left) and y (right) position of the start of the electron tracks entering the ν_e selection.

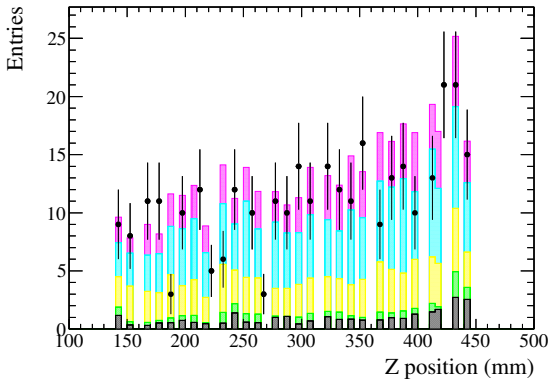


Figure 4.31: Reconstructed z position of the start of the electron tracks entering the ν_e selection.

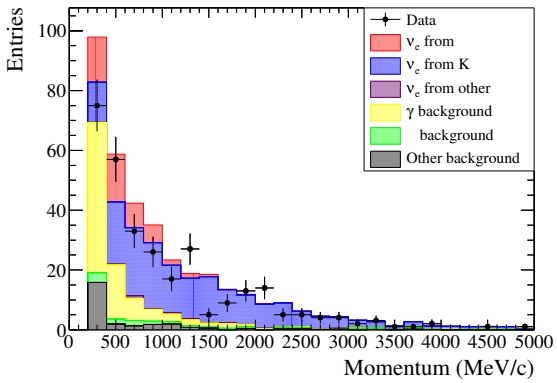


Figure 4.32: Momentum distribution of the final ν_e sample, broken down by the particle that decayed to create the ν_e.

Category	Fraction of events (%)	MC expected events
ν_e CCQE	25.2	95.0
ν_e CCnonQE	39.8	150.0
γ background	23.8	89.8
μ background	4.1	15.5
Other	7.1	26.8

Table 4.5: Composition of the final selected sample. In this table the ν_e component is broken down by interaction type.

The purity of the sample is detailed in Table 4.5. The CCQE branch is 45.1% true ν_e CCQE, and the CCnonQE branch is 54.7% true ν_e CCnonQE. The overall sample is 66.5% pure in ν_e CC interactions, and the breakdown of the interaction modes is shown in Figure 4.6.

The ν_e interactions selected in this analysis mainly come either from a μ decay or from a K decay in the decay tunnel. The former comes from muons produced by the same pion producing the ν_μ beam and mainly populate the low energy region. The latter, instead, directly come from kaons produced by the proton beam interactions with the target and mainly populate the high energy region. Figure 4.32 shows the final selected sample with the ν_e component broken down by the parent particle that decayed to create the electron neutrino, and the purity of the sample is detailed in Table 4.7. Note the very different momentum spectra of particles coming from μ and K decay in Figure 4.32.

Category	Fraction of events (%)	Expected events
CCQE	38.8	95.0
CCRES	32.3	79.1
CCCOH	3.7	9.1
CCDIS	25.2	61.8

Table 4.6: Interaction types of true CC ν_e events selected in the ν_e sample.

Category	Fraction of events (%)	MC expected events
ν_e from μ	13.3	50.3
ν_e from K	50.9	191.8
ν_e from other	0.8	3.0
γ background	23.8	89.8
μ background	4.1	15.5
Other	7.1	26.8

Table 4.7: Composition of the final selected sample. In this table the ν_e component is broken down by particle type of the parent that decayed to create the ν_e.

The overall efficiency of the selection is shown in Figure 4.33 as a function of true neutrino energy, and Figure 4.34 shows the efficiency as a function of true electron momentum and true electron $\cos(\theta)$. The overall efficiency of selecting true CCQE events is 28.9%, and the overall efficiency of selecting true CCnonQE events is 25.4%. The overall efficiency of selecting any ν_e CC interaction is 26.7%.

If only events that have the reconstructed highest momentum negative track starting in the FGD1 FV are considered, then the efficiency of selecting true ν_e CC events is 44.3%.

The efficiency and purity of the selection as each cut is applied is shown in Figure 4.35, and Table 4.8 shows the reduction in the number of events in both data and MC, and shows that the relative number of events surviving each cut agrees very well between the data and MC.

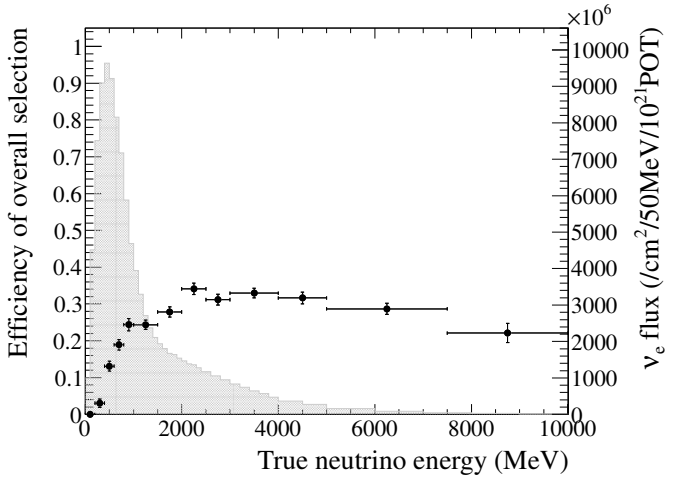


Figure 4.33: Efficiency of selecting true ν_e events as a function of true neutrino energy, with the expected true ν_e flux at ND280 shown for reference.

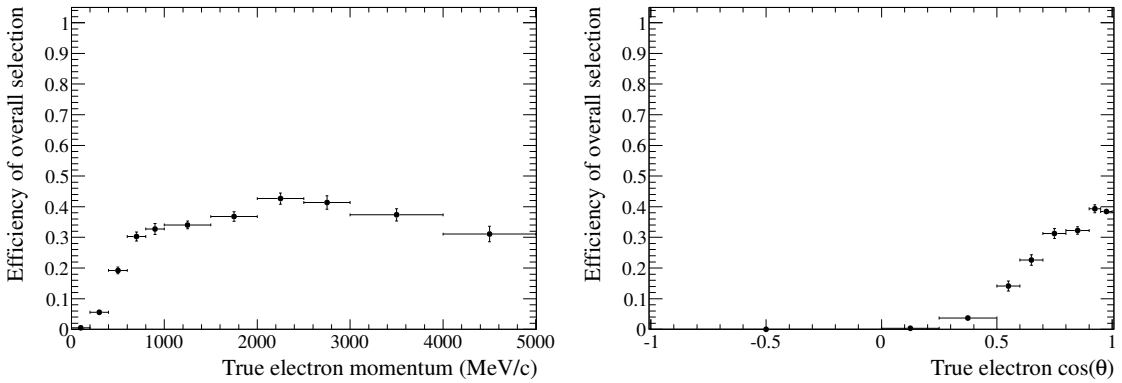


Figure 4.34: Efficiency of selecting true ν_e events as a function of true electron momentum (left) and true electron $\cos(\theta)$ (right).

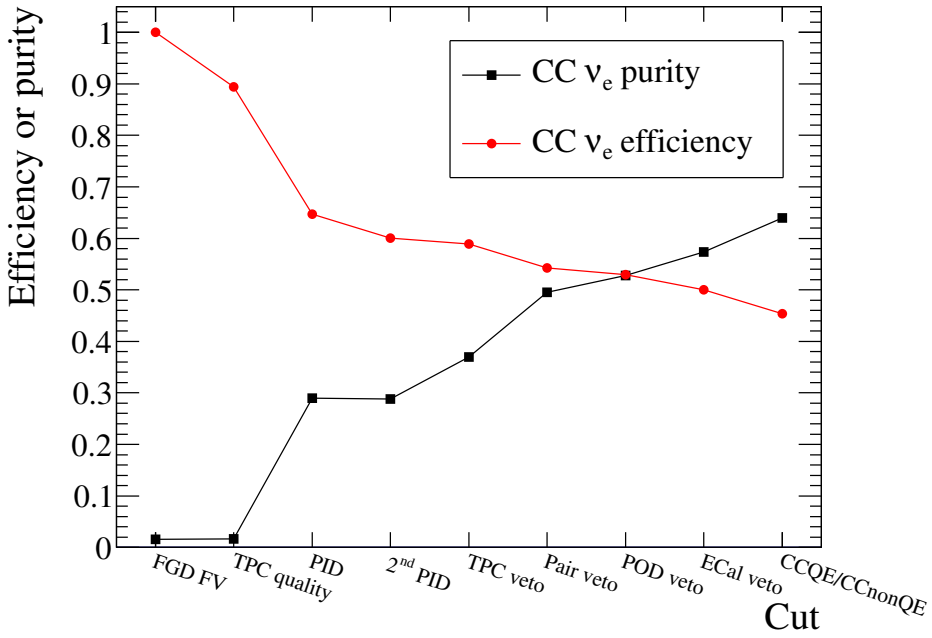


Figure 4.35: Efficiency and purity of selecting true ν_e events as each cut is applied. The efficiency is defined relative to the number of events that have the highest momentum negative track starting in the FGD FV.

	Events		Relative ratio (%)	
Cut	Data	MC	Data	MC
Good quality TPC track, $p > 200\,\mathrm{MeV}/c$	26 231	29 065.6	100.0	100.0
PID	1011	1209.0	3.9	4.2
PID in second TPC	934	1125.8	92.4	93.1
TPC1 veto	693	862.9	74.2	76.7
Pair veto	480	592.0	69.3	68.6
PØD veto	447	543.7	93.1	91.8
Upstream ECal veto	386	470.7	86.4	86.6
CCQE: No Michel electrons	343	427.9	88.9	90.9
CCQE: One track	112	137.7	32.7	32.2
CCnonQE: Michel or other tracks	203	239.4	52.6	50.9

Table 4.8: Reduction in the number of events selected in data and MC. The MC numbers are scaled to the data POT. The Poisson uncertainties on the relative ratios are 4–6% for both the data and MC for all cuts, except for the PID cut (0.1% for both the data and MC).

4.2 γ conversion control sample

The previous section showed that over 20% of the events selected in the ν_e selection come from photon conversions in the FGD. These photons primarily come from the decay of a π^0 produced in a ν_μ interaction outside the FGD. The π^0 decays almost immediately to two photons, which then pass through ND280 and, in some cases, convert in the FGD FV. If the electron produced in the conversion enters the TPC, then the event can enter the ν_e sample. To constrain this background, a control sample of γ conversions is selected. This selection was primarily developed by other members of the ND280 ν_e group [74].

4.2.1 Selecting $\gamma \rightarrow e^+e^-$ conversions

The γ selection is based on identifying e^+e^- pairs, where both the particles enter the TPC. The cuts that are applied are listed below.

- **Event quality:** select events with good data quality and compatible with one of the 8 bunches (6 bunches for Run 1).
- **FGD FV:** require that the highest momentum negative track starts in the FGD1 fiducial volume. This is referred to as the primary track.
- **Secondary track:** require that at least one other track uses the TPC, is positively-charged, and starts in the FGD1 FV.
- **TPC quality:** require the primary track to have more than 18 clusters in the TPC.
- **Secondary proximity:** require that the primary and secondary tracks start within 10 cm of each other.
- **Invariant mass:** require the invariant mass of the primary and secondary tracks to be less than 50 MeV.
- **Electron PID:** require the primary track to be electron-like in TPC2, with $|\delta_e| < 2$.

Figure 4.36 shows a typical event that passes the γ selection criteria.

The fiducial volume and the bunch definitions are the same as used in the ν_e selection, as described in Section 4.1.1. Figure 4.37 shows the number of secondary tracks for events that pass the first two cuts. As a reminder of the event categorisation used for the Monte Carlo, the "γ background" is defined by the true particle of the primary track

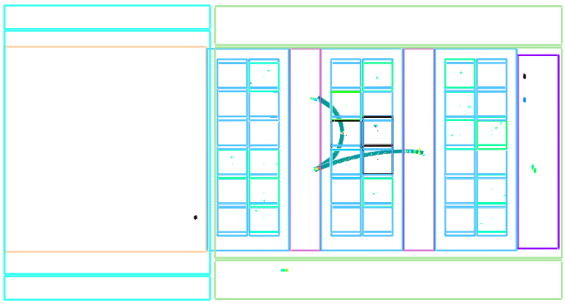


Figure 4.36: Event display of a typical data event that passes the γ selection. The electron candidate starts in FGD1, travels through TPC2 and stops in FGD2. The positron candidate starts in FGD1, travels through TPC2 and curves backwards to stop in FGD1.

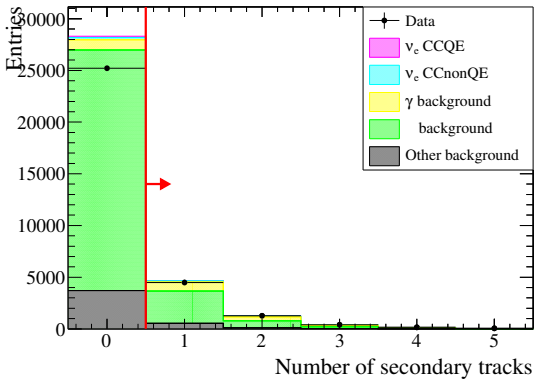


Figure 4.37: Number of secondary tracks for tracks starting in the FGD FV.

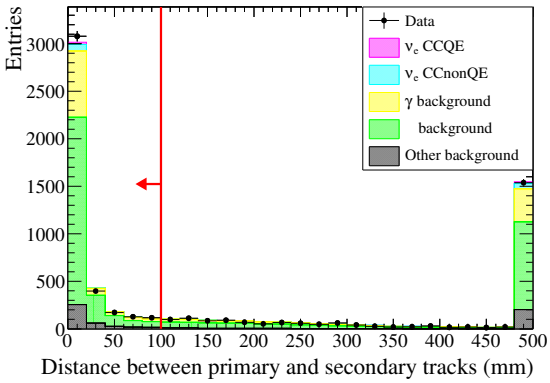


Figure 4.38: Distance between primary and secondary tracks, for tracks passing the TPC quality cut. The last bin is an overflow bin.

being an electron or positron that starts in FGD1 and has a photon as its parent. This category is the signal for this selection.

As TPC PID will later be used to enhance the e^+e^- purity of the sample, the primary track is required to have at least 18 clusters. This is less strict than the 36 clusters requirement of the ν_e selection, as the PID is less critical to selecting a clean sample in this selection.

If the primary and secondary tracks come from the same vertex, then they should be reconstructed as starting near each other. The distance between the two tracks is shown in Figure 4.38, and a cut requiring the tracks to start within 10 cm of each other is imposed.

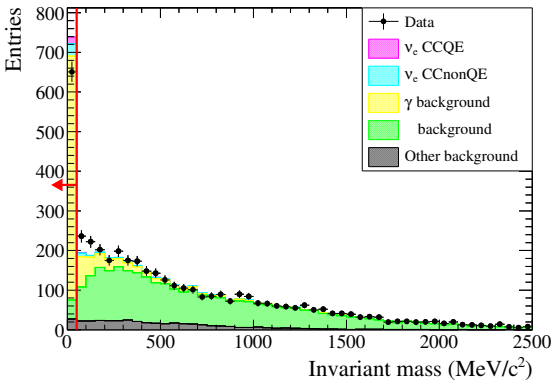


Figure 4.39: Invariant mass distribution of tracks passing the proximity cut.

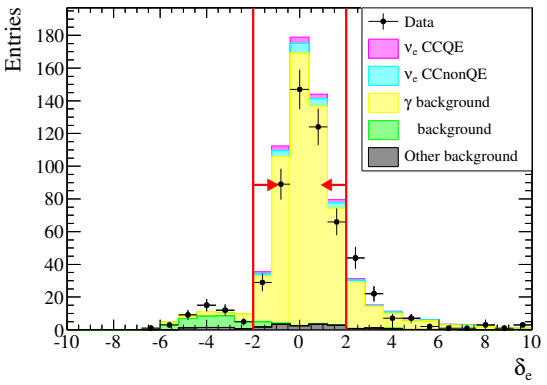


Figure 4.40: Electron pull of the primary track, δ_e, of tracks passing the invariant mass cut.

At this stage, the majority of events contain muons and protons from CC ν_μ interactions, rather than the electrons and positrons that are being searched for. To reject the ν_μ events, the invariant mass of the pair is calculated in the same way as in the ν_e selection, assuming both particles have the mass of an electron. In the ν_e selection, events with a low invariant mass were rejected, whereas in the γ selection they are selected. The invariant mass distribution is shown in Figure 4.39, and by requiring that the invariant mass is less than $50\,\mathrm{MeV}/c^2$ the majority of the muons and protons are rejected.

The e^+e^- purity at this stage is 90%, and this sample is used to measure TPC PID systematics, as it selects a high-purity sample of electrons and positrons without using any PID information. For the γ selection, however, TPC PID information is used to increase the purity further. The TPC electron pull for the primary track, δ_e, is shown in Figure 4.40, and only events with $|\delta_e| < 2$ are selected. In future analyses it may be possible to relax the upper δ_e constraint to improve the selection efficiency. The cut should not be relaxed too far, however, to preserve similar δ_e characteristics in the ν_e and γ samples.

The momentum and angular distributions of the electrons selected in the final γ conversion sample are shown in Figure 4.41. The final electron purity is 98%, and the composition of the sample broken down by reaction type is shown in Table 4.9.

The efficiency and purity of selecting γ conversions in the FGD as each cut is applied is shown in Figure 4.42. For events in which the highest momentum track starts in the FGD FV, the efficiency of selecting γ conversion is 7.8%

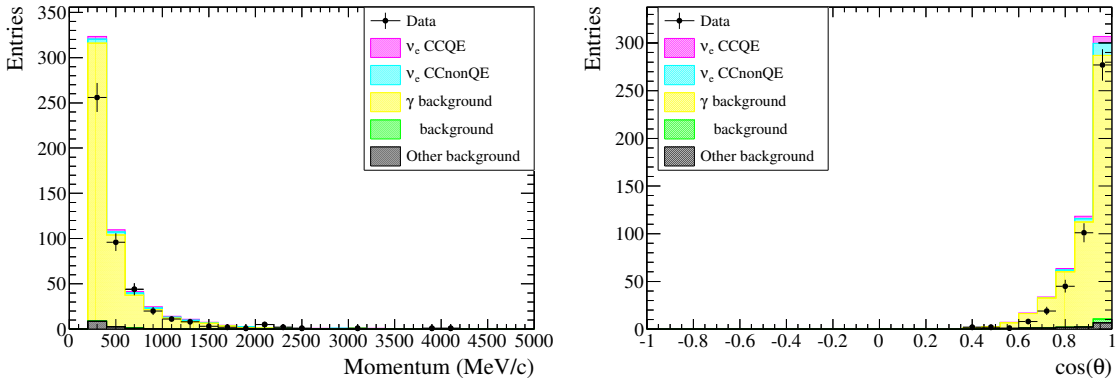


Figure 4.41: Momentum (left) and cos(θ) (right) distributions of the electron selected in the final γ conversion sample.

Category	Fraction of events (%)	MC expected events
ν_e CCQE	2.3	12.9
ν_e CCnonQE	3.4	18.9
γ background	90.9	500.5
μ background	0.8	4.5
Other	2.5	13.9

Table 4.9: Composition of the final γ sample.

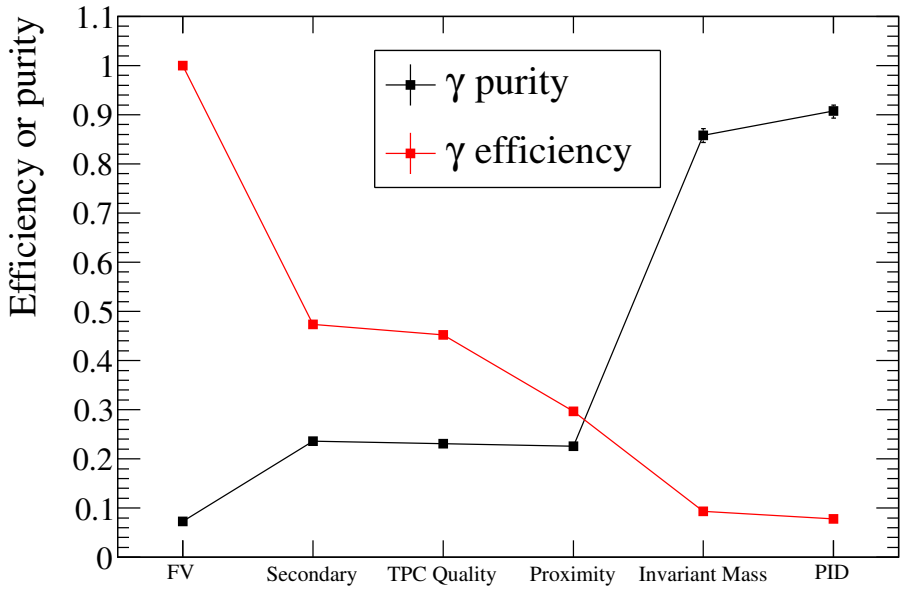


Figure 4.42: Efficiency and purity of selecting γ conversions in the FGD, as each cut is applied. The efficiency is relative to the number of events that have the highest momentum negative track starting in the FGD1 FV.

4.2.2 Comparison of the ν_e and γ samples

The γ sample was designed to constrain the dominant background in the ν_e selection, which is $\gamma \to e^+e^-$ conversions in the FGD. However, the composition of the γ conversions in the ν_e and γ samples are not expected to be exactly the same, as the different selection criteria mean they preferentially select interactions from different sources. The γ sample requires both the e^- and e^+ to enter the TPC and be reconstructed, so is preferentially more likely to select higher-energy and more forwards-going photons. Conversely, the ν_e sample requires the e^+ to not be seen, so is more likely to select higher-angle or lower-momentum photons, where the e^+ is less likely to escape the FGD.

The location of the neutrino interactions that cause γ interactions to be selected in the ν_e and γ samples is shown in Table 4.10. As expected, the γ sample contains more very forwards-going photons from the PØD compared to the ν_e sample. Note that although 10% of interactions occur in TPC1, these are generally interactions in the supporting structure of the TPC, rather than in the gas itself.

Category	γ in ν_e sample	γ in γ sample
FGD1	59.4	30.6
TPC1	9.7	11.6
PØD	13.4	44.9
Barrel ECal	2.1	1.7
PØD ECal	7.6	6.4
Other	7.8	4.8

Table 4.10: Fraction of γ component interactions (in %) occurring in each module, for the ν_e and γ samples.

Category	γ in ν_e sample	γ in γ sample
CCQE	0.9	0.4
CCRES	11.3	7.5
CCDIS	36.0	44.0
NC1π^0	19.2	11.9
NC other	23.9	25.2
Other	8.7	11.0

Table 4.11: Fraction of γ component interactions (in %) caused by different interaction modes, for the ν_e and γ samples.

Table 4.11 shows further evidence of the differences of the γ component of the ν_e and γ samples, and shows the neutrino interaction modes that lead to the γ events.

In summary, although the γ selection provides a clean sample of $\gamma \rightarrow e^+e^-$ conversions in the FGD, it does not select a sample with exactly the same composition as the background it is trying to constrain. Care must therefore be taken when using the γ sample, as will be described in Section 6.4.2.

Chapter 5

Systematic uncertainties

This chapter describes the calculation of systematic uncertainties that affect the analysis. The detector, flux and cross-section systematics are described in this chapter. The effect of propagating these uncertainties to the analysis result is described in Chapter 6, and that chapter also describes the implementation of the data and Monte Carlo statistical uncertainties, and the use of the γ sample to constrain the background in the ν_e sample.

The detector systematics will be described in more detail than the flux and cross-section systematics, as a significant amount of my time was spent implementing code for propagating their effect to the event selection, as well as computing several of the systematics themselves (notably the ECal-TPC matching efficiency systematic and the PØD, ECal and Upstream ECal pile-up corrections). Thanks must go to the ND280 ν_μ group who computed most of the other detector systematics.

5.1 Detector corrections and systematics

ND280 is a complex detector, consisting of many separate sub-detectors. As such, there are a large number of sources of systematic uncertainty. It was found that the nominal Monte Carlo did not match the data for several sources of uncertainty, and in these cases a correction is applied—such that the data and Monte Carlo agree—in addition to computing the systematic uncertainty. In other cases, the Monte Carlo is left as it is, with just a systematic uncertainty computed.

The corrections and systematics mainly fall into two separate categories: weighting systematics and migration systematics. Migration systematics cause events to migrate between momentum bins, between selections, or into and out of a selection entirely.

Weighting systematics do not alter the observables of an event, but re-weight it to have more or less contribution to the selection.

The 26 detector systematics in this analysis are listed in Table 5.1, which also shows the type of each systematic, whether they are also corrections, and which selections they affect. All of the detector systematics are standard T2K systematics that are also used for other analyses [74]. The rest of this section outlines the method used to compute each systematic. For reference, the most important systematics are the TPC momentum resolution, the TPC PID and the pile-up corrections.

5.1.1 TPC systematics

B-field distortion

In an ideal TPC, the electric and magnetic fields are parallel to each other, and the ionisation electrons produced by charged particles drift to the readout plane to create an image of the track. Any distortions in the magnetic field will cause the image to be distorted, affecting the reconstructed momentum of the track.

Distortions of the ND280 magnetic field are measured in two ways.

- **Field correction:** the magnetic field in the Tracker region was measured using a Hall probe before the detectors were installed, and the reconstruction accounts for these measured deviations from the ideal field.
- **Distortion correction:** the central cathode of the TPC has small circles of aluminium covering it, and a laser is shone at these to create photoelectrons which then drift to the readout plane. Comparing the location of the photoelectrons after they have drifted to the readout plane with the known position of the calibration targets allows extra variations in the magnetic field to be observed.

For ND280 data, it was observed that the field correction alone gives a more accurate reconstructed momentum than the combined field and distortion corrections. The ND280 data is therefore processed with the field correction alone, and the effect of the distortion correction is taken as a systematic. The ND280 MC uses a uniform magnetic field, and is processed twice: once with the distortion correction disabled and once with it enabled. The difference between the two reconstructed momenta for each track becomes the systematic uncertainty on the momentum due to B-field distortions.

Group	Systematic	Type	Corr.	Selections	Detail
TPC	B-field distortion	Migrate	No	All	Page 104
	TPC charge confusion	Weight	Yes	All	Page 105
	TPC momentum resolution	Migrate	Yes	All	Page 106
	TPC momentum scale	Migrate	No	All	Page 108
	TPC cluster efficiency	Weight	Yes	All	Page 108
	TPC track efficiency	Weight	Yes	All	Page 108
	TPC PID scale ($e^{\pm}$)	Migrate	Yes	All	Page 109
	TPC PID bias ($e^{\pm}$)	Migrate	Yes	All	Page 109
	TPC PID scale ($\mu^{\pm}$ and $\pi^{\pm}$)	Migrate	Yes	All	Page 109
	TPC PID bias ($\mu^{\pm}$ and $\pi^{\pm}$)	Migrate	Yes	All	Page 109
	TPC PID scale (p)	Migrate	Yes	All	Page 109
	TPC PID bias (p)	Migrate	Yes	All	Page 109
FGD	FGD mass uncertainty	Weight	No	All	Page 111
	FGD track efficiency	Weight	Yes	Not γ	Page 111
	Michel electron efficiency	Weight	Yes	Not γ	Page 113
	TPC-FGD matching efficiency	Weight	Yes	All	Page 113
	Pion secondary interactions	Weight	Yes	All	Page 113
ECal	ECal PID	Migrate	No	All	Page 115
	ECal energy resolution	Migrate	No	All	Page 116
	ECal energy scale	Migrate	No	All	Page 116
	TPC-ECal matching efficiency	Migrate	No	All	Page 116
External	Entering the selection	Other	No	All	Page 118
	TPC1 pile-up	Weight	Yes	Not γ	Page 118
	PØD pile-up	Weight	Yes	Not γ	Page 118
	ECal pile-up	Weight	Yes	CCQE	Page 118
	Upstream ECal pile-up	Weight	Yes	CCnonQE	Page 118

Table 5.1: Detector systematics in this analysis. Unless otherwise noted, each systematic affects the ν_e CCQE, ν_e CCnonQE and γ selections. The "Corr." column indicates whether the systematic also acts as a correction to the nominal Monte Carlo.

TPC charge confusion

The charge confusion systematic is calculated using a sample of tracks that cross all three TPCs. The probability of all the TPCs reconstructing the same charge, P_{same} is

related to the probability of mis-identifying the charge, P_{wrong}, through

$$P_{\text{same}} = (1 - P_{\text{wrong}})^3 + P_{\text{wrong}}^3 \quad (5.1)$$

$$P_{\text{wrong}} = \frac{1}{2}\left(1 - \sqrt{\frac{1}{3}(4P_{\text{same}} - 1)}\right). \quad (5.2)$$

The charge confusion probability, P_{wrong}, can therefore be calculated directly from the number of global tracks in which all three TPC segments have the same charge. The confusion probability is much less than 1% for low-momentum tracks, but increases to 10% for tracks above $5\,\text{GeV}/c$ as these straight tracks are harder to fit a curve to. As an example of the systematic uncertainty, tracks in the $600\,\text{MeV}/c$ to $1000\,\text{MeV}/c$ range have a data−MC correction of $(0.18 \pm 0.20)\%$ applied.

TPC momentum resolution

The momentum resolution of the TPCs is studied by looking at tracks that cross multiple TPCs. The error on the inverse transverse momentum, $1/p_{\text{T}}$, is expected to be Gaussian, and this is what is computed for this systematic. For tracks that cross TPC1 and TPC2, the difference between the two measured $1/p_{\text{T}}$ values is computed, after accounting for the expected energy loss in FGD1. This is referred to as $\Delta 1/p_{\text{T}}$, and the resolution of this as a function of p_{T} is shown in Figure 5.1.

Figure 5.1 clearly shows that the resolution on $1/p_{\text{T}}$ is better in MC than data, and a large smearing factor must be applied to the MC to make the resolutions agree. The smearing factor α is defined such that $\sigma^{\text{data}}_{1/p_T} = (1 + \alpha)\sigma^{\text{MC}}_{1/p_{\text{T}}}$. There is a significant dependence on momentum, and the found values of α in three momentum bins are shown in Table 5.2. The table shows that at high p_T a 31% correction must be made. The uncertainties shown in the table are purely statistical, and are inappropriate to be used as the systematic uncertainty on the correction when the cause of the discrepancy is unknown. Conversely, assigning a correction of 31% with an uncertainty of 31% overestimates the uncertainty. As such, an uncertainty of 0.10 is used for all three momentum bins.

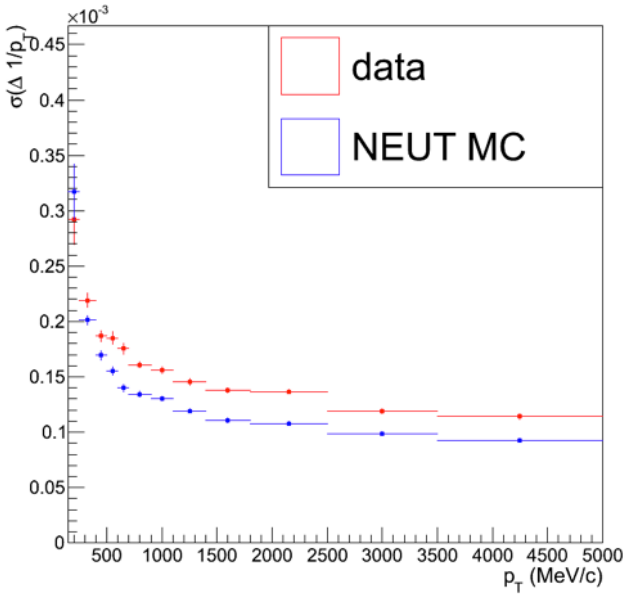


Figure 5.1: Resolution on $\Delta 1/p_{\mathrm{T}}$ as a function of p_{T}. $\Delta 1/p_{\mathrm{T}}$ is the difference between the transverse momenta measured in TPC1 and TPC2, after accounting for the expected energy loss in FGD1. Plot provided by A. Cervera.

p_{T} range (MeV/c)	α
0−500	0.11 ± 0.04
500−1400	0.24 ± 0.02
1400−5000	0.31 ± 0.02

Table 5.2: Smearing factor α to be applied to the MC to make the momentum resolution agree with data. The uncertainties shown are statistical only; a 0.10 uncertainty is actually used for each bin.

TPC momentum scale

Uncertainties on the overall momentum scale of reconstructed tracks arise from uncertainty in the overall magnetic field strength. The magnetic field map study was mentioned in Section 5.1.1 to compute the distortion correction, and uncertainties on this correction include misalignment of the Hall probes, the intrinsic uncertainty of the Hall probes, movement of the mapping device, and conversion of the raw data to B-field values. The overall uncertainty on the momentum scale is 0.57%.

TPC cluster efficiency

The ν_e selection requires TPC tracks to have at least 36 clusters, each consisting of MicroMEGAS pads in the same vertical column. This systematic looks at the efficiency of reconstructing a cluster where one is expected. Separate uncertainties are computed for pads on the edge of MicroMEGAS modules and those in the centre, as the outer pads are affected by edge effects. The extra inefficiencies added to the MC are $0.097 \pm 0.001\%$ for the inner pads and $2.8 \pm 0.2\%$ for the outer pads.

TPC track efficiency

The TPC track-finding efficiency relates to how well a TPC reconstructs particles passing through it. Due to the finite resolution of the TPC MicroMEGAS, the reconstruction efficiency is different for cases where there is a single track or two close tracks.

For the single track efficiency, a sample of through-going muons is used. If the event has objects in TPC1 and TPC3, for example, then it is assumed that a track should have been reconstructed in TPC2. However, this sample only covers a small angular range (the tracks are necessarily very forwards-going and straight) and only considers long tracks (the tracks have to cross the entire TPC). To counter this, a selection of tracks that use a TPC and enter the Barrel ECal are also used. For example, if an event has a track in TPC2 and the Barrel ECal, and the projection of the track crosses TPC3, then the expected number of MicroMEGAS pads crossed can be estimated. The short track efficiency as a function of the number of expected clusters is shown in Figure 5.2. There is no dependency on the number of clusters for tracks with at least 16 expected clusters, and the integrated efficiency over all momenta, angles and lengths is $(99.8^{+0.2}_{-0.4})\%$.

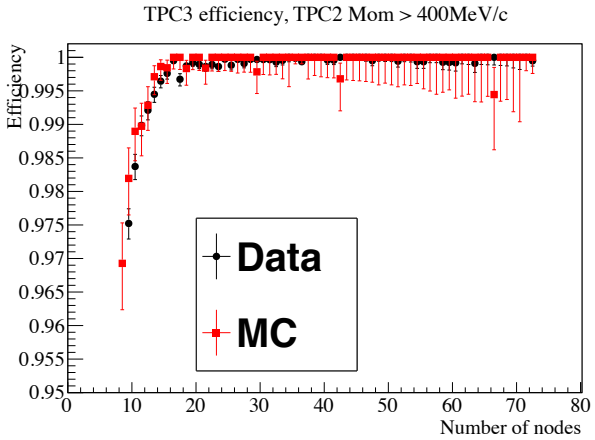


Figure 5.2: TPC3 short track efficiency as a function of the number of expected clusters (nodes). Plot provided by A. Hillairet.

TPC PID

As was explained in Section 3.3.1, the TPC PID is based on measuring the dE/dx of a particle. The expected energy loss depends on β, and is therefore different for particles with different masses but the same momentum. Systematics must be calculated for both the energy resolution and energy scale of the TPCs, and are calculated separately for electrons, muons/pions (which have a similar mass so are treated together) and protons, giving a total of six separate systematics. The general principle for computing the systematics is to select a high-purity sample of a specific particle type without using any TPC PID information, and then examine its dE/dx characteristics. The TPC pull, δ_α, for particle type α should then be a Gaussian distribution of mean 0 and width 1. Any difference in the mean reflects a systematic on the energy scale, and any difference in the width reflects a systematic on the energy resolution.

The electron sample is based on the γ conversion sample detailed in Section 4.2.1, but without the δ_e requirement. Although in principle the systematic uncertainties can vary as a function of momentum, the low statistics in the γ conversion sample mean that all tracks are considered in a single momentum bin. The TPC electron pull is shown in Figure 5.3 for data and MC, and the extracted systematic uncertainties are -0.12 ± 0.12 for the energy bias (where a bias of 0 would be perfect) and 1.02 ± 0.07 for the scale (where a scale of 1 would be perfect).

The muon PID systematics are calculated using a sample of through-going sand muons. Particles that enter through the front face of the PØD and pass through all the TPCs are generally muons, so this provides a high-purity sample without having to use

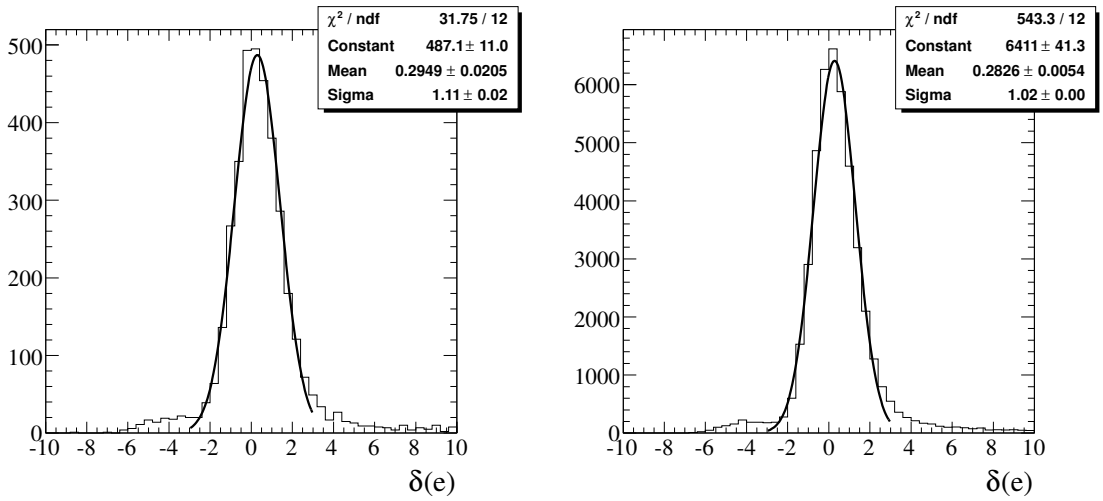


Figure 5.3: TPC electron pull, δ_e, for electron sample in data (left) and MC (right). Plots provided by J. Caravaca.

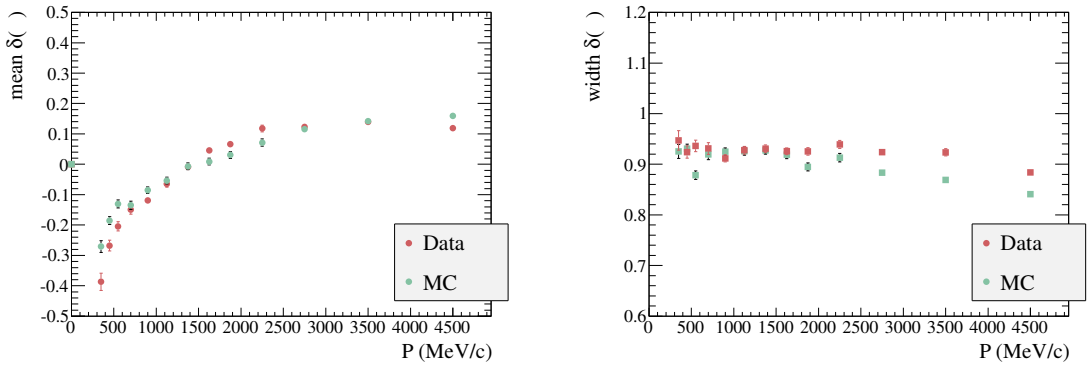


Figure 5.4: TPC muon pull means (left) and widths (right) for muon sample in TPC2 in data and MC. Plots provided by C. Giganti.

the TPC PID. Due to the much larger statistics available, the systematic uncertainties are computed in 13 momentum bins. The measured mean and width of the muon pull in TPC2 are shown in Figure 5.4.

The sample of protons for computing the proton PID systematics is made by selecting the highest-momentum positive track in an event that starts in the FGD FV, and requiring it to have a momentum between 300 MeV/c and 1.1 GeV/c. Tracks below this momentum window tend to be positrons, and the purity decreases above the window. By also requiring the tracks to deposit a large amount of energy in the TPC, a 98% pure sample of protons is selected. The measured mean and width of the proton pull in TPC2 are shown in Figure 5.5.

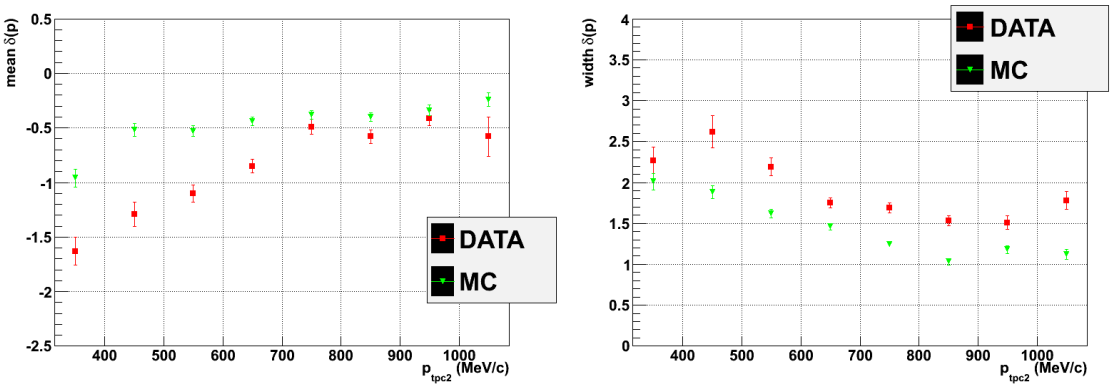


Figure 5.5: TPC proton pull means (left) and widths (right) for proton sample in TPC2 in data and MC. Plots provided by L. Magaletti.

Taking example points from Figure 5.5, the mean of the proton pull in the 400 MeV/c to 500 MeV/c range is -1.29 ± 0.11 in data and -0.52 ± 0.06 in MC. The data−MC difference is 0.77, so the bias correction that must be applied to the MC is 0.77, with an uncertainty of 0.77. The width of the pull distributions are 2.62 ± 0.2 in data and 1.88 ± 0.08 in MC, so the scaling correction that must be applied to the MC is $2.62/1.88 = 1.39$, with an uncertainty of 0.39.

5.1.2 FGD systematics

FGD mass uncertainty

FGD1 is composed of 15 xy modules, and the elemental composition of these modules is shown in Table 5.3. The uncertainties are computed from uncertainties on the measured sizes and composition of the FGD components. The overall expected density of the FGD is 2.147 ± 0.0144 g/cm^2, which is within 0.5% of the measured value of 2.120 g/cm^2.

FGD track efficiency

The FGD-only track efficiency is computed using a sample of proton-like tracks that stop in FGD1. Proton-like tracks that pass through TPC1 and appear to enter the FGD are selected, and if an FGD track is reconstructed within 10 cm of the expected start position, the reconstruction is considered to have been successful.

Element	Areal density (g/cm^2)
Carbon	1.8490 ± 0.0092
Hydrogen	0.1579 ± 0.0021
Oxygen	0.0794 ± 0.0048
Titanium	0.0355 ± 0.0059
Silicon	0.0218 ± 0.0043
Nitrogen	0.0031 ± 0.0012
Total	2.1470 ± 0.0144

Table 5.3: Calculated elemental composition of an FGD xy module. The measured density is $2.120\,g/cm^2$.

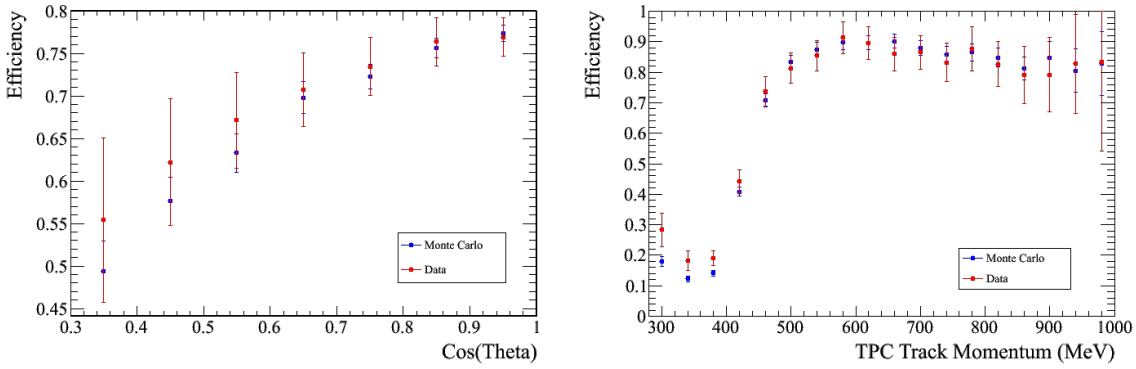


Figure 5.6: FGD-only track reconstruction efficiency for protons stopping in FGD1.

The efficiency with which FGD-only tracks are reconstructed is shown in Figure 5.6 as a function of track angle and momentum. The data/MC ratio is flat above $400\,MeV/c$, but is strongly dependent on the track angle. The systematic uncertainty is therefore computed in bins of $\cos(\theta)$, and as an example the MC/data ratio is 0.99 ± 0.03 in the $0.9 < \cos(\theta) < 1.0$ range, with efficiencies of 77% in both data and MC.

The systematic studied so far only related to single tracks in the FGD. To test the effect of there being multiple tracks, a hybrid sample is used. In this sample, data or MC events in which there is only 1 FGD track have a simulated proton or pion added to them, with the energy and angle of the extra particle chosen at random from uniform and isotropic distributions. The efficiency of reconstructing this extra track is then calculated. A 4.44% data/MC difference is seen in the proton sample, and a 3.32% difference in the pion sample. These values are added in quadrature to the uncertainty found in the single-track study, for events in which the true particle is a proton or a pion.

Michel electron efficiency

The Michel electron tagging efficiency is studied using a sample of cosmic muons that stop in FGD1. Cuts are applied to select a clean sample of muons, free from cosmic electrons. This clean sample of stopping muons should then all produce Michel electrons, and the efficiency with which the FGD finds a Michel electron is computed. The tagging efficiency is found to be 62% in MC and 58% in data, and the systematic uncertainty is taken as 4%..

TPC-FGD matching efficiency

The TPC-FGD matching systematic considers the efficiency of a TPC track being matched to hits in the FGD. A sample of through-going muons that use TPC1 and TPC2 is used to compute the systematic, where the assumption is made that the presence of TPC1 and TPC2 tracks results from a single long track that passes through FGD1. If the event contains a FGD1-TPC2 segment then it is considered to contain a good match.

For tracks with momentum greater than 200 MeV/c, the matching efficiency is very high and uniform, at $99.9^{+0.1}_{-0.2}$%. The efficiency drops sharply below 200 MeV/c, but as these tracks are not used in the ν_e or γ selections, this is not a concern.

It is possible that the TPC-FGD matching efficiency is different for muons and electrons. To check this, the γ selection was altered slightly by not requiring the tracks to start in FGD1, but requiring that the TPC segment points towards the FGD. All the selected tracks had an associated FGD component, so there is no indication of a different TPC-FGD matching efficiency for electrons and muons.

Pion secondary interactions

A pion secondary interaction refers to an interaction a pion undergoes outside the nucleus it was created in. Although these interactions are modelled in Geant4, the model used differs significantly from external data. The external pion scattering measurements cover momenta from 50 MeV/c to 1 GeV/c, and the results are extrapolated to increase the momentum range covered [75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85]. As an example, Figure 5.7 shows the Geant4 π^+ absorption cross section on ^{12}C, along with external data.

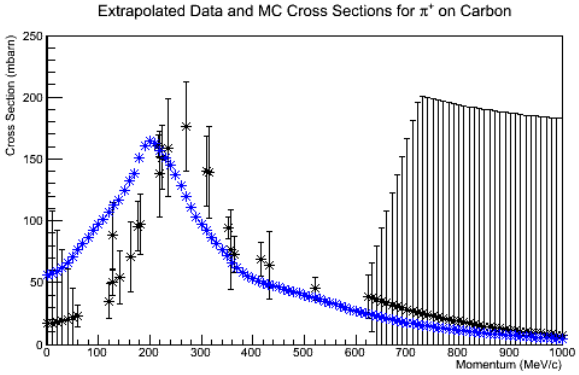


Figure 5.7: π^+ absorption cross-section on ^{12}C. The Geant4 model is shown in blue, and the external data in black [75, 77, 80, 81, 82]. The external data is extrapolated in the $p < 100$ MeV/c and $p > 600$ MeV/c regions. Plot provided by J. Myslik.

The three most significant interactions considered for this systematic are

- **absorption:** the pion is completely absorbed by the nucleus
- **charge exchange:** the pion interacts with the nucleus to create a π^0 and no other pions
- **quasi-elastic scattering:** the pion scatters off the nucleus, and other non-pion particles are ejected.

Absorption and charge exchange can cause the pion to disappear before it can be detected, so events could be mis-categorised as CCQE rather than CCnonQE. Scattering can cause the pion to lose momentum or change direction, which, among other effects, could complicate the reconstruction.

The systematic uncertainty is computed by changing the weight of each event. For each event, all the pions in FGD1 are considered, with the trajectory each pion takes broken down into multiple steps. The event weight is calculated in two components: a correction weight to bring the data and MC cross-sections into agreement, and a variation weight to account for the uncertainty in the data.

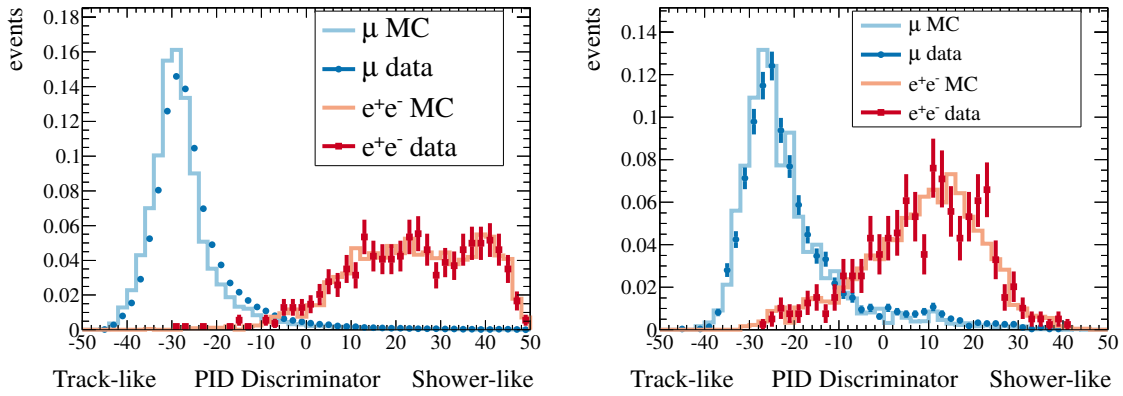


Figure 5.8: $R_{\mathrm{MIP/EM}}$ for electron and muon control samples entering the DsECal (left) and Barrel ECal (right). Plots provided by D. Hadley.

Sample	Data (%)	MC (%)	Data−MC (%)	Systematic (%)
DsECal $e^{\pm}$	87.3 ± 1.3	88.9 ± 0.6	-1.6 ± 1.4	2.1
Barrel ECal $e^{\pm}$	55.1 ± 2.2	56.5 ± 1.1	-1.4 ± 2.5	2.9
DsECal $\mu^{\pm}$	2.35 ± 0.04	2.90 ± 0.04	-0.5 ± 0.1	0.6
Barrel ECal $\mu^{\pm}$	4.2 ± 0.2	4.9 ± 0.2	-0.7 ± 0.3	0.7

Table 5.4: ECal PID systematic uncertainty calculation, from the efficiency of applying $R_{\mathrm{MIP/EM}} > 0$ to the samples shown in Figure 5.8.

5.1.3 ECal systematics

ECal PID

The systematic uncertainty on the ECal PID is calculated using high-purity samples of electrons and muons. The electron sample is based on γ conversions in the FGD, and the muon sample uses through-going muons from upstream interactions. The $R_{\mathrm{MIP/EM}}$ distributions for the two samples are shown in Figure 5.8, for particles entering the DsECal and Barrel ECal.

By comparing the efficiency of cutting at $R_{\mathrm{MIP/EM}} > 0$ in data and MC, the systematic uncertainties shown in Table 5.4 are found, where the data−MC difference is added in quadrature with the statistical uncertainty.

ECal energy resolution and scale

To test the reconstructed energy performance of the ECals, two samples of electrons are used. For low-energy electrons, the same sample as used for evaluating the ECal PID systematics is used. For higher-energy electrons this sample does not have sufficient statistics, so testbeam data taken with the DsECal at CERN is used [86]. The systematics are computed by comparing the energy reconstructed in the ECal, E_{EM}, to a known momentum measurement, p_{known}. For the low-energy sample, the momentum at the end of the TPC track (just before it enters the ECal) is used, while for the testbeam sample, the momentum of the electrons is known.

The fractional difference between E_{EM} and p_{known},

$$\Delta E = \frac{E_{\mathrm{EM}} - p_{\mathrm{known}}}{p_{\mathrm{known}}}, \tag{5.3}$$

should follow a Gaussian distribution, and any data−MC difference between the mean of the ΔE distribution represents a systematic on the energy scale, and any difference on the width represents a systematic on the energy resolution.

The largest systematic uncertainties are found at the higher energies of the testbeam data, so to be conservative these values are taken as the systematic. The systematic uncertainty is 6% for the energy scale and 15% for the energy resolution.

TPC-ECal matching and ECal reconstruction efficiency

A combined systematic uncertainty is applied to the efficiency of reconstructing an ECal object and matching it to a TPC track. The efficiency is defined as

$$\text{Efficiency} = \frac{\text{Number of tracks entering the ECal with an ECal object attached}}{\text{Total number of tracks entering the ECal}} \tag{5.4}$$

Samples of electrons that appear to enter an ECal module are used to calculate this systematic, with separate selections for the DsECal and Barrel ECal. The tracks are required to start in the FGD FV, be electron-like in the TPC, and appear to be entering either the DsECal or Barrel ECal, based on the end position and direction of the TPC tracks. The purity of the sample decreases at higher momentum, so only tracks below 800 MeV/c are considered. Tracks must also be above 300 MeV/c, as the ν_e analysis does not use ECal information for lower-momentum tracks. The selected samples are shown

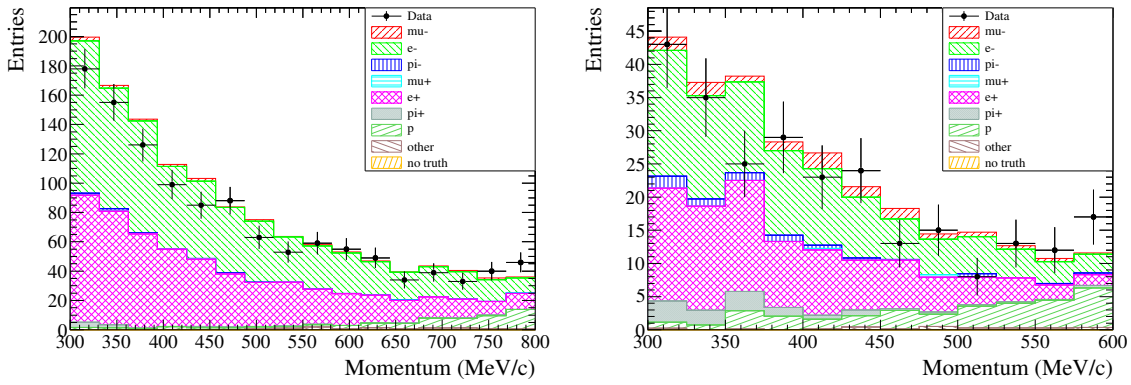


Figure 5.9: Tracks selected as being electron-like according to the TPC PID, originate in the FGDs, and that appear to enter the Downstream ECal (left) and the Barrel ECal (right).

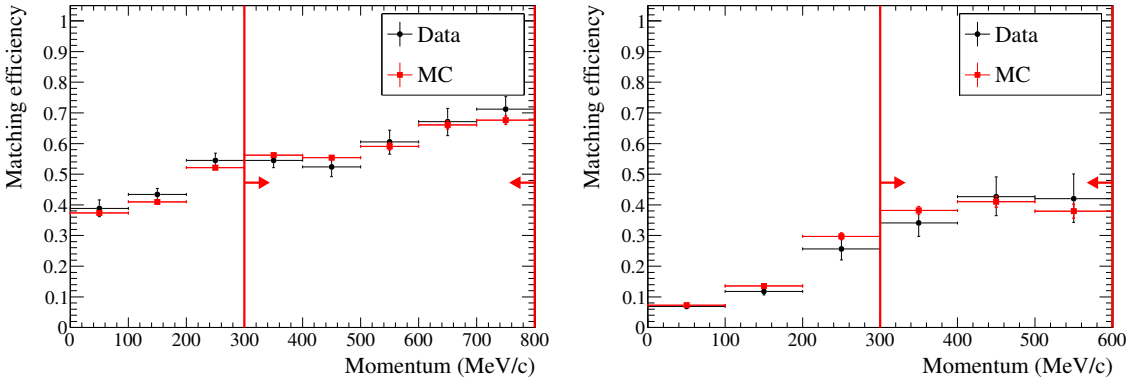


Figure 5.10: Efficiency of matching to an ECal object as a function of the track's momentum for the Downstream ECal (left) and Barrel ECal (right). Only the region indicated by the cut lines is used for computing the systematic.

in Figure 5.9, and the overall purity of selecting electrons and positrons is 92% for the DsECal, and 75% for the Barrel ECal.

The efficiency of matching to an ECal object is shown in Figure 5.10 as a function of the track's momentum, and the integrated efficiency for data and MC is shown in Table 5.5. The systematic uncertainty is calculated by adding in quadrature the data−MC difference and the statistical uncertainty on the efficiency. The systematic is calculated to be 1.6% for tracks entering the DsECal, and 3.4% for tracks entering the Barrel ECal.

Module	Downstream	Barrel
Momentum range	300–800 MeV/c	300–600 MeV/c
Data efficiency	0.582 ± 0.015	0.381 ± 0.032
MC efficiency	0.586 ± 0.004	0.389 ± 0.009
Data−MC	-0.004 ± 0.015	-0.008 ± 0.033
Systematic (%)	1.6	3.4

Table 5.5: Combined systematic uncertainty for reconstructing an ECal object and matching it to a TPC track.

5.1.4 External systematics

The standard ND280 Monte Carlo only models neutrino interactions occurring in the detector itself. Interactions which occur outside the magnet, but that then produce particles which enter ND280, are accounted for using a separate Monte Carlo sample. These particles—referred to as sand muons—have two possible impacts: they can either be mis-identified as ν_e events, or can trigger one of the veto cuts in the analysis (TPC1, PØD, Upstream ECal or ECal vetoes).

Sand muons entering the selection

The sand muons themselves should not enter the selection, as for them to be selected would require the reconstruction to break the muon track in the FGD FV, and then mis-identify the track as an electron. However, secondary particles from interactions of the muon could cause a sand muon event to pass the selection.

When running the analysis on the sand muon Monte Carlo, no events entered the CCnonQE sample, and 1.7 events are expected in the CCQE sample. The effect of sand muons directly entering the analysis is negligible when compared to the rest of the detector systematics.

Pile-up

Sand muons can also trigger one of the veto cuts, and this is not simulated in the Monte Carlo. As the sand muon events are independent of the ν_e interactions, a correction can be applied to re-weight each event to account for this possible pile-up.

The correction is computed separately for each run, as the pile-up depends on the beam intensity. The intensity for each run, I_d, is derived from the data as $I_d = \mathrm{POT}/N_{\mathrm{spills}}$, where N_{spills} is the number of spills in that run. The sand muon MC is then used to calculate the pile-up correction for a given veto cut, x, as

$$C^{x}_{\text{pile-up}} = \frac{N_x \times I_d}{\mathrm{POT}_{\mathrm{sand}} \times N_b}, \tag{5.5}$$

where N_x is the number of bunches in the sand muon MC that trigger the veto cut, $\mathrm{POT}_{\mathrm{sand}}$ is the POT in the sand muon MC, and N_b is the number of bunches per spill (6 for Run 1, 8 otherwise).

As four separate veto cuts are applied, care must be taken to not double-count the pile-up probabilities. For example, tracks that would trigger both the TPC1 and PØD vetoes should not be counted twice. The correction for the TPC1 veto is calculated by selecting events that would trigger the TPC1 veto. For the PØD veto, only events that trigger the PØD veto and do not trigger the TPC1 veto are selected. For the ECal veto, two numbers are derived:

- ECal: event triggers the ECal veto, but not the PØD or TPC1 vetoes
- ECal, $z <$ FGD1: event has an ECal object that is more than 150 mm upstream of the middle of FGD1 ($z < 142$ mm), and does not trigger the PØD or TPC1 vetoes

The purpose of the last number is to find an average pile-up correction that should be applied to events that pass the CCnonQE selection, as the final cut in that selection requires there to not be an ECal object more than 150 mm upstream of where the electron candidate starts.

The overall pile-up correction for each event depends on which sample the event enters, due to the different veto cuts applied in each selection. If the existing MC weight is defined as w_{MC}, then the new weights due to pile-up are

$$w^{\gamma}_{\mathrm{MC}} = w_{\mathrm{MC}} \tag{5.6}$$

$$w^{\mathrm{CCQE}}_{\mathrm{MC}} = w_{\mathrm{MC}} \times (1 - C^{\mathrm{TPC1}}_{\text{pile-up}}) \times (1 - C^{\mathrm{P\emptyset D}}_{\text{pile-up}}) \times (1 - C^{\mathrm{ECal}}_{\text{pile-up}}) \tag{5.7}$$

$$w^{\mathrm{CCnonQE}}_{\mathrm{MC}} = w_{\mathrm{MC}} \times (1 - C^{\mathrm{TPC1}}_{\text{pile-up}}) \times (1 - C^{\mathrm{P\emptyset D}}_{\text{pile-up}}) \times (1 - C^{\mathrm{ECal},z<\mathrm{FGD1}}_{\text{pile-up}}) \tag{5.8}$$

The uncertainty on this systematic is evaluated by comparing N_x in data and MC. The MC is weighted to the data intensity, and the sum of the beam and sand MC is then

Sample	$C^{\text{TPC}}_{\text{pile-up}}$	$C^{\text{PØD}}_{\text{pile-up}}$	$C^{\text{ECal}}_{\text{pile-up}}$	$C^{\text{ECal},z<\text{FGD1}}_{\text{pile-up}}$
Run 1 (water)	0.00510(086)	0.00072(012)	0.00861(679)	N/A
Run 2 (water)	0.00800(110)	0.00117(031)	0.01404(666)	0.00737(145)
Run 2 (air)	0.00990(140)	0.00155(017)	0.01671(610)	0.00887(121)
Run 3b (air)	0.00960(096)	0.00149(119)	0.01608(672)	0.00853(297)
Run 3c (air)	0.01100(150)	0.00168(017)	0.01811(440)	0.00961(096)
Run 4 (water)	0.01200(160)	0.00166(030)	0.01984(480)	0.01042(104)
Run 4 (air)	0.01300(130)	0.00215(054)	0.02313(503)	0.01227(228)

Table 5.6: Pile-up correction (and uncertainty) for each veto cut for each run period. For example, 0.01227(228) means 0.01227 ± 0.00228. There is no pile-up systematic for the Barrel ECal in Run 1, as the Barrel ECal modules were installed between Runs 1 and 2.

compared to the data. The data−MC difference is taken as the uncertainty. As there is a 10% normalisation uncertainty on the sand muon MC, the uncertainty is taken as $0.1 \times C^{x}_{\text{pile-up}}$ if the data−MC difference is less than 10% of $C^{x}_{\text{pile-up}}$.

The extracted values of $C^{x}_{\text{pile-up}}$ are shown in Table 5.6. As expected, the pile-up corrections increase as the beam power increases, with the largest corrections for data taken at the end of Run 4. The largest correction comes from the ECal pile-up, and is 2.3%.

5.2 Flux and cross-section systematics

5.2.1 The beam and ND280 flux task force (BANFF) fit

The nominal NEUT model is not used in T2K oscillation analyses. Instead, the "beam and ND280 flux task force" (BANFF group) uses external and ND280 data to tweak parameters describing the flux and cross-section model. These tweaks are applied using the T2KReWeight tool [87], which gives each MC event a weight for a given set of parameter tweaks. This approach allows the model to be varied without re-running the full Monte Carlo processing chain.

The flux and cross-section systematic parameters that are varied are:

- ND280 flux (11 ν_μ energy bins, 5 $\overline{\nu}_\mu$ bins, 7 ν_e bins, 2 $\overline{\nu}_e$ bins)

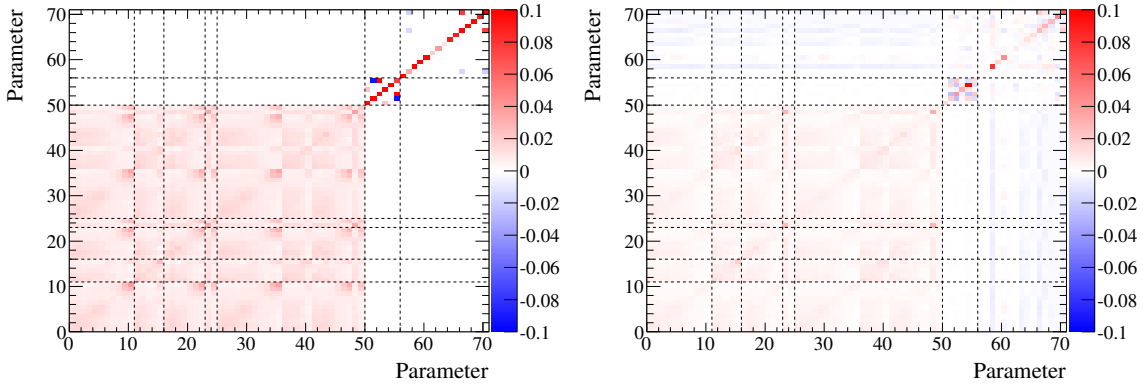


Figure 5.11: BANFF pre-fit (left) and post-fit (right) covariance matrices. Parameter numbering is: 0–10 ND280 ν_μ flux, 11–15 ND280 $\overline{\nu}_\mu$ flux, 16–22 ND280 ν_e flux, 23–24 ND280 $\overline{\nu}_e$ flux, 25–49 SK flux, 50–55 FSI, 56–70 cross-section.

- Super-Kamiokande flux (irrelevant for this analysis)
- final state interactions (6 parameters)
- CCQE cross-section (4 parameters)
- pion production cross-section (8 parameters)
- nuclear model (3 parameters).

External data constraints from the MiniBooNE experiment are used to alter three of the pion production cross-section parameters to form the BANFF pre-fit prediction, and ND280 ν_μ data is then used to constrain all the flux and cross-section parameters, to form the BANFF post-fit prediction [88].

For both the pre-fit and post-fit predictions, the BANFF group provide a set of T2KReWeight tweaks and a covariance matrix describing the uncertainties. T2KReWeight is then used to re-weight all the signal and background events that enter the selection, and all signal events that are missed. For reference, the pre-fit and post-fit covariance matrices are shown in Figure 5.11, and the effect of the BANFF fit on the parameter values is shown in Figure 5.12. Note that before the BANFF fit, the flux parameters are not correlated with the cross-section parameters. The BANFF post-fit introduces significant correlations, as neutrino detectors such as ND280 are only able to measure the rate of interactions, which is a product of the flux and cross-section of neutrinos.

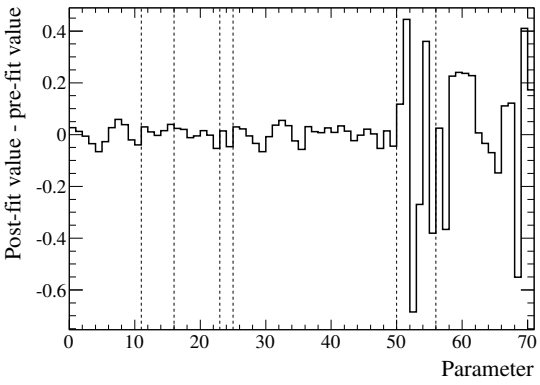


Figure 5.12: Fractional change to the flux and cross-section parameters due to the BANFF fit. Parameter numbering is: 0–10 ND280 ν_μ flux, 11–15 ND280 $\overline{\nu}_\mu$ flux, 16–22 ND280 ν_e flux, 23–24 ND280 $\overline{\nu}_e$ flux, 25–49 SK flux, 50–55 FSI, 56–70 cross-section.

This analysis will use the BANFF pre-fit covariance matrix and parameter values to define the systematic uncertainties. The rest of this section summarises how these uncertainties were quantified.

5.2.2 Flux uncertainties

The neutrino flux prediction is affected by uncertainties on

- hadron production, which are constrained using NA61 and other external data, as explained in Section 3.1.1
- properties of the proton beam, including the beam profile and axis alignment, which are accounted for by varying the MC simulation by the uncertainty measured in real T2K data for each run
- alignment of the target and magnetic horns, which are studied by rotating the target and shifting the horns in the MC simulation
- the horn current and magnetic field, which are again varied in the simulation.

A covariance matrix is produced for each source of uncertainty, binned in neutrino energy, neutrino flavour and detector (ND280 and SK). The total uncertainty is simply the sum of the individual covariance matrices.

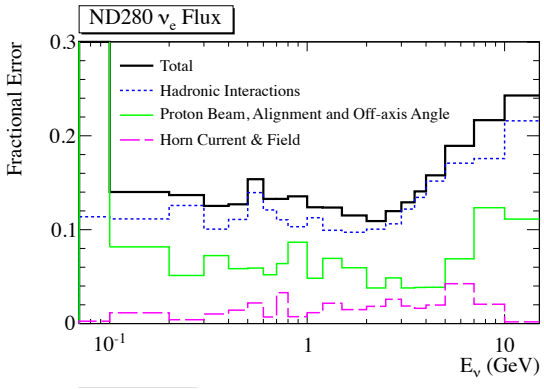


Figure 5.13: Fractional uncertainty on the ND280 ν_e flux prediction [54].

Figure 5.13 shows the ND280 ν_e flux uncertainty as a function of neutrino energy, and shows that the hadron production uncertainty is the dominant error source. In future T2K analyses, this source of error will be significantly reduced by using higher-statistics NA61 data taken with both the thin target and T2K replica target. Note that the BANFF fit only uses 7 ν_e bins to parameterise the flux, compared to the 20 bins shown in Figure 5.13. This is due to the limited power of the BANFF fit to constrain the ν_e flux, as it uses a very pure selection of ν_μ interactions.

5.2.3 Cross-section uncertainties

Cross-section model uncertainties are broadly split into four categories: final state interactions, the CCQE model, the pion production model, and the nuclear model. The same uncertainties affect all T2K analyses and are described in detail in Reference [89]. A summary of all the cross-section model parameters is shown in Table 5.7.

Final state interactions uncertainties

The NEUT FSI model includes parameters which alter the pion interaction probabilities in the nucleus. The six parameters affect inelastic scattering at low and high energies ($F^{\text{Inel}}_{1,2}$), pion production and absorption ($F^{\text{Prod,Abs}}$) and charge exchange at low and high energies ($F^{\text{CX}}_{1,2}$). The central value of these parameters and their uncertainties are determined from fits to pion scattering data [90].

Parameter	Nominal value	Uncertainty	Category
F_1^{Inel}	0.00	0.41	FSI
F_2^{Inel}	0.00	0.34	FSI
F^{Prod}	0.00	0.57	FSI
F^{Abs}	0.00	0.28	FSI
F_1^{CX}	0.00	0.50	FSI
F_2^{CX}	0.00	0.41	FSI
M_A^{QE} (GeV/c^2)	1.21	0.45	CCQE
M_A^{RES} (GeV/c^2)	1.41	0.11	Pion production
$x^{\text{CC other}}$ (GeV)	0.00	0.40	Pion production
x^{SF}	0.00	1.00	Nuclear
E_B (MeV)	25.00	9.00	Nuclear
p_F (MeV/c)	217.00	30.00	Nuclear
$x^{\pi\text{-less}}$	0.20	0.20	Pion production
x_1^{QE}	1.00	0.11	CCQE
x_2^{QE}	1.00	0.30	CCQE
x_3^{QE}	1.00	0.30	CCQE
$x_1^{\text{CC1}\pi}$	1.15	0.43	Pion production
$x_2^{\text{CC1}\pi}$	1.00	0.40	Pion production
$x^{\text{CC coh.}}$	1.00	1.00	Pion production
$x^{\text{NC other}}$	1.00	0.30	Pion production
$x^{\text{NC1}\pi^0}$	0.96	0.43	Pion production

Table 5.7: Nominal values and uncertainties of NEUT FSI and cross-section model parameters. The ordering in this table matches the ordering in the BANFF covariance matrices shown in Figure 5.11.

CCQE model uncertainties

As mentioned in Section 1.2, there is a large discrepancy in the value of M_A^{QE} measured by MiniBooNE and other experiments. T2K therefore uses a large prior uncertainty on this parameter ($\sigma_{M_A^{\mathrm{QE}}} = 0.43\,\mathrm{GeV}$) to cover the NEUT nominal value and the NEUT best fit to MiniBooNE data.

One possible explanation for the discrepancy between MiniBooNE and NOMAD is that the flux of one experiment may be incorrectly modelled. Three independent CCQE normalisation factors ($x_{1,2,3}^{\mathrm{QE}}$) are therefore introduced that cover different neutrino energy regions. The MiniBooNE flux uncertainty of 11% is assigned to x_1^{QE}, which covers energies up to 1.5 GeV. x_2^{QE} ($1.5 < E_\nu < 3.5\,\mathrm{GeV}$) and x_3^{QE} ($E_\nu > 3.5\,\mathrm{GeV}$) are assigned uncertainties of 30% to cover the difference between MiniBooNE and NOMAD.

Pion production uncertainties

Pion production uncertainties are set using a joint fit to MiniBooNE CC1π^+ [91], CC1π^0 [92] and NC1π^0 [93] data, varying the following NEUT parameters:

- M_A^{RES}, the axial mass in the Rein and Sehgal model of pion production [94]
- CC1π normalisation for low energy neutrinos ($x_1^{\mathrm{CC1}\pi}$ for $E_\nu < 2.5\,\mathrm{GeV}$)
- NC1π^0 normalisation ($x^{\mathrm{NC1}\pi^0}$).

The best fit of these three parameters are the BANFF pre-fit tweaks to the nominal NEUT model.

The MiniBooNE samples only contain a small fraction of CC multi-pion, NC coherent, NC charged pion and NC multi-pion interactions, and large prior uncertainties are therefore assigned to these modes. For CC multi-pion and CC DIS interactions, an energy dependent uncertainty, $x^{\mathrm{CC\ other}}$, applies a weight w with the form $w = 1 + x^{\mathrm{CC\ other}}/E_\nu(\mathrm{GeV})$, where the parameter is allowed to vary around a nominal value of 0 with a prior uncertainty of 0.4. For MC charged pion production and all other NC interactions a 30% normalisation uncertainty is used ($x^{\mathrm{NC\ other}}$).

K2K [95] and SciBooNE [96] did not observe charged current coherent pion production, so a 100% normalisation uncertainty is assigned to this mode ($x^{\mathrm{CC\ coh.}}$).

Finally, a 100% uncertainty is assigned to the fraction of Δ that de-excite without emitting pions ($x^{\pi\text{-less}}$).

Nuclear model uncertainties

As mentioned in Section 1.2, NEUT uses the relativistic Fermi gas (RFG) model of nuclei. The Fermi surface momentum, p_F, and binding energy of nucleons, E_B, in carbon are determined from electron scattering data, and uncertainties of 30 MeV/c and 9 MeV are applied, respectively.

The RFG model is very simplistic, and alternatives are considered by comparing to a "spectral function" nuclear model implemented in the NuWro neutrino interaction generator [97]. The discrepancy between the models is assigned as the uncertainty, and represented by the parameter x^{SF} which linearly interpolates between the predicted lepton kinematics with the RFG ($x^{\text{SF}} = 0$) and spectral function ($x^{\text{SF}} = 1$) models. Due to the interaction models implemented in NEUT, the nuclear model uncertainties only affect CCQE interactions.

5.3 Other systematics

Other systematics that are specific to the analysis technique will be described in the next chapter. These systematics are the data statistics, Monte Carlo statistics and the effect of using the γ sample to constrain the background in the ν_e sample.

Chapter 6

ν_e cross-section measurement

This chapter describes the measurement of the ν_e charged-current cross-section on carbon using ND280. Due to the low efficiency of selecting ν_e interactions in which the electron is emitted with a low momentum or at a high angle, two separate results are presented. In the first result, the low efficiency is corrected for, and the cross-section for the full electron kinematic phase-space is found. This method introduces a significant model dependency in the regions of low efficiency, so a second result is also presented in which only events with $p_e > 550\,\mathrm{MeV}/c$ and $\cos(\theta_e) > 0.72$ are selected. The result of this restricted phase-space analysis has reduced systematic uncertainties and is less model-dependent.

6.1 Event selections

Before describing the cross-section analysis method, the key features of the event selections will be repeated for convenience.

The reconstructed momentum, angle and Q^2 distributions of the electrons selected in the ν_e sample are shown in Figure 6.1. Recall that tracks with momentum below $200\,\mathrm{MeV}/c$ are not selected, as the efficiency of the selection is very low, and the γ background dominates the very small ν_e signal.

Particular mention must be made of the Q^2 definitions used. For $Q^{2,\mathrm{true}}$ (where the superscript $^{\mathrm{true}}$ indicates that a true quantity is being referred to) the actual four-momentum transfer of the interaction is used, as:

$$Q^{2,\mathrm{true}} = -|\mathbf{p}_e^{\mathrm{true}} - \mathbf{p}_{\nu_e}^{\mathrm{true}}|^2, \tag{6.1}$$

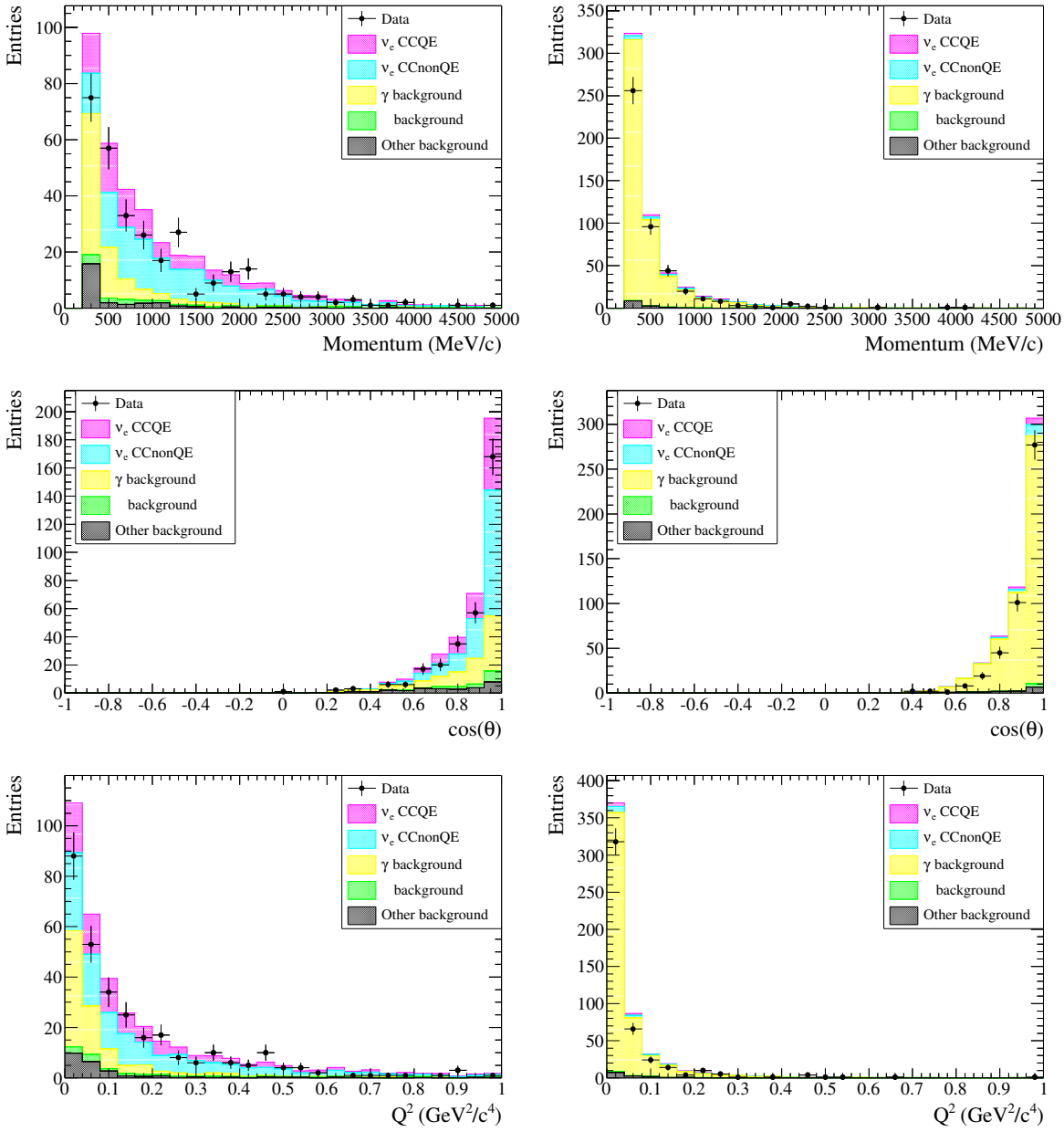


Figure 6.1: Reconstructed momentum (top), angle (middle) and Q^2 (bottom) distributions of electrons selected in the ν_e sample (left) and γ sample (right).

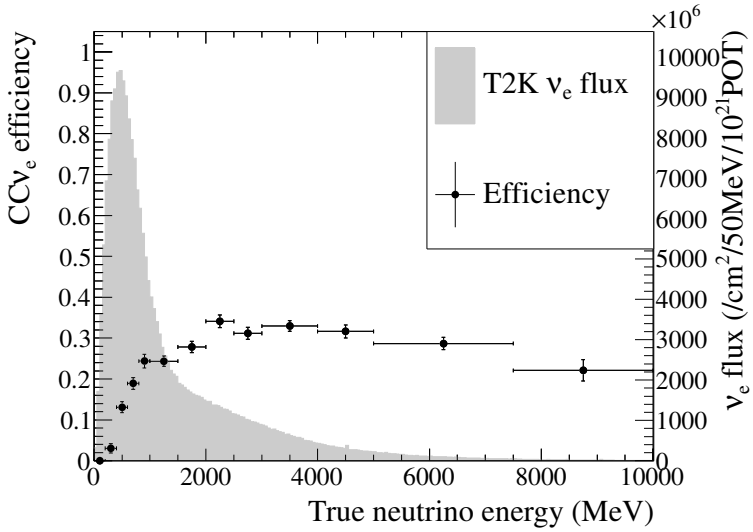


Figure 6.2: Efficiency as a function of neutrino energy for the ν_e sample.

where **p** is the four-momentum of the particle.

$Q^{2,\text{reco}}$ (where the superscript $^{\text{reco}}$ indicates that a reconstructed quantity is being referred to) is calculated from the three-momentum of the reconstructed electron, and assuming CCQE kinematics. The neutrino is assumed to be travelling in the $+z$ direction, which is not quite true for ND280 events but will be accounted for by the unfolding preocedure. A nuclear potential of 25 MeV is assumed, which is the potential used in NEUT for carbon. In the equations below, V is the nuclear potential, $m_{e,n,p}$ are the masses of the electron, neutron and proton, and p_e is the magnitude of the three-momentum of the electron. The reconstructed $Q^{2,\text{reco}}$ is then defined as:

$$E_e^{\text{rec}} = \sqrt{m_e^2 + p_e^2} \tag{6.2}$$

$$E_{\nu_e}^{\text{rec}} = -\frac{1}{2}\frac{(m_p^2 - m_e^2) + 2E_e^{\text{rec}}(m_n - V) - (m_n - V)^2}{E_e^{\text{rec}} + (m_n - V) + p_e\cos(\theta_e)} \tag{6.3}$$

$$Q^{2,\text{reco}} = 2E_{\nu_e}^{\text{rec}}(E_e^{\text{rec}} - p_e\cos(\theta_e)) - m_e^2 \tag{6.4}$$

The efficiency of the selection as a function of true neutrino energy is shown in Figure 6.2. Although this analysis is described as a flux-averaged cross-section measurement, it is important to note that the selection efficiency varies as a function of neutrino energy, and in particular is very low at low neutrino energy. The overall efficiency of the selection is 26.7%.

A high-purity sample of $\gamma \to e^+e^-$ interactions is used to constrain the background in the ν_e sample. The sample is 95% pure in $\gamma \to e^+e^-$ conversions, and the momentum and angular distributions of the electrons selected are also shown in Figure 6.1.

6.1.1 Out of fiducial volume background

Although the dominant background in the ν_e sample is $\gamma \to e^+e^-$ conversions, it is the Out Of Fiducial Volume (OOFV) portion of this that is of most concern. These are events that occur outside the FGD FV, but that produce a γ which then enters the FV and converts into an e^+e^- pair. The reasons why the OOFV background is of particular concern are:

1. 41.8% of OOFV events in the ν_e sample occur in the magnet, Barrel ECals or PØD ECals. There was no magnetic field simulated in the magnet flux return in the Monte Carlo used in this analysis, so the MC incorrectly simulates what happens to particles in this region. Studies in preparation for future analyses have shown that the lack of magnetic field simulation seriously affects the MC prediction in the ECals, but further detailed study is not yet complete.

2. 30.2% of OOFV events in the ν_e sample are caused by neutrino interactions on heavy targets (iron and lead), for which there is a large uncertainty on the interaction modelling.

3. 35.2% of OOFV events in the ν_e sample are CC DIS events and 31.2% are NC events, for which again there is large uncertainty on the accuracy of the interaction modelling.

4. The flux, cross-section and detector systematic uncertainties on interactions in the FGD are well-studied, so a constraint on the background from in the fiducial volume (the In FV background) is not needed. Far less study has been devoted to OOFV interactions.

For these reasons, the γ selection will be used to constrain the OOFV background, rather than the whole $\gamma \to e^+e^-$ background. The re-weighting method will be explained in Section 6.4.4.

6.2 ν_e cross-section predictions

6.2.1 Cross-section definition

In this analysis, the differential cross-section is computed as a function of electron momentum (p_e), electron angle ($\cos(\theta_e)$) and the four-momentum transfer of the interaction (Q^2). In this section any such variable will be referred to as X.

The total number of true interactions in bin t_k of variable X is given by

$$\begin{aligned} N_{t_k} &= T\phi \int_{X_{t_k}} \left\langle \frac{\partial \sigma}{\partial X} \right\rangle_{\phi} dX \\ &= T\phi \int_{X_{t_k}} \frac{\partial \left\langle \sigma \right\rangle_{\phi}}{\partial X} dX, \end{aligned} \tag{6.5}$$

where

- T is the number of target nucleons
- ϕ is the total integrated flux
- $\langle \cdots \rangle_{\phi}$ indicates that the quantity is averaged over the flux
- the second step proceeds as the flux-averaging is independent of the derivatives.

If ΔX_{t_k} is now defined as the width of bin t_k, equation (6.5) can be rearranged to give the flux averaged differential cross-section per nucleon in bin t_k as

$$\left\langle \frac{\partial \langle \sigma \rangle_{\phi}}{\partial X} \right\rangle_{t_k} = \frac{N_{t_k}}{T \cdot \phi \cdot \Delta X_{t_k}}. \tag{6.6}$$

The total flux averaged cross-section per nucleon is found by integrating over all X and is given by

$$\langle \sigma \rangle_{\phi} = \frac{N_{\text{total}}}{T \cdot \phi}, \tag{6.7}$$

where N_{total} is the sum of N_{t_k} in all bins t_k.

6.2.2 Predicted ν_e cross-sections

Equation (6.6) shows that as well as measuring the number of events in each bin, the number of target nucleons, T, and the total integrated flux, ϕ, must also be known.

T is 5.5×10^{29} nucleons, and is calculated from the known composition and mass of the components in the FGD [51].

ϕ is calculated from flux histograms provided by the beam group. Separate flux histograms are provided for each T2K run, and are then weighted by the good quality POT for each run, and summed to give the total flux histogram shown in Figure 6.2. The total integrated ν_e flux for this analysis is 1.35×10^{11} cm^{-2}.

NEUT

NEUT provides a tool to extract predicted cross-sections on the elemental constituents of the FGD: carbon, hydrogen, oxygen, titanium, silicon and nitrogen. For each element, the tool produces histograms of the cross-section as a function of energy for each mode (CCQE, DIS etc). Bin widths of 50 MeV are chosen to match the flux histograms, and the cross-sections are evaluated at the centre of each bin. All the CC modes are then summed for each element, and the total cross-section per nucleon for the FGD is computed using

$$\sigma_{\mathrm{FGD}} = \sum_{e=\mathrm{C,H,O,Ti,Si,N}} \frac{\sigma_e f_e}{A_e}, \tag{6.8}$$

where e denotes the element, σ_e the CC cross-section for that element, f_e the fraction by mass of the FGD composed of that element, and A_e the number of nucleons per atom of that element. For reference, f_e and A_e are listed for each element in Table 6.1.

Figure 6.3 shows the predicted cross-section as a function of energy for the FGD, and for carbon alone. Although the FGD is 86% carbon, there is a significant difference between the cross-section on carbon and the average cross-section on the FGD. This means that although this measurement is described as the ν_e CC inclusive cross-section on carbon, this is only an approximate statement.

Also shown in Figure 6.3 are the flux-averaged cross-section predictions. These are calculated by multiplying the contents of the flux and cross-section histograms, and dividing by the total flux. The predicted flux-averaged CC ν_e cross-section on the FGD

Element	f_e (%)	A_e
Carbon	86.1	12
Hydrogen	7.4	1
Oxygen	3.7	16
Titanium	1.7	48
Silicon	1.0	28
Nitrogen	0.1	14

Table 6.1: Composition of the FGD. f_e is the fraction by mass of the FGD composed of that element. A_e is the number of nucleons in that element.

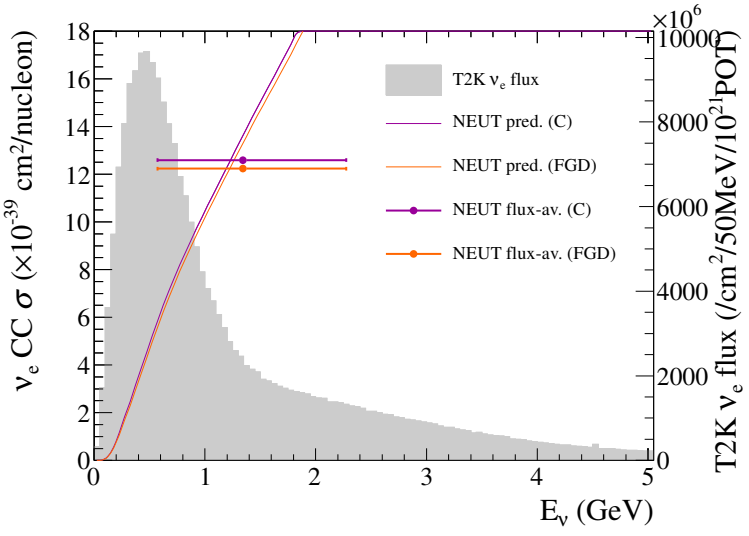


Figure 6.3: CC inclusive ν_e cross-section predictions from NEUT on carbon and the FGD as a whole. The curves are the cross-sections as a function of neutrino energy. The points are the flux-averaged cross-sections.

is 1.224×10^{-38} cm^2/nucleon. The mean neutrino energy, E_{mean}, is used for the location of the "dot", and the horizontal error bars follow the same convention as used for the T2K ν_μ CC inclusive measurement in 2013 [51]: variables E_{lower} and E_{upper} are defined such that $\phi(E_{\text{lower}}) = \phi(E_{\text{upper}})$ and

$$\frac{\int_{E_{\text{lower}}}^{E_{\text{upper}}} \frac{\partial\phi(E)}{\partial E} dE}{\int_0^\infty \frac{\partial\phi(E)}{\partial E} dE} = 90\%, \tag{6.9}$$

and the variances in the intervals $[E_{\text{lower}}, E_{\text{mean}}]$ and $[E_{\text{mean}}, E_{\text{upper}}]$ are then calculated to obtain the errors, through

$$\sigma_-^2 = \frac{\int_{E_{\text{lower}}}^{E_{\text{mean}}} (E - E_{\text{mean}})^2 \frac{\partial\phi(E)}{\partial E} dE}{\int_{E_{\text{lower}}}^{E_{\text{mean}}} \frac{\partial\phi(E)}{\partial E} dE} \tag{6.10}$$

$$\sigma_+^2 = \frac{\int_{E_{\text{mean}}}^{E_{\text{upper}}} (E - E_{\text{mean}})^2 \frac{\partial\phi(E)}{\partial E} dE}{\int_{E_{\text{mean}}}^{E_{\text{upper}}} \frac{\partial\phi(E)}{\partial E} dE}, \tag{6.11}$$

where σ_- and σ_+ are the extents of the lower and upper error bars. For the T2K ν_e flux, $E_{\text{mean}} = 1.343$ GeV, $E_{\text{lower}} = 0.025$ GeV, $E_{\text{upper}} = 3.175$ GeV, $\sigma_- = 0.772$ GeV and $\sigma_+ = 0.933$ GeV.

There is no simple way in NEUT to calculate the flux-averaged differential cross-section as a function of the interaction kinematics. Instead, the ND280 Monte Carlo files are used. The number of generated events in each bin are counted, and divided by the number of target nucleons and the total simulated flux. As a cross-check to validate the method, the total flux-averaged cross-section is computed from the ND280 MC events. This gives a cross-section of 1.231×10^{-38} cm^2/nucleon, compared to the 1.224×10^{-38} cm^2/nucleon found using NEUT directly. This small discrepancy can be explained by the fact that the direct NEUT prediction was calculated by summing the contents of bins of width 50 MeV. Although the flux histogram is an average over the 50 MeV range, the cross-section is just evaluated at the central value. As the flux and cross-section can vary rapidly across each bin, the true flux-averaged cross-section of the bin is not necessarily the same as that found by multiplying the two histograms, and as the difference between the calculations is small, it shows that there is no problem with the method used to calculate the cross-sections from the ND280 MC files.

One final subtle point relates to applying the detector systematics. The nominal set of detector systematics affects which events enter each sample, and the weight of those events. These weights just reflect the detector efficiency, rather than the physics that

is being simulated. To keep the same predicted ν_e cross-section, events entering the ν_e sample with a detector systematic weight w_d must also be counted as events that are missed with a weight $1 - w_d$. Applying this procedure means that the same predicted ν_e cross-sections are found when turning detector systematics off and when setting them to their nominal values.

GENIE

GENIE is the alternative neutrino interaction generator used by T2K, and is also used by many other neutrino experiments. Comparing the T2K data to the GENIE prediction is of more relevance to the neutrino community as a whole than comparing to NEUT, which is mostly used by the T2K and Super-Kamiokande collaborations.

Generating cross-section predictions for GENIE proceeds in a similar way to NEUT. Cross-section splines directly from the generator are used to compute the total CC flux-averaged cross-section, and differential cross-section predictions are found from ND280 MC files. As a cross-check, the total cross-section from the MC files is compared to the direct prediction from the generator.

The GENIE cross-section splines give a total predicted CC ν_e cross-section on the material composition of the FGD of 1.072×10^{-38} cm^2/nucleon.

GENIE MC files are used to compute the differential cross-sections. Fewer GENIE files are available than NEUT MC files, and a total of 28.04×10^{20} simulated POT is available, just 4.75 times the data POT. More POT would be desirable, but as the GENIE MC is only used to draw predicted cross-sections (and not used as part of the analysis) it is less critical that there are large statistics. As with the NEUT files, true ν_e CC interactions that occur in the FGD FV are selected, and a total flux-averaged cross-section of 1.083×10^{-38} cm^2/nucleon is found from the GENIE MC files. As was the case for NEUT, the prediction from the MC files is slightly different to the prediction from the cross-section splines (1.083×10^{-38} cm^2/nucleon rather than 1.072×10^{-38} cm^2/nucleon).

Predictions

The flux-averaged predictions for the NEUT nominal, BANFF pre-fit, BANFF post-fit and GENIE nominal models are shown in Figure 6.4 (see Section 5.2.1 for details of the BANFF fit). Predictions for both the full electron kinematic phase-space and the

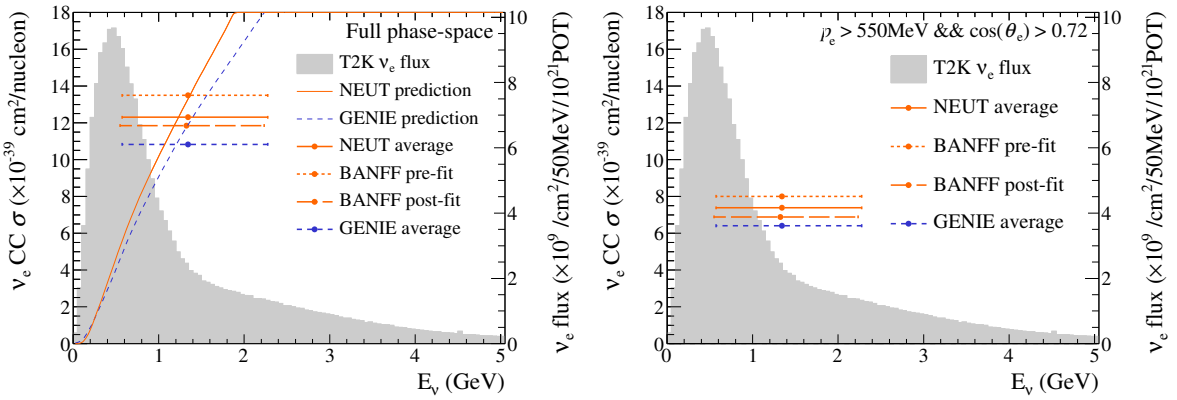


Figure 6.4: Flux-averaged cross-section predictions from ND280 MC. The left plot shows the prediction for the full electron kinematic phase-space; the right plot shows the restricted phase-space. The overflow and underflow bins (represented by $>$ and $<$ markers) are normalised to the bin width shown on the plot.

Model	$\sigma^{\mathrm{CC}}_{\nu_e}(\times 10^{-39}\ \mathrm{cm}^2/\mathrm{nucleon})$	
	Full phase-space	Restricted phase-space
NEUT	12.31	7.38
BANFF pre-fit	13.49	8.01
BANFF post-fit	11.85	6.89
GENIE	10.83	6.41

Table 6.2: Flux-averaged cross-section predictions from the NEUT and GENIE interaction generators. The restricted phase-space is defined by events with $p_e > 550\,\mathrm{MeV}/c$ and $\cos(\theta_e) > 0.72$.

restricted phase-space are shown. The restricted phase-space corresponds to events with $p_e > 550\,\mathrm{MeV}/c$ and $\cos(\theta_e) > 0.72$. Note that the BANFF post-fit re-weighting changes the shape of the ν_e flux as well as the cross-section parameters, which can be seen in the shifted mean energy and horizontal error bars. The predictions are also tabulated in Table 6.2. The differential cross-section predictions are shown in Figure 6.5.

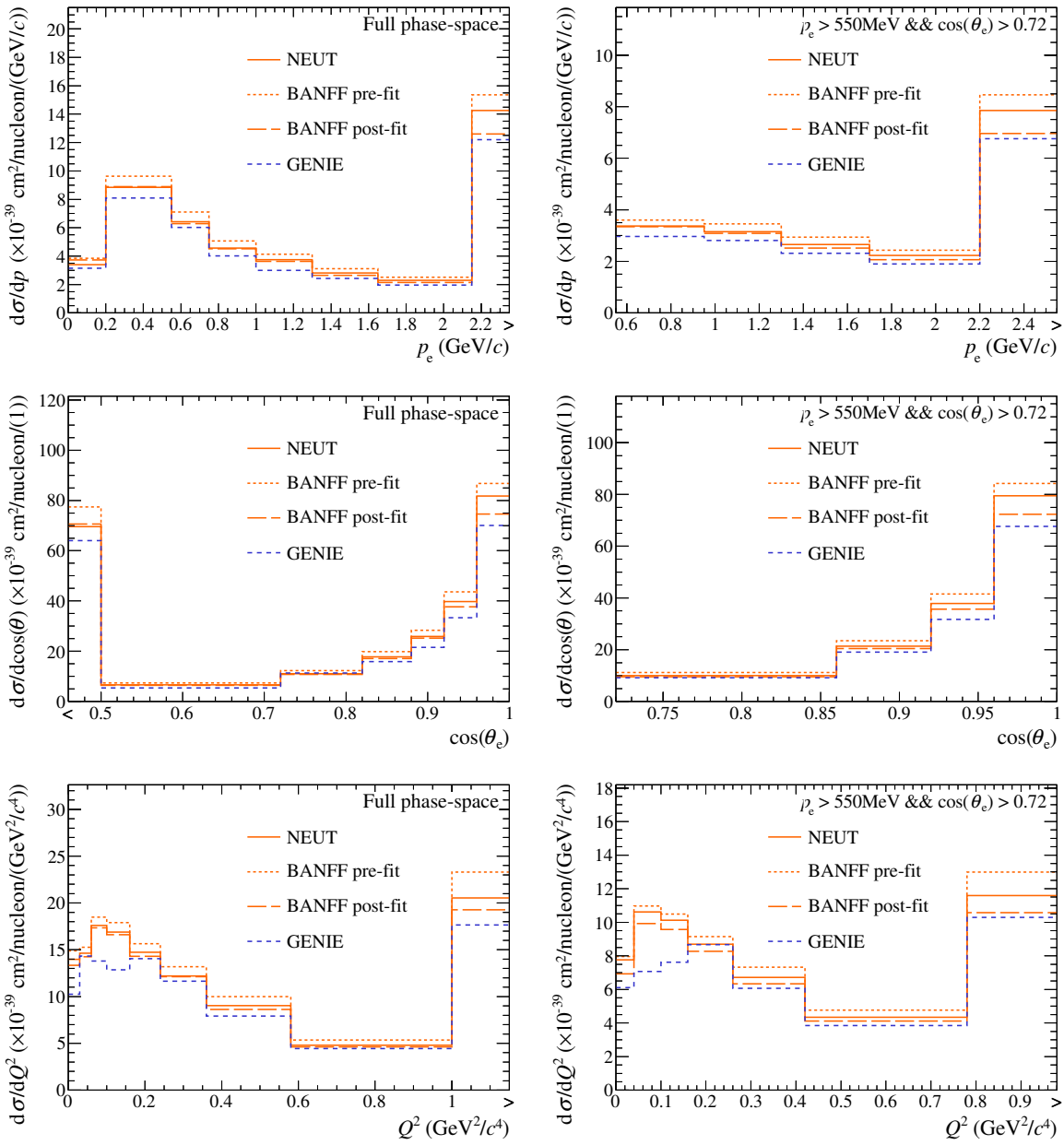


Figure 6.5: Flux-averaged differential cross-sections in p_e (top) and $\cos(\theta_e)$ (middle) and Q^2 (bottom) from ND280 MC. The left column shows the prediction for the full electron kinematic phase-space; the right column shows the restricted phase-space. Underflow and overflow bins are marked by $<$ and $>$ respectively, and are normalised to the width shown on the plot.

6.3 Unfolding procedure

6.3.1 The Bayesian unfolding method

The Bayesian unfolding technique was used for the T2K ν_μ CC inclusive cross-section measurement in 2013 [51], and is based on the method proposed by d'Agostini [98]. The aim of the method is to unfold (or unsmear) from a reconstructed dataset to the true distribution that caused it, with the unsmearing accounting for detector inefficiencies and mis-reconstruction.

Before discussing the unfolding method, a few brief definitions will be made.

- n_t is the number of true bins for a given observable, n_r is the number of reconstructed bins, and t_k and r_j are the true and reconstructed bins.
- $S_{r_j t_k}$ is referred to as the *signal matrix* and is the number of true simulated signal events in true bin t_k that were reconstructed in bin r_j.
- M_{t_k} is referred to as the *missed vector*, and is the number of simulated signal events in true bin t_k that were not selected.
- B_{r_j} is referred to as the *background vector*, and is the number of simulated non-signal events that were selected in each reconstructed bin.

$S_{r_j t_k}$, M_{t_k} and B_{r_j} are the only MC inputs required for the Bayesian unfolding technique. Given these basic definitions, the following quantities can then be derived.

- N_{t_k} is the *initial estimator* of the number of signal events in each true bin and is defined as

$$N_{t_k} = \sum_{j=1}^{n_r} S_{r_j t_k} + M_{t_k}. \tag{6.12}$$

- $P(r_j|t_k)$ is the *smearing matrix* and is defined as

$$P(r_j|t_k) = \frac{S_{r_j t_k}}{N_{t_k}}. \tag{6.13}$$

- ϵ_{t_k} is the overall *efficiency* of reconstructing events in each true bin and is defined as

$$\epsilon_{t_k} = \sum_{j=1}^{n_r} \frac{S_{r_j t_k}}{N_{t_k}}. \tag{6.14}$$

- $P_0(t_k)$ is the prior probability of observing a true signal event in true bin t_k and is defined as

$$P_0(t_k) = \frac{N_{t_k}}{\sum_{\alpha=1}^{n_t} N_{t_\alpha}}. \tag{6.15}$$

The unfolding technique chosen is based on Bayes' theorem. In this case, Bayes' theorem is used to find the probability that an event is observed in true bin t_k given it was reconstructed in bin r_j, as

$$P_m(t_k|r_j) = \frac{P(r_j|t_k)P_m(t_k)}{\sum_{\alpha=1}^{n_t} P(r_j|t_\alpha)P_m(t_\alpha)}. \tag{6.16}$$

Here, $P_m(t_k|r_j)$ is referred to as the *unsmearing matrix*. m is the iteration index which is initially set to 0 and will be explained shortly.

Given a dataset $N_{r_j}^{\text{meas}}$, the estimate of the number of events in each true bin is given by

$$N_{t_k}^{m+1} = \frac{1}{\epsilon_{t_k}} \sum_{j=1}^{n_r} P_m(t_k|r_j)(N_{r_j}^{\text{meas}} - B_{r_j}). \tag{6.17}$$

This is often referred to as the *unfolded* distribution.

The Bayesian unfolding can be applied iteratively, as denoted by the index m. $P_0(t_k)$ was defined previously, and comes from the MC prediction. After the unfolding has been performed, it can be updated using the output of equation (6.17) through

$$P_m(t_k) = \frac{N_{t_k}^m}{\sum_{\alpha=1}^{n_t} N_{t_\alpha}^m}. \tag{6.18}$$

A new unsmearing matrix can then be computed with equation (6.16), and a new unfolded distribution found by applying equation (6.17) again.

In the limit of high statistics, the literature suggests approximately 3 iterations should be performed. Performing more than one iteration allows for the unfolded distribution to become less biased if the MC does not closely resemble the data. However, in the limit of m tending to infinity, the unfolding matrix contains large fluctuations, so a small m should be chosen. Further, the statistical error on the number of data and MC events is propagated through each iteration, and can be significantly amplified if there are low statistics. Studies on the number of iterations that should be performed for this analysis are shown in Section 6.4.3.

For clarity, the unfolded distribution after all iterations have been performed will be referred to as $\widehat{N}_{t_k}$. After unfolding the number of events in each true bin, the differential cross-section simply follows as

$$\left\langle \frac{\partial \langle \sigma \rangle_\phi}{\partial X} \right\rangle_{t_k} = \frac{\widehat{N}_{t_k}}{T \cdot \phi \cdot \Delta X_{t_k}}, \tag{6.19}$$

and the total flux-averaged cross-section is

$$\langle \sigma \rangle_\phi = \frac{\widehat{N}_{\text{total}}}{T \cdot \phi}, \tag{6.20}$$

where $\widehat{N}_{\text{total}}$ is the sum of $\widehat{N}_{t_k}$ in all true bins t_k.

For this analysis, the RooUnfold package is used to do the unfolding [99].

6.3.2 Binning

The binning used for the different kinematic variables are shown in Table 6.3 for the full phase-space analysis and in Table 6.4 for the restricted phase-space analysis. There are extra true bins in the p_e^{true} and $\cos(\theta_e^{\text{true}})$ distributions in the full phase-space analysis in regions of very low efficiency. Unfolding into these bins relies on the model prediction, so this analysis is not model-independent. The restricted phase-space analysis does not try to unfold into these regions, and is therefore less model-dependant.

The bin distributions were chosen such that there are at least 25 true CC ν_e events expected in each reconstructed bin, and the bins are multiples of 50 MeV/c for p_e, 0.02 for $\cos(\theta_e)$ and 0.01 GeV$^2/c^4$ for Q^2. The requirement of 25 events is chosen as a compromise between

Variable	Truth or reco	Bins	Units
p_e^{true}	Truth	[0, 200, 550, 750, 1000, 1300, 1650, 2150, 10^{20}]	MeV/c
p_e^{reco}	Reco	[200, 550, 750, 1000, 1300, 1650, 2150, 10^{20}]	MeV/c
$\cos(\theta_e^{\text{true}})$	Truth	[−1.0, 0.5, 0.72, 0.82, 0.88, 0.92, 0.96, 1.0]	
$\cos(\theta_e^{\text{reco}})$	Reco	[0.5, 0.72, 0.82, 0.88, 0.92, 0.96, 1.0]	
$Q^{2,\text{true}}$	Truth	[0, 0.03, 0.06, 0.1, 0.16, 0.24, 0.36, 0.58, 1, 10^{20}]	GeV$^2/c^4$
$Q^{2,\text{reco}}$	Reco	[0, 0.03, 0.06, 0.1, 0.16, 0.24, 0.36, 0.58, 1, 10^{20}]	GeV$^2/c^4$

Table 6.3: Binning used for the kinematic variables in the full phase-space analysis.

Variable	Truth or reco	Bins	Units
p_e^{true}	Truth	[550, 950, 1300, 1700, 2200, 10^{20}]	MeV/c
p_e^{reco}	Reco	[550, 950, 1300, 1700, 2200, 10^{20}]	MeV/c
$\cos(\theta_e^{\text{true}})$	Truth	[0.72, 0.86, 0.92, 0.96, 1.0]	
$\cos(\theta_e^{\text{reco}})$	Reco	[0.72, 0.86, 0.92, 0.96, 1.0]	
$Q^{2,\text{true}}$	Truth	[0, 0.04, 0.1, 0.16, 0.26, 0.42, 0.78, 10^{20}]	GeV$^2/c^4$
$Q^{2,\text{reco}}$	Reco	[0, 0.04, 0.1, 0.16, 0.26, 0.42, 0.78, 10^{20}]	GeV$^2/c^4$

Table 6.4: Binning used for the kinematic variables in the restricted phase-space analysis.

- there being enough bins to show the shape of the differential distributions, particularly at low Q^2 which is most interesting from a theoretical standpoint
- there being enough events in each bin to make the results statistically significant
- the overall statistical uncertainty on the flux-averaged distributions being approximately 10%. The calculation of these uncertainties will be described in Section 6.4.4.

Events that are reconstructed outside of the accepted regions (for example events with $\cos(\theta_e^{\text{reco}}) < 0.5$ in the full phase-space analysis) are ignored regardless of which variable is being used for the unfolding. That is, these events are treated as being missed, rather than reconstructed. This is to ensure that the same dataset is used for all three differential cross-section measurements. In the restricted phase-space analysis, any true ν_e CC interactions that are truly $p_e^{\text{true}} < 550\,\text{MeV}/c$ and $\cos(\theta_e^{\text{true}}) < 0.72$, but that are reconstructed as in the signal region, are treated as background. There are very few such events however, and this introduces a negligible bias to the result.

6.3.3 Inputs to unfolding

Figure 6.6 shows the BANFF pre-fit smearing matrices $P(r_j|t_k)$ for unfolding in p_e, $\cos(\theta_e)$ and Q^2. These matrices combine the reconstruction efficiency and detector confusion, and for p_e and $\cos(\theta_e)$ would be diagonal for a perfect detector (apart from the extra true bins in the momentum and angle matrices). The Q^2 smearing matrix would not be expected to be perfectly diagonal, as nuclear effects smear the reconstructed neutrino energy and perfect CCQE kinematics have been assumed. The angular reconstruction is seen to be very accurate, but there is a significant bias and poor resolution for the momentum reconstruction. The nominal set of detector systematics are applied when creating these plots, including corrections for the momentum bias and resolution, but note that there are also large systematic uncertainties on these corrections. The majority of events that have very poor momentum reconstruction are those in which bremsstrahlung photons are emitted soon after the electron is produced. Without significantly improved reconstruction algorithms which can match these photons to the original electron, it is not possible to improve the momentum bias.

The prior distribution $P_0(t_k)$ is the other MC input to Bayes' theorem in equation (6.16). These distributions are shown in Figure 6.7 for each variable. For reference, the ν_e selection efficiency is shown in Figure 6.8. For brevity, these plots are only shown for the full phase-space analysis. For the restricted phase-space analysis, only the final differential and total flux-averaged cross-section plots will be shown.

6.3.4 OOFV re-weighting method

The γ sample is used to re-weight the OOFV portion of the background in the ν_e sample. As explained in Section 4.2.2, the two samples preferentially select photons from different sources. The different photon kinematic distributions in turn mean that the electron kinematics are different. Figure 6.9 shows area-normalised p_e^{reco}, $\cos(\theta_e^{\text{reco}})$ and $Q^{2,\text{reco}}$ distributions for electrons from OOFV interactions selected in each sample.

Due to the different kinematic phase-spaces probed, the OOFV re-weighting does not simply take the data/MC ratio of the γ sample in p_e^{reco}, $\cos(\theta_e^{\text{reco}})$ or $Q^{2,\text{reco}}$ bins and apply that scaling to the OOFV MC prediction in the ν_e sample. Instead, the re-weighting procedure is:

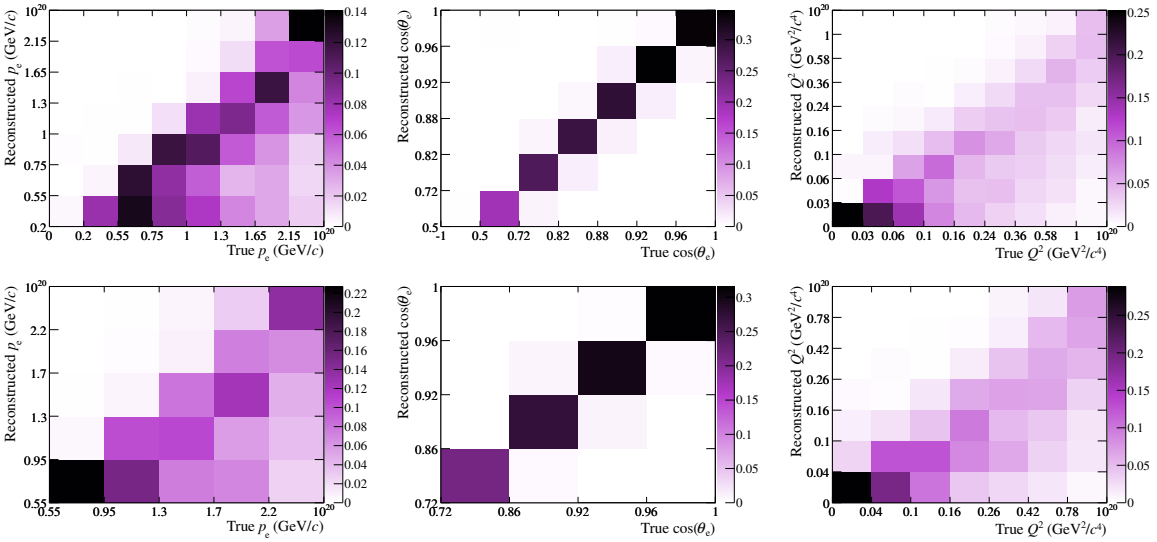


Figure 6.6: Smearing matrices $P(r_j|t_k)$ for p_e (left), $\cos(\theta_e)$ (middle) and Q^2 (right). The top row shows the full phase-space analysis and the bottom row the restricted phase-space analysis. Note that the bins are represented as if the axes were in bin number, to allow all bins to be easily seen. The axis labels are provided so the reader does not have to consult Table 6.3 and Table 6.4 to check the bin boundaries.

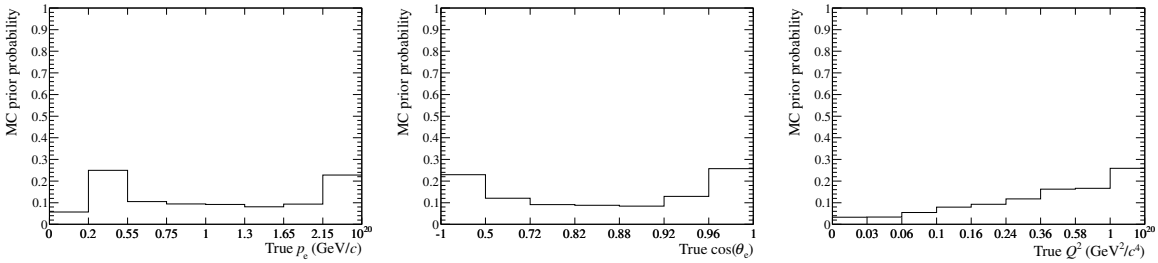


Figure 6.7: Prior probabilities $P_0(t_k)$ for p_e^{true} (left), $\cos(\theta_e^{\text{true}})$ (middle) and $Q^{2,\text{true}}$ (right).

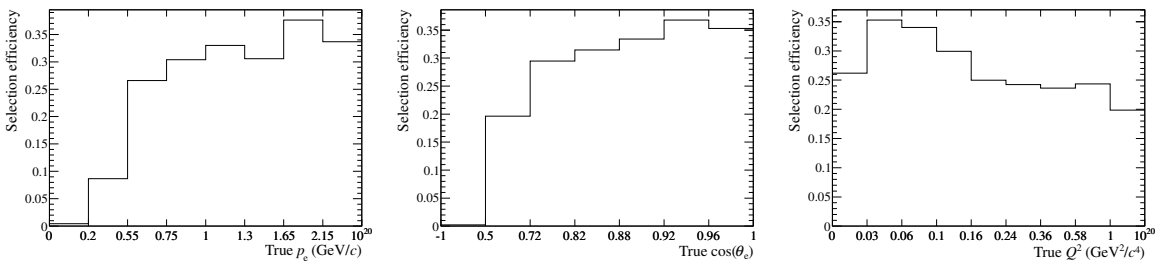


Figure 6.8: ν_e selection efficiency ϵ_{t_k} for p_e^{true} (left), $\cos(\theta_e^{\text{true}})$ (middle) and $Q^{2,\text{true}}$ (right).

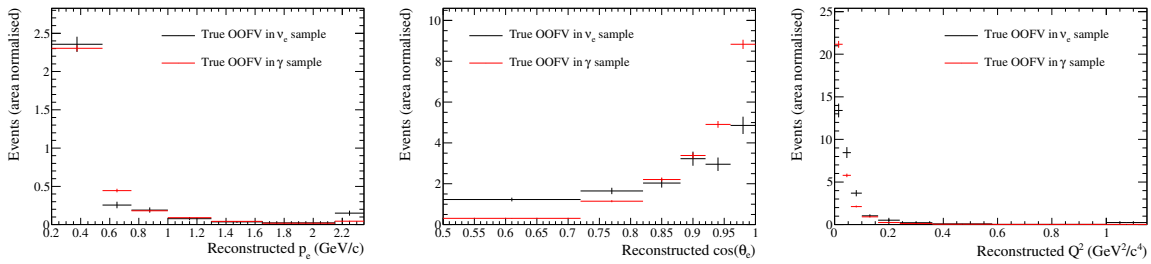


Figure 6.9: Shape comparison of the OOFV background in the ν_e and γ samples for p_e^{reco} (left), $\cos(\theta_e^{\text{reco}})$ (middle) and $Q^{2,\text{reco}}$ (right).

- "Background-subtract" the γ sample data, to remove the MC prediction of ν_e and In FV background events.
- Calculate the OOFV data/MC ratio in (p_e^{reco}, $\cos(\theta_e^{\text{reco}})$) bins, calling this 2D distribution R_γ^{2D}. A finer binning is used for this re-weighting than for the unfolded p_e^{reco} and $\cos(\theta_e^{\text{reco}})$ distributions, with the first p_e^{reco} bin being split into [200, 250, 300, 400, 550] MeV/c, and the last two $\cos(\theta_e^{\text{reco}})$ bins split into [0.92, 0.94, 0.96, 0.97, 0.98, 0.99, 1.0]. This finer binning is used as the OOFV background varies rapidly in these regions, and there are significant statistics of OOFV events. Further, the binning constraints described in Section 6.3.2 are not applicable in this context.
- Calculate the weighted average OOFV data/MC ratio for each p_e^{reco}, $\cos(\theta_e^{\text{reco}})$ and $Q^{2,\text{reco}}$ bin in the ν_e sample, where the weights are from the number of true OOFV events in each (p_e^{reco}, $\cos(\theta_e^{\text{reco}})$) bin in the ν_e sample. Call this 1D distribution $R_{\nu_e}^X$ for variable X. This is constrained to not be negative, even if the background-subtracted data suggests that it should be, with such bins set to 0 instead.
- Re-weight the OOFV prediction in the ν_e sample by $R_{\nu_e}^X$.

Note that the γ sample contains a small contamination of ν_e events. If the ν_e cross-section is significantly different from the MC prediction, this would affect the OOFV re-weighting and introduce a bias. However, as the contamination is small, this effect will not significantly affect the unfolded cross-section result. Estimates show that if the true ν_e cross-section was 10% lower than the model, a 0.1% bias would be introduced due to the ν_e contamination in the γ sample.

Plots to help visualise the re-weighting method, along with fake data studies to validate it, are detailed in Section 6.4.2.

6.4 Fake data studies

6.4.1 Fake datasets used

Three main fake datasets are used for testing the unfolding.

- The BANFF pre-fit. This is the distribution of events that enter the ν_e sample when detector systematics are applied, and events are re-weighted by the BANFF pre-fit tweaks. The BANFF pre-fit distribution will be used to 'generate' the unfolding,[1] so using this also as the fake dataset tests for pathological problems in the unfolding—if the unfolded result does not exactly agree with the input, then there is a bug in the code.
- The BANFF post-fit. This is the same as the BANFF pre-fit, but with the BANFF post-fit flux and cross-section weights applied. This fake dataset is compared to the BANFF pre-fit MC prediction in Figure 6.10.
- The "crazy signal" dataset. This is the same as the BANFF post-fit, but the shape and normalisation of the signal is significantly modified. Signal events with $Q^{2,\text{true}} < 1\,\text{GeV}^2/c^4$ are weighted down by a factor $0.5(1 + Q^{2,\text{true}})$. That means that events at $Q^{2,\text{true}} = 1\,\text{GeV}^2/c^4$ have the normal weight, but those at $Q^{2,\text{true}} = 0\,\text{GeV}^2/c^4$ are weighted by a factor of 0.5, with a linear scaling between 0 and $1\,\text{GeV}^2/c^4$. This fake dataset is compared to the BANFF pre-fit MC prediction in Figure 6.11

6.4.2 OOFV re-weighting

The method for re-weighting the OOFV background was explained in Section 6.3.4, and a special fake dataset (the "reduced OOFV" dataset) is used to test the method. The BANFF pre-fit distribution is used, but with the OOFV background changed. Specifically, the $x^{\text{NC1}\pi^0}$ and $x^{\text{NC other}}$ dials in T2KReWeight are tweaked to be -3σ from their nominal value, but only for OOFV events. The p_e^{reco} distributions of the ν_e and γ samples with this fake dataset are shown in Figure 6.12. R_γ^{2D} for this fake dataset is shown in Figure 6.13, and the extracted $R_{\nu_e}^X$ for the p_e^{reco} distribution of the ν_e sample is shown in Figure 6.14. Also shown in Figure 6.14 is the effect of applying the re-weighting to the ν_e sample.

[1] The phrase 'generate the unfolding' will be used as shorthand to mean generating the signal matrix, missed vector and background vector.

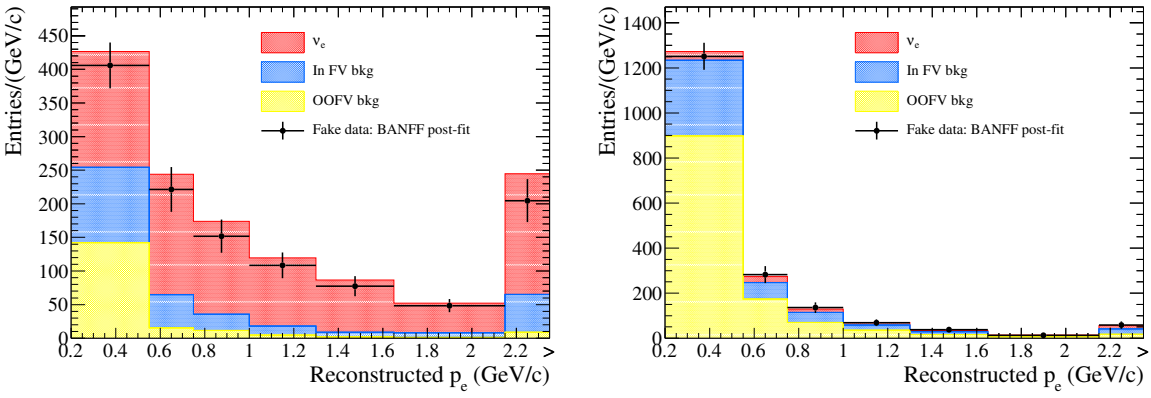


Figure 6.10: p_e distributions for the ν_e (left) and γ samples (right) comparing the BANFF pre-fit MC and the BANFF post-fit fake dataset. The BANFF fit reduces the predicted ν_e cross-section and flux, but does not significantly reduce the OOFV background.

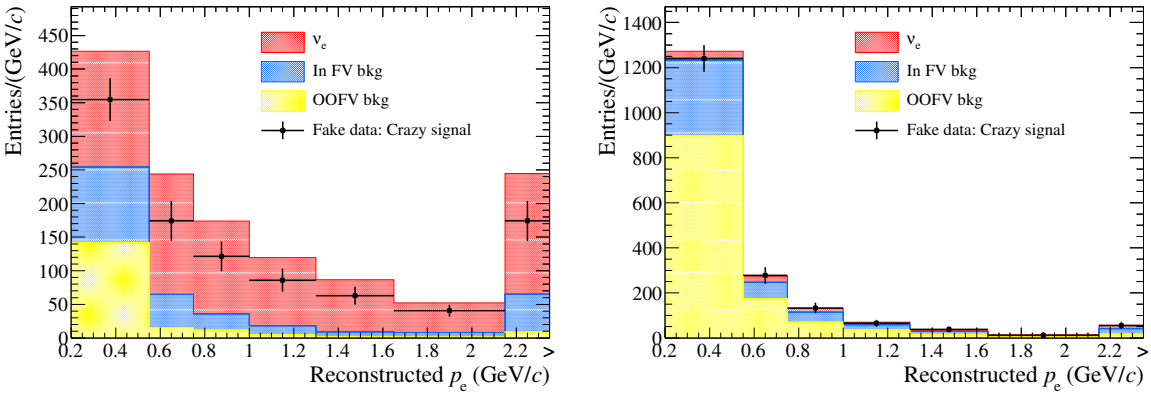


Figure 6.11: p_e distributions for the ν_e (left) and γ samples (right) comparing the BANFF pre-fit MC and the "crazy signal" fake dataset. The fake dataset reduces the ν_e cross-section, but does not affect the In FV or OOFV backgrounds.

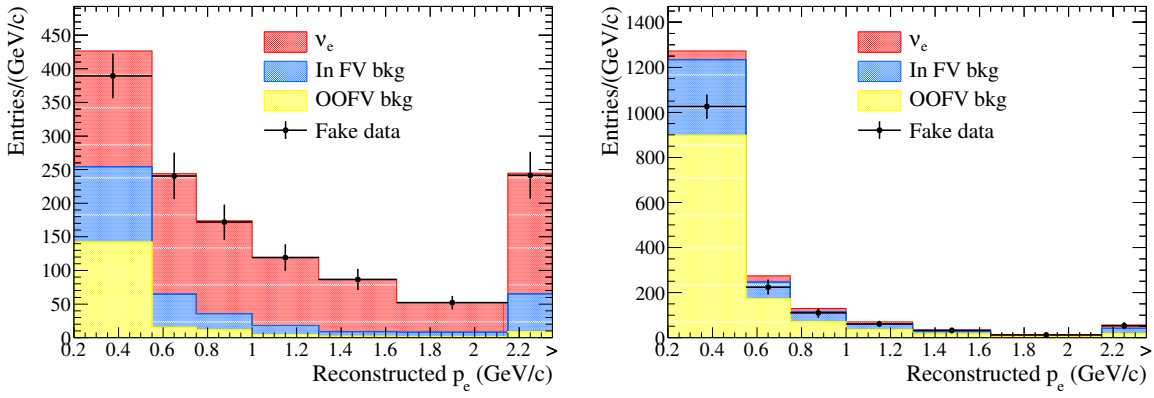


Figure 6.12: Comparison between BANFF pre-fit MC and fake dataset use to test the OOFV re-weighting technique for the ν_e sample (left) and γ sample (right). In this fake dataset, the OOFV background is reduced by tweaking the $x^{\mathrm{NC1}\pi^0}$ and $x^{\mathrm{NC\ other}}$ dials in T2KReWeight.

For completeness, the corresponding plots for the γ sample are shown in Figure 6.15. Comparing the re-weighting distributions for the two samples shows that they are indeed different, highlighting the different (p_e^{reco}, $\cos(\theta_e^{\mathrm{reco}})$) phase space probed. The data/MC comparisons show that the re-weighting procedure works well, with the re-weighted MC prediction agreeing with the fake data.

The error bars shown in the $R_{\nu_e}^X$ distributions are the propagated data statistical errors from the γ sample (p_e^{reco}, $\cos(\theta_e^{\mathrm{reco}})$) distribution. Although the errors appear to be very large in some bins, these are also the bins in which there are few OOFV events. These large errors will therefore not have such a significant impact on the unfolded number of ν_e interactions. Conversely, the error in the first bin is very small compared to how far the OOFV ratio is from 1. The method used to determine the systematic uncertainty on the OOFV re-weighting will be explained in Section 6.4.4.

As well as p_e^{reco} and $\cos(\theta_e^{\mathrm{reco}})$, the selection efficiency depends on the start position of the electron. There is significant correlation between the variables however, and to check that re-weighting in (p_e^{reco}, $\cos(\theta_e^{\mathrm{reco}})$) space is sufficient, the x_e^{reco}, y_e^{reco} and z_e^{reco} distributions—the x, y and z components of the reconstructed start position of the track—before and after re-weighting can be examined. Figure 6.16 shows the distributions in the γ sample before and after re-weighting. The MC distributions after re-weighting agree well with the data, so re-weighting in (p_e^{reco}, $\cos(\theta_e^{\mathrm{reco}})$) space is sufficient, and re-weighting in x_e^{reco}, y_e^{reco}, or z_e^{reco} is unnecessary.

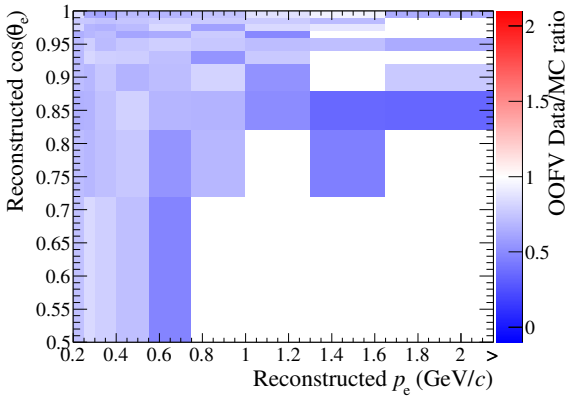


Figure 6.13: R_{γ}^{2D} when MC is BANFF pre-fit, and fake dataset has OOFV background reduced.

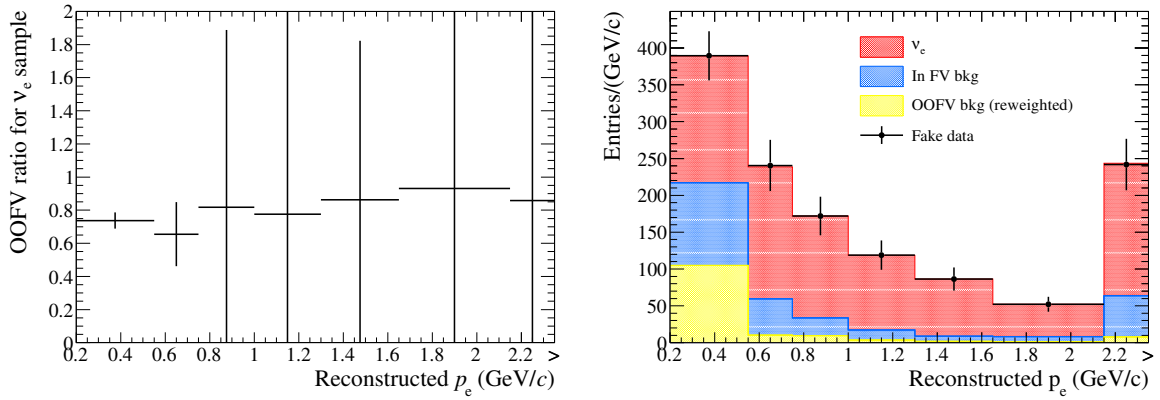


Figure 6.14: $R_{\nu_e}^{X}$ for the ν_e sample (left) and the re-weighted MC prediction (right). The error bars in the left plot are the propagated uncertainty on the number of data events, and are shown for reference. The actual error associated with the OOFV re-weighting is explained in Section 6.4.4. The right plot shows that the MC is successfully re-weighted to the fake dataset.

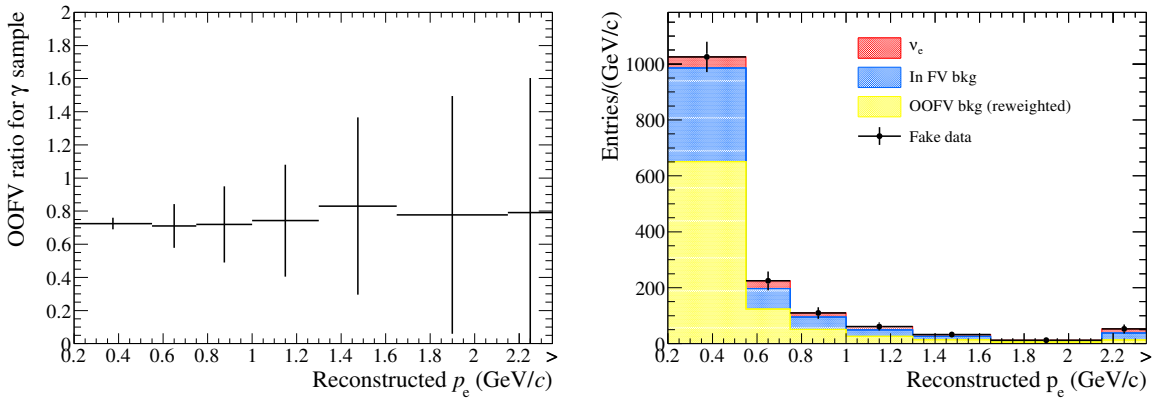


Figure 6.15: R_γ^X for the γ sample (left) and the re-weighted MC prediction (right). The error bars in the left plot are the propagated data statistics.

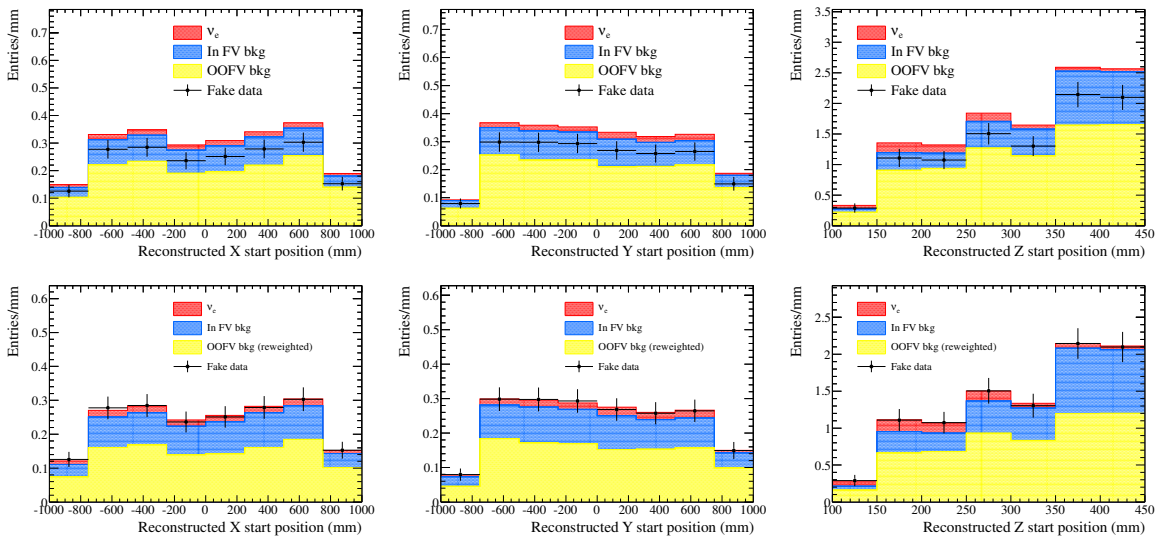


Figure 6.16: x_e^{reco}, y_e^{reco} and z_e^{reco} distributions of the γ sample before (top) and after (bottom) re-weighting, when the "reduced OOFV" fake dataset is used.

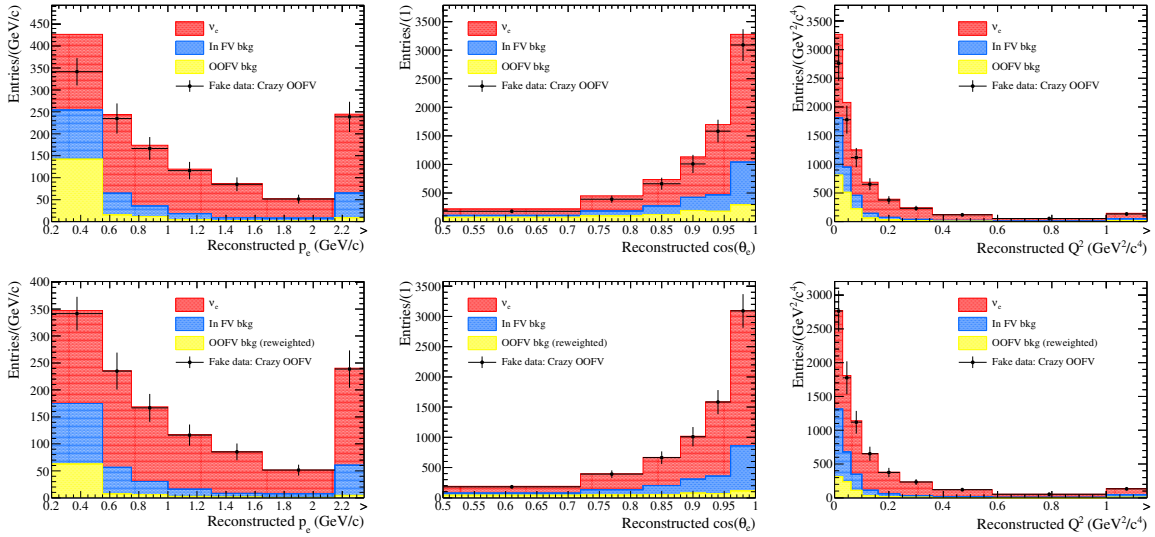


Figure 6.17: ν_e sample before and after OOFV re-weighting when the fake data is the "crazy OOFV" model.

When re-weighting in $(p_e^{\text{reco}}, \cos(\theta_e^{\text{reco}}))$ space, the assumption is made that events in those bins come from the same origins in the ν_e and γ samples. One way to test this assumption is to make another fake dataset where the OOFV is re-weighted based on the true momentum and angle of the photon. This fake dataset will be referred to as the "crazy OOFV" model. In this fake dataset, events are weighted to the BANFF pre-fit, then true OOFV events have the following weights applied multiplicatively:

- If the true photon momentum, p_γ^{true}, is less than 2000 MeV/c, the weight is $1 - (2000 - p_\gamma^{\text{true}})/4000$. True photons at 0 MeV/$c$ have a weight of 0.5; true photons above 2000 MeV/c have a weight of 1.
- If the true photon angle, $\cos(\theta_\gamma^{\text{true}})$, is greater than 0.5, the weight is $1.5 - \cos(\theta_\gamma^{\text{true}})$. True photons below 0.5 have a weight of 1; true photons at +1 have a weight of 0.5.

The effect of the re-weighting on this fake dataset is shown for the ν_e and γ samples in Figure 6.17 and Figure 6.18, respectively. The re-weighting procedure is successfully able to re-weight the ν_e sample.

6.4.3 Bias and number of iterations

As mentioned in Section 6.3.1, the literature recommends approximately 3 unfolding iterations should be run in the case of infinite statistics. 1 iteration was found to be

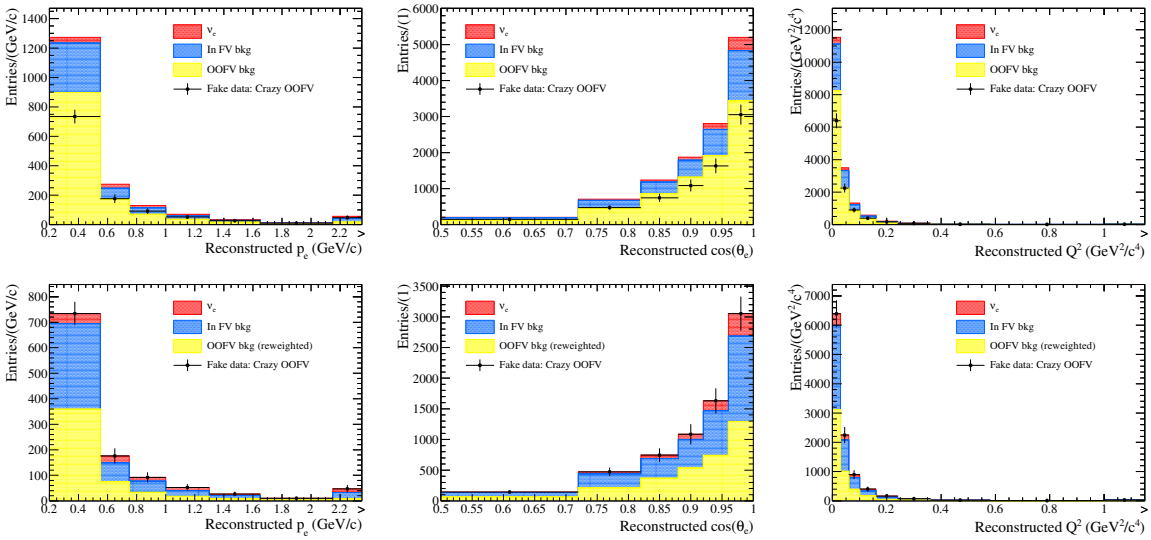


Figure 6.18: γ sample before and after OOFV re-weighting when the fake data is the "crazy OOFV" model.

optimum for the T2K ν_μ CC inclusive measurement due to the low statistics in that dataset. For this analysis, which has even fewer events than the ν_μ analysis, 1 iteration would again be expected to be optimal.

The choice of how many iterations to perform is based on studying the bias of the unfolding method and the fractional statistical error as the number of iterations is increased. The bias is the fractional deviation of the unfolded differential cross-section from the true cross-section of the fake dataset,

$$\text{Bias} = \frac{N^m_{t_k} - N^{\text{true}}_{t_k}}{N^{\text{true}}_{t_k}}, \tag{6.21}$$

where $N^{\text{true}}_{t_k}$ is the true number of events in bin t_k for the fake dataset being tested. The statistical error is the quadratic sum of the data statistical and MC statistical errors, which are described in Section 6.4.4. In Figure 6.19, the BANFF pre-fit is used for both generating the unfolding and as the fake dataset. The negligible bias ($10^{-10}\%$) shows that there is no pathological bug in the unfolding code, and the correct cross-section is extracted when using the same data for generation and unfolding. As expected, the statistical error increases with the number of iterations.

A more thorough test of the unfolding routine is shown in Figure 6.20, where the BANFF post-fit is used as the fake dataset. The statistical error increases with the number of iterations, as expected, but there is also a slight bias, which again increases

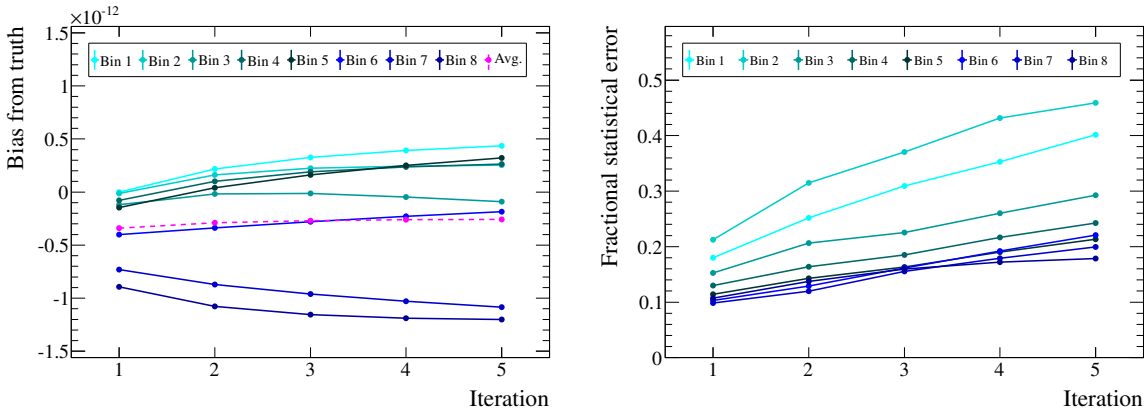


Figure 6.19: Bias from the true BANFF pre-fit (left) and fractional statistical error (right) in each p_e^{true} bin when generating with the BANFF pre-fit, and using BANFF pre-fit as the fake dataset. Note the y-axis scale on the bias plot.

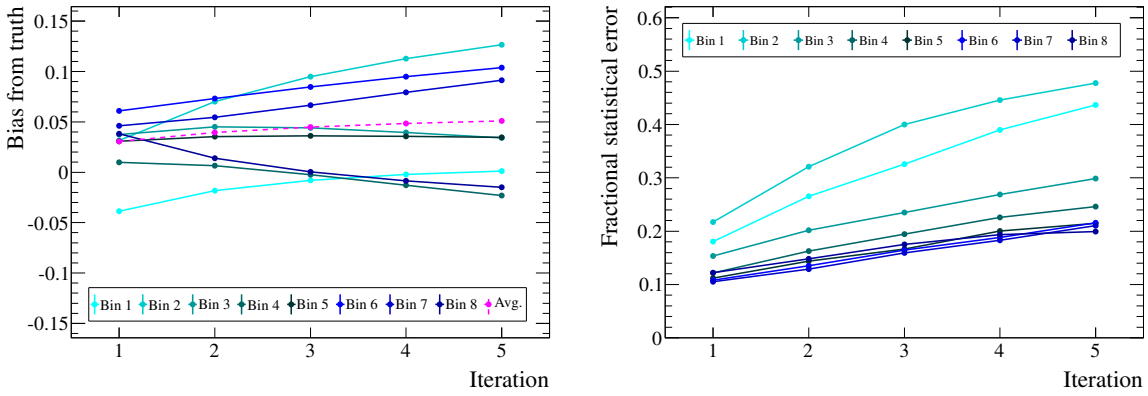


Figure 6.20: Bias from the true BANFF post-fit (left) and fractional statistical error (right) in each p_e^{true} bin when generating with the BANFF pre-fit, and using BANFF post-fit as the fake dataset.

with more iterations. Here, the bias is the fractional difference from the BANFF post-fit cross-section prediction, and is approximately 3% for 1 iteration. 1% of this bias is expected as the BANFF post-fit changes the ν_e flux by 1%, whilst the unfolding assumes the BANFF pre-fit flux is correct. A further bias is also expected as the BANFF post-fit re-weights the background contribution in the fake data. As there is less background in the fake data than in the MC, the unfolded ν_e cross-section is expected to be biased slightly high. Finally, as the 3% bias is small compared to any systematic uncertainty, it is not a concern.

It is also interesting to examine how necessary the unfolding procedure is. Figure 6.21 shows the same information as Figure 6.20, but also includes a "0 iteration" result.

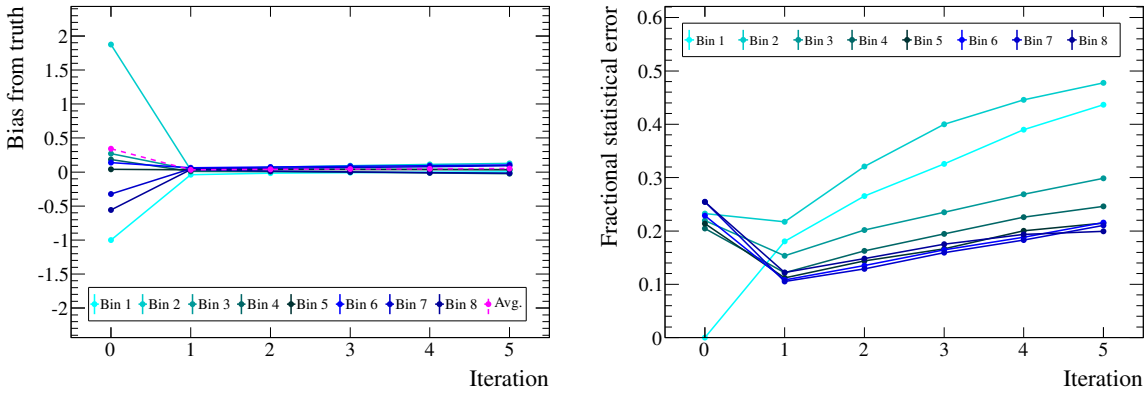


Figure 6.21: Bias from the true BANFF post-fit (left) and fractional statistical error (right) in each p_e^{true} bin when generating with the BANFF pre-fit, and using BANFF post-fit as the fake dataset. These plots include a "zero-iteration" result, where no unfolding is done and the background-subtracted data is simply efficiency-corrected.

This result is found by simply taking the background-subtracted data and correcting for the efficiency in each bin. For the p_e case shown in the figure, this means that the cross-section in the 0–200 MeV/c bin is zero, as there is no data in that bin. Significant biases are present in the "0 iteration" result, indicating that the unfolding procedure is absolutely necessary.

A further test of the method is performed by using the "crazy signal" fake dataset, in which the ν_e signal shape and normalisation is significantly modified. Figure 6.22 shows the bias and statistical error as a function of the number of iterations, and again the bias is not significantly reduced by applying more iterations, whilst the statistical error still increases. This further justifies the choice of using a single iteration. The size of the bias will be discussed in more detail in Section 6.4.5, where it is explained that the bias is small compared to the difference between the model predictions, and is also small compared to the uncertainties.

6.4.4 Uncertainties

All statistical and systematic uncertainties are computed through the covariance matrix method. M throws are performed, in which all the inputs to the unfolding—such as the smearing matrix and initial estimator—are recomputed each time. The unfolded number of events is recomputed for each throw, and the fractional covariance matrix is

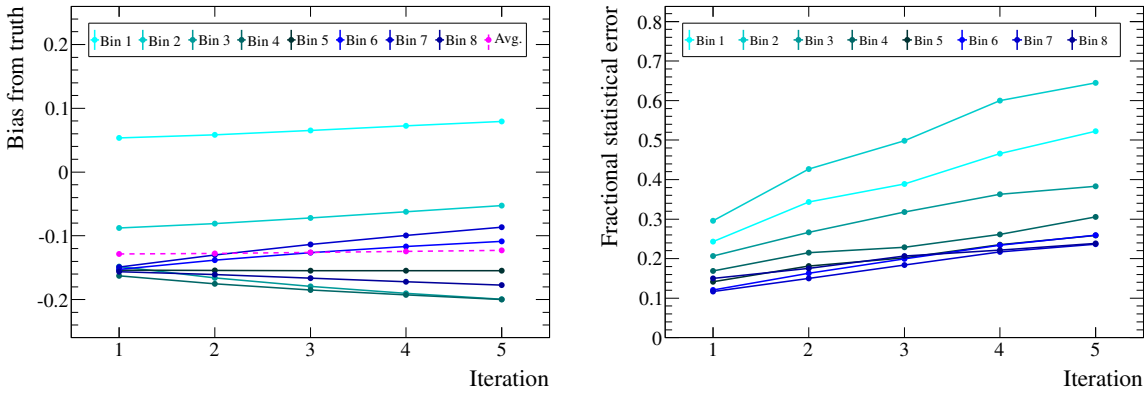


Figure 6.22: Bias (left) and statistical error (right) when generating with the BANFF pre-fit, and unfolding using the "crazy signal" model as the fake dataset.

then given by

$$V_{kl}^{(s)} = \frac{1}{M}\sum_{i=1}^{M}\frac{(\widehat{N}_{t_k}^{(s_i)} - \widehat{N}_{t_k}^{(\mathrm{nom})})(\widehat{N}_{t_l}^{(s_i)} - \widehat{N}_{t_l}^{(\mathrm{nom})})}{\widehat{N}_{t_k}^{(\mathrm{nom})}\,\widehat{N}_{t_l}^{(\mathrm{nom})}}, \tag{6.22}$$

where (s) is the current source of uncertainty being considered, (s_i) is the throw, $\widehat{N}_{t_k}^{(s_i)}$ is the unfolded number of events in bin t_k for that throw, and $\widehat{N}_{t_k}^{(\mathrm{nom})}$ is the nominal number of unfolded events in bin t_k. For plots of the differential cross-sections, the square-root of the diagonal elements of the covariance matrices are plotted for the error bars. In this analysis, 999 throws are performed for each source of uncertainty.

The uncertainty on the total flux-averaged cross-section, which includes information from all t_k bins, is calculated as

$$\sigma_{\widehat{N}_{\mathrm{total}}} = \sqrt{\sum_{k=1}^{n_t}\left(\sum_{l=1}^{k-1} 2V_{kl}\widehat{N}_{t_k}\widehat{N}_{t_l}\right) + V_{kk}\widehat{N}_{t_k}\widehat{N}_{t_k}}. \tag{6.23}$$

It is important to note that in general the systematic throws affect both the ν_e and γ samples. The only exception is the OOFV systematic, which is calculated using the γ sample and applied to the ν_e sample.

Data statistics

Data statistics are simply considered by varying the contents of each p_e^{reco}, $\cos(\theta_e^{\text{reco}})$ and $Q^{2,\text{reco}}$ bin in the ν_e sample according to Poisson statistics. The γ sample is Poisson-varied in the finer (p_e^{reco}, $\cos(\theta_e^{\text{reco}})$) binning used for the OOFV re-weighting.

MC statistics

The following MC inputs are separately Poisson-varied to evaluate the MC statistical error.

- The signal matrix.
- The background vector.
- The missed vector.
- The ν_e, OOFV and "other" portions of the γ sample prediction.

Detector systematics

The effect of the detector systematics described in Section 5.1 are evaluated by varying the low-level parameters (such as the TPC momentum bias) for each throw, propagating the effects to each event, and re-performing the selection. Events can migrate between bins, between samples, and into and out of samples completely. The uncertainty on the number of nucleons in the FGD FV is included as part of the detector systematics, and is 0.67%.

As mentioned previously, true ν_e events entering the ν_e sample with a detector systematic weight w_d are also counted as missed events with a weight $1 - w_d$, in order to conserve the total number of simulated ν_e interactions.

OOFV systematic

The systematic on the OOFV re-weighting is chosen such that the data/MC scaling required in each p_e^{reco}, $\cos(\theta_e^{\text{reco}})$ or $Q^{2,\text{reco}}$ bin defines the 3σ uncertainty. For example, if a specific bin is nominally weighted down by -30%, the OOFV throws come from a Gaussian with mean -30% and width 10%. Note that the statistical uncertainties on

the data and MC in the γ sample are already accounted for, as described in previous sections, and this is simply an extra systematic to account for the fact that the source of the discrepancy is not well understood. The same logic was used to define the size of the momentum resolution systematic uncertainty in the ND280 detector systematics.

Flux and cross-section systematics

The BANFF pre-fit uncertainty matrix is used to define the uncertainty on the flux and cross-section parameters, as described in Section 5.2.1. All parameters are thrown together, and each event is re-weighted accordingly. All background events that enter the selection are re-weighted, along with all signal events regardless of whether they enter the selection. This means that the flux and cross-section systematics affect the efficiency and other Monte Carlo inputs to the unfolding, with the BANFF pre-fit uncertainties constraining how much these inputs can be varied. The flux and cross-section parameters cannot cause events to migrate between bins, just how much each event contributes to the sample.

Due to the implementation of the FSI re-weighting, it is possible for some events with many final state particles to be given absurdly large weights for certain parameter throws. For example, if the FSI charge exchange parameter is tweaked to 0, each quasi-elastic final state vertex in an interaction is given a weight of 2.59. An interaction with 9 such vertices is therefore given a weight of over 5000. To prevent these rare edge cases from significantly affecting the covariance matrix calculation, the flux and cross-section weight for each event is constrained to not be larger than 50.

Remembering that the cross-section is proportional to $\widehat{N}_{t_k}/\phi$ (see equation (6.6)), the covariance matrix defined in equation (6.22) is not appropriate when there are systematics that can alter the integrated flux. Instead, the following fractional covariance matrix is used

$$
\begin{aligned}
V_{kl}^{(s)} &= \frac{1}{M}\sum_{i=1}^{M} \frac{(\frac{\widehat{N}_{t_k}^{(s_i)}}{\phi^{(s_i)}} - \frac{\widehat{N}_{t_k}^{(\mathrm{nom})}}{\phi^{(\mathrm{nom})}})(\frac{\widehat{N}_{t_l}^{(s_i)}}{\phi^{(s_i)}} - \frac{\widehat{N}_{t_l}^{(\mathrm{nom})}}{\phi^{(\mathrm{nom})}})}{\frac{\widehat{N}_{t_k}^{(\mathrm{nom})}}{\phi^{(\mathrm{nom})}}\frac{\widehat{N}_{t_l}^{(\mathrm{nom})}}{\phi^{(\mathrm{nom})}}} \\
&= \frac{1}{M}\sum_{i=1}^{M} \frac{(\widehat{N}_{t_k}^{(s_i)}\frac{\phi^{(\mathrm{nom})}}{\phi^{(s_i)}} - \widehat{N}_{t_k}^{(\mathrm{nom})})(\widehat{N}_{t_l}^{(s_i)}\frac{\phi^{(\mathrm{nom})}}{\phi^{(s_i)}} - \widehat{N}_{t_l}^{(\mathrm{nom})})}{\widehat{N}_{t_k}^{(\mathrm{nom})}\widehat{N}_{t_l}^{(\mathrm{nom})}},
\end{aligned}
\tag{6.24}
$$

where $\phi^{(\mathrm{nom})}$ is the nominal integrated flux, and $\phi^{(s_i)}$ is the integrated flux for throw (s_i).

6.4.5 Fake data results

Full phase-space analysis

Differential and total flux-averaged cross-sections when generating with the BANFF pre-fit and unfolding with the BANFF post-fit fake dataset are shown in Figure 6.23. The fractional uncertainties and correlation matrices are shown in Figure 6.24. Overall, an error of approximately 20% is expected on the total flux-averaged cross-section measurement.

The unfolded flux-averaged results are slightly different depending on which variable was used in the unfolding. The BANFF post-fit prediction is 1.19×10^{-38} cm^2/nucleon, and the unfolded results are

- $\langle\sigma\rangle_\phi = 1.22 \pm 0.24 \times 10^{-38}$ cm^2/nucleon when unfolding through p_e
- $\langle\sigma\rangle_\phi = 1.21 \pm 0.24 \times 10^{-38}$ cm^2/nucleon when unfolding through $\cos(\theta_e)$
- $\langle\sigma\rangle_\phi = 1.20 \pm 0.22 \times 10^{-38}$ cm^2/nucleon when unfolding through Q^2.

The slight disagreement between the three results is expected, as although the total number of events are the same in each distribution, they are distributed among bins with different efficiencies.

Differences in the uncertainties are also to be expected. To see whether the smaller error in the Q^2 unfolding is significant, rather than just due to the precise fake dataset being studied, 100 fake datasets were made based on the BANFF post-fit covariance matrix. The fractional error for each of these datasets is shown in Figure 6.25, and the key features are discussed below.

- Unfolding through Q^2 gives a smaller statistical error, and this is consistent throughout all 100 toys. This lower statistical error is due to the fact that the $Q^{2,\text{true}}$ distribution has no bins with very low efficiency, unlike the low momentum and high angle bins. In such bins the statistical error is magnified by the $1/\epsilon_{t_k}$ factor in the unfolding. As 245 ν_e events are expected to be selected, a $1/\sqrt{245} = 6.4\%$ uncertainty would be expected for a simple counting experiment. The 8–13% uncertainties from the unfolding method are larger than this, as expected due to the use of the γ control sample and the binned analysis.

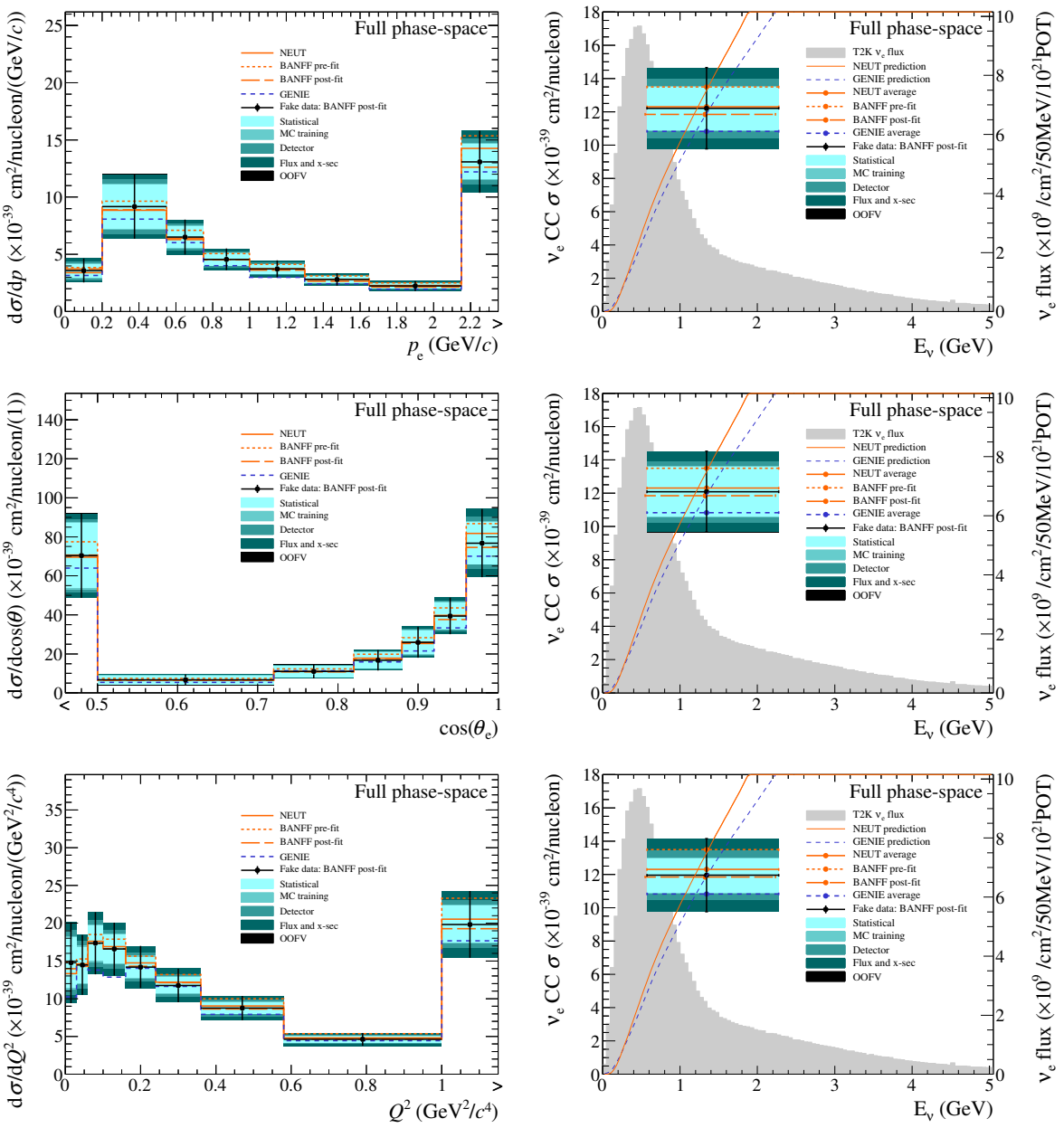


Figure 6.23: Differential cross-section (left) and resulting total flux-averaged cross-section (right), when generating with the BANFF pre-fit and using BANFF post-fit as the fake data, for p_e (top), $\cos(\theta_e)$ (middle) and Q^2 (bottom). The separate sources of uncertainty are added in quadrature.

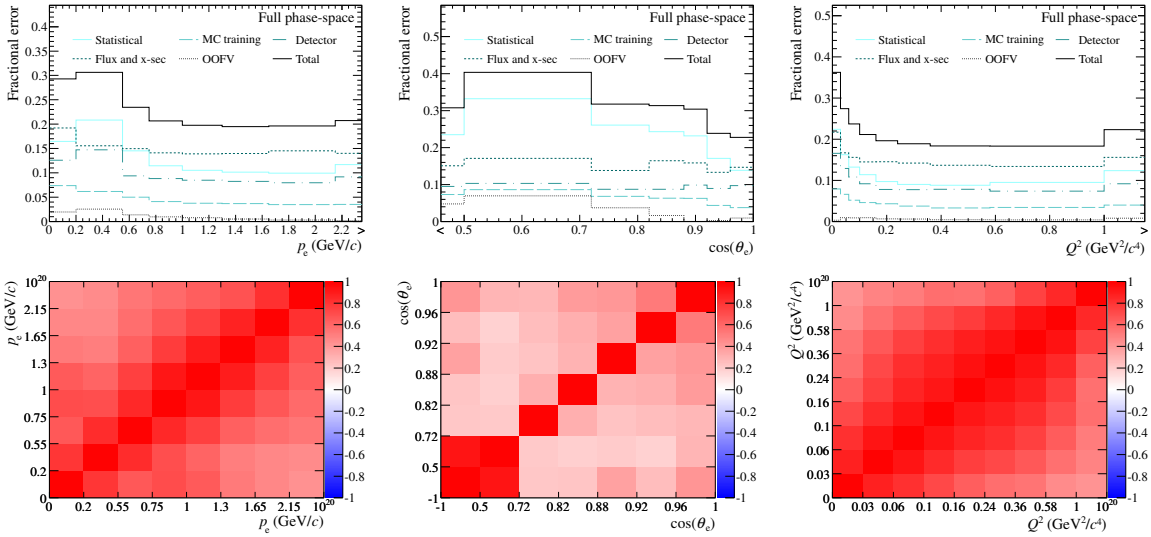


Figure 6.24: Fractional uncertainties (top) and correlation matrices (bottom) when generating with the BANFF pre-fit and using BANFF post-fit as the fake data, in p_e (left), $\cos(\theta_e)$ (middle) and Q^2 (right).

- As the Monte Carlo sample is 11 times larger than the data sample, the Monte Carlo statistical uncertainty would naively be expected to be $1/\sqrt{11} = 0.3$ times the size of the data statistical uncertainty. The actual uncertainty is found to be slightly larger than this, as the separate categories—true ν_e events, true γ events and so on—are fluctuated separately.

- The p_e unfolding gives a much larger detector systematic error than the other distributions, which is due to the detector systematics affecting the low momentum events more than any others, and the p_e^{reco} distribution concentrating all these events in one bin.

- The flux and cross-section uncertainty is found to be approximately 13%, and this is dominated by the overall flux uncertainty. The differences for unfolding through the different variables is due to the E_ν parameterisation used for the flux uncertainty and some of the cross-section uncertainties.

- The OOFV systematic is small ($< 3\%$) for the BANFF post-fit fake datasets, as the OOFV background is not significantly affected by the BANFF fit. In the real data, where a large disagreement is observed between data and MC in the γ sample, this systematic would be expected to be much larger.

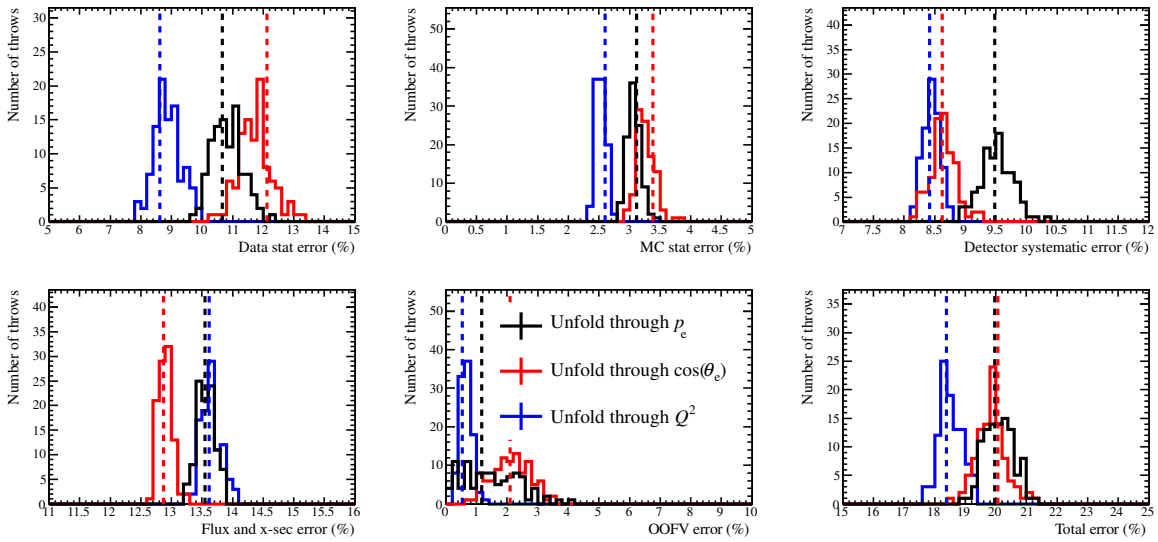


Figure 6.25: Histograms showing the percentage error on the flux-averaged cross-section for 100 toy datasets thrown from the BANFF post-fit covariance matrix. The dashed lines show the error on the standard BANFF post-fit fake dataset.

It is preferable to report only one number in certain situations, such as in conference summary talks. For this, an averaging of the three numbers is not necessary, nor is adding an extra systematic error to account for the differences. The differences are expected, and the 2% scale is much smaller than the current systematic uncertainties. Due to the lower expected systematic uncertainty on the Q^2 measurement, and the fact that Q^2 is the most interesting kinematic variable, the Q^2 measurement will be presented as the main result.

Section 6.4.3 showed that the result was biased significantly low when unfolding using the "crazy signal" fake dataset. The unfolded distributions are shown in Figure 6.26, along with the BANFF pre-fit and "crazy signal" model predictions, which are different in both shape and normalisation. Looking particularly at the Q^2 distribution, the new shape is not completely recovered by the unfolding, especially in the Q^2 >1 GeV$^2/c^4$ bin. However, due to the large uncertainties in the measurement, the vast majority of the unfolded data points agree with the input model within 1σ. As was shown previously, applying more than one iteration does not improve the bias. As the unfolded result agrees with the input model within errors, and this model is significantly different to anything expected in nature, the conclusion is made that the unfolding method works well.

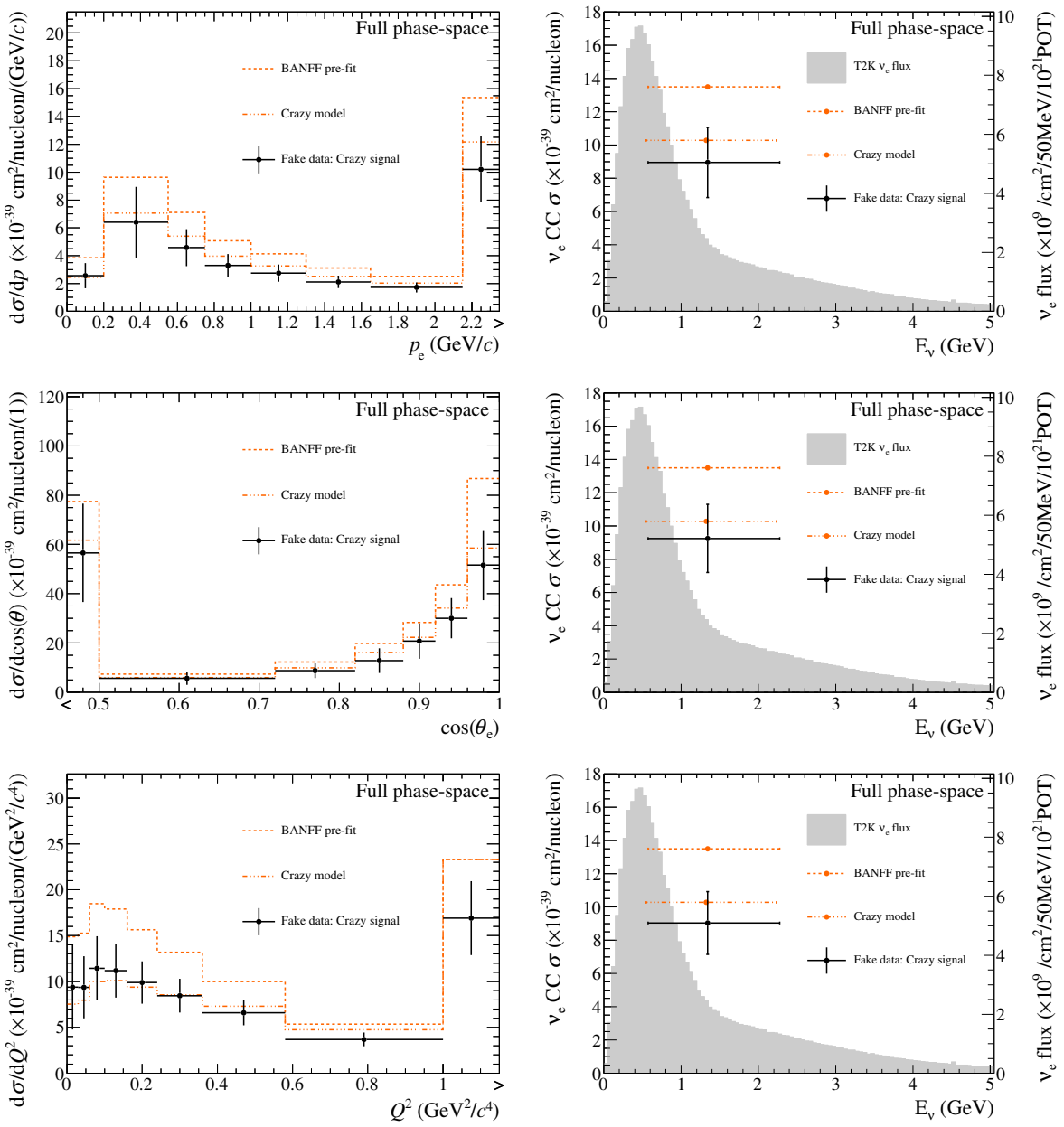


Figure 6.26: Unfolded differential cross-sections when the "crazy signal" fake dataset is analysed, when unfolding through p_e (top), $\cos(\theta_e)$ (middle) and Q^2 (bottom). The left column shows the differential cross-sections, and the right column the total flux-averaged cross-sections.

Restricted phase-space analysis

The results of unfolding the BANFF post-fit fake dataset in the restricted phase-space analysis are shown in Figure 6.27. The BANFF post-fit predicted cross-section is 6.89×10^{-39} cm^2/nucleon, and the unfolded results are

- $\langle\sigma\rangle_\phi = 6.94 \pm 1.29 \times 10^{-39}$ cm^2/nucleon when unfolding through p_e
- $\langle\sigma\rangle_\phi = 6.91 \pm 1.31 \times 10^{-39}$ cm^2/nucleon when unfolding through $\cos(\theta_e)$
- $\langle\sigma\rangle_\phi = 6.93 \pm 1.29 \times 10^{-39}$ cm^2/nucleon when unfolding through Q^2.

The result is again biased slightly high, but not significantly so. The fractional uncertainty on the measurement is similar to the result found when examining the full phase-space, as the reduction in systematic uncertainty has been offset by the increase in statistical uncertainty.

6.5 Results

This section describes the results of looking at real T2K data from Runs 1–4. The effect of the OOFV re-weighting is first described in Section 6.5.1, followed by the differential cross-section results in Section 6.5.2 and the total flux-averaged cross-section results in Section 6.5.3. The results are then discussed in Section 6.5.4.

6.5.1 OOFV re-weighting

The kinematic distributions of the data have already been shown in Figure 6.1, and a large deficit is seen in the low-momentum bins in both the ν_e and γ samples. The data/MC ratio for the OOFV re-weighting, R_γ^{2D}, is shown in Figure 6.28, and the effect of the re-weighting on the γ sample is shown in Figure 6.29 and Figure 6.30.

The effect of the re-weighting on the ν_e sample is shown in Figure 6.31. Although the data/MC disagreement is reduced, the data still sees fewer events than the Monte Carlo prediction, so the unfolded cross-sections are expected to be lower than those of the Monte Carlo. It is interesting to note that $R_{\nu_e}^X$ is 0 for some bins, suggesting that the data prefers for the In FV background to be reduced, as well as the OOFV. As explained previously, there are many systematics covering the In FV background.

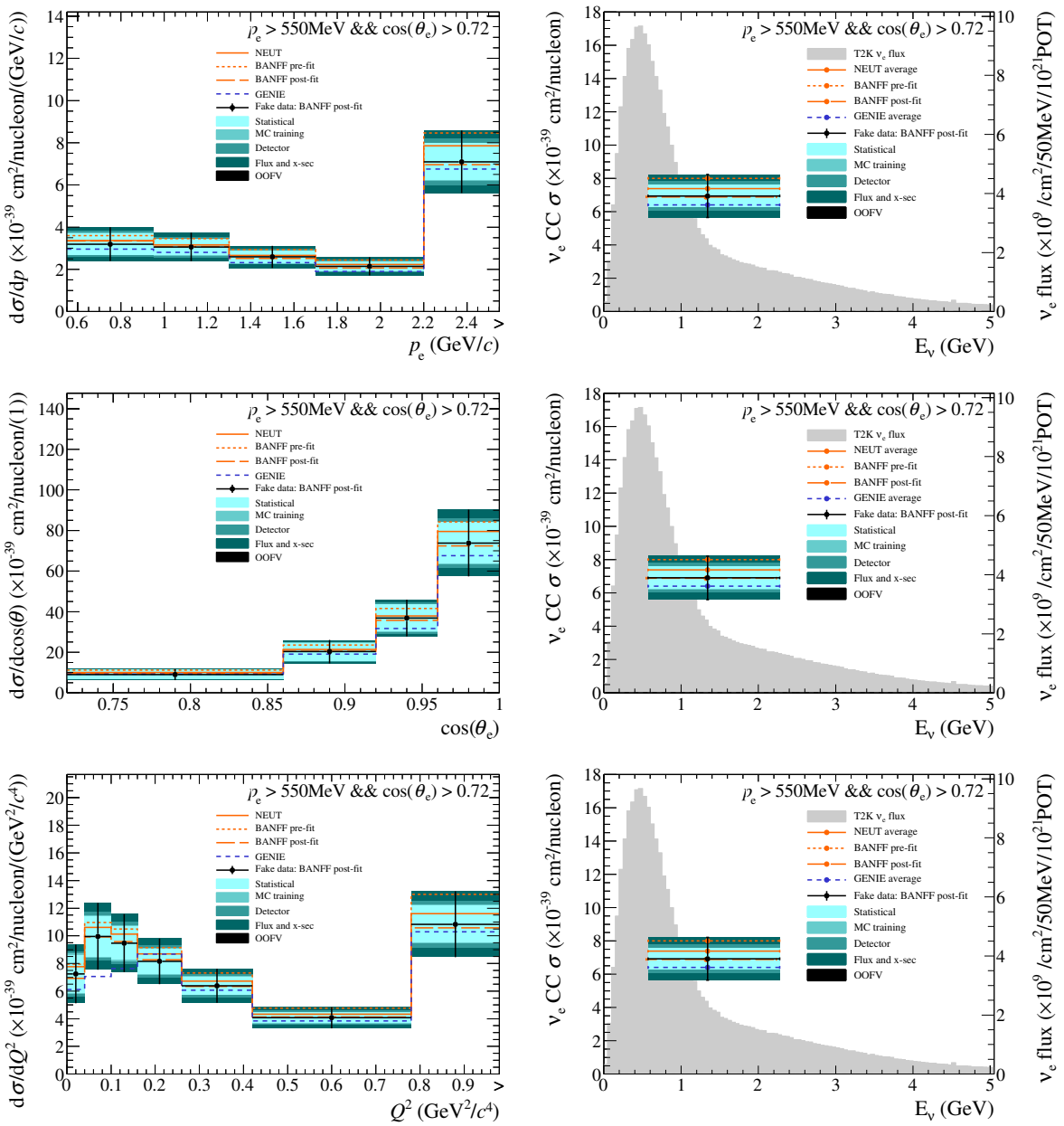


Figure 6.27: Differential cross-sections (left) and resulting total flux-averaged cross-sections (right), when generating with the BANFF pre-fit and using BANFF post-fit as the fake data, for p_e (top), cos(θ_e) (middle) and Q^2 (bottom), for the restricted phase-space analysis.

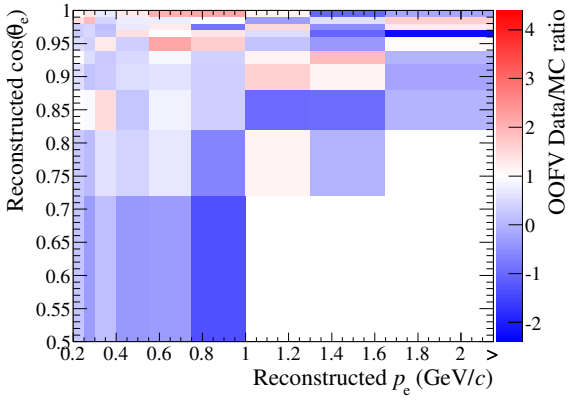


Figure 6.28: R_γ^{2D} when MC is BANFF pre-fit and real data is analysed.

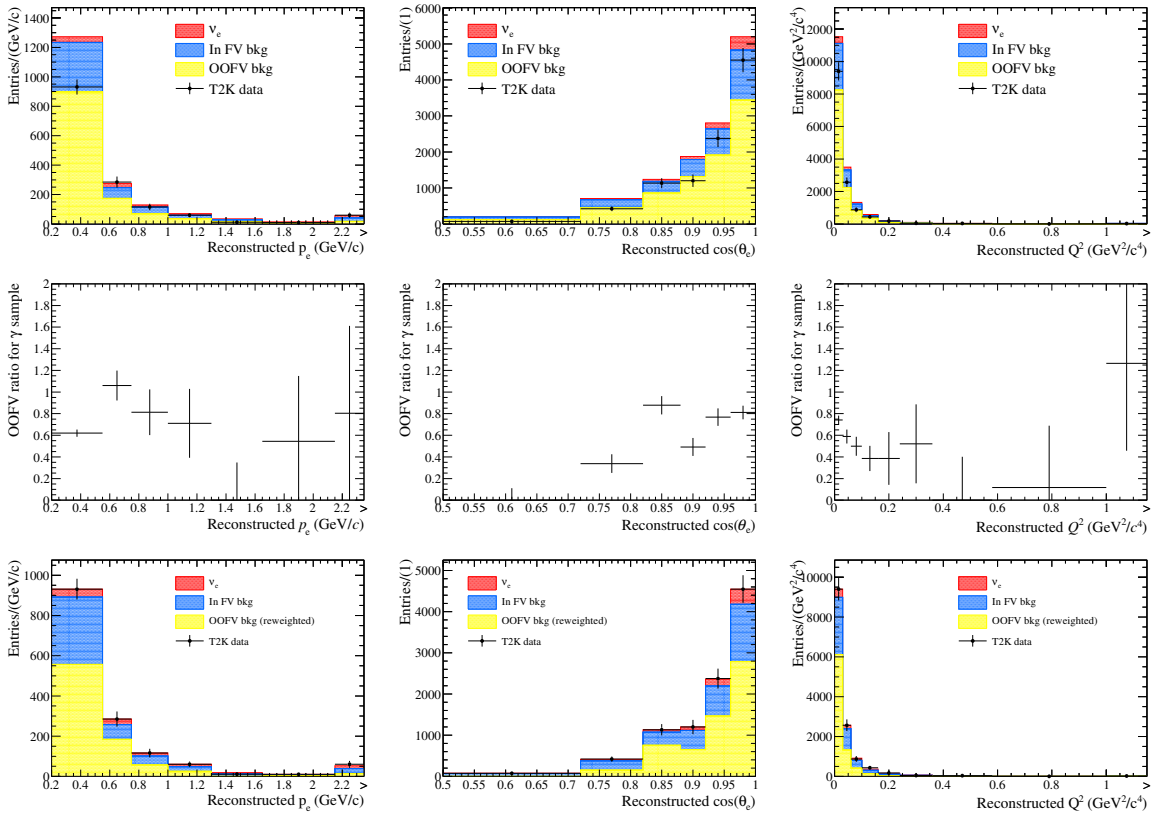


Figure 6.29: γ sample before (top) and after (bottom) the OOFV re-weighting, and R_γ^X (middle row) when MC is BANFF pre-fit and real data is analysed, for p_e^{reco} (left), $\cos(\theta_e^{\text{reco}})$ (middle column) and $Q^{2,\text{reco}}$ (right). The error bars in the middle row are the propagated data statistics, and are shown for reference.

The re-weighting method prevents these bins from going negative; if this restriction is removed, the unfolded cross-section results change by less than 1%.

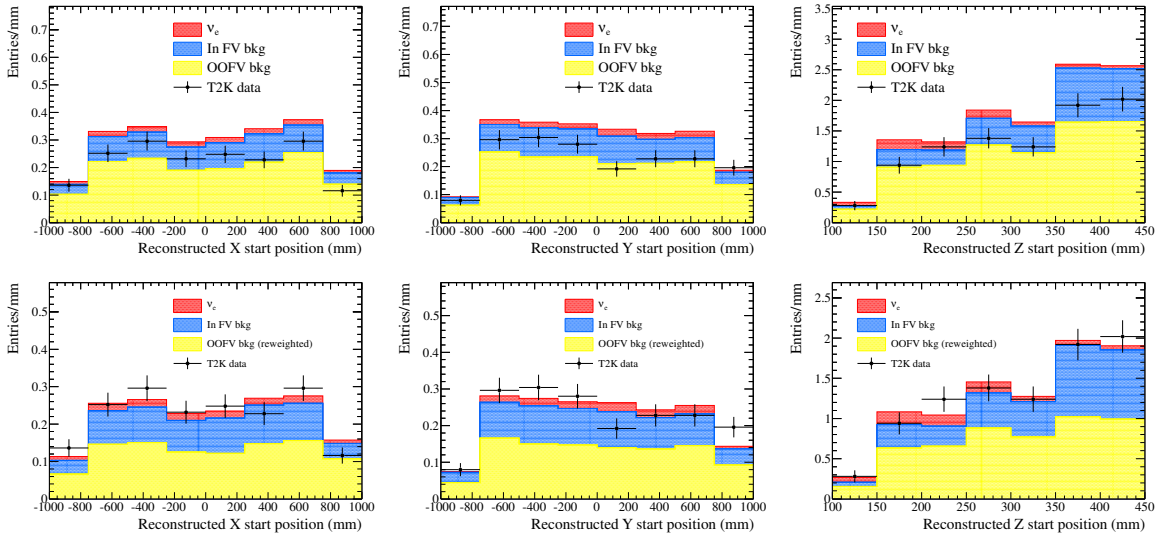


Figure 6.30: x_e^{reco}, y_e^{reco} and z_e^{reco} distributions of the γ sample before (top) and after (bottom) re-weighting.

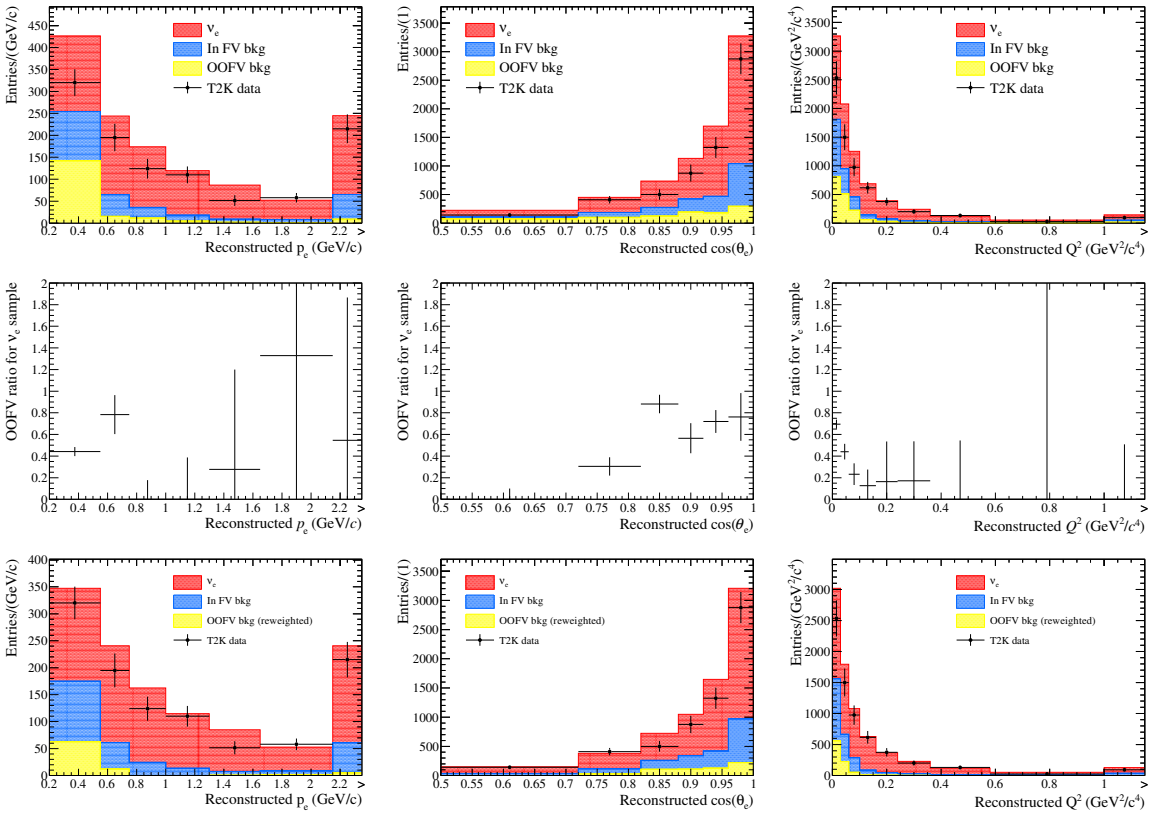


Figure 6.31: ν_e sample before (top) and after (bottom) the OOFV re-weighting, and $R^X_{\nu_e}$ (middle row) when MC is BANFF pre-fit and real data is analysed, for p_e^{reco} (left), $\cos(\theta_e^{\text{reco}})$ (middle column) and $Q^{2,\text{reco}}$ (right). The error bars in the middle row are the propagated data statistics, and are shown for reference.

6.5.2 Differential cross-section results

Full phase-space analysis

The differential cross-sections as a function of p_e, $\cos(\theta_e)$ and Q^2 are shown in Figure 6.32, and the fractional uncertainties and correlation matrices are shown in Figure 6.33. Tabulations of the fractional uncertainty in each bin are shown in Tables 6.5–6.7, and in these tables the flux and cross-section systematics have been separated.

It is interesting to note that there is a data deficit at low Q^2 compared to NEUT, which is theoretically the most interesting kinematic area, although each bin is within the 1σ uncertainty.

Restricted phase-space analysis

The differential cross-sections for the restricted phase-space analysis are shown in Figure 6.34, and it is interesting to note that the data deficit compared to the NEUT prediction at low Q^2 is still evident. The fractional uncertainties and correlation matrices are shown in Figure 6.35.

6.5.3 Flux-averaged cross-section results

Full phase-space analysis

The total flux-averaged CC inclusive ν_e cross-sections when unfolding through p_e, $\cos(\theta_e)$ and Q^2 are shown in Figure 6.36. The results for each variable are

- $\langle\sigma\rangle_\phi = 1.11 \pm 0.24 \times 10^{-38}$ cm^2/nucleon when unfolding through p_e
- $\langle\sigma\rangle_\phi = 1.13 \pm 0.23 \times 10^{-38}$ cm^2/nucleon when unfolding through $\cos(\theta_e)$
- $\langle\sigma\rangle_\phi = 1.11 \pm 0.20 \times 10^{-38}$ cm^2/nucleon when unfolding through Q^2.

The T2K data therefore agrees well with both the NEUT prediction of 1.23×10^{-38} cm^2/nucleon and the GENIE prediction of 1.08×10^{-38} cm^2/nucleon.

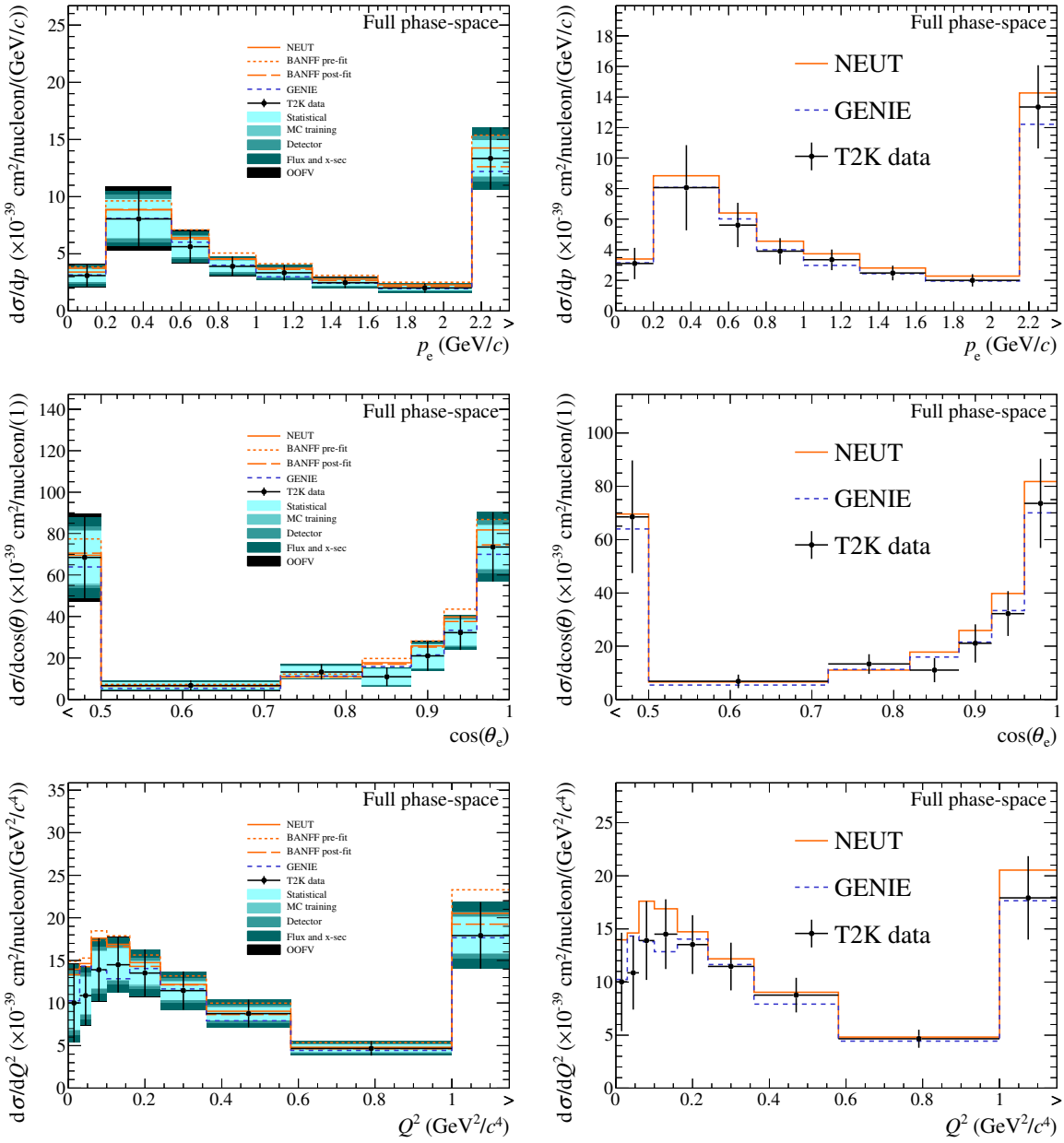


Figure 6.32: Differential cross-sections when generating with the BANFF pre-fit and using real data, for p_e (top), $\cos(\theta_e)$ (middle) and Q^2 (bottom). The right-hand plots are identical to the left-hand ones, but only show the NEUT and GENIE nominal predictions, and the data points with their total errors. These are the final results of the full phase-space differential cross-section.

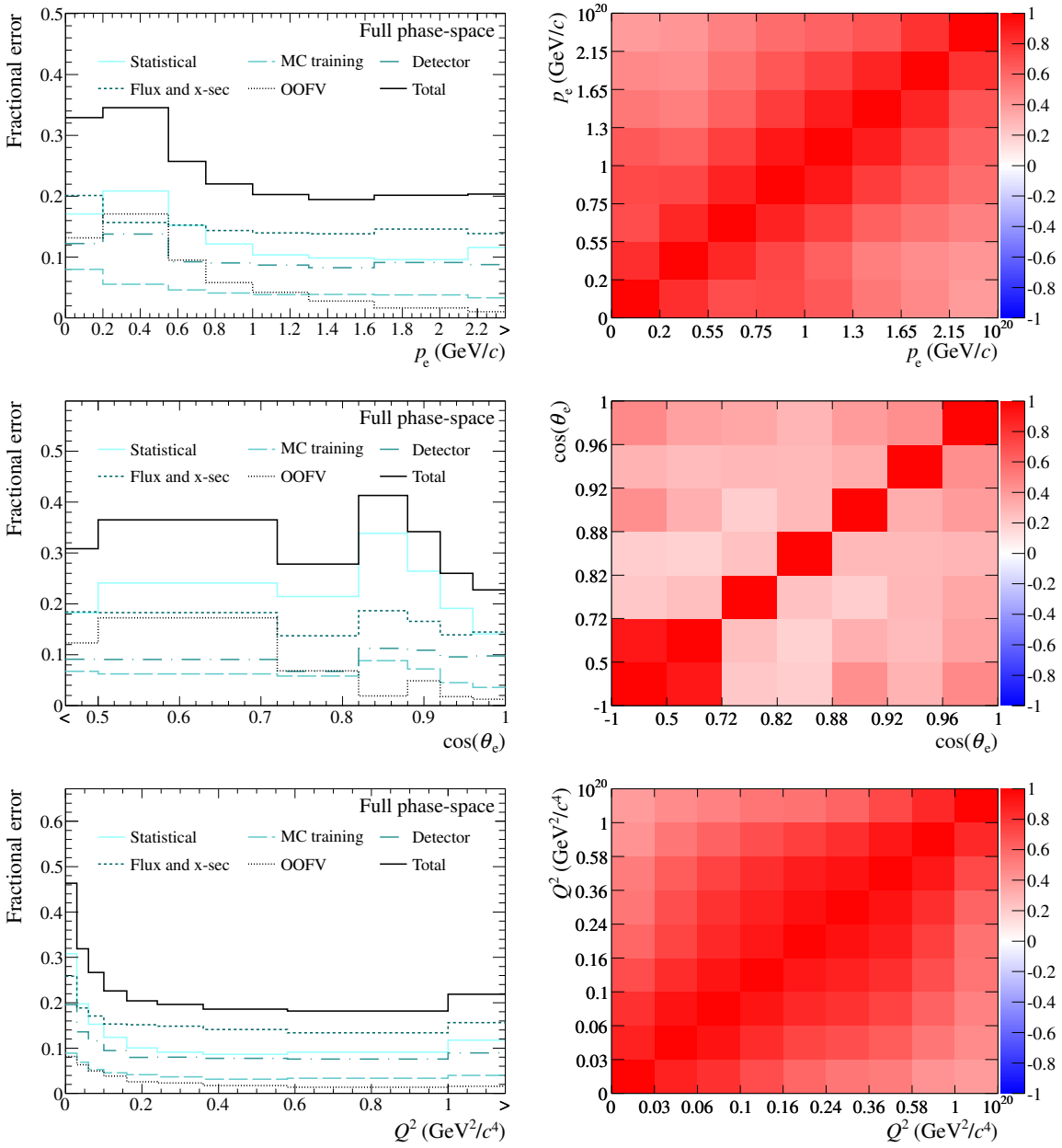


Figure 6.33: Fractional uncertainties (left) and correlation matrices (right) in bins of p_e (top), $\cos(\theta_e)$ (middle) and Q^2 (bottom) for the real data.

p_e (MeV/c)	Data stat.	MC stat.	Detector	Flux	X-sec	OOFV	Total
0–200	17.1	8.0	12.2	13.3	15.6	13.1	32.9
200–550	20.8	5.5	13.8	13.3	8.3	17.1	34.5
550–750	15.2	4.6	9.2	13.7	6.9	9.5	25.7
750–1000	12.1	4.1	9.0	13.7	5.2	5.8	22.0
1000–1300	10.4	3.8	8.7	13.4	4.8	4.2	20.3
1300–1650	9.8	3.9	8.2	13.1	4.8	2.8	19.5
1650–2150	9.6	3.8	9.1	12.8	6.1	1.6	20.1
> 2150	11.6	3.3	8.7	12.6	5.3	1.1	20.4
Total	10.5	2.8	9.2	13.0	7.6	7.5	21.1

Table 6.5: Fractional error (in %) when unfolding through p_e.

$\cos(\theta_e)$	Data stat.	MC stat.	Detector	Flux	X-sec	OOFV	Total
< 0.50	18.2	6.7	9.1	14.1	12.3	12.3	30.8
0.50–0.72	24.1	6.2	9.1	14.1	11.5	17.3	36.5
0.72–0.82	21.4	5.8	6.7	12.3	6.2	6.8	27.8
0.82–0.88	33.9	8.9	11.3	14.6	12.2	1.9	41.3
0.88–0.92	26.4	7.2	10.9	13.4	9.3	4.9	34.2
0.92–0.96	19.1	4.5	9.6	13.0	6.1	1.8	26.0
0.96–1.00	14.1	3.6	9.8	13.1	5.2	1.3	22.8
Total	10.4	2.9	8.5	13.3	5.7	5.5	20.6

Table 6.6: Fractional error (in %) when unfolding through $\cos(\theta_e)$.

Q^2 (GeV^2/c^4)	Data stat.	MC stat.	Detector	Flux	X-sec	OOFV	Total
0.00–0.03	30.8	8.9	19.6	17.6	18.9	8.2	46.3
0.03–0.06	19.7	6.9	13.6	15.1	11.3	6.3	31.9
0.06–0.10	15.3	5.3	11.6	14.2	9.7	4.9	26.7
0.10–0.16	12.4	4.6	9.4	13.3	7.5	3.8	22.6
0.16–0.24	10.0	4.1	7.9	12.6	8.2	2.5	20.4
0.24–0.36	9.1	3.7	8.0	12.6	8.1	2.3	19.6
0.36–0.58	8.6	3.2	7.7	12.5	6.4	1.7	18.6
0.58–1.00	9.1	3.3	7.6	12.5	4.9	1.4	18.2
> 1.00	11.7	4.0	8.9	13.2	7.9	1.6	21.9
Total	8.7	2.3	8.4	12.9	5.3	2.1	18.7

Table 6.7: Fractional error (in %) when unfolding through Q^2.

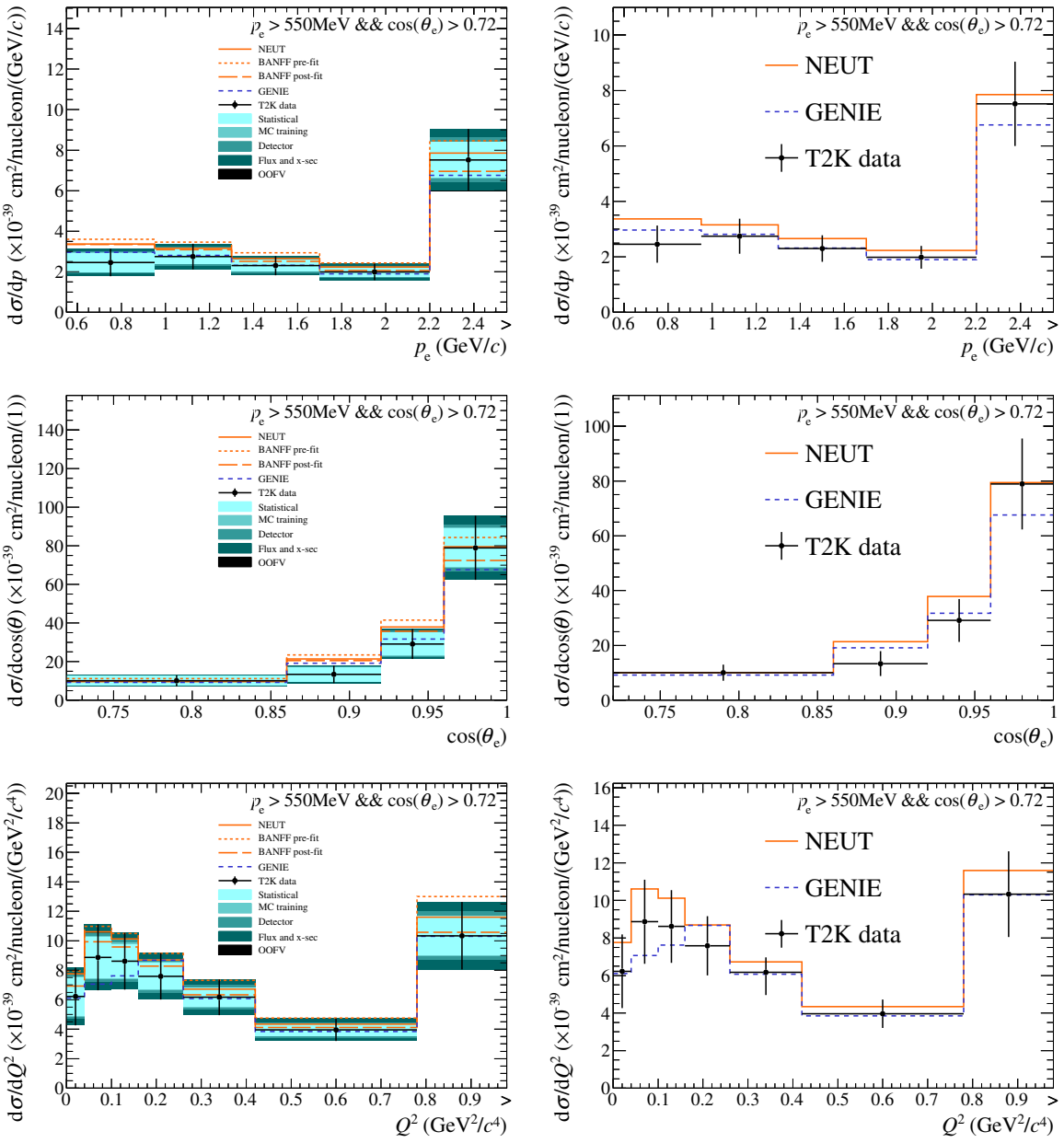


Figure 6.34: Differential cross-sections when generating with the BANFF pre-fit and using real data, for p_e (top), cos(θ_e) (middle) and Q^2 (bottom), for the restricted phase-space analysis. The right-hand plots are identical to the left-hand ones, but only show the NEUT and GENIE nominal predictions, and the data points with their total errors. These are the final results of the restricted phase-space differential cross-section.

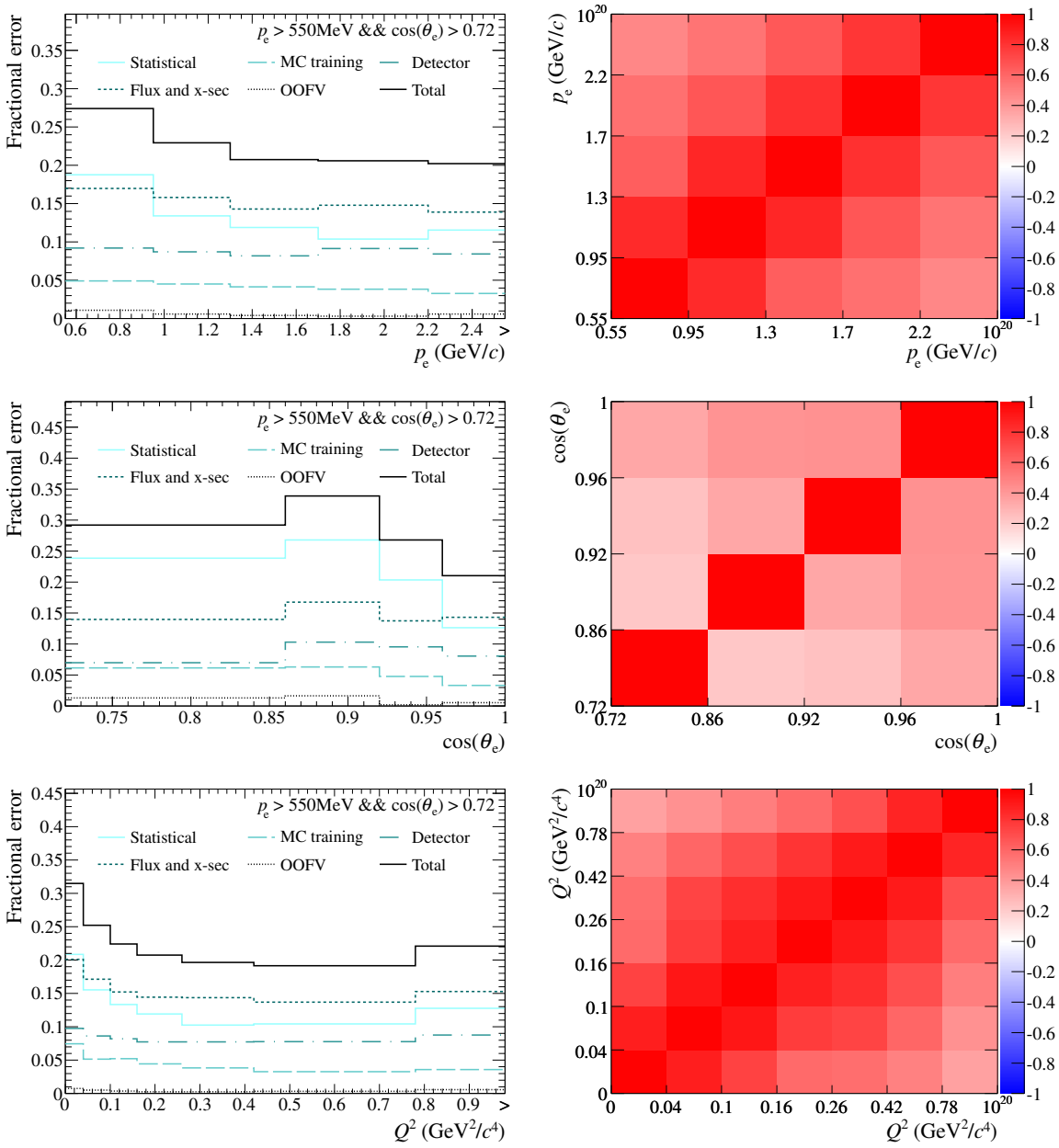


Figure 6.35: Fractional uncertainties (left) and correlation matrices (right) in bins of p_e (top), $\cos(\theta_e)$ (middle) and Q^2 (bottom) for the restricted phase-space analysis.

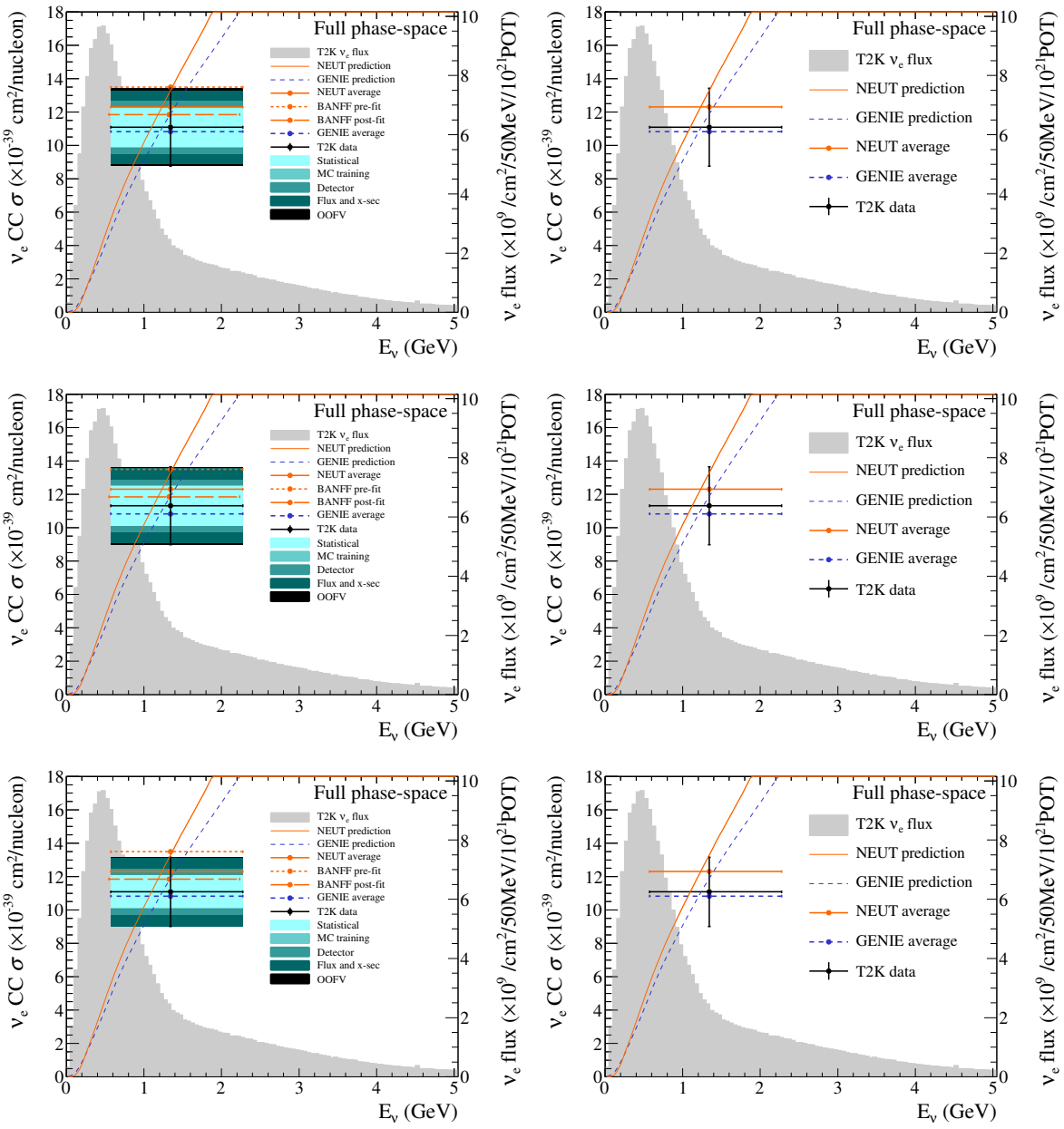


Figure 6.36: Unfolded flux-averaged cross-sections when unfolding through p_e (top), $\cos(\theta_e)$ (middle) and Q^2 (bottom) for the full phase-space analysis. The right-hand plots are identical to the left-hand ones, but only show the NEUT and GENIE nominal predictions, and the data points with their total errors. These are the final results of the full phase-space total flux-averaged cross-section.

Restricted phase-space analysis

The total flux-averaged CC inclusive ν_e cross-sections when unfolding through p_e, $\cos(\theta_e)$ and Q^2 are shown in Figure 6.37. The results for each variable are

- $\langle\sigma\rangle_\phi = 6.49 \pm 1.21 \times 10^{-39}$ cm^2/nucleon when unfolding through p_e
- $\langle\sigma\rangle_\phi = 6.53 \pm 1.21 \times 10^{-39}$ cm^2/nucleon when unfolding through $\cos(\theta_e)$
- $\langle\sigma\rangle_\phi = 6.54 \pm 1.22 \times 10^{-39}$ cm^2/nucleon when unfolding through Q^2.

Again, the T2K data agrees well with both the NEUT prediction of 7.38×10^{-39} cm^2/nucleon and the GENIE prediction of 6.41×10^{-39} cm^2/nucleon.

6.5.4 Discussion

As expected, the unfolded cross-sections differ slightly depending on which variable is used to do the unfolding. As discussed in Section 6.4.5, these differences are small compared to the systematic uncertainties.

Figure 6.38 compares the fractional errors when looking at the real data in the full phase-space analysis to the expected errors when looking at fake datasets thrown from the BANFF post-fit covariance matrix (see Section 6.4.5 for details). The OOFV systematic is larger for data than in the toy throws, as is expected given the larger data/MC discrepancy in the γ sample. The only other discrepancy between the real data and MC studies is for the flux and cross-section systematics, particularly for the $\cos(\theta_e)$ unfolding. No suitable explanation has been found for this behaviour, but as the discrepancy is small (especially when added in quadrature with all the other systematics), it is not a big concern.

Concentrating on the uncertainties for the Q^2 result, the flux uncertainty dominates at 12.9%, followed by the data statistical uncertainty (8.7%) and the detector systematic uncertainty (8.4%). In future analyses, the flux uncertainty should be reduced as more data from NA61 is used in the flux prediction. The data statistical uncertainty will also naturally improve as more data is collected, but if the selection efficiency can also be improved then the reduction will be accelerated. The detector systematic uncertainties are currently very large and are dominated by the TPC performance. Significant study is currently being devoted to better understanding theses issues, and the detector systematic uncertainty should hopefully be reduced in future analyses, especially if the momentum

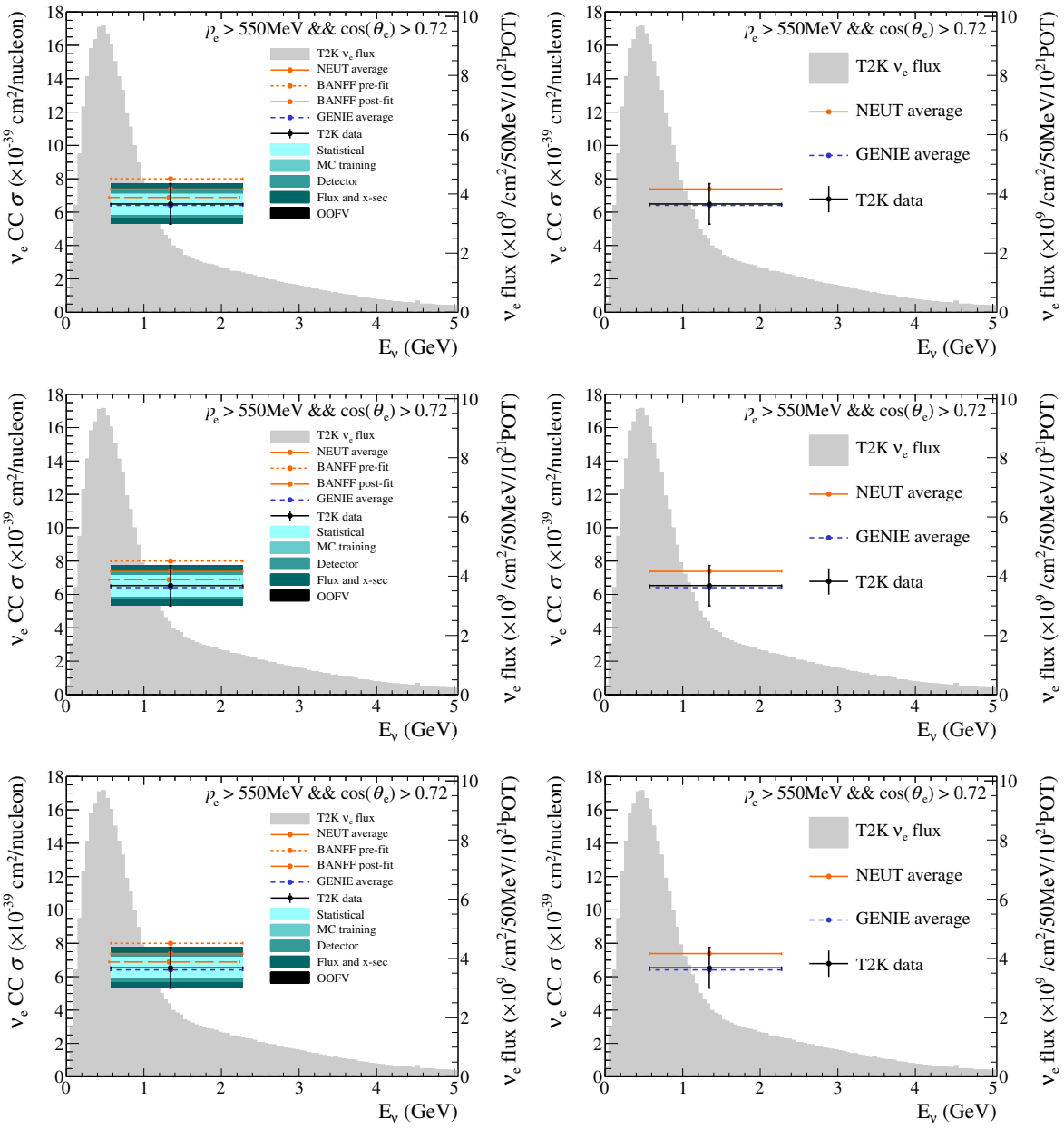


Figure 6.37: Unfolded flux-averaged cross-sections when unfolding through p_e (top), $\cos(\theta_e)$ (middle) and Q^2 (bottom) for the restricted phase-space analysis. The right-hand plots are identical to the left-hand ones, but only show the NEUT and GENIE nominal predictions, and the data points with their total errors. These are the final results of the restricted phase-space total flux-averaged cross-section.

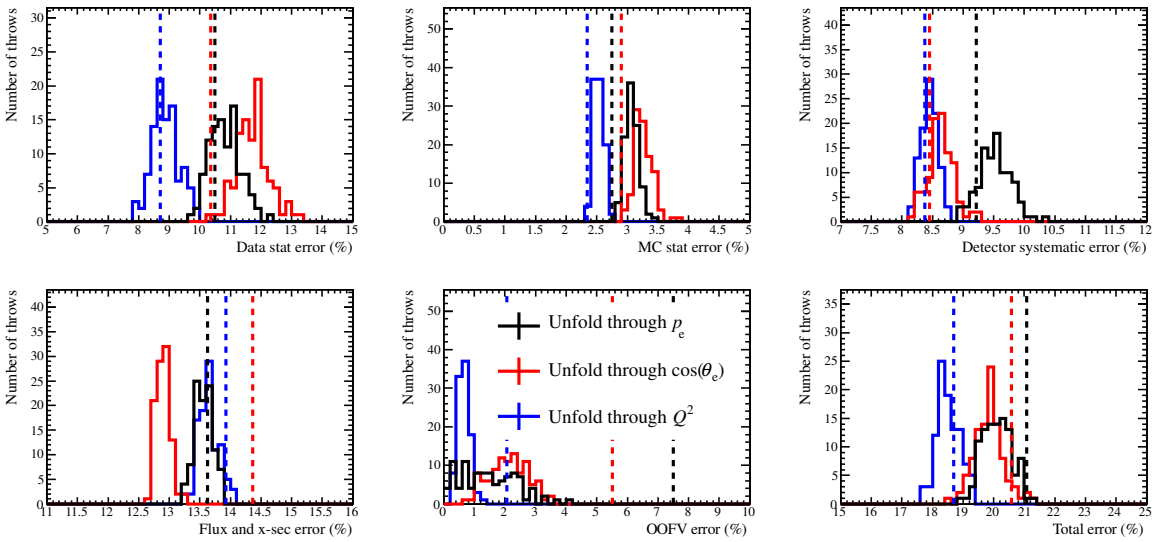


Figure 6.38: Histograms showing the percentage error on the flux-averaged cross-section for 100 toy datasets thrown from the BANFF post-fit covariance matrix. The dashed lines show the error on the real data.

resolution uncertainty can be reduced. The other uncertainties are small, but can all also be reduced in future. The cross-section model and uncertainties are being redeveloped by T2K to have fewer ad-hoc parameters and better agreement with external data. The OOFV systematic can be reduced if the selection is improved to reduce the background from $\gamma \to e^+e^-$ conversions, although this will be a challenging task. Finally, the Monte Carlo statistical error can be reduced by producing and analysing more Monte Carlo files, but the analysis framework may need to be improved such that this does not become computationally infeasible.

As mentioned in Section 6.3.2, neither the full phase-space nor the restricted phase-space results are model-independent. To check the effect of generating with a different model, the BANFF post-fit was also used to generate the unfolding. The effect of using the different model is shown in Table 6.8 for the full phase-space analysis, and Table 6.9 for the restricted phase-space analysis. The largest difference is 4%, which is well within the systematic uncertainties of the result. It is recommended that the results found when generating with the BANFF pre-fit are reported, as this MC has a consistent set of cross-section model parameter values and uncertainties.

Finally, Section 1.2 explained that the only other CC inclusive ν_e cross-section measurement at T2K energies is from the Gargamelle experiment, which measured a

Variable	$\sigma^{CC}_{\nu_e}$ ($\times 10^{-38}$ cm^2/nucleon)	
	Pre-fit	Post-fit
p_e	1.11	1.07
$\cos(\theta_e)$	1.13	1.15
Q^2	1.11	1.10

Table 6.8: Comparison of unfolded flux-averaged cross-sections when looking at real data in the full phase-space analysis, but generating the unfolding with the BANFF pre-fit and the BANFF post-fit predictions.

Variable	$\sigma^{CC}_{\nu_e}$ ($\times 10^{-39}$ cm^2/nucleon)	
	Pre-fit	Post-fit
p_e	6.49	6.41
$\cos(\theta_e)$	6.53	6.50
Q^2	6.54	6.46

Table 6.9: Comparison of unfolded flux-averaged cross-sections when looking at real data in the restricted phase-space analysis, but generating the unfolding with the BANFF pre-fit and the BANFF post-fit predictions.

cross-section of $\sigma = (0.7 \pm 0.2) E_\nu \times 10^{-38}$ cm^2/nucleon on CF_3Br. Using this model with the T2K ν_e flux predicts a cross-section of $0.95 \pm 0.27 \times 10^{-38}$ cm^2/nucleon, where the uncertainty only comes from the Gargamelle result, not the T2K flux. This agrees well with the $1.11 \pm 0.20 \times 10^{-38}$ cm^2/nucleon measured in this analysis. Of course, as the two experiments have different target elements, and the interactions occur in different nuclear environments, such a comparison is not rigorous. It does, however, serve as a useful cross-check of the result. Finally, Figure 6.39 shows a comparison of this result with the Gargamelle data, which was digitised from the plot shown in Figure 1.7.

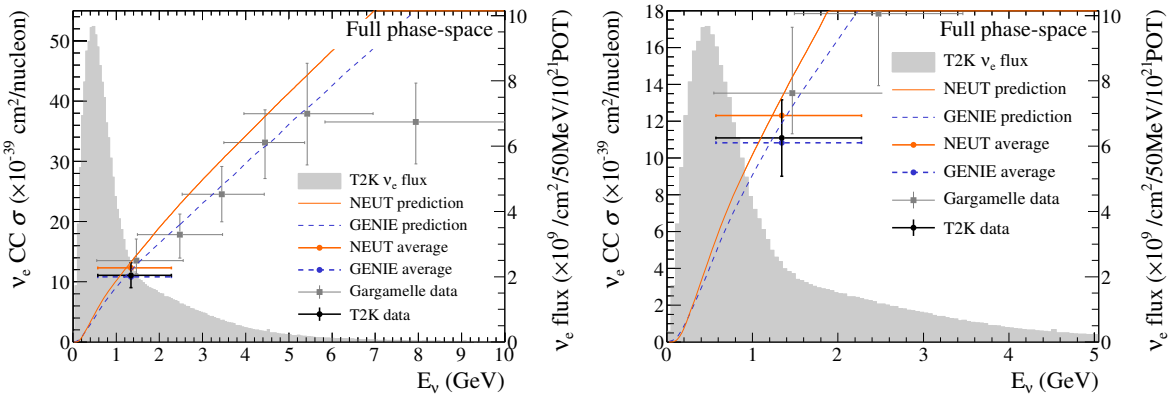


Figure 6.39: Comparison of the result of this analysis with data from the Gargamelle experiment, digitised from Figure 1.7. The left plot shows the full energy range covered by Gargamelle; the right plot shows the energy range most relevant to T2K.

Chapter 7

Conclusions and outlook

The Bayesian unfolding technique has been used to extract the ν_e CC inclusive flux-averaged cross-section at ND280, the T2K near detector. Differential cross-sections in p_e, $\cos(\theta_e)$ and Q^2 are also produced. The flux-averaged cross-section when unfolding through Q^2 is $1.11 \pm 0.20 \times 10^{-38}$ cm^2/nucleon, which is in good agreement with both the NEUT prediction of 1.23×10^{-38} cm^2/nucleon and the GENIE prediction of 1.08×10^{-38} cm^2/nucleon, as shown in Figure 6.36. The differential cross-sections all also agree well with the NEUT and GENIE predictions, as shown in Figure 6.32. A deficit is seen at low Q^2 compared to the NEUT prediction, although this is not significant.

A restricted phase-space analysis is also performed, in which only events with $p_e > 550\,\text{MeV}/c$ and $\cos(\theta_e) > 0.72$ are selected. In this case, the flux-averaged cross-section when unfolding through Q^2 is $6.54 \pm 1.22 \times 10^{-39}$ cm^2/nucleon, which is in good agreement with both the NEUT prediction of 7.38×10^{-39} cm^2/nucleon and the GENIE prediction of 6.41×10^{-39} cm^2/nucleon. The differential cross-section results are shown in Figure 6.34 and the flux-averaged cross-section results in Figure 6.37.

Measuring ν_e cross-sections is very important for the neutrino community, especially in the hunt for CP violation in the lepton sector. The results in this thesis are an important milestone towards constraining differences between ν_e and ν_μ cross-sections, and are the first measurement of the ν_e CC cross-section on carbon at energies relevant to long baseline oscillation experiments. This work will provide valuable input as models of neutrino interactions with nuclei are developed. Improved cross-section models, with reduced systematic uncertainties, are now vital if the precision of neutrino oscillation parameters is to be improved.

As the differences between ν_μ and ν_e cross-sections are important, an excellent next goal for T2K is to make a precise measurement of the ν_μ and ν_e cross-section ratio as a function of energy. Many of the systematic uncertainties should be significantly cancelled in this measurement (especially the flux, cross-section and detector systematics), and with more data the statistical uncertainty should also be reduced. A total uncertainty of less than 10% should be achievable in the immediate future.

Another goal for T2K should be to develop a CCQE-enhanced selection. The CCQE channel dominates at low energies, and it is in this region that the largest differences between the ν_μ and ν_e cross-sections are expected. It is also in this channel that the theoretical differences have been studied by Day and McFarland [41].

Other experiments should also be able to contribute to this area, in particular MINERνA, the dedicated neutrino cross-section experiment, and NOνA, a long-baseline oscillation experiment similar to T2K. Both these experiments use Fermilab's NuMI neutrino beam, which, like T2K's beam, is dominated by muon neutrinos. These experiments will face similar challenges to this analysis, with particle identification being key to rejecting the large number of ν_μ interactions. These experiments will provide datasets with different systematic uncertainties to this analysis, and will provide independent confirmation of the results.

Overall, this analysis should be reassuring to the long baseline neutrino oscillation community. The electron neutrino cross-section has been measured, and found to agree with two of the main interaction generators, NEUT and GENIE. With more data, and reduced systematic uncertainties, T2K should soon be able to provide more stringent tests of the differences between ν_μ and ν_e charged-current interactions at the GeV-scale.

Bibliography

[1] W. Pauli. Open letter to the Gauverein meeting in Tubingen, 1930.

[2] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956.

[3] R. Davis, D. Harmer, and K. Hoffman. Search for neutrinos from the sun. *Phys. Rev. Lett.*, 20(21):1205–1209, 1968.

[4] SAGE Collaboration. Measurement of the solar neutrino capture rate with gallium metal. *Phys. Rev. C*, 60(5), 1999.

[5] GALLEX Collaboration. GALLEX solar neutrino observations: complete results for GALLEX II. *Phys. Lett. B*, 357(1), 1995.

[6] Z. Mäki, M. Nakagawa, and S. Sakata. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 28:870–880, 1962.

[7] B. M. Pontecorvo. Mesonium and antimesonium. *JETP Lett.*, 33:549–551, 1957.

[8] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81(8):1562–1567, 1998.

[9] C. Giunti and C. W. Kim. *Fundamentals of Neutino Physics and Astrophysics.* Oxford University Press, 2007.

[10] C. Giunti. No effect of Majorana phases in neutrino oscillations. *Phys. Lett. B*, 686(1), 2010.

[11] E. K. Akhmedov et al. Series expansions for three-flavor neutrino oscillation probabilities in matter. *Journal of High Energy Physics*, 2004(04):078, 2004.

[12] B. Aharmim et al. Electron energy spectra, fluxes, and day-night asymmetries of ^{8}B solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory. *Phys. Rev. C*, 72:055502, Nov 2005.

[13] M. Sanchez et al. Measurement of the L/E distributions of atmospheric ν in Soudan 2 and their interpretation as neutrino oscillations. *Phys. Rev. D*, 68:113004, Dec 2003.

[14] P. Adamson et al. Measurement of neutrino and antineutrino oscillations using beam and atmospheric data in MINOS. *Phys. Rev. Lett.*, 110:251801, Jun 2013.

[15] S. Abe et al. Precision measurement of neutrino oscillation parameters with kamland. *Phys. Rev. Lett.*, 100:221803, Jun 2008.

[16] Daya Bay Collaboration. Observation of electron-antineutrino disappearance at Daya Bay. *Phys. Rev. Lett.*, 108(17), 2012.

[17] RENO Collaboration. Observation of reactor electron antineutrinos disappearance in the RENO experiment. *Phys. Rev. Lett.*, 108(19), 2012.

[18] Y. Abe et al. Indication of reactor $\overline{\nu}_e$ disappearance in the Double Chooz experiment. *Phys. Rev. Lett.*, 108:131801, Mar 2012.

[19] D. Ayres et al. NOvA proposal to build a 30 kiloton off-axis detector to study neutrino oscillations in the Fermilab NuMI beamline. http://arxiv.org/abs/hep-ex/0503053, 2005.

[20] J. Beringer et al. 2012 Review of particle physics. *Phys. Rev. D*, 86, 2012.

[21] D. Larson et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: power spectra and WMAP-derived parameters. *The Astrophysical Journal Supplement Series*, 192(2):16, 2011.

[22] P. Huet and E. Sather. Electroweak baryogenesis and standard model cp violation. *Phys. Rev. D*, 51:379–394, Jan 1995.

[23] K. Abe et al. Observation of electron neutrino appearance in a muon neutrino beam. *Phys. Rev. Lett.*, 112:061802, Feb 2014.

[24] KATRIN Collaboration. The KATRIN neutrino mass experiment. *Nucl. Instrum. Meth. A*, 623(1), 2010.

[25] K. N. Abazajian et al. Cosmological and astrophysical neutrino mass measurements. *Astroparticle Physics*, 35(4):177 – 184, 2011.

[26] H. Ohsumi et al. SuperNEMO project. *Journal of Physics: Conference Series*, 120(5):052054, 2008.

[27] M. C. Chen. The SNO liquid scintillator project. *Nuclear Physics B - Proceedings Supplements*, 145(0):65 – 68, 2005. NOW 2004 Proceedings of the Neutrino Oscillation Workshop.

[28] A. Aguilar et al. Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam. *Phys. Rev.*, D64:112007, 2001.

[29] S. Schael et al. Precision electroweak measurements on the z resonance. *Phys.Rept.*, 427:257–454, 2006.

[30] A.A. Aguilar-Arevalo et al. Improved search for $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations in the MiniBooNE experiment. *Phys.Rev.Lett.*, 110:161801, 2013.

[31] G. Mention et al. The reactor antineutrino anomaly. *Phys. Rev. D*, 83:073006, 2011.

[32] Carlo Giunti and Marco Laveder. Statistical significance of the gallium anomaly. *Phys. Rev. C*, 83:065504, 2011.

[33] J Kopp et al. Sterile neutrino oscillations: The global picture. *JHEP*, 1305:050, 2013.

[34] F. Dydak et al. A search for ν_μ oscillations in the δm^2 range $0.3 - 90\,\mathrm{eV}^2$. *Phys. Lett. B*, 134:281, 1984.

[35] Samoil M. Bilenky et al. Four neutrino mass spectra and the super-kamiokande atmospheric up-down asymmetry. *Phys. Rev. D*, 60:073007, 1999.

[36] P. Adamson et al. Active to sterile neutrino mixing limits from neutral-current interactions in MINOS. *Phys.Rev.Lett.*, 107:011802, 2011.

[37] Maria Archidiacono, Nicolao Fornengo, Carlo Giunti, and Alessandro Melchiorri. Testing 3+1 and 3+2 neutrino mass models with cosmology and short baseline experiments. *Phys. Rev. D*, 86:065028, 2012.

[38] J. Nieves, I. Ruiz Simo, and M.J. Vicente Vacas. The nucleon axial mass and the MiniBooNE quasielastic neutrino-nucleus scattering problem. *Physics Letters B*, 707(1):72 – 75, 2012.

[39] J. A. Formaggio and G. P. Zeller. From eV to EeV: Neutrino cross sections across energy scales. *Rev. Mod. Phys.*, 84:1307–1341, Sep 2012.

[40] A. A. Aguilar-Arevalo et al. First measurement of the muon neutrino charged current quasielastic double differential cross section. *Phys. Rev. D*, 81:092005, May 2010.

[41] M. Day and K. S. McFarland. Differences in quasi-elastic cross-sections of muon and electron neutrinos. *Phys. Rev. D*, 86:053003, 2012.

[42] R. Maschuw et al. Neutrino spectroscopy with KARMEN. *Progress in Particle and Nuclear Physics*, 40(0):183 – 192, 1998.

[43] C. Athanassopoulos et al. Measurements of the reactions $^{12}C(\nu_e, e^-)^{12}N_{\mathrm{g.s.}}$ and $^{12}C(\nu_e, e^-)^{12}N^*$. *Phys. Rev. C*, 55:2078–2091, Apr 1997.

[44] D. A. Krakauer et al. Experimental study of neutrino absorption on carbon. *Phys. Rev. C*, 45:2450–2463, May 1992.

[45] Y. Hayato. NEUT. *Nuclear Physics B-Proceedings Supplements*, 112(1-3):171–176, 2002.

[46] C. Andreopoulos et al. The GENIE neutrino monte carlo generator. *Nucl. Instrum. Meth.*, A614:87–104, 2010.

[47] J. Blietschau et al. Total cross sections for ν_e and $\overline{\nu}_e$ interactions and search for neutrino oscillations and decay. *Nuclear Physics B*, 133(2):205 – 219, 1978.

[48] Y. Ajima et al. Tokai-to-Kamioka (T2K) long baseline neutrino oscillation ex-

periment proposal. http://j-parc.jp/NuclPart/pac_0606/pdf/p11-Nishikawa.pdf, April 2006.

[49] K. Abe et al. Indication of electron neutrino appearance from an accelerator-produced off-axis muon neutrino beam. *Phys. Rev. Lett.*, 107(4), 2011.

[50] K. Abe et al. Measurement of neutrino oscillation parameters from muon neutrino disappearance with an off-axis beam. *Phys. Rev. Lett.*, 111:211803, 2013.

[51] K. Abe et al. Measurement of the inclusive ν_μ charged current cross section on carbon in the near detector of the T2K experiment. *Phys. Rev. D*, 87:092003, 2013.

[52] T. Kikawa. Recent results from T2K on neutrino interaction. http://indico.cern.ch/event/234536/material/slides/0?contribId=10&sessionId=1. Summary talk at WIN 2013 conference.

[53] K. Abe et al. The T2K experiment. *Nucl. Instrum. Meth. A*, 659, 2011.

[54] K. Abe et al. T2K neutrino flux prediction. *Phys. Rev. D*, 87:012001, Jan 2013.

[55] K. Matsuoka et al. Design and performance of the muon monitor for the T2K neutrino oscillation experiment. *Nucl. Instrum. Meth. A*, 624(3):591 – 600, 2010.

[56] N. Abgrall et al. Time projection chambers for the T2K near detectors. *Nucl. Instrum. Meth. A*, 637(1):25 – 46, 2011.

[57] P.-A. Amaudruz et al. The T2K fine-grained detectors. *Nucl. Instrum. Meth. A*, 696(0):1 – 31, 2012.

[58] S. Assylbekov et al. T2K ND280 off-axis pi-zero detector. *Nucl. Instrum. Meth. A*, 686(0):48 – 63, 2012.

[59] D Allan et al. The electromagnetic calorimeter for the T2K near detector ND280. *Journal of Instrumentation*, 8(10):P10019, 2013.

[60] S. Aoki et al. The T2K side muon range detector (SMRD). *Nucl. Instrum. Meth. A*, 698(0):135 – 146, 2013.

[61] R. Brun and F. Rademakers. ROOT - an object oriented data analysis framework. *Nucl. Inst. and Meth. in Phys. Res. A*, 389:81–86, 1997.

[62] S. Agostinelli et al. Geant4, a simulation toolkit. *Nucl. Instrum. Meth. A*, 506(3):250 – 303, 2003.

[63] G. Battistoni et al. The FLUKA code: description and benchmarking. *AIP Conference Proceedings*, 896(1):31–49, 2007.

[64] R. Brun, F. Carminati, and S. Giani. *CERN-W5013*, 1994.

[65] N. Abgrall et al. Measurements of cross sections and charged pion spectra in proton-carbon interactions at 31 GeV/c. *Phys. Rev. C*, 84:034604, 2011.

[66] T. Eichten et al. Particle production in proton interactions in nuclei at 24 GeV/c. *Nucl. Phys. B*, 44, 1972.

[67] J. V. Allaby et al. High-energy particle spectra from proton interactions at 19.2 GeV/c. Technical Report 70-12, CERN, 1970.

[68] C. H. Llewellyn Smith. *Phys. Rep. C*, 3:261, 1972.

[69] R. A. Smith and E. J. Moniz. *Nucl. Phys. B*, 43:605, 1972.

[70] R. A. Smith and E. J. Moniz. *Nucl. Phys. B*, 101:547(E), 1975.

[71] D. Rein and L. M Sehgal. *Ann. Phys.*, 133:79, 1981.

[72] T. Sjostrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006.

[73] A. Cervera-Villanueva et. al. "RecPack" a reconstruction toolkit. *Nuclear Instruments and Methods in Physics Research A*, 534:180, 2004.

[74] K. Abe et al. Measurement of the intrinsic electron neutrino component in the T2K neutrino beam with the ND280 detector. http://arxiv.org/abs/1403.2552, 2014.

[75] Giannelli et al. Multiproton final states in positive pion absorption below the $\Delta(1232)$ resonance. *Physical Review C*, 61:054615, 2000.

[76] A. Saunders et al. Reaction and total cross sections for low energy π^+ and π^- on isospin zero nuclei. *Physical Review C*, 53(4):1745, 1996.

[77] I. Navon et al. True absorption and scattering of 50 MeV pions. *Physical Review C*, 28(6):2548, 1983.

[78] T. J. Bowles et al. Inclusive ($\pi^{\pm}$, π^0) reactions in nuclei. *Physical Review C*, 23(1):439, 1981.

[79] F. Binon et al. Scattering of negative pions on carbon. *Nuclear Physics B*, 17:168, 1969.

[80] D. Ashery et al. True absorption and scattering of pions on nuclei. *Physical Review C*, 23(5):2173, 1981.

[81] D. Ashery et al. Inclusive pion single-charge-exchange reactions. *Physical Review C*, 30(3):946, 1984.

[82] M. K. Jones et al. Pion absorption above the $\Delta(1232)$ resonance. *Physical Review C*, 48(6):2800, 1993.

[83] M. Crozon et al. Etude de la diffusion π-noyau entre 500 et 1300 MeV. *Nuclear Physics*, 64:567, 1964.

[84] T. Takahashi et al. $\pi^- - {}^{12}$C elastic scattering above the Δ resonance. *Physical*

Review C, 51(5):2542, 1995.

[85] K. Aoki et al. Elastic and inelastic scattering of π^+ and π^- on ^{12}C at 995 MeV/c. *Physical Review C*, 76:024610, 2007.

[86] D. Allan et al. The electromagnetic calorimeter for the T2K near detector ND280. *Journal of Instrumentation*, 8(10):P10019, 2013.

[87] J. Dobson. *Neutrino Induced Charged Current π^+ Production at the T2K Near Detector*. PhD thesis, Imperial College London, 2012.

[88] K. Abe et al. Evidence of electron neutrino appearance in a muon neutrino beam. *Phys. Rev. D*, 88:032002, Aug 2013.

[89] K. Abe et al. Evidence of electron neutrino appearance in a muon neutrino beam. *Phys. Rev. D*, 88:032002, 2013.

[90] P. de Perio. NEUT pion FSI. *AIP Conf.Proc.*, 1405:223–228, 2011.

[91] A. A. Aguilar-Arevalo et al. 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.

[92] A. A. Aguilar-Arevalo et al. Measurement of ν_μ-induced charged-current neutral pion production cross sections on mineral oil at $E_\nu \in 0.5 \sim 2.0$GeV. *Phys. Rev. D*, 83:052009, 2011.

[93] A. A. Aguilar-Arevalo et al. Measurement of ν_μ and $\overline{\nu}_\mu$ 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.

[94] D. Rein and L.M. Sehgal. Neutrino-excitation of baryon resonances and single pion production. *Annals of Physics*, 133(1):79–153, 1981.

[95] M. Hasegawa et al. Search for coherent charged pion production in neutrino-carbon interactions. *Phys. Rev. Lett.*, 95:252301, 2005.

[96] K. Hiraide et al. Search for charged current coherent pion production on carbon in a few-GeV neutrino beam. *Phys. Rev. D*, 78:112004, 2008.

[97] T. Golan, J. T. Sobczyk, and J. Zmuda. NuWro: the Wroclaw monte carlo generator of neutrino interactions. *Nucl.Phys.Proc.Suppl.*, 229-232:499, 2012.

[98] G. D'Agostini. A multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth. A*, 362(2–3):487 – 498, 1995.

[99] T. Adye et al. RooUnfold. http://hepunx.rl.ac.uk/~adye/software/unfold/htmldoc/RooUnfold.html.