Imperial College London
Department of Physics

Neutrino Induced Charged Current π^+ Production at the T2K Near Detector

James Edward Young Dobson

June 2012

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

Abstract

A study of ν_μ-induced charged current (CC) π^+ production at the T2K off-axis near detector (ND280) is presented. Using Monte Carlo (MC) data studies event selections for both CC-inclusive and enriched CC-π^+ samples have been developed using the ND280 tracker-region and surrounding ECals. Two types of CC-π^+ selections were developed—one using the TPC to identify the pion and the other using a new ECal PID based on the deposited charge per unit length. Data/MC ratios are calculated and compared with the associated detector, neutrino interaction and flux simulation systematics.

The predicted neutrino interaction rate was based on v2.6.2 of the GENIE MC generator and on T2Ks tuned 11a JNUBEAM flux simulation. The data used was collected between Nov. 2010 and March 2011 during the Run 2 data taking period and corresponds to a total integrated POT of 7.83×10^{19}. For the ν_μ-CC-Inclusive selection which selects ν_μ-CC interactions with a purity of 88.1% we find:

$$N_{\text{Data}}^{\text{CCIncl}}/N_{\text{MC}}^{\text{CCIncl}} = 1.021 \pm 0.015(\text{stat})^{+0.032}_{-0.031}(\text{det})^{+0.112}_{-0.097}(\text{xsec})^{+0.093}_{-0.093}(\text{flux}).$$

For the TPC-based ν_μ-CC-$\pi^\pm$ selection which selects ν_μ-CC interactions with at least one π^+ in the final state with a purity of 81.3% we find:

$$N_{\text{Data}}^{\text{CC}\pi^\pm}/N_{\text{MC}}^{\text{CC}\pi^\pm} = 1.041 \pm 0.057(\text{stat})^{+0.044}_{-0.044}(\text{det})^{+0.125}_{-0.127}(\text{xsec})^{+0.208}_{-0.190}(\text{flux}).$$

For the ECal-based ν_μ-CC-$\pi^\pm$ selection which selects ν_μ-CC interactions with at least one π^+ in the final state with a purity of 67.9% we find:

$$N_{\text{Data}}^{\text{CC}\pi^\pm}/N_{\text{MC}}^{\text{CC}\pi^\pm} = 0.985 \pm 0.070(\text{stat})^{+0.074}_{-0.1}(\text{det})^{+0.119}_{-0.118}(\text{xsec})^{+0.205}_{-0.187}(\text{flux}).$$

These show that the current measured and predicted rates for both the inclusive rate of ν_μ neutrino interactions and those with at least one π^+ in the final state agree to within the systematic uncertainties associated with neutrino interaction and flux simulation. Moreover, these selections lay the groundwork for future analyses, using larger data sets, that can be used to constrain these sources of uncertainty.

Declaration

As is typical for a thesis performed within the context of a large collaborative physics experiment the work presented here relies critically on the output of a large body of work, which in the case of T2K was performed by around 500 people over the last 10 years. The purpose of this declaration is to make clear my personal contributions.

Chapter 1 gives an overview of the current status of neutrino oscillation physics and describes the modelling of neutrino interactions in the few-GeV region using the GENIE Monte Carlo generator. This information was synthesised from a number of sources including textbooks, review articles and publications and is referenced appropriately. Throughout my time on T2K I have also been an active member of the GENIE collaboration. This involved acting as the contact point between T2K and GENIE as well as work on the validation and optimisation of the T2K specific event generation programs, a short summary of this work is given in § 1.3.5.

Chapter 2 describes T2K and is largely based upon the published descriptions of the experiment. The motivation for studying neutrino induced charged current π^+ production at the T2K near detector discussed in § 2.5 is based on the output of internal studies performed as part of the T2K disappearance measurement and is also referenced appropriately. I contributed to these studies through work on the validation and development of GENIE specific reweighting libraries. In addition to my work within the GENIE collaboration my own personal contributions to T2K come mainly from my involvement with the UK designed and built ND280 ECal and my role as manager, inherited from my supervisor Dr Yoshi Uchida[1], of the oaAnalysis software package. oaAnalysis is the important final stage in the ND280 software chain which distills and collates the information necessary to allow individual users to perform ND280-based analyses. My work on the ND280 ECal included an early test bench, performed under the supervision of Dr Antonin Vacheret[2], which provided the first demonstration of the complete scintillator bar readout chain, as well as work on the *in situ* calibration of the ECal using cosmic tracks resulting in the provision of routines to equalise the response of all channels on the Ds- and Barrel-ECal. I also contributed to the testing of the Ds-ECal in the CERN T9 test beam and its commissioning in the ND280 detector in J-PARC.

The main body of my own work is presented in chapters 3 and 4. In chapter 3 the event selections I developed to study neutrino induced π^+ production at the T2K near detector are described. These are based on the output of the sub-detector and global reconstruc-

[1] Of Imperial College London.
[2] At the time at Imperial College London but now at the University of Oxford.

tion, and in particular the global vertexing algorithm, developed by others. The data-MC comparisons and validation work presented is my own. In general I use the particle identification (PID) variables provided by the various sub-detectors, the only exception is the ECal-based dQ/dL PID which I developed specifically for this analysis. The evaluation of the detector systematics associated with these selections was performed by myself but draws heavily on the output of internal studies performed within the T2K NuMu physics working group.

In chapter 4 I evaluate the systematics associated with neutrino flux and neutrino interaction simulation using the T2KReWeight package. T2KReWeight was originally developed by my co-supervisor, Dr Costas Andreopoulos[3], and myself with the aim of providing a GENIE specific reweighting tool. The generalisation of T2KReWeight to act as an interface to not only the GENIE reweighting libraries but also to the NEUT MC generator and JNUBEAM flux reweighting libraries was my own work. This involved providing the general interface and implementations for both the NEUT and JNUBEAM reweighting libraries in the form of example templates which were then filled out by the relevant experts.

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

James Edward Young Dobson

[3]Of the Rutherford Appleton Laboratory, UK.

Acknowledgments

Over the past four years I had more than my fair share of the knowledge, guidance and support that access to not one, but two, excellent mentors brings. I would like to thank my supervisor at Imperial, Dr. Yoshi Uchida—who, in addition to providing the expertise and advice only possible from someone with such a broad background in particle physics, also encouraged me to work on multiple aspects of the experiment, which helped me develop as a researcher. I would especially like to express my thanks to him for the generous level of support and encouragement I received towards the end of my PhD.

I would also like to thank Dr. Costas Andreopoulos, my co-supervisor at the Rutherford Appleton Laboratory (RAL). Through frequent trips to RAL I benefited from his extensive knowledge of both neutrino cross section and oscillation physics. The invaluable physics discussions, often philosophical, on the many lifts to and from Didcot train station were a delight, and contributed significantly to my motivation over the course of my doctorate.

The T2K Imperial group has been an excellent place to study. The feedback received at the weekly group meetings was central to developing the work presented in this thesis. Many thanks to all the students and staff for their contributions. In particular I would like to acknowledge the guidance and suggestions of Dr. Morgan Wascko and Dr. Antonin Vacheret.

T2K is a large experiment and it almost goes without saying that the work presented here would not have been possible without the effort of many people. I am very grateful for the feedback and suggestions I received from the NuMu physics group regarding the near detector selections I developed: in particular I would like to acknowledge suggestions and clarifications I received from Dr. Thomas Lindner, Dr. Federico Sanchez, Dr. Anthony Hillairet, Javier Caravaca, Prof. Scott Oser, Dr. Mark Hartz and Dr. Anselmo Cervera when evaluating the detector and flux systematics. I am grateful to Gustav Wikstrom who developed the global vertexing algorithm upon which this thesis relies so heavily. In addition I thank my GENIE and T2K collaborator Prof. Steve Dytman for all the useful discussions and suggestions with regards to charged current π production and the methodology of my analysis.

Finally I would like to thank all of my family and close friends who have been so supportive over the last four years. In particular I would like to thank my mother Dor, stepmom Biddy and my wonderful girlfriend Martina for their dedicated support and understanding over the last year, without which I would not have been able to finish.

Contents

List of Tables

List of Figures

1 Neutrino Oscillation Physics

There is now overwhelming evidence that contrary to the standard model of particle physics[1] neutrinos have a non-zero mass. This evidence manifests itself in the form of neutrino oscillations which arise as a consequence of this non-zero mass and the fact that the weak eigenstates differ to the mass eigenstates. At the time of writing the results of neutrino oscillation experiments constitute the only conclusive experimental evidence of physics beyond the standard model [2].

In 1968 the Homestake radiochemical experiment led by Ray Davis provided the first evidence for neutrino flavour change. In a landmark paper they presented results [3] showing a deficit of approximately 70% in the measured rate of solar ν_e compared to that predicted by the standard solar model [4] (SSM). This deficit was coined the solar neutrino problem and was confirmed in the late 1980s by the measurements [5] of the Kamiokande experiment, led by Masatoshi Koshiba. Kamiokande, in addition to measuring a deficit in the total rate also demonstrated that the neutrinos originated preferentially from the direction of the sun, showing conclusively that they were in fact solar neutrinos. The deficit was further confirmed in the early 1990s by the GALLEX/GNO [6] and SAGE [7] experiments[2]. Table 1.1 summarises these result showing the deficit of solar neutrinos from the various experiments as well as the reaction channel and energy threshold for detection. As all of these measurements were only sensitive to the flux of electron neutrinos it was impossible to decide whether the disagreement was coming from the SSM, our understanding of neutrinos, or both.

Although here we focus on neutrino oscillations, it should be mentioned that the first achievements of these solar neutrino experiments were the detection of solar and astrophysical neutrinos, see also the detection of neutrinos from supernova SN 1987a [8] in Kamiokande, and the confirmation of thermonuclear energy generation in stars. It was 'for pioneering contributions to astrophysics, in particular for the detection of cosmic neutrinos' that Ray Davis and Masatoshi Koshiba shared the 2002 Nobel Prize in Physics.

The atmospheric neutrino anomaly was first observed in the late 1980s by the Kamiokande [10] and IMD [11] experiments. They measured a deficit in the flux of atmospheric ν_μ compared to that predicted by Monte Carlo (MC) simulations and at the same time measured the rate of atmospheric ν_e to be consistent with predictions. This was resolved in 1998 by the high precision measurements of the Super-Kamiokande experiment [12], a large water

[1] Where we take the standard model to be defined through its gauge symmetries, matter content and the mechanism of spontaneous symmetry breaking and require it to be renormalizable [1].

[2] These newer experiments were sensitive to the low energy *pp* solar neutrinos (< 0.5 MeV) as opposed to the previous experiments which only measured the high energy (~ 1 to 10 MeV) B^8 neutrinos.

Experiment	Reaction	E_ν^{Thresh}(MeV)	R_{SSM}	n_σ
HOMESTAKE	$\nu_e + {}^{37}Cl \rightarrow {}^{37}Ar + e^-$	0.814	0.301 ± 0.027	3.3
GALLEX/GNO	$\nu_e + {}^{71}Ga \rightarrow {}^{71}Ge + e^-$	0.233	0.529 ± 0.042	5.4
SAGE	$\nu_e + {}^{71}Ga \rightarrow {}^{71}Ge + e^-$	0.233	0.540 ± 0.040	5.0
Kamiokande	$\nu_\alpha + e^- \rightarrow \nu_\alpha + e^-$	6.7	0.484 ± 0.066	2.2
SK	$\nu_\alpha + e^- \rightarrow \nu_\alpha + e^-$	4.7	0.406 ± 0.014	2.6

Table 1.1: Summary of results from solar neutrino experiments before resolution of the solar neutrino problem showing the reaction channel, detection threshold and ratio, R_{SSM}, of the measured flux to that predicted by the BP04 SSM [9]. Also shown is the experimental significance of the difference. Table adapted from [2].

Cherenkov detector in Japan and now also the T2K far detector, which measured a deficit of atmospheric upward going ν_μs with a characteristic angular dependence consistent with mass induced oscillations. These measurements were then confirmed by the Soudan 2 [13] and MACRO [14] experiments. Independent results from the K2K long baseline neutrino experiment [15] were consistent with these earlier findings and more recently precision measurements by the MINOS [16] long baseline experiment have been made.

The SNO experiment provided the solution to the solar neutrino problem in 2002 [17]. Using an ultra-pure heavy water Cherenkov detector they were sensitive to both the rate of solar ν_e, through the charged current reaction $\nu_e + n \rightarrow e^- + p$, as well as the total flux from all neutrino flavours, through the neutral current reaction $\nu_\alpha + d \rightarrow p + n + \nu_\alpha$ (where the neutron was identified by the 6.25 MeV γ from the neutron capture on deuterium)[3]. Their results confirmed the deficit in the ν_e flux but showed that the total flux ($\nu_e + \nu_\mu + \nu_\tau$) was consistent with that predicted by the SSM: with the conclusion being that the missing ν_es had been converted into ν_μ or ν_τs on their way to the earth. These results were later combined with $\bar{\nu}_e$ data from the KamLAND reactor experiment, which provided clear evidence [18] of $\bar{\nu}_e$ oscillations, to fully constrain the solar mixing parameters.

All of these results are well explained within the framework of three-flavour neutrino mixing. Before discussing this and latest experimental results in more detail it is helpful to first give a description of the formalism of neutrino oscillations. In the next section we give an overview of the standard theory of neutrino oscillations in a vacuum and its extension to oscillations in matter.

1.1 Neutrino Oscillation Formalism

Neutrino oscillation is a purely quantum mechanical phenomenon which arises if neutrinos are massive and mixed. Consider a neutrino with flavour α ($= \mathrm{e}, \mu, \tau$) and momentum $\vec{p}$ created in a charged current weak interaction, either from a charged lepton l_α^- or together

[3] For simplicity here we do not discuss the additional elastic scattering process which was dominantly sensitive to the ν_e flux due to their $\sim 6\times$ larger cross section.

with a charged antilepton l_α^+, and described by the flavour state

$$|\nu_\alpha\rangle = \sum_k U_{\alpha k}^* |\nu_k\rangle \tag{1.1}$$

where, in the case of three active neutrinos[4], U is the 3×3 unitary PMNS matrix (Named after Pontecorvo, Maki, Nakagawa and Sakata, the four people credited with the development of the formalism) which expresses the fact that the flavour states $|\nu_\alpha\rangle$ are not necessarily the same as the mass eigenstates $|\nu_k\rangle$, i.e. that they are mixed. The PMNS matrix can be parameterised by three mixing angles and three complex phases[5]. If neutrinos are Dirac particles (distinct particle anti-particle states) then two of these phases are unphysical and if they are Majorana particles (identical particle anti-particle states) then it can be shown [23] that the additional two phases have no physical effect on neutrino or anti-neutrino oscillations ($\nu_\alpha \to \nu_\beta$ and $\bar{\nu}_\alpha \to \bar{\nu}_\beta$). Thus in what follows the PMNS matrix effectively depends only on three mixing angles θ_{12}, θ_{13}, and θ_{23} and one complex phase δ. A convenient representation of the PMNS matrix is

$$\begin{aligned} U_{\alpha k} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{1.2}$$

where the abbreviations $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$ have been used. The upper equality makes clear the factorisation in terms of the three distinct rotations: In this representation these correspond to the mixing parameters accessible to atmospheric/beam, reactor/beam, and solar experiments respectively. For example the c_{23} and s_{23} terms are dominant in atmospheric/beam experiments whereas the c_{12} and s_{12} terms are dominant for solar oscillations. Which mixing parameters an experiment is sensitive to depends on the neutrino type, detection method, baseline/energy and the mass splittings between the neutrinos.

By definition the massive neutrino states $|\nu_k\rangle$ are eigenstates of the Hamiltonian and satisfy $\mathscr{H}|\nu_k\rangle = E_k|\nu_k\rangle$ with energy eigenvalues

$$E_k = \sqrt{\vec{p}^2 + m_k^2}. \tag{1.3}$$

Their evolution is determined by the time dependent Schrödinger equation such that for plane wave solutions

$$|\nu_k(t)\rangle = e^{-iE_k t}|\nu_k\rangle \tag{1.4}$$

[4]The LEP experiments at CERN [19, 20, 21, 22] constrained the number of active neutrinos to be three but if there exist non-interacting sterile neutrinos then the dimension of U is not constrained.

[5]In general an $n \times n$ unitary matrix can be parameterised by $n(n-1)/2$ Euler angles and $n(n+1)/2$ complex phases.

so that the time evolution of a neutrino created in a pure flavour state α is given by

$$|\nu_\alpha(t)\rangle = \sum_k U^*_{\alpha k} e^{-iE_k t} |\nu_k\rangle. \tag{1.5}$$

Inverting eqn (1.1) to get $|\nu_k\rangle = \sum_\alpha U_{\alpha k}|\nu_\alpha\rangle$ and substituting into the above yields the flavour composition as a function of time

$$|\nu_\alpha(t)\rangle = \sum_{\beta=e,\mu,\tau} \left(\sum_k U^*_{\alpha k} e^{-iE_k t} U_{\beta k} \right) |\nu_\beta\rangle. \tag{1.6}$$

This shows that a pure flavour state at time $t = 0$ evolves into a superposition of different flavour states at a time $t > 0$ as long as the mixing matrix is not diagonal. Thus the transition probability of a neutrino produced in a flavour state α at time t = 0 to be detected as a flavour state β at a time t later is

$$P_{\nu_\alpha \to \nu_\beta}(t) = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} e^{-i(E_k - E_j)t}. \tag{1.7}$$

For ultra-relativistic neutrinos with $|\vec{p}| \gg m_k$ the dispersion relation in eqn (1.3) can be approximated to $E_k \simeq |\vec{p}| + m_k^2/2$ so that

$$E_k - E_j \simeq \frac{m_k^2 - m_j^2}{2|\vec{p}|} = \frac{\Delta m^2_{kj}}{2|\vec{p}|} \tag{1.8}$$

where we have define the squared-mass difference $\Delta m^2_{kj} \equiv m_k^2 - m_j^2$. Combining this with the approximation that for ultra-relativistic neutrinos $t = L$ and denoting $|\vec{p}| = E$ gives the oscillation probability in a form useful for experiments capable of measuring L and E

$$P_{\nu_\alpha \to \nu_\beta}(L, E) = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} \exp\left(-i \frac{\Delta m^2_{kj} L}{2E} \right). \tag{1.9}$$

This shows that the phase of the oscillations depends on the source-detector distance L, the neutrino energy E and the squared-mass differences whereas the amplitude of the oscillations are governed solely by the PMNS mixing matrix U. The squared-mass differences[6] and the components of the PMNS matrix are physical constants of nature open to measurement by neutrino oscillation experiments. It is convenient to express this oscillation probability in terms of its real and imaginary components:

[6] Although neutrino oscillations by themselves imply massive neutrinos all that can be inferred about the absolute masses m_k or m_j is that at least one is greater than or equal to $|\Delta m^2_{kj}|$ and, if matter effects are present, the sign of Δm^2_{kj}.

$$P_{\nu_\alpha\to\nu_\beta}(L,E) = \delta_{\alpha\beta} - 4\sum_{k>j}\mathfrak{Re}[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]\sin^2\left(\frac{\Delta m^2_{kj}L}{4E}\right) + 2\sum_{k>j}\mathfrak{Im}[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]\sin\left(\frac{\Delta m^2_{kj}L}{2E}\right). \tag{1.10}$$

Equation (1.10) describes the oscillation probability for massive neutrinos produced and detected as flavour eigenstates. For antineutrinos, created in a charged current weak interaction from a charged antilepton l^+_α or together with a charged lepton l^-_α, the relationship between flavour and mass states is given by the hermitian conjugate of (1.1)

$$|\bar{\nu}_\alpha\rangle = \sum_k U_{\alpha k}|\bar{\nu}_k\rangle. \tag{1.11}$$

Using this the oscillation probability for anti-neutrinos $\bar{\nu}_\alpha$ ($\alpha = e, \mu, \tau$) can be derived in a similar manner to that for neutrinos, giving the result

$$P_{\bar{\nu}_\alpha\to\bar{\nu}_\beta}(L,E) = \delta_{\alpha\beta} - 4\sum_{k>j}\mathfrak{Re}[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]\sin^2\left(\frac{\Delta m^2_{kj}L}{4E}\right) - 2\sum_{k>j}\mathfrak{Im}[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]\sin\left(\frac{\Delta m^2_{kj}L}{2E}\right). \tag{1.12}$$

This differs to the neutrino case by a sign factor in the imaginary contribution from the quartic product of U. A charge parity (CP) transformation relates neutrinos and antineutrinos (interchanges neutrino and antineutrino and reverses helicity)

$$\nu_\alpha \to \nu_\beta \overset{\text{CP}}{\longleftrightarrow} \bar{\nu}_\alpha \to \bar{\nu}_\beta.$$

By measuring the CP asymmetry[7]

$$A^{CP}_{\alpha\beta} = P_{\nu_\alpha\to\nu_\beta} - P_{\bar{\nu}_\alpha\to\bar{\nu}_\beta}$$

neutrino oscillations are sensitive to CP violation in the mixing matrix. Substituting eqns (1.11) and (1.12) this becomes

$$A^{CP}_{\alpha\beta}(L,E) = 4\sum_{k>j}\mathfrak{Im}[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]\sin\left(\frac{\Delta m^2_{kj}L}{2E}\right).$$

Thus A^{CP} can only be measured for flavour transitions $\alpha \to \beta$ where $\alpha \neq \beta$, as the imaginary part of the quartic product of U vanishes in the case $\alpha = \beta$. CPT invariance

[7] Here we adopt the definition of CP asymmetry for neutrino oscillations used by the Particle Data Group [23] instead of the standard definition of an asymmetry: $A_{AB} = \frac{P_A - P_B}{P_A + P_B}$.

implies that $A^{CP}_{\alpha\beta} = -A^{CP}_{\beta\alpha}$ and using the PMNS parameterisation from Equation (1.2) we see that $A^{CP}_{\alpha\beta}(L, E)$ becomes:

$$A^{CP}_{e\mu} = -A^{CP}_{e\tau} = A^{CP}_{\mu\tau} = \\ 4c_{13}s_{12}c_{12}s_{23}c_{23}s_{13}c_{13}\sin\delta \times \left[\sin\left(\frac{\Delta m^2_{32}}{2E}L\right) + \sin\left(\frac{\Delta m^2_{21}}{2E}L\right) + \sin\left(\frac{\Delta m^2_{13}}{2E}L\right)\right], \tag{1.13}$$

where, due to the unitary properties of U, the $\mathfrak{Im}[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]$ terms coincide up to a sign for all combinations of α, β. From this we can see that, as we would expect, CP violating effects in neutrino oscillations vanish in the limit $\delta = 0$ but also if any of θ_{12}, θ_{13} or θ_{23} are zero or 90°. In addition, as the mass squared differences satisfy $\Delta m^2_{21} + \Delta m^2_{32} + \Delta m^2_{13} = 0$ it follows that the CP violating effects vanish if any of the neutrino masses are degenerate, i.e. that any of Δm^2_{ij} are equal to zero.

1.1.1 Matter Effects

The previous expression for the transition probability assumed that the neutrinos propagate in a vacuum. In 1978 Wolfenstein [24] proposed that neutrinos propagating in matter were subject to an effective potential, due to coherent scattering with the particles in the medium, leading to a modification to the evolution of flavour states. In 1985 Mikheev and Smirnov generalised this and discovered [25] that neutrinos propagating through a medium with varying density can experience a resonant flavour transition resulting in an effective maximal mixing angle of $\pi/4$, this was coined the Mikheev-Smirnov-Wolfenstein (MSW) effect.

For simplicity we consider the MSW effect for the two flavour neutrino case[8]. To leading order in the Fermi coupling constant G_{F} the CC and NC scattering channels shown in Fig. 1.1 give effective contributions to the potential of

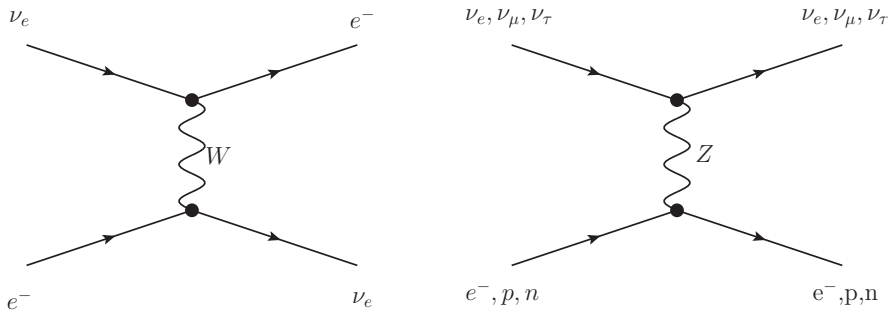
$$V_{CC} = \sqrt{2}G_{\mathrm{F}}N_e$$

and

$$V_{NC} = -\frac{1}{2}\sqrt{2}G_{\mathrm{F}}N_n,$$

where N_e and N_n are the electron and nucleon number density in matter respectively. When considering active neutrino flavours only the CC term contributes to the modification of the oscillation probability as each flavour state receives the same contribution from the NC term. This leads to an effective Hamiltonian which results in a modification to the evolution equation of the flavour states. This effective Hamiltonian has different energy

[8] A treatment in the full three flavour case adds unnecessary complexity and is not relevant to the application to solar neutrino oscillations, which turn out to be an effective two flavour system.



(a) (b)

Figure 1.1: Leading order Feynman diagrams for CC (a) and NC coherent (b) forward elastic scattering of neutrinos in matter that contribute to the MSW effect though the generation of the CC and NC potentials.

eigenstates which leads to effective mass-splitting and mixing angle in matter given by

$$\Delta m^2_M = \sqrt{(\Delta m^2 \cos 2\theta - A_{CC})^2 + (\Delta m^2 \sin 2\theta)^2}$$

and

$$\tan 2\theta_M = \frac{\tan 2\theta}{1 - (A_{CC}/\Delta m^2 \cos 2\theta)}. \tag{1.14}$$

respectively, where $A_{CC} = 2EV_{CC}$ and Δm^2 and θ are the vacuum equivalents. We can see that for $\tan(2\theta_M)$ there is a resonance at $A^R_{CC} = \Delta M^2 \cos 2\theta$ corresponding to an electron number density of

$$N^R_e = \frac{\Delta m^2 \cos 2\theta}{2\sqrt{2}EG_F}.$$

This leads to an effective mixing angle of $\pi/4$, which is maximal. For normal matter the sign of A_{CC} is positive meaning eqn (1.14) will only diverge for values of θ for which $\cos 2\theta > 0$, i.e. $\theta < \pi/4$. An important feature of the MSW effect is that now the effective mixing angle for neutrinos in matter depends on the absolute value of Δm^2. This means that experiments sensitive to the matter effect will also be sensitive to the sign of Δm^2.

1.2 Overview of Current Knowledge

We have now reached a precision age of neutrino physics [23]. We know that there are two dominant mass scales $|\Delta m^2_{21}| \ll |\Delta m^2_{31}| \simeq |\Delta m^2_{32}|$. Atmospheric and beam experiments have determined the magnitude of Δm^2_{32}, and the absolute value of Δm^2_{21} is know from the results of solar and reactor experiments, where knowledge of the sign of Δm^2_{21} comes from the sensitivity of solar experiments to the MSW effect discussed in the previous section.

Fig. 1.2 is a schematic of the current knowledge of neutrino squared-mass splittings. It shows how the uncertainty in the sign of Δm^2_{32} leads to either a normal ($\Delta m^2_{32} > 0$) or inverted ($\Delta m^2_{32} < 0$) hierarchy. Although the labelling of the massive neutrinos is arbitrary [23] we can see that with the current convention the labelling of the mass

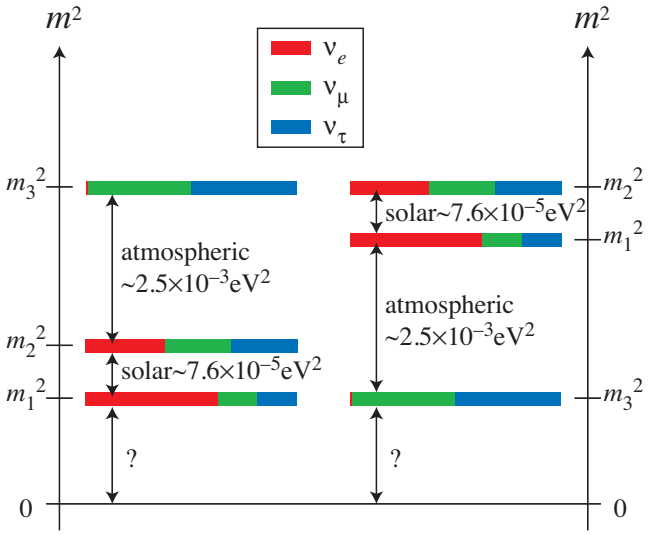


Figure 1.2: Schematic of current knowledge of neutrino squared-mass splitting, taken from [27]. Because the sign of Δm^2_{32} is unknown both the normal (left) and inverted (right) hierarchies are allowed, corresponding to positive or negative Δm^2_{32} respectively. Each neutrino mass eigenstate is coloured according to its approximate flavour content.

eigenstates 1, 2 and 3 correspond to decreasing fractions of ν_e content. As indicated on the schematic, neutrino oscillation experiments cannot tell us about the absolute mass of neutrinos. Direct mass measurements [26] based on the kinematical analysis of the end-point of the Tritium β-decay spectrum have set an upper limit on the mass of the $\bar{\nu}_e$ to be $\lesssim 2$ eV/c^2 which, given the much smaller scale for the mass splitting in turn sets an upper limit for the lightest neutrino.

Fig. 1.3 shows a summary of the current experimental constraints on the squared-mass splitting and mixing angles describing neutrino oscillations. It shows the experiments which constrain the solar (Δm^2_{12}, θ_{12}) and the atmospheric (Δm^2_{32}, θ_{23}) mixing parameters and those putting limits on the size of the only unknown mixing angle θ_{13} on the same plot.

In the central region of Fig. 1.3 we see that the intersection of the experimental results from the Super-Kamiokande (solar), SNO[9] and KamLAND experiments put tight constraints on the both the mass difference and the mixing angle describing the solar mixing parameters. They show a mass splitting scale of the order 8×10^{-5} eV^2 and the value of $\tan^2\theta_{12} \sim 0.45$ corresponds to a value of $\sin^2 2\theta_{12} \sim 0.86$. Above this the intersection of the results from the MINOS and Super-Kamiokande (atmospheric) experiments show a mass splitting scale of $|\Delta m^2_{32}| \sim 2.4 \times 10^{-3}$ eV^2 and the values of $\tan^2\theta_{23} \sim 1$ indicate near maximal mixing with a value of $\sin^2 2\theta_{23}$ close to 1. Finally we see that the CHOOZ experimental results exclude values of $\tan^2\theta_{13}$ around 1 which corresponds to putting a limit on the size of the third and only unknown mixing angle. At the time of writing the latest three-flavour fits performed by the particle data group [23] give the following values

[9] These solar results are consistent with this mixing angle only when the effect of resonant flavour transitions via the MSW effect are taken into account.

and limits:

- $\sin^2 2\theta_{12} = 0.861 + 0.026 - 0.022$
- $\Delta m^2_{12} = (7.59 \pm 0.21) \times 10^{-5}$ eV^2
- $\sin^2 2\theta_{23} > 0.92$, CL=90%
- $|\Delta m^2_{32}| = 0.00243 \pm 0.00013$ eV^2
- $\sin^2 2\theta_{13} < 0.15$, CL=90%

As shown in Equation (1.13) if any of the mixing angles are zero then it will not be possible to observe CP violation in the lepton sector. Thus determining whether θ_{13} is non-zero is one of the most intriguing questions in neutrino physics at the moment.

T2K recently released its first physics results showing an indication of $\nu_\mu \rightarrow \nu_e$ appearance in a ν_μ beam [30]. We observe 6 events passing all ν_e selection cuts compared to the 1.5 ± 0.3 expected for a value of $\sin^2 2\theta_{13} = 0$ and assuming the standard atmospheric oscillation parameters. This represents a 2.5σ deviation. At 90% the data is consistent with $0.03(0.04) < \sin^2 2\theta_{13} < 0.28(0.34)$ for $\delta_{CP} = 0$ and a normal (inverted) hierarchy. Soon after this exciting result MINOS reported similar evidence for $\nu_\mu \rightarrow \nu_e$ appearance [31] consistent with that from T2K. By combining these results with the existing reactor, solar and atmospheric data, global 3-flavour neutrino fits indicate a non-zero value of θ_{13} at the 3σ level [32].

There are still a number of unanswered questions namely:

- Is $\theta_{13} > 0$ and if it is what is its value?
- Is θ_{23} mixing maximal, i.e. is $\theta_{23} = \pi/4$?
- Is there CP violation in neutrino oscillations, i.e. is $\delta \neq 0$?
- What is the sign of Δm^2_{23}?

T2K has already started to shed light on the first two of these and is expecting to collect new data at the start of 2012 with the aim of providing conclusive evidence on the size of θ_{13}.

1.3 Neutrino Interactions: The GENIE Monte Carlo Generator

In the current era of high precision oscillation measurements, the uncertainties on the models describing neutrino interactions with nuclear targets in the few-GeV region are significant. Understanding neutrino cross sections in this energy regime is vital for both T2K in its next phase of increased statistics and also for many other current and future neutrino experiments. There are many challenges with modelling interactions in

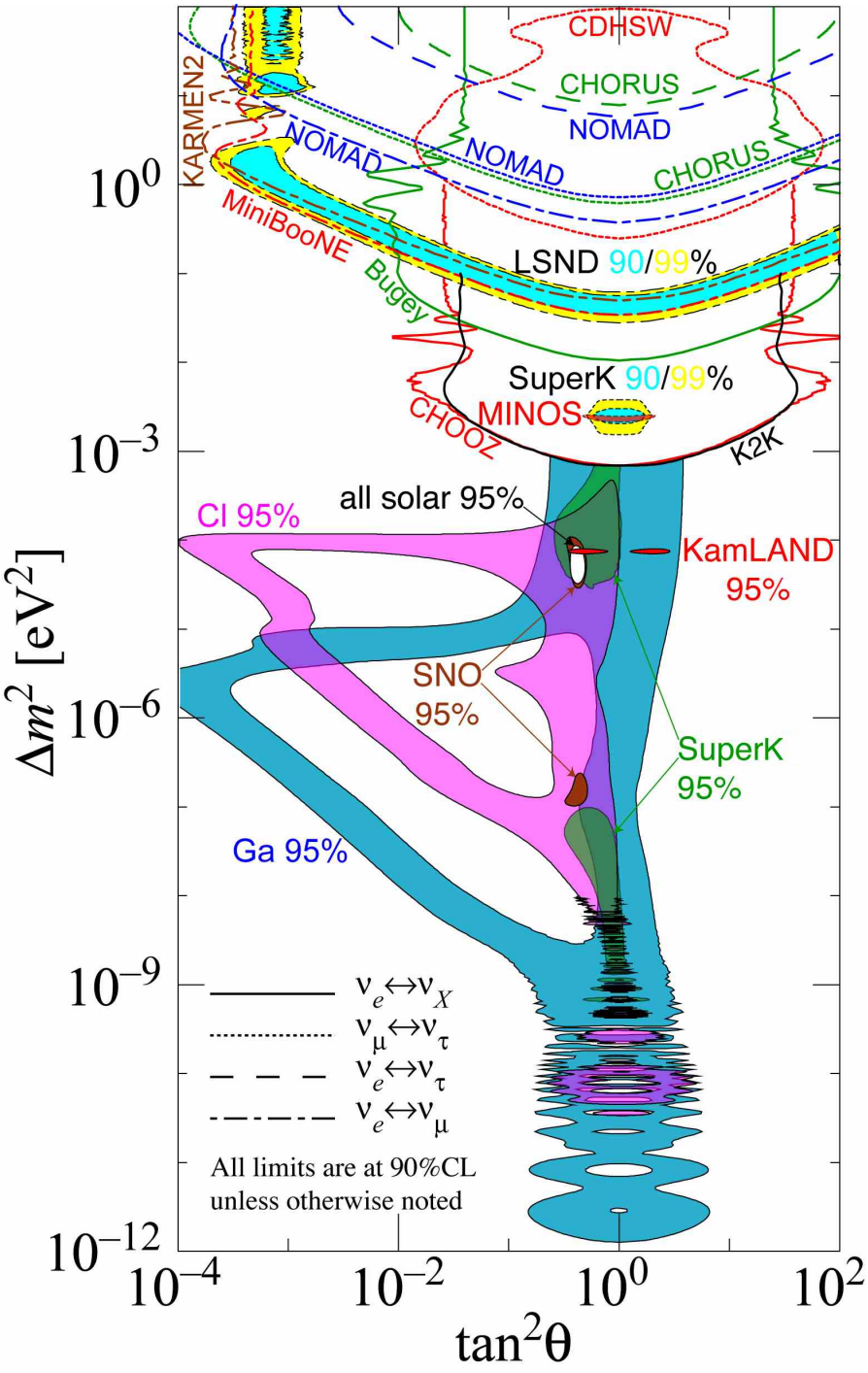


Figure 1.3: Summary of experimental knowledge of neutrino oscillation parameters, taken from [28]. All results are based on analysis of data in the relevant two-flavour approximation. The plot shows the allowed (filled) or excluded regions (un-filled and excluding their central region) for the values of mass-squared difference and mixing angle (as a function of the tangent squared) set by each experiment. A description of the various experimental limits is given in the text. For vacuum oscillation all contours are symmetrical about the $\tan^2(\theta) = 1$ vertical, where values of $\tan^2(\theta) < 1$ and $\tan^2(\theta) > 1$ correspond to positive and negative Δm^2 respectively. The solar neutrino experiments can be seen to break this symmetry as they are sensitive to matter effects, and hence the sign of Δm^2. A full discussion of the motivation for plotting in the $\tan^2(\theta)$ space is given in [29].

this transition region between perturbative and non-perturbative scattering and there is a worldwide effort to solve these which includes dedicated neutrino cross section experiments and a large degree of communication between the experimental and theoretical communities [33]. T2K is fortunate enough to have a number of independent Monte Carlo (MC) generators integrated with its software simulation framework. In this thesis we focus on one of these, the GENIE (Generates Events for Neutrino Interaction Experiments) MC generator [34]. We now give an outline of the main areas of importance when simulating neutrino interactions and the models which GENIE uses for each of these.

Broadly speaking, the modelling of neutrino interactions is categorised into three areas:

- The description of the nuclear environment in which the interaction takes place.
- The primary cross section models describing the scattering processes.
- The formation and propagation of hadrons produced in the primary interaction.

An important aspect of the simulation is ensuring that boundaries between the regions of different validity for the variety of models are treated in such a manner as to avoid theoretical inconsistencies, discontinuities in the generated observables and double-counting in the regions of overlap.

1.3.1 The Nuclear Model

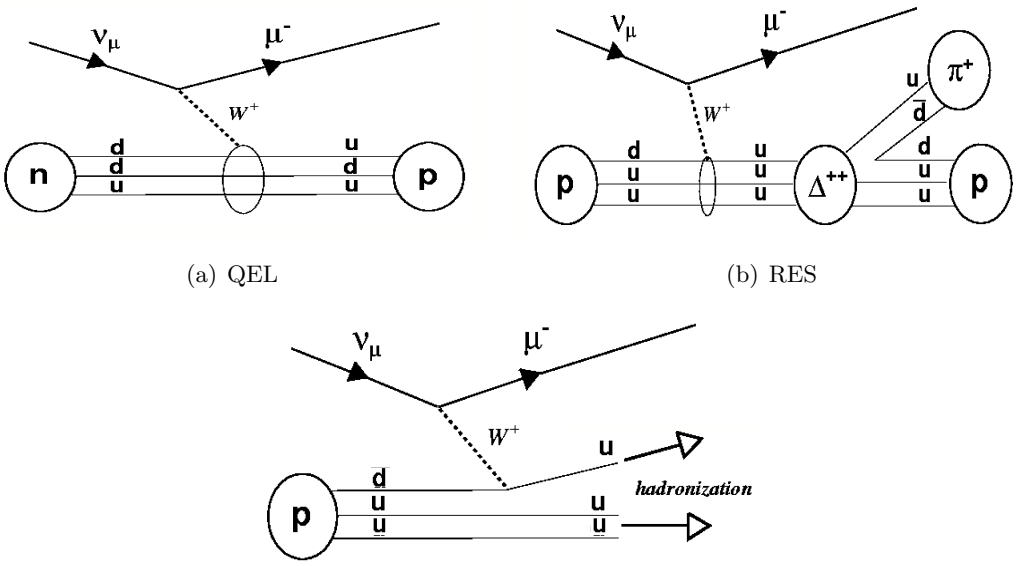
In order to simulate scattering off the nuclear targets that typically make up the fiducial mass of the detectors, the effect of the nuclear environment must be taken into account. In GENIE the Impulse Approximation (IA) is used in which the bound nature of the struck nucleon is ignored and where the effects of the nuclear environment are incorporated using the Relativistic Fermi Gas (RFG) model. In the RFG the nucleons are treated as non-interacting fermions where the nucleon occupancy is characterised using a simple step function: $n(k) = \theta(k_F - k)$. The A-dependence of different targets are taken into account by choosing the relevant fermi momentum k_F and an average nucleon binding energy (the energy needed to remove the nucleon from the nuclear environment). The values for these are taken from electron scattering experiments [35]. In particular GENIE uses a modified version of the RFG model by Bodek and Ritchie which takes into account short range nucleon-nucleon correlations [36].

The location of the neutrino interaction within the nucleus and the density profile of the nucleus are important from the point of view of hadronic final state interactions, which are discussed in § 1.3.4. For nuclei with $A < 20$, a modified Gaussian density parameterisation is used whereas for heavier nuclei the Woods-Saxon [37] density function is used. The mass densities themselves are taken from review articles on electron scattering data [38]. The interaction is then placed at a location inside the nucleus with a probability based on the density profile. This approach is used for all nuclei where the fit parameters known for certain nuclei are used with simple A-dependent interpolations to work for others.

Other effects due to the nuclear environment such as Pauli blocking and differences between the nuclear and free nucleon structure functions are included in the calculation of the cross sections themselves.

1.3.2 Cross Section Processes

Many different scattering processes are important in the few-GeV energy region. These describe neutrino scattering off a variety of targets including the nucleus itself, the individual nucleons, the quarks within the nucleons and atomic electrons. Fig. 1.4 shows the dominant charged current scattering modes at T2K energies. Fig. 1.5 shows the GENIE default cross section prediction for ν_μ charged current scattering from an isoscalar target compared to data from a wide range of experiments. Also shown is the estimated uncertainty on the total cross section.



(a) QEL

(b) RES

(c) DIS

Figure 1.4: Feynman diagrams for the dominant charged current scattering processes [39].

We now outline the cross section models used in GENIE that are relevant at T2K energies (~ 0.5 GeV).

Quasi-Elastic Scattering (QEL): Quasi-elastic scatting ($\nu_\mu + n \rightarrow \mu^- + p$) is based on an implementation of the Llewellyn-Smith model [40]. In this model the most general Lorentz invariant form for the hadronic weak current is reduced (using parity arguments) to two vector form factors, a pseudo-scalar form factor and an axial form factor. The vector form factors are related to electromagnetic form factors using the conserved vector current hypothesis (CVC), these are then taken from recent fits to electron scattering data [41]. The pseudo-scalar form factor is given the form

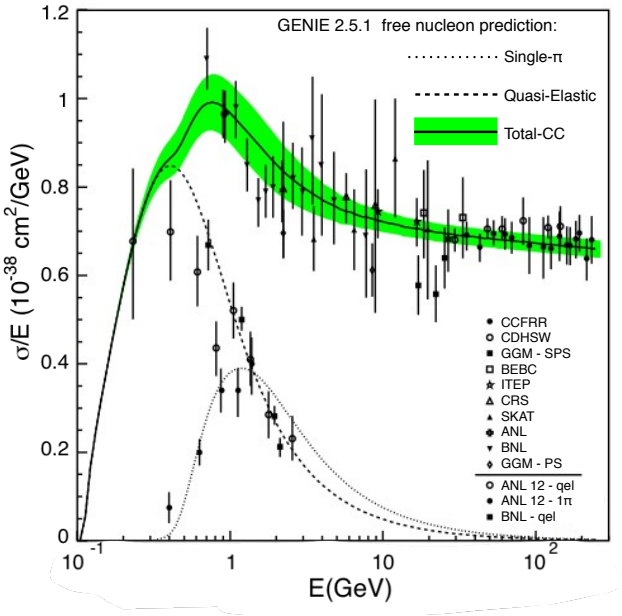


Figure 1.5: The default GENIE prediction for charged current ν_μ scattering from an isoscalar target. The quasi-elastic, single-π, and total charged current (with estimated uncertainty in green) predictions are compared to data from a wide range of experiments [34].

suggested by the partially conserved axial vector current (PCAC) hypothesis. This leaves the axial form factor $\mathrm{F}_A(Q^2)$ as the only unknown quantity, where Q^2 is the 4-momentum transfer of the neutrino to the target nucleon. In GENIE a dipole form is assumed for the axial form factor $\mathrm{F}_A(Q^2) \propto (1 + Q^2/M_{\mathrm{A}}^2)^{-2}$, with the only unknown parameter being the quasi-elastic axial mass M_{A}. The value of this is determined from fits to existing neutrino data and is currently set to a value 0.99 GeV/c^2.

Baryon Resonance Production (RES): Neutrino induced production of baryon resonances for both neutral current (NC) and charged current (CC) scattering is based on the Rein-Sehgal model [42] which employs the Feynman-Kislinger-Ravndal (FKR) description of resonance excitations [43]. In their original paper Rein and Sehgal consider 18 resonances and GENIE uses the 16 of these listed as unambiguous in the 2010 Review of Particle Physics [44]. Interference between the different resonances is not taken into account. There are many free parameters in the Rein-Sehgal model and in § 4.1.3 we consider uncertainties coming from the values of the axial and vector form factors for both CC and NC scattering ($M_{\mathrm{A}}^{\mathrm{CCRES}}$, $M_{\mathrm{V}}^{\mathrm{CCRES}}$, $M_{\mathrm{A}}^{\mathrm{NCRES}}$ and $M_{\mathrm{V}}^{\mathrm{NCRES}}$) assuming a similar dipole form as in the case of QEL scattering.

Coherent Neutrino-Nucleus Scattering (COH): Coherent neutrino-nucleus scattering is defined as any interaction where the nucleus is left in its ground state after scattering. This typically results in the production of forward-going pions for both neutral current ($\nu_\mu + A \to \nu_\mu + \pi^0 + A$) and charged current ($\nu_\mu + A \to \mu^- + \pi^+ + A$)

scattering. These are implemented using a recent revision of the Rein-Sehgal [45] model of coherent production which includes lepton mass corrections leading to suppression of reactions with forward-going μs due to destructive interference between form factors. The most important parameters in this model are the axial mass $M_{\mathrm{A}}^{\mathrm{COH}}$ and the nuclear size parameter R_0, a parameter that controls the amount of pion absorption included in the calculation of the cross section. Both of these are considered when calculating the effect of neutrino cross section uncertainties in § 4.1.3.

Deep Inelastic Scattering (DIS): Neutrino induced DIS scattering is calculated using an effective leading order (LO) model with modifications suggested by Bodek and Yang [46] to improve description of scattering at low Q^2 and high Bjorken x, where in the parton model, x is the fraction of the nucleons momentum carried by the struck quark. The parameter values in the Bodek-Yang model are set through fits to low-Q^2 electron-scattering data. Their model is an effective LO model as it includes corrections to account for the neglected higher-order terms—these corrections modify the parton distribution functions (PDFs) and are applied on top of the GRV98 PDF set [47]. In § 4.1.3 we consider only the A_{HT} and B_{HT} higher-twist parameters and the $C_{V1\mu}$ and $C_{V2\mu}$ corrections to the parton distribution functions.

Other scattering mechanisms: A number of other scattering mechanisms are also modelled including elastic neutral current scattering (NCEL), QEL- and DIS-Charm production, inverse muon decay (IMD), and neutrino-electron elastic scattering.

Full details of the neutrino cross section models used can be found in the GENIE physics and user manual [48].

1.3.3 Transition Region Tuning

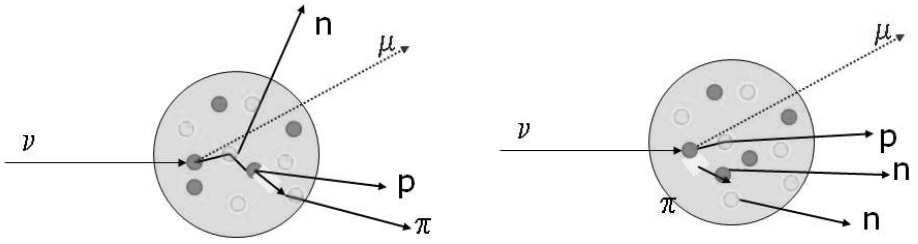
There is overlap in the transition region between resonance production of hadrons and those produced through DIS scattering. Both processes can produce similar final states which from the point of view of an experimental signal are indistinguishable. This can lead to double counting where the RES and DIS production models both account for hadron production in this region when they are fit separately to data. To avoid this, GENIE includes a set of suppression factors on the production of hadrons through non-RES (DIS) processes in this region. There is a separate suppression factor for all combinations of CC and NC scattering, hit nucleon type, and number of pions in the final state. These are tuned so that the total rate from both RES and non-RES production replicates both inclusive cross section data and exclusive 1-pion and 2-pion cross section data. These suppression factors are the R_{bkg} terms considered in § 4.1.3.

1.3.4 Hadronisation and Final State Interactions

GENIE uses the Andreopoulos-Gallagher-Kehayias-Yang (AGKY) model [49] to describe the formation of hadrons in DIS scattering events. This was developed for the MINOS experiment and combines an empirical Koba-Nielsen-Olesen (KNO) based low-hadronic invariant mass model [50], anchored to bubble chamber data for $\nu/\bar{\nu}$ interactions on hydrogen and deuterium, with the PYTHIA-6 fragmentation model [51] at higher invariant masses. To avoid discontinuities, there is a smooth transition between the two models over an adjustable region with a default range of W from 2.3 to 3.0 GeV/c^2.

It is well known that hadrons produced in the nuclear environment do not immediately re-interact with their full cross section. This is because during the time it takes for the quarks to fully materialise as hadrons they propagate through the nucleus with a dramatically reduced interaction probability. In GENIE this is implemented as a free-step given to all hadrons before they are passed to the intranuclear rescattering model. The free step or 'formation zone' is determined based on a characteristic formation time of 0.523 fm/c as determined by data from the SCAT experiment [52].

The hadronisation models describe particle production for free nucleons and are primarily tuned to bubble chamber data on hydrogen or deuterium. However, hadrons produced in the nuclear environment of a heavier nuclear target have to travel through the nucleus and may rescatter before escaping. This is important as the propagation through the nuclear environment degrades energies and alters event topologies and so directly affects the observables that a detector will see and its ability to reconstruct exclusive cross section processes. Fig. 1.6 shows a schematic for final state interactions which lead to pion production and pion absorption.



(a) Pion production (b) Pion absorption

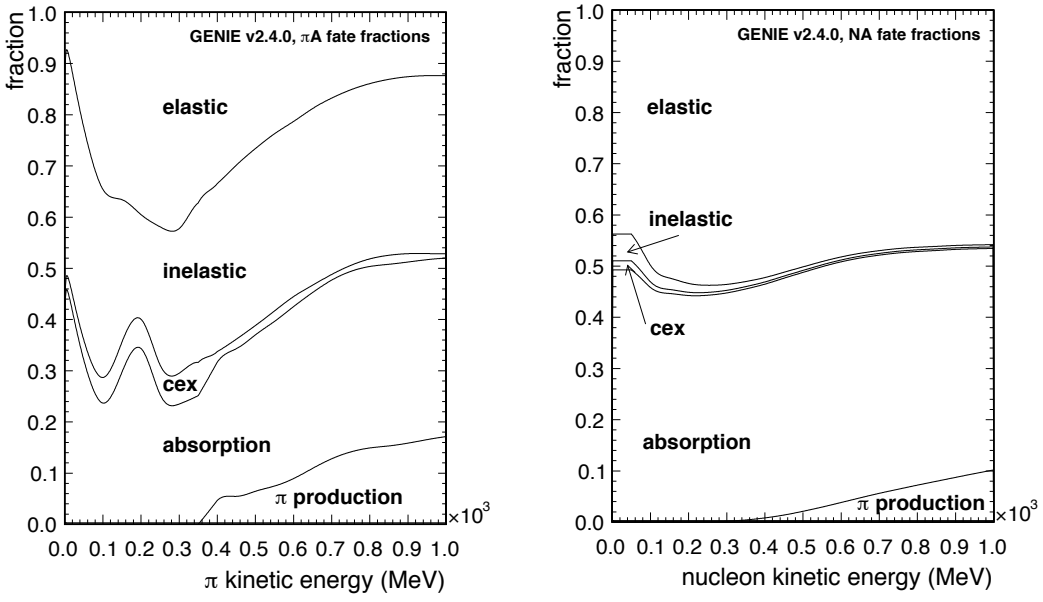
Figure 1.6: Schematic of hadronic final state interactions showing pion production from a nucleon and pion absorption in the nucleus [48].

There are currently two models implemented in GENIE to simulate hadronic final state interactions (FSIs). The first is the INTRANUKE/hA model [34, 53]. It is a simple empirical model. Rather than treat the propagation of each hadron as a cascade of many successive interactions it uses the total inclusive cross section for each possible type of nuclear scattering for pions and nucleons up to 1.2 GeV. It is called the hA model because the relative and total rate for each rescattering process are based on data for pion-nucleus

and proton-nucleus interactions. In the model the total reinteraction rate and the type of rescattering given a reinteraction does occur are factorised. To determine if a hadron reinteracts the mean free path (λ) is calculated based on the local density of nucleons and analysis of a large body of hadron-nucleon cross sections [54]:

$$\lambda(r, E_h) = \frac{1}{\sigma_{hN}(E_h)\rho(r)}.$$

The hadron is then stepped through the nucleus and the probability for each step is calculated based on $\lambda(r, E_h)$. A MC method is used to decide whether the hadron reinteracts in the next step. This is repeated until the hadron either rescatters or escapes the nucleus. If the hadron does rescatter then the type of scattering is chosen based on the relative probabilities for the rescattering processes as determined from hadron nucleus scattering data for exclusive final states. The following types of rescattering are considered: elastic, inelastic, charge exchange, absorption and pion production. Here we collectively refer to these as the rescattering fates. Fig 1.7 shows the relative probability of for each rescattering fate as a function of momentum for both pions and nucleons. Each rescattering fate represents many types of scattering modes. Because of its simplicity, reweighting schemes for the INTRANUKE/hA model can be developed which allow, within the context of the model, full evaluation of the systematics associated with hadronic final state interactions.



(a) Scattering of pions

(b) Scattering of nucleons

Figure 1.7: The default fate fractions for pions and nucleons in the INTRANUKE/hA model. The area given to each fate reflects the probability for that fate to occur given a rescattering takes place and is given as a function of kinetic energy [55].

The INTRANUKE/hN [53] is an alternative model for simulating hadronic final state interactions. It is a full intranuclear cascade model. Here the pion and nucleon are

allowed to rescatter many times before exiting the nucleus. The probability for each type of rescatter is based on partial wave analysis of pion nucleon and nucleon-nucleon scattering data. The INTRANUKE/hN model is not the default INTRANUKE model in the version of GENIE currently used by T2K.

1.3.5 Simulating Event Generation For Realistic Flux and Detector Geometries

An important aspect of event generation for modern neutrino experiments is the need to simulate neutrino interactions over realistic detector geometries, using detailed flux predictions as input. This is computationally demanding and non-trivial given the large number of atomic targets present in a modern neutrino detector and the complex geometrical shapes. In addition particular care has to be taken as any mistakes at this stage would lead to incorrect predictions on the event rate for a given neutrino flux. GENIE has a comprehensive set of libraries with which to do this and provides experiment specific applications tailored to accept the required input format from the flux simulation and provide the necessary output format required by the experiment software. Fig. 1.8 show the distribution of vertices over the tracker-region of the T2K near detector. It was generated using the GENIE libraries and the T2K specific event generation program which interfaces with the output of the experiment specific flux simulation (see § 2.1.4) and shows a non-trivial distribution of vertices which pick out the regions of high-density.

During my time on T2K I have been an active member of the GENIE collaboration. In particular I acted as the contact point between T2K and GENIE. My work included validation of the T2K specific event generation software and optimisation of the code which simulates neutrino interactions over the complex geometries of the T2K near detectors.

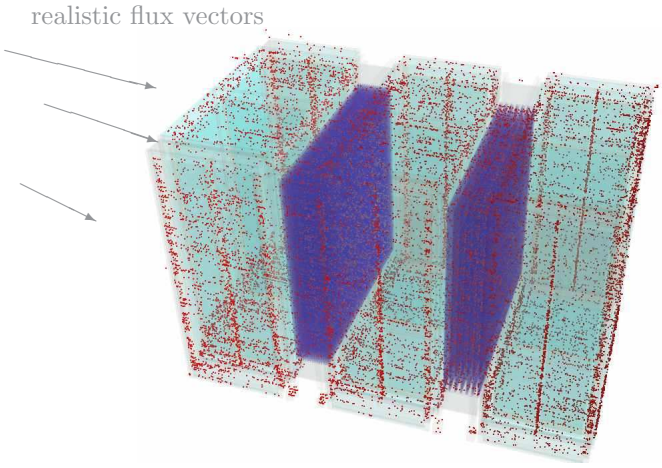


Figure 1.8: The distribution of GENIE vertices (red dots) generated over the tracker-region of the T2K near-detector. Event generation performed using the *gT2Kevgen* custom driver for T2K which takes as input the T2K vector-level flux output and complex ND280 ROOT geometry. The high density regions of vertices can be seen to pick out the two FGD volumes and the mechanical support structure of the FGDs [56].

2 The T2K Experiment

The T2K (Tokai-to-Kamioka) experiment is a long baseline neutrino oscillation experiment designed to study the mixing of muon neutrinos using a high intensity off-axis muon neutrino beam. The primary goals are to look for electron neutrino appearance in the muon neutrino beam, thereby measuring the only unknown mixing angle θ_{13}, and to make precision measurements of the parameters controlling muon neutrino disappearance Δm^2_{23} and $\sin^2 2\theta_{23}$ to within $\delta(\Delta m^2_{23}) \sim 10^{-4}\text{eV}^2$ and $\delta(\sin^2 2\theta_{23}) \sim 0.01$ respectively.

Fig. 2.1 shows a schematic of the experimental setup. A beam of muon neutrinos is produced at the newly-constructed J-PARC accelerator complex on the east coast of Japan. The beam is directed through a suite of near-detectors, 280 m downstream of the production point, towards the far-detector 295 km to the west. T2K employs an off-axis method—by deliberately directing the neutrino beam so that it is at an angle of 2.5° with respect to the line of sight between the production point and the far detector, the energy spread of the beam is greatly reduced, creating a narrow-band neutrino beam with a peak energy of about 0.6 GeV. This energy is at the oscillation maximum for ν_μ disappearance, which maximises sensitivity to θ_{23} and θ_{13} and reduces backgrounds to the ν_e appearance measurement.

The Super-Kamiokande [58] water Cherenkov detector acts as the far detector for T2K and measures the beam's flavour composition after oscillation to search for $\nu_\mu \to \nu_e$ appearance and ν_μ disappearance. The near detectors sample the beam prior to oscillation and measure the neutrino energy spectrum and flavour content. There are two near detectors: the Interactive Neutrino GRID (INGRID) which measures the on-axis beam profile

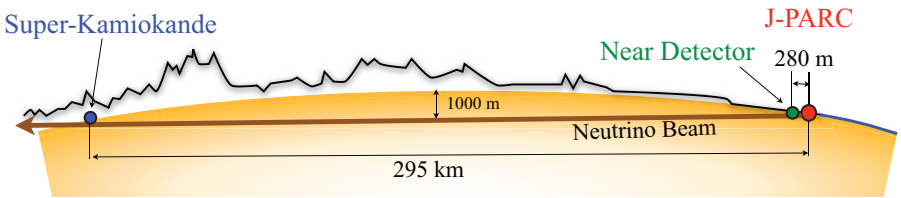


Figure 2.1: A neutrinos journey from the production point at J-PARC, through the near detector suite at 280 m and then across 295 km underneath the main Island of Japan before reaching the far detector, Super-Kamiokande [57].

and intensity, and a magnetised off-axis tracking detector (ND280) which is located along the off-axis angle and measures the interaction rate, intensity and flavour composition of the beam before oscillation.

The main purpose of this section is to give an overview of the components that make up T2K. We then go on to describe T2K's first ν_μ-disappearance result and discuss the motivation for studying neutrino induced charged current π^+ production at the off-axis near detector. A full description of the T2K experiment has been published [57].

2.1 The T2K Neutrino Beam

The T2K neutrino beam is produced at the newly-constructed J-PARC research facility. A high-intensity proton beam impinges on a fixed graphite target to produce a slew of secondary hadrons. A series of magnetic horns is then used to sign-select[1] and focus those with positive charge. This focused beam, dominated by π^+s but also with a significant fraction of K^+s and other mesons, then enters a 96 m decay volume. Whilst the π^+s decay almost exclusively (99.99%) to $\mu^+ + \nu_\mu$, approximately 1/3 of the K^+ and other hadron decays will instead produce ν_e and $\bar{\nu}$ as well as tertiary hadrons (which in turn can decay to produce ν_es) resulting in a small ν_e contamination, of about 1.5%, of the ν_μ beam. The focusing of the parent particles and the relativistic boost given to the decay products produces a highly collimated beam of neutrinos. The leptons and any other particles are then absorbed by a graphite beam dump leaving only the neutrino beam.

The T2K beam differs from existing neutrino beams in two important ways. Firstly, the high intensity of the proton beam combined with the high current three-horn focusing system leads to a much higher intensity beam than previously possible, far exceeding that of previous accelerator beams. Secondly, it employs an off-axis design with an off-axis angle chosen to produce a narrow beam enhanced at the specific neutrino energy at which oscillations are expected at the Super-Kamiokande detector. We now describe the various components of the neutrino beam, its simulation and its current status.

2.1.1 The J-PARC Accelerator

To produce the 30 GeV intense proton beam, J-PARC uses three stages of acceleration: a linear accelerator (LINAC) followed by a Rapid-Cycling Synchrotron (RCS) and the Main Ring synchrotron (MR). In the first stage a beam of H^- anions are accelerated to up to 400 MeV (181 MeV at present) using the LINAC before being converted to a H^+ beam using charge stripping foils prior to injection into the RCS. The RCS then accelerates this proton beam up to 3 GeV, with a 25 Hz cycle rate and with 2 bunches per cycle. About 5% of these bunches are fed into the MR and accelerated up to 30 GeV. The MR has a harmonic number of 9 and stores 8 bunches (6 for the initial run period before June 2010). When run in fast extraction mode, all 8 bunches are extracted within a single cycle using a

[1] Focus particles of a particular sign charge and defocus those of opposite sign—see §2.1.2.

Table 2.1: Machine design parameters of the J-PARC MR for the fast extraction [57].

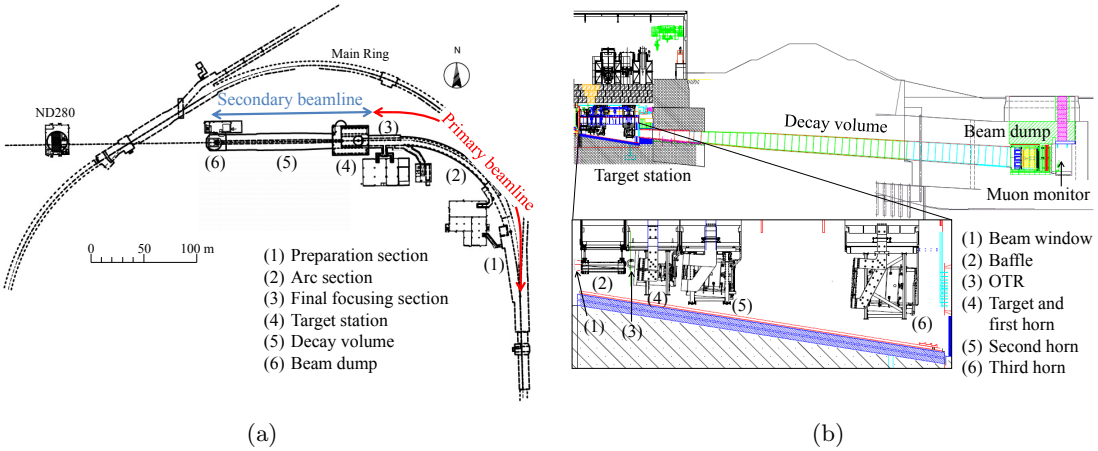
Circumference	1567 m
Beam power	∼750 kW
Beam kinetic energy	30 GeV
Beam intensity	$\sim 3 \times 10^{14}$ p/spill
Spill cycle	∼0.5 Hz
Number of bunches	8/spill
RF frequency	1.67 – 1.72 MHz
Bunch interval	581 ns
Bunch width	58 ns
Spill width	∼5 μsec

set of five kicker magnets, producing a spill containing 8 bunches, each of width 58 ns, over a total duration of approximately 5 μs. This time structure is key to reducing neutrino backgrounds, such as coincident cosmic rays, in the various detectors. The specific design parameters of the proton beam at extraction are shown in Table 2.1.

2.1.2 The Neutrino Beamline

After fast extraction from the MR the proton beam enters the T2K beamline which is composed of two sequential sections: the primary section, which serves to align the beam so that it points in the correct direction, and the secondary section, which contains the graphite target, focusing horns, decay volume and beam dump used to produce the neutrinos. Fig. 2.2(a) shows a schematic of their arrangement. The primary beamline consists of a preparation, arc, and final focusing section. After being tuned in the preparation section, using 11 normal conducting magnets (NCMs), the beam enters the arc section and is bent by 80.7°, using 14 doublets of superconducting combined function magnets (SCFMs) in addition to three pairs of horizontal and vertical steering superconducting magnets for correcting the beam orbit. Then in the final section 10 NCMs focus and guide the beam to the target as well as directing it at an angle of 3.637° to the horizontal that is necessary to produce the 2.5° offset at the far detector.

The secondary beamline consists of the target station, the decay volume and the beam dump. Fig. 2.2(b) shows their arrangement. After entering the target station the beam passes through a baffle, to reduce exposure of the horn to beam loss, and then through the optical transition radiation monitor (OTR), which measures the beam profile by collecting transition radiation produced as the beam passes through a thin layer of titanium-alloy foil. The beam then impinges on the target, a 91.4 cm long, 2.6 cm diameter graphite rod of density 1.8 g/cm^3. The target sits within the first of three magnetic horns. Pions produced in the target are sign-selected by the first horn and then focused by the second and third horns. A magnified view of the target station is shown in the bottom half of Fig. 2.2(b). Each horn consists of two coaxial aluminium alloy conductors encompassing



(a) (b)

Figure 2.2: Shows an overview of primary and secondary sections of the T2K beamline in (a) and a side view of the secondary beamline in (b), where the decay volume is ~ 96 m in length and the beam enters from the left [57].

a closed cylindrical volume. When pulsed (in time with the beam spill) with 320 kA, the horns generate a maximum magnetic field of 2.1 T. Before installation the magnetic field of the horns was measured using a Hall probe to within an uncertainty of 2% for the first horn and 1% for the second and third horns. The horns were designed to maximise the neutrino flux at the far detector—with the current design, the flux at Super-Kamiokande is increased by a factor of about 16 compared to that when the horns are switched off.

The focused beam of pions[2] then enter the decay volume, a 96 m-long steel tunnel surrounded by 6 m of concrete shielding. In total, the distance from the centre of the target and the start of the beam dump is 109 m, meaning that almost all of the π^+s will have decayed to $\mu^+ + \nu_\mu$ pairs before reaching the beam dump. The beam dump is composed of a graphite core (with a total thickness of 3.174 m) and 15 iron plates (total thickness 2.4m), through which only muons above about 5 GeV can penetrate.

2.1.3 Beamline Monitoring

The primary beamline contains a number of different monitoring systems to ensure a stable proton beam. Five current transformers (CTs) measure the proton beam intensity to within 2% absolute and 0.5% relative intensity, and also measure the beam timing to within 10 ns. The beam position is measured non-destructively to better than 450 μm (the design requirement is 500 μm) using 21 electrostatic monitors (ESMs). The beam profile is measured using 19 segmented secondary emission monitors (SSEMs)[3] giving a beam width measurement to 200 μm. The optical parameters of the beam (Twiss parameters and emittance) are reconstructed from the profiles. Because the SSEMs make a destructive measurement causing 0.005% beam loss, they are only inserted during beam tuning. There are also 29 beam loss monitors (BLMs) installed around the beam pipe and in the final

[2] In addition to this there will be a contamination from accidentally focused kaons.
[3] See [59] for description of the workings of an SSEM.

focusing section. Each BLM signal is integrated during a spill and if it exceeds a critical value the beam is aborted to prevent overexposure to personnel and damage to machines.

Immediately after the beam dump is the muon monitor. This measures the muon flux of the penetrating muons using two different detector arrays, oriented with their normal along the beam direction and separated by 1.2 m. Each array covers an area of 1.5 m by 1.5 m and is instrumented with 49 sensors. The upstream array is instrumented with ionisation chambers whereas the downstream array uses silicon PIN photodiodes. The muon monitor can be used to measure the neutrino beam intensity and direction (based on the line vector between the target centre and the beam profile centre at the muon monitor surface) on a per-bunch basis. The resolution is less than 0.1% for the beam intensity and 0.3 cm for the beam profile centre (this is 10 times better than that needed to reduce the directional uncertainty of the beam to less than the required 0.25 mrad).

After every beam run the beam working group analyse the data collected by the various monitors in both the primary and secondary beamlines and provide spill-by-spill statistics on quantities such as the beam power, alignment and synchronisation, as well as the number of protons on target (POT) making each bunch. The integrated number of POT is used to normalise MC simulations of the beam to that in data. This information is summarised and incorporated into the Super-Kamiokande and ND280 data chains to allow beam quality cuts and POT counting on a spill-by-spill basis.

2.1.4 Beam Simulation and Composition

An accurate prediction of the neutrino beam at both ND280 and Super-Kamiokande is integral to the success of T2K and the JNUBEAM [60] simulation is used to do this. JNUBEAM is a mixture of custom-built and external simulation packages which models all aspects of the beam production, from the injection of the primary proton beam and its interaction with the target to the tracking and subsequent decays of the secondary and tertiary hadrons that produce the neutrinos. A key element of this simulation is the use of external hadron production data to constrain the predictions on the neutrino flux. In particular T2K uses data from the NA61/SHINE experiment [61] at CERN which produced dedicated hadron production cross section measurements for proton-Carbon interactions. The beam simulation is then tuned to reproduce these results before being used to predict the flux for T2K. NA61/SHINE includes both a thin target, for measuring hadron production with a reduced rate of secondary interaction within the target, and a replica of the T2K target which provides coverage of the region of phase space important for T2K. The complete JNUBEAM simulation consists of a number of steps:

- 30 GeV protons are generated upstream of the baffle and their hadronic interactions with the target and the subsequent hadronic chains (secondary interaction) inside the target are simulated using the FLUKA hadron production model [62], which was chosen because it best replicated the NA61 measurements.
- The particles exiting the target (mainly pions and kaons) are tracked using a GEANT3-

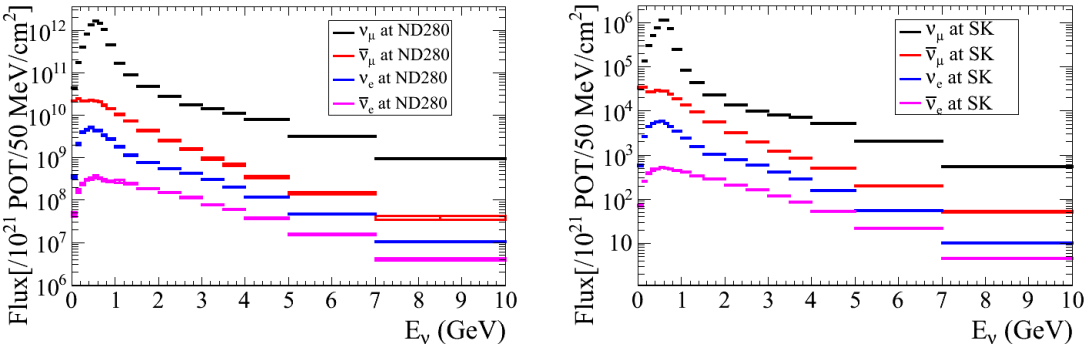


Figure 2.3: Neutrino flux expected at (b) ND280 and (b) Super-Kamiokande broken down into the neutrino type and as a function of energy [65].

based [63] simulation where subsequent hadronic interactions are modelled using the GCALOR hadron production model [64]. When tracking the particles, JNUBEAM uses a complete description, including simulation of the magnetic fields, of the secondary beamline which includes the target, horn magnets, decay volume, beam dump and muon monitor.

- The particles are then traced into the decay volume where they decay to produce neutrinos. In JNUBEAM decays from $\pi^{\pm}$, $K^{\pm}$, K^0_L and $\mu^{\pm}$ are considered using the current best knowledge of the various branching ratios and decay rates from the Particle Data Group [44]. The path of each neutrino is projected to see if it intersects with either the near detector or far detector locations, if it does then it is saved. For each simulation file a record of the number of protons on target (POT) used in the simulation is stored to allow normalisation to data samples in analyses.

- After running the default simulation the results are tuned to reproduce the available hadron production data. This includes varying the pion production multiplicity in $p + \mathrm{C} \rightarrow \pi^{\pm} + X$ reactions to match that measured by NA61 and varying the interaction rate for $p+\mathrm{C}$, $\pi^{\pm}+\mathrm{C}$, $K^{\pm}+\mathrm{C}$, $p+\mathrm{Al}$, $\pi^{\pm}+\mathrm{Al}$ and $K^{\pm}+\mathrm{Al}$ reactions to fit existing data. The final tuning is then made available and applied at the analysis stage.

The MC simulation of the unoscillated neutrino flux at the ND280 and Super-Kamiokande detector locations is shown in Fig. 2.3, broken down into neutrino type. This is based on the most recent version of the JNUBEAM simulation (11a). As expected the flux is dominated by ν_μ with a peak at around 600 MeV. We can see the dominant contamination of the beam comes from $\bar{\nu}_\mu$ and ν_e. This intrinsic beam ν_e component is one of the dominant backgrounds to the ν_e appearance search. Reducing the uncertainties on the predicted rate at Super-Kamiokande will hence be important for T2K to reach its final sensitivity, and measurements of the ν_e component of the beam at ND280 will help constrain these.

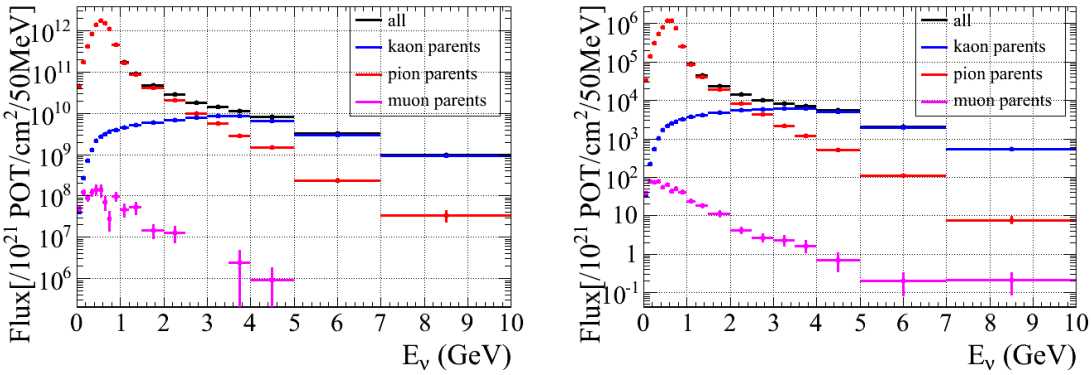


Figure 2.4: ν_μ component of the flux predicted at (a) ND280 and (b) Super-Kamiokande broken down by parent hadron type. Neutrinos from the decay of kaons can be seen to dominate at energies above approximately 3 GeV [65].

Fig. 2.4 shows the ν_μ component of the beam broken down by neutrino parent hadron type. From this we can see that for energies less than $\sim$ 3 GeV, neutrinos produced by the decay of pions are dominant, while at higher energies those from the decay of kaons dominate. There is more uncertainty in the rate and multiplicity of kaon production in the JNUBEAM simulation, as at present only the pion production rates have been tuned to NA61/SHINE data. The effect of this will be seen when evaluating the systematic uncertainty from the flux simulation on the selections developed in § 3.

2.1.5 Commissioning and Current Status

The accelerator and neutrino beamline were successfully commissioned during 2009 and accumulation of neutrino beam data for physics analysis started in January 2010. Since then there have been two continuous physics runs: Run 1 (January to June 2010) and Run 2 (Nov 2010 to March 2011). Over the course of these runs the MR proton beam power was continually increased reaching 145 kW, corresponding to 9×10^{13} protons per pulse. The period between the two runs was used to make improvements to the beam hardware and to install the Barrel-ECal, the last remaining sub-detector, into ND280. Fig. 2.5 shows the accumulated POT and a measure of the instantaneous beam intensity (protons per bunch) throughout the two run periods. Because of the continually increasing power of the beam the accumulated POT for the Run 2 period (see after Aug/10) dominates that for the whole period. The Run 2 period was stopped prematurely because of the tragic events of the March 2011 Tohoku earthquake. The neutrino beam is expecting to be re-commissioned by the beginning of 2012.

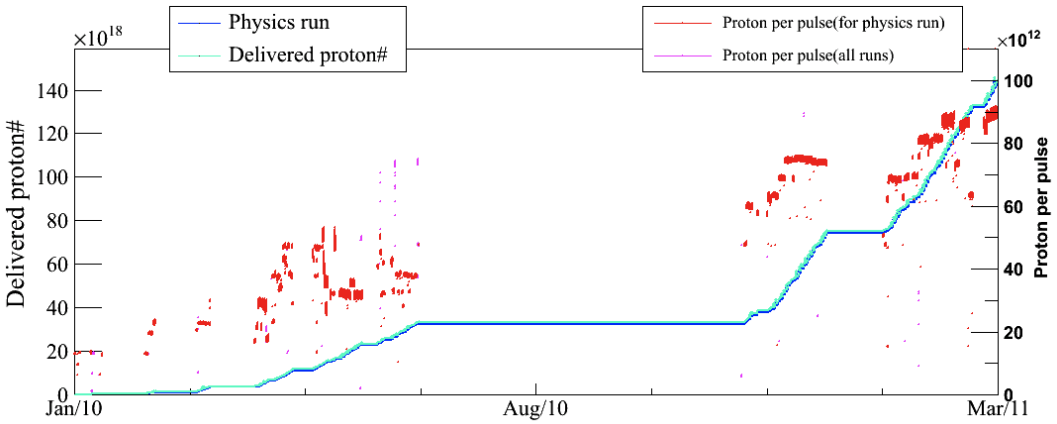


Figure 2.5: The accumulated POT and instantaneous beam intensity (protons per pulse) for Run 1 and 2 beam periods [65].

2.2 The INGRID On-axis Detector

The Interactive Neutrino GRID (INGRID) is designed to measure the on-axis neutrino flux. It consists of 14 identical modules arranged in a cross with its plane centred on and perpendicular to the nominal beam direction (defined by the primary proton beam direction). It forms an important component of the beam monitoring system, especially given the off-axis nature of the beam and the sensitivity of the neutrino energy spectrum at Super-Kamiokande to changes in the off-axis angle and beam profile.

In total the detector spans a transverse section of 10 m × 10 m. The arrangement of the modules can be seen in Fig. 2.6. In addition to the 14 horizontal and vertical modules, two diagonal modules were installed to enable checks of the axial symmetry of the 2D beam profile. By comparing the neutrino interaction rates in each of the modules the beam intensity, direction and spread can be inferred. The current setup yields a measurement of the beam centre to a precision of better than 10 cm which corresponds to an angular resolution of less than 0.4 mrad with respect to the beam origin.

The tracking region of each module consists of 11 scintillator planes interleaved with 9 iron plates[4]. The dimensions of the iron plates are 124 cm × 124 cm in the x and y directions and 6.5 cm along the beam direction, giving a total target mass of iron of 7.1 tons per module. The fiducial mass of the modules is such that for the nominal T2K beam there will be sufficient statistics to make a daily intensity measurement. Each scintillator tracking plane is made up of 24 scintillator bars in the horizontal direction, and another 24 bars glued perpendicular to these in the vertical direction, where each bar has dimensions of 1.0 cm × 5.0 cm × 120.3 cm. The design and readout of the scintillator bars is the same as for the scintillator-based sub-detectors in the ND280 off-axis detector, and are discussed in § 2.3.2. In addition to the main tracking region, each module is surrounded by veto planes to enable rejection of particles produced outside the module. A typical INGRID

[4]No iron plate was placed between then 10th and 11th scintillator layers because of weight restrictions, but this does not affect the tracking capabilities.

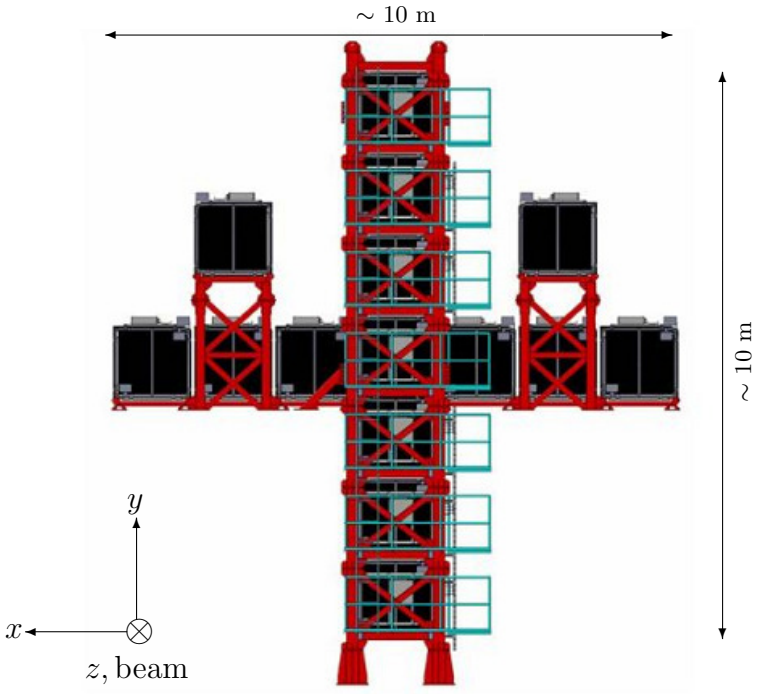


Figure 2.6: A profile of the INGRID on-axis detector where the beam direction is into the page. The 14 modules sample the neutrino beam over a $\sim$ 10 m $\times$ 10 m transverse area. The red cross indicates the approximate location of ND280 and line of sight to Super-Kamiokande [57].

event from T2K beam data can be seen in Fig. 2.7. Here the neutrino has entered from the left and has most likely interacted in the fifth iron plane producing a charged particle which leaves a clearly visible track in both views.

INGRID has been operational since the start of physics data taking in January 2010 and has been used both to check the beam profile and measure the beam centre. Fig. 2.8 shows the beam position as measured by INGRID throughout the Run 1 and 2 data taking periods, demonstrating the sub-mrad stability of the beam direction. In the future the INGRID module will be used in combination with the off-axis detector to constrain the flux systematics.

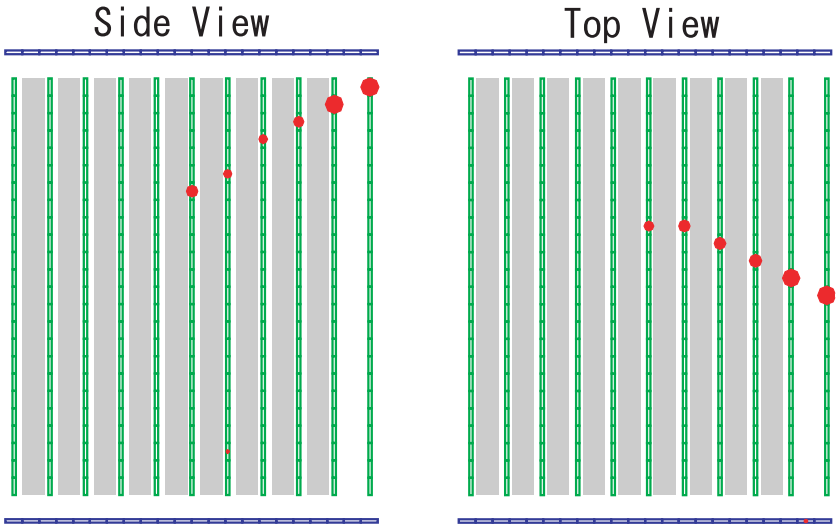


Figure 2.7: An example neutrino event in one of the 14 INGRID modules. Each green cell represents the end of a scintillator bar and the red circles represent deposited charge (the size is proportional to the charge). In this event the neutrino is likely to have entered from the left and interacted to produce the charged track visible [57].

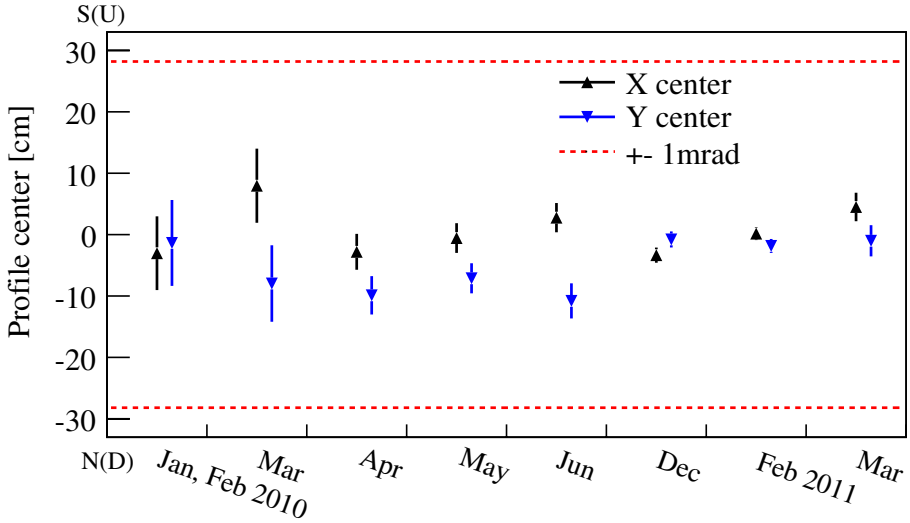


Figure 2.8: Measured beam centre position by INGRID in the horizontal (x, south-north) and vertical (y, up-down) directions. The dashed lines correspond to a change in beam direction of $\pm$1mrad. This shows that the beam direction was stable at the sub mrad throughout the entire Run 1 and 2 data taking periods [30].

2.3 The Off-axis Near Detector

The off-axis near detector (ND280) is located 280 m downstream of the beam production point on the upper level of the near detector pit. It is positioned along the off-axis angle such that it samples the beam over a solid angle that covers that seen by the Super-Kamiokande far detector. As mentioned, ND280 is a magnetised tracking detector designed to measure the neutrino interaction rates[5] at the off-axis angle. In particular ND280 must provide measurements of the following:

- The ν_e contamination of the beam, which enters as an irreducible background in the appearance measurement.
- The inclusive ν_μ flux, which can be extrapolated to Super-Kamiokande and will be used to reduce the effect of systematic uncertainties in the beam simulation.
- Interaction rates for ν_μ processes which cause backgrounds in Super-Kamiokande event selections, in particular neutral current π^0 production, a dominant background in the appearance measurement, and charged current π^+ production, a dominant background in the disappearance measurement—the latter of which is the topic of this thesis.

To meet these requirements, the detector needs to be able to reconstruct and make spectral measurements of exclusive neutrino interaction channels such as ν_μ and ν_e charged current quasi-elastic and inelastic scattering (as described in § 1.3), and in particular neutral and charged current single pion production. Because the interactions at Super-Kamiokande occur on a water target it is also necessary to make measurements at ND280 on both carbon and water to allow correct extrapolation between the different nuclear targets. These considerations drove the design of the off-axis detector.

ND280 is made up of a number of different sub-detectors all housed within the recycled magnet from the UA1 experiment at CERN. Fig. 2.9 shows an exploded view of the detector. At the upstream end is the pi-zero detector (P0D) which, as its name suggests, is specifically designed to make measurements of the rate of neutrino-induced neutral current π^0 production. The P0D also contains removable water modules which allow comparisons of the interaction rates on water and carbon. Downstream of this is the tracker region, composed of two scintillator-based fine grained detectors (FGDs) sandwiched between three time projection chambers (TPCs). The FGDs act as the fiducial target for neutrino interactions in the tracker region and provide vertexing and particle identification of the particles produced by these interactions. Like the P0D, the downstream FGD also contains some water target modules to allows comparison with interaction rates on carbon. The

[5]The interaction rate is a product of flux × cross section × target density. This is important when extrapolating to the expected rates at Super-Kamiokande, which sees a different neutrino flux and contains a different target material, as well as when aiming to make cross section measurements that are of use outside of T2K, where the effect of the flux and target material need to be removed (using their predicted values based on independent measurements/models and marginalising over the associated uncertainty).

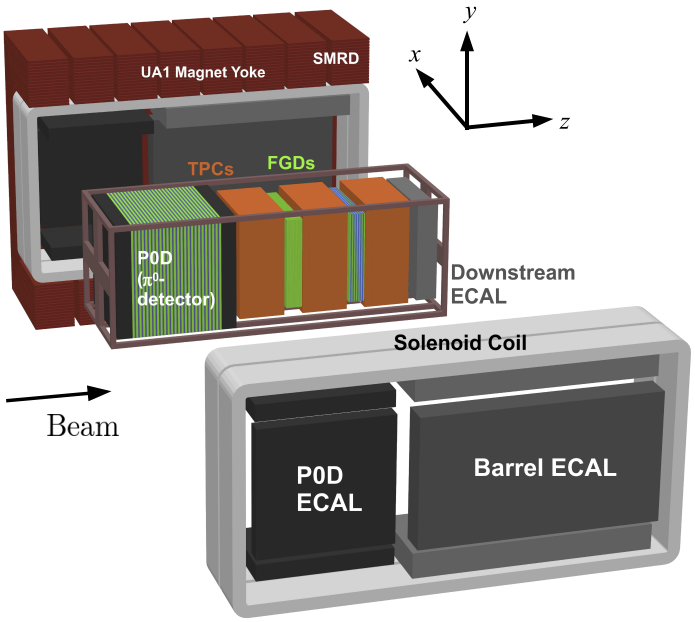


Figure 2.9: An exploded view of the off-axis detector. The basket, containing the P0D, the TPCs, the FGDs and the Ds-ECal, has dimensions of 6.5 m × 2.6 m × 2.5 m (length × width × height). Surrounding this is the P0D- and Barrel-ECal. This is all contained within the recycled UA1 Magnet capable of producing a 0.2 T field along the x-direction. The magnet is instrumented with scintillator detector strips to act as a muon range detector (SMRD) [30].

TPCs measure the momentum and charge of charged particles exiting the FGDs and are used for particle identification, in particular they are able to separate electrons and muons coming from CC ν_e and ν_μ interactions respectively. The P0D and the tracker regions are then surrounded by a set of electromagnetic calorimeters, known as the P0D-ECal, Barrel-ECal and Downstream(Ds)-ECal respectively. They are arranged to provide hermetic coverage of the outgoing particles produced by neutrino interactions in the P0D and FGDs. They can reconstruct high angle particles missed by the tracker and provide e/μ separation as well as reconstruction of photons, produced as the decay products of π^0s. The ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ analyses presented later in this thesis are based around the tracker and surrounding Barrel-ECals. An event display showing data collected during a Run 2 beam spill is shown in Fig. 2.10. It shows two tracks consistent with two CC interactions occurring in separate time bunches within the same spill and demonstrates the successful operation of all sub-detectors making up the tracker and surrounding ECals. Here we give a general overview of the components that make up ND280 and in § 2.6 we describe in more detail the elements of reconstruction and particle identification pertinent to the tracker- and ECal-based analyses developed in § 3.

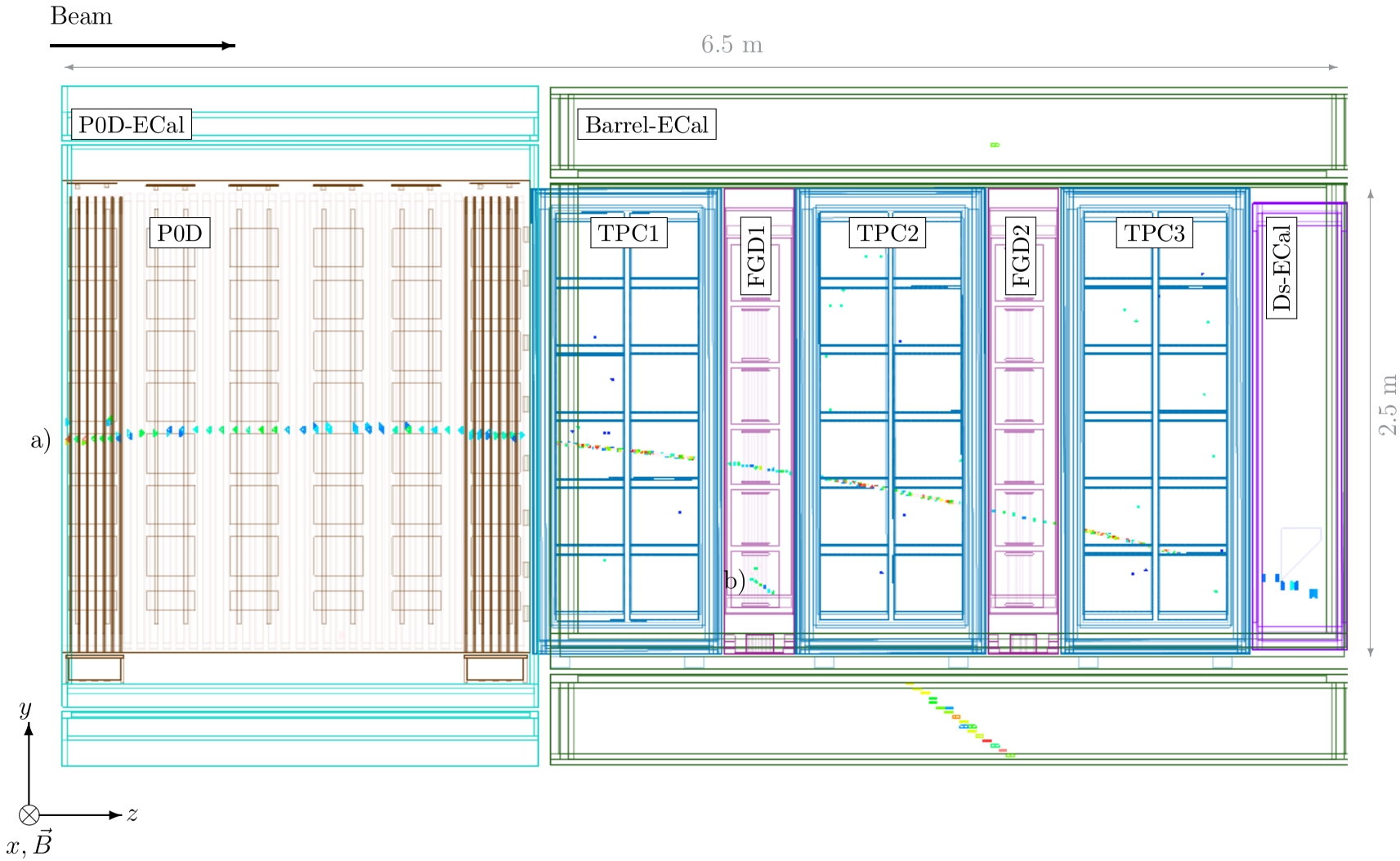


Figure 2.10: ND280 event display showing hit-level information for a Run 2 beam spill. The beam enters from left to right with an average direction $\sim 1.5°$ below the horizontal. Two tracks from separate beam bunches, and hence different neutrino interactions, are visible: a) A through-going muon-like track entering via the front face of the P0D and traversing the full length of the detector with hits in all sub-detectors, and b) a muon-like track starting in FGD1 and travelling into the Barrel-ECal without crossing the TPC2 active volume. The former is typical of a muon from a CC neutrino interaction upstream of ND280, the downwards curvature expected for a negatively charged particle is clearly visible. The latter is consistent with a CC neutrino interaction in FGD1 and represents a class of events inaccessible to current event selections due to the reliance on the TPC information to identify the muon candidate.

2.3.1 The UA1 Magnet

The UA1 magnet surrounds the P0D, tracker and ECal sub-detectors and provides the 0.2 T dipole magnetic field (oriented along the x direction) that is necessary to measure the momentum and sign of charged particles produced by neutrino interactions in the P0D and FGDs. The inner and outer dimensions of the magnet are 7.0 m $\times$ 3.5 m $\times$ 3.6 m and 7.6 m $\times$ 5.5 m $\times$ 6.1 m respectively. It is made up of a set of water-cooled aluminium coils, which carry the 2900 A current necessary to produce the required magnetic field, and a set of 8 low-carbon return yokes with a combined weight of 850 tons.

After instillation of the magnet in August 2009 there was a dedicated commissioning and mapping period where the magnetic field was measured *in situ* using a computer-controlled movable device holding three orthogonal Hall probes. Because of limited electrical power available at the time this was performed for a magnetic field strength of 0.07 T, which is lower than the nominal 0.2 T field used for neutrino data taking. The measurements were therefore re-scaled using a quadratic function as a first order correction. This results in a final magnetic field uncertainty of less than 2 G for each of the field components at the nominal field strength. Knowing the magnetic field map to this precision is necessary to reduce the systematic uncertainty in the momentum measurements of charged particles traversing the TPCs to below the 2% precision target.

Another important aspect is the magnet control system (MCS) which monitors various operation parameters such as the temperature, water flow, current, and voltage drop across the coils. This information is processed several times per second and if any sub-system exceeds its operational values the MCS will switch off the Magnet and log the corresponding information for later diagnostics. The MCS also interfaces with the global slow control (GSC). This allows remote control of the magnet as well as recording of the operational parameters for use later. In particular, the measured current and the operational status of the magnet during neutrino beam data taking is used in offline data analyses to define a "magnet on" flag and to allow comparison to the magnetic field in the simulation.

2.3.2 Scintillator Bar Instrumentation

The scintillator-based sub-detectors use a common readout system. Scintillation light is collected using a wavelength-shifting (WLS) fibre running through the centre of each scintillator bar which is then read-out and converted into an electrical signal using a Hamamatsu Multi-Pixel Photon Counter (MPPC). MPPCs are a novel silicon-based photosensor technology that, unlike vacuum photo multiplier tube (PMT)-based sensors traditionally used in scintillator-based neutrino detectors, are able to function without degradation in the 0.2 T magnetic field that most of the ND280 detectors reside in. Each MPPC consists of a square array of 667 independent pixels which each act as a Geiger micro-counter with a gain comparable to that of a vacuum PMT. Fig. 2.11 shows a close up of the active area of an MPPC as well it mounted into its ceramic housing. This shows the compactness of

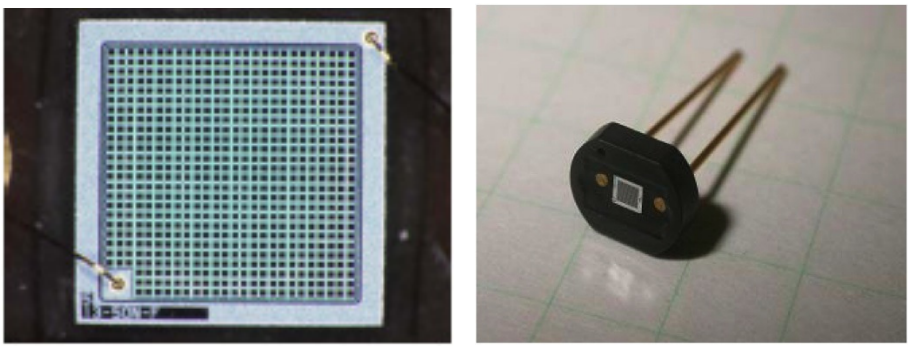


Figure 2.11: Images of the Hamamatsu Multi-Pixel Photon Counter (MPPC) used for readout of ND280 scintillator-based sub-detectors. On the left is a close up showing the 667 individual pixels that make up the 1.3×1.3 mm^2 active area. The MPPC loaded into its ceramic package is shown on the right [30].

the whole device which was another design requirement given the total number of MPPCs needed for ND280 and INGRID is about 64,000.

The MPPC gain depends on the accumulated charge on each pixels capacitor $Q_{pixel} = C_{pixel} \times \Delta V$ where the overvoltage ΔV is the difference between the applied voltage and the breakdown voltage of the device. These MPPCs are operated at about 70 V, at approximately 1 V above their breakdown voltage. Given the pixel capacitance of 90 fF, this results in a gain of around 1.0×10^6. A photoelectron produced in a given pixel will produce a Geiger avalanche meaning that the amplitude of the signal is independent of the number of photoelectrons created in a given pixel. Thus each pixel acts as a binary device and the signal is given by the sum of all pixels. The MPPC as a whole acts as an analogue device with an active range determined by the finite number of pixels and the area of illumination provided by the optical fibre [66]. A detailed description of the characterisation and calibration procedure can be found in [67, 68].

All electronic signals from the scintillator-based sub-detectors, with the exception of the FGDs, are read out using a custom-designed Trip-T-based front-end readout board (TFB). Each TFB houses 4 Trip-T ASIC chips which are each capable of reading out the signal from 16 MPPCs, where for each the signal is capacitively split (1:10) into high- and low-gain channels to increase the dynamic range of the digitisation for small and large signals. In total, each TFB reads out 64 MPPCs. The charge collected on each channel of the Trip-T chip is integrated over a programmable window which is timed so that the start of each integration cycle is timed to overlap with the arrival time of each beam spill. There is a reset time of 50 ns after each integration cycle and in total there are 23 cycles which easily accommodate the 8 bunch beam spill structure [69].

The calibration of the bars and MPPCs, both ex situ and in situ, are critical for the performance of the scintillator-based detectors, and this work formed a significant part of my contributions to T2K during my PhD studentship.

2.3.3 The Pi-zero Detector (P0D)

The main purpose of the P0D is to measure neutrino-induced neural current single π^0 production $\nu_\mu + N \rightarrow \nu_\mu + \pi^0 + N'$ on a water target and for a similar neutrino flux as that experienced by Super-Kamiokande. To do this it must be able to reconstruct the photons from the π^0 decay as well as charged particles such as muons and electrons to enable background rejection. To meet these requirements it is composed of planes of x- and y-orientated scintillator bars interleaved with fillable water target bags and brass and lead sheets. The water targets can be emptied so that comparisons of the interaction rate on water and on carbon can be made and to enable a subtraction measurement of the absolute cross section on water.

An x and y layer together make up a single module (P0Dule) consisting of 134 vertical bars (2.2 m long) and 126 horizontal bars (2.34 m long). Each bar has a triangular (isosceles) cross section with a base of 33 mm and a height of 17 mm and is coated in a thin layer of polystyrene with 20% TiO_2 to prevent light escaping from the bar and to increase the capture efficiency. These are then arranged in a tessellating pattern to form a single layer. As with the other scintillator-based sub-detectors each bar is read out using WLS fibres which, in the case of the P0D, are mirrored on one end and read out using a single MPPC on the other. The scintillator bars provide sufficiently fine segmentation to enable reconstruction of charged particle tracks and electromagnetic showers. In total 10,400 channels are read out using the same TFB-based system as described in § 2.3.2. Fig. 2.12 shows a schematic of the P0D. The upstream and central water targets can be seen, each composed of a sandwich of 13 P0Dules alternating with water bags and 1.5 mm thick brass sheets which help induce electromagnetic showers. These are surrounded by two regions with no water but where each P0Dule is instead interleaved with stainless steel-clad lead sheets (4 mm thick) which act as an ECal in order to contain electromagnetic showers (not to be confused with the ECals enclosing the P0D and tracker). The total weight of the P0D with and without the water targets filled is 16.1 and 13.3 tons respectively.

2.3.4 The Fine Grained Detectors

The fine grained detectors (FGDs) act as the target mass for neutrino interactions in the tracker region whilst also providing tracking, and hence vertexing, of charged particles coming from the interaction. Each FGD is made up of layers of alternating horizontal and vertical bars. A single bar is a 9.61 mm $\times$ 9.61 mm $\times$ 1864.3 mm piece of extruded polystyrene scintillator coated with a reflective layer of TiO_2 and with a WLS fibre going through a hole in the centre of the bar. A layer is made of 192 bars arranged in parallel and orientated along the x and y directions for horizontal and vertical layers respectively. Together, a horizontal and vertical layer make up a single 'XY-module'. The FGDs each have outer dimensions of 2300 mm $\times$ 2400 mm $\times$ 356 mm in the x, y and z directions and contain approximately 1.1 tons of target material.

The upstream FGD, known as FGD1, consists of 15 XY-modules meaning the target

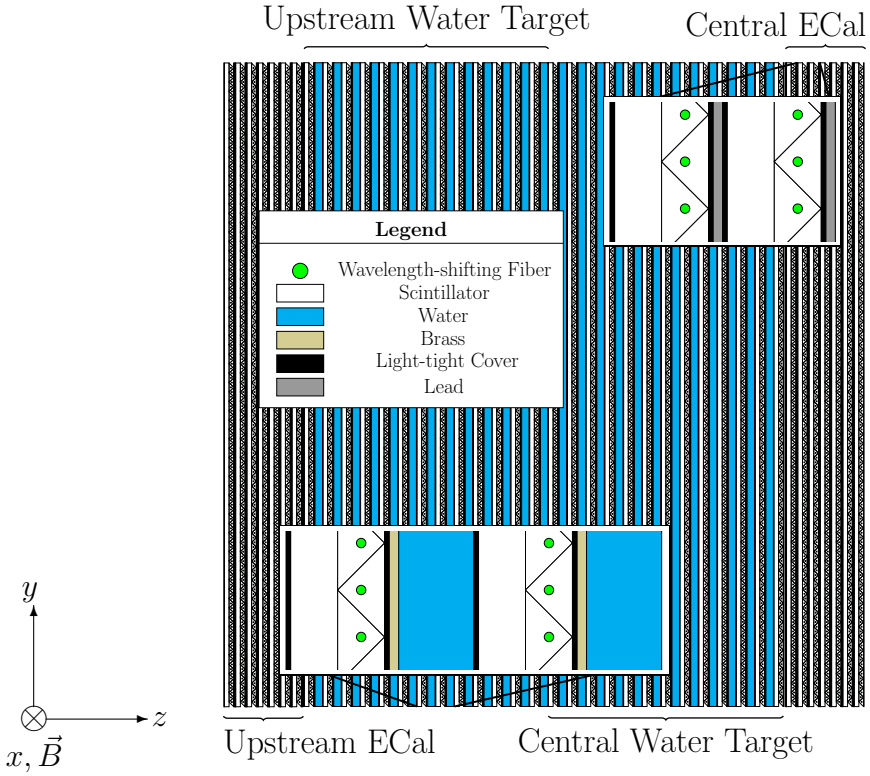


Figure 2.12: Schematic of the P0D where the beam enters from the left. The dimensions of the active region are 2.1 m × 2.4 m × 2.4 m. The zoomed in regions show the various P0Dules, each made up of a combined x and y layer of scintillator bars: In the central and upstream target regions these are interleaved with the 28 mm thick water modules as well as 1.5 mm thick brass sheets to induce electromagnetic showers. In the central and upstream ECal regions each P0Dule is interleaved with 4 mm thick stainless steel to contain as much of the showers as possible. The mass of the P0D with and without the water modules filled is 16.1 tons and 13.3 tons respectively [30].

material is almost entirely scintillator, resulting in a predominantly carbon target. The downstream FGD, known as FGD2, consists of seven XY-layers of scintillator interleaved with six 2.5 cm thick water modules. These provide a total water thickness of 15 cm which corresponds to about 50% of the target mass of FGD2. By comparing the interaction rates in the two FGDs, separate cross section measurements for interactions on carbon and water are possible, as with the P0D. This is important when extrapolating interaction rates measured at ND280 to those expected at Super-Kamiokande. Both FGDs have identical external dimensions, mountings and readout systems to allow them to be switched between their upstream and downstream positions which will allow consistency studies to be performed and ultimately lead to a reduction of flux variation systematics between the two locations. Fig. 2.13 shows FGD1 prior to installation in the ND280 basket.

As with the other scintillator-based detectors the photons from the WLS fibre are captured using MPPC photosensors, each fibre is read out at a single end (alternates ends for neighbouring bars in the same layer). Using a light injection system it is possible to illuminate the free end of a fibre which allows in-situ calibration of the MPPC and

Figure 2.13: The upstream FGD1 detector held in a storage cart prior to installation in the pit. The x-y plane corresponding to the orientation of the XY-modules is superimposed. In an x layer the bars are oriented in the y direction and vice versa. When installed inside the tracker the XY-modules hang down perpendicular to the beam direction [30].

readout chain. Unlike the other scintillator-based detectors, the electronic readout of the photosensors is performed using custom-designed application-specific integrated circuits called 'AFTER-ASICs' which are specially designed to allow continuous readout during a beam spill. Having a continuous readout with no dead time is important for the FGDs to ensure that delayed signals, such as Michel electrons from μ decays, are not lost [70].

In addition to tracking and vertexing, the FGD also provides information that can be used to identify particle types. Using the custom designed electronics readout it can look for the presence of a delayed cluster indicating the presence of a Michel electron from the decay of a muon or charged pion. It can also use the deposited charge per unit length of track to distinguish minimally-ionising particles (MIPs) such as muons or charged pions with more heavily ionising particles such as protons. Unfortunately, because of presently not understood systematics, these particular PID variables are not suitable for use in the current round of analyses.

Much work has gone into the relative timing calibration of the FGDs so that the directionality of tracks traversing both FGDs can be calculated based on the difference in times. Fig. 2.14 shows the FGD1-FGD2 times for FGD-triggered cosmics that travel through both FGDs. After applying all timing calibrations the two peak structure, for cosmics going either through FGD1 first or FGD2 first, is clearly visible. The width of the peaks indicate a timing resolution of 1.47 ns and the clear separation shows that this information can be used to infer the directionality of tracks that pass through both FGDs. The FGDs are a key element in the selections described in § 3 and will be discussed further

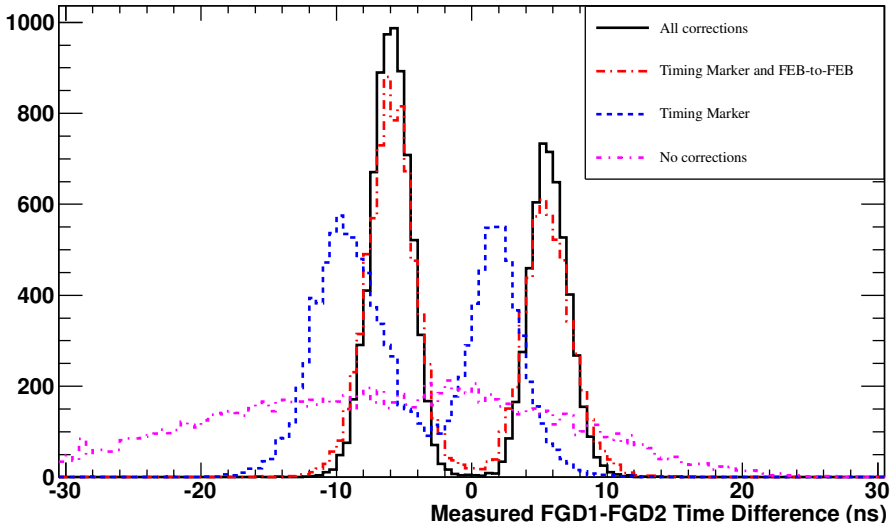


Figure 2.14: Difference between time recorded by FGD1 and FGD2 for FGD-triggered cosmic-rays showing a two peak structure from cosmics travelling first through either FGD1 or FGD2. The various corrections applied to reach the final resolution are also shown. The final width of each peak is 1.47 ns [71].

in § 2.6.

2.3.5 The Time Projection Chambers

Three separate time projections chambers (TPCs) are located on either side of the FGDs. Their locations can be seen in Fig. 2.9. They are known as TPC1, TPC2 and TPC3, based on their upstream, central and downstream positions respectively. They satisfy the requirements of the tracking detector in three important ways:

- Firstly, the fine grained and three-dimensional nature of their imaging capabilities allows many simultaneous tracks to be separately reconstructed. This means the number and orientation of charged tracks originating from a neutrino interaction in the FGD can be determined. This is important for reconstruction of exclusive neutrino interaction processes where the number of tracks is one of the key handles on the type of interaction.

- Secondly, because they operate in a magnetic field the momentum of charged particles can be inferred based on the curvature of the reconstructed track. Measuring the momentum of charged particles originating from neutrino interactions is central to the spectral measurements required to constrain both flux and neutrino interaction uncertainties. In addition to the momentum, the charge of the particle can be inferred if the direction is known. This is important for isolating particular types of neutrino interaction as it allows a charge cut to be put on the track identified as the outgoing lepton.

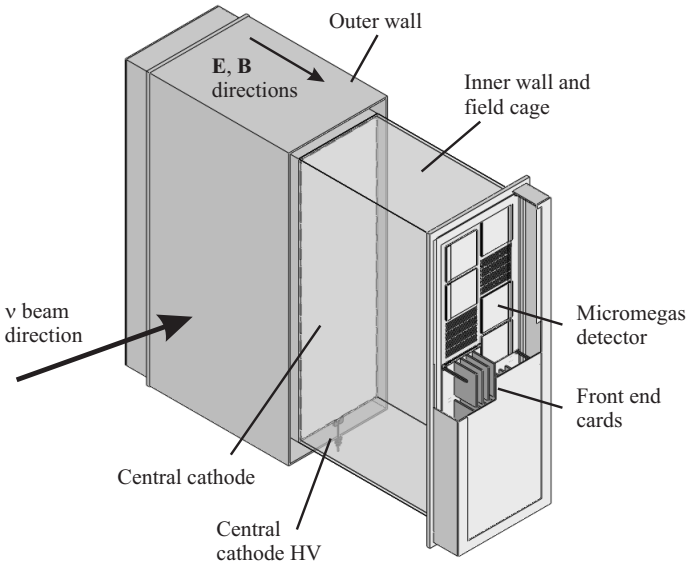


Figure 2.15: The main components of a single TPC detector shown as a simplified cut-away drawing. The coordinate convention is the same as ND280, a right handed coordinate system with z in the horizontal plane along the neutrino beam direction, and y in the vertical direction [72].

- Finally, comparing, as a function of momentum, the measured amount of ionisation energy to that expected for various particle types provides a powerful particle identification variable. In particular, the separation between electrons and muons that is necessary to measure the ν_e beam contamination is possible.

Each TPC is made up of an inner box containing an argon-based drift gas surrounded by an outer box filled with CO_2. This insulates the two boxes electrically and prevents atmospheric oxygen from entering the internal volume. The inner box walls are made from copper-clad G10 composite panels[6]. The inner box acts as a high voltage field cage, with the inner walls precisely machined to form an 11.5 mm pitch copper strip pattern which, when combined with a central cathode panel, gives rise to a uniform electric drift field that is aligned with the nominal magnetic field (along the x direction in the ND280 coordinates). A simplified drawing of a TPC with the right hand cut away to reveal the inner volume is shown in 2.15.

Charged particles traversing the gas region leave a trail of ionisation electrons. These are carried by the electric drift field away from the central cathode and towards the readout plane located on the side walls of the inner box. Each readout plane consists of an array of 12 bulk MicroMegas[7] [72]. Each bulk MicroMegas covers an area of 342 mm $\times$ 359 mm with 1726 rectangular anode pads with a segmentation of 7.0 mm $\times$ 9.8 mm. The 72

[6] G10 is a form of glass-based epoxy resin (fibreglass) with extremely good insulating properties and mechanical strength.

[7] These are arranged in two vertical columns which are slightly misaligned so that the inactive regions between modules are not aligned.

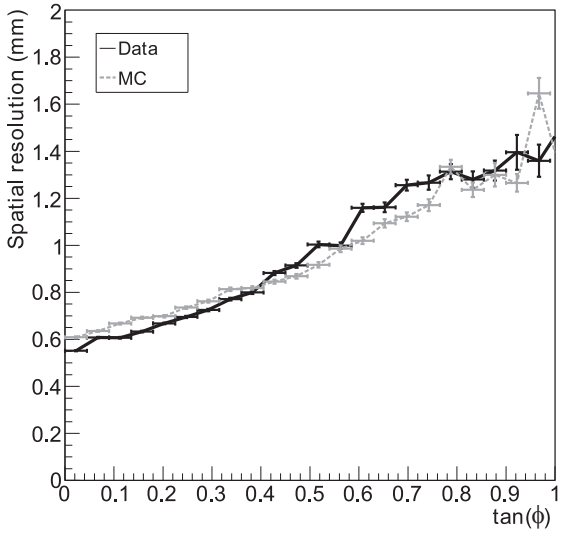
MicroMegas that tile the readout planes of the three TPCs provide a total active surface area of almost 9 m^2. The bulk MicroMegas technology was invented in 2004 by a CERN-Saclay collaboration. Like the FGDs, the TPCs are read out using the specially designed 'AFTER-ASIC' chips. Using the pattern of signals in the pad plane and the arrival time of the electrons (combined with their predicted drift speed) the TPCs are able to reconstruct complete 3D images of the paths of traversing charged particles.

In order to meet the requirements on the momentum resolution and absolute momentum scale, it is necessary to correct for any distortions in the electric or magnetic field and any differences in the drift velocities of the electrons. A calibration system capable of producing a known controlled pattern of electrons on the central cathode is used for this purpose. By shining a diffuse pulse of 266 nm light onto thin aluminium discs glued to the copper surface of the central cathode, photo-electrons are emitted at a specific time and location, allowing the drift velocity and any distortions in the electric or magnetic field to be measured. Full details on the calibration procedures are available elsewhere [72].

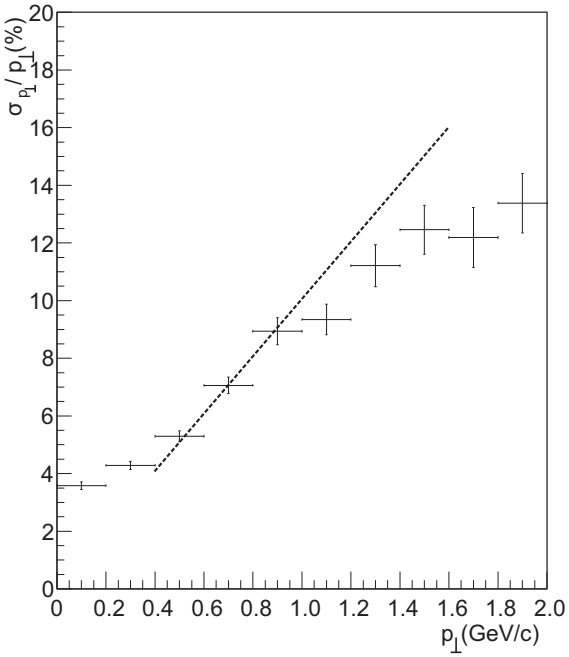
The track reconstruction is performed using separate methods for track finding and track fitting. In the track finding stage, signals from neighbouring pads which are consistent with coming from a single ionising track are grouped together to form a cluster of hits. In the fitting stage, the likelihood of the observed charge sharing between tracks is maximised to estimate the track parameters and the width of the ionisation track. When doing the likelihood fit the diffusion constant for the mean drift distance of the track is allowed to vary in the fit. By comparing the track parameters extracted from a single column to those from a global fit, the spatial resolution of a single column is found to be typically 0.7 mm, this is sufficient to achieve the momentum resolution goals of the detectors. The spatial resolution degrades for higher angle tracks and for tracks with a larger drift distance. Fig. 2.16(a) shows the spatial resolution as function of the tangent of the angle to the horizontal plane for both data and MC. It shows good agreement between the two. The momentum resolution as a function of the track momentum is shown in Fig. 2.16(b). Superimposed is a conservative estimate of the required resolution necessary to meet T2Ks physics requirements. This shows that the TPCs meet the requirement.

Once the momentum has been measured, particle identification can be performed. The TPC uses a truncated mean[8] of measurements of energy loss of charged particles in the gas, normalised to the track length, and compares this to the energy loss expected for various particle hypothesis at the given momentum. A pull value is then calculated which is simply the difference between the measured and expected energy loss per unit length divided by the expected width of the energy loss measurement for the current momentum. For each track this essentially represents the number of standard deviations away from the expected energy loss it is given a particular particle hypothesis. For each reconstructed TPC track, a set of 5 pulls are calculated (corresponding to muon, electron, pion, proton and kaon particle hypotheses) for use at the final analysis stage. The TPCs form an integral part of the selections described in § 3 and will be discussed more in § 2.6.

[8] Only the lowest 70% of the collected charge signals are used based on a MC optimisation of performance.

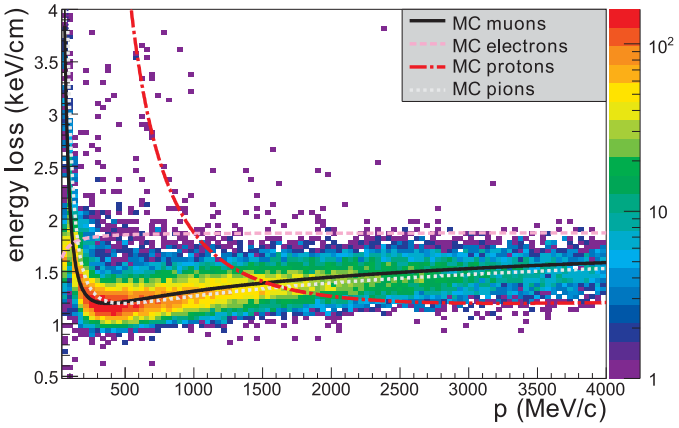


(a)

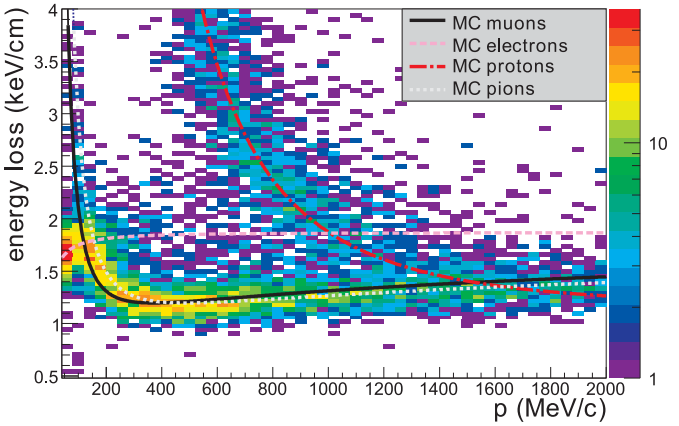


(b)

Figure 2.16: TPC spatial resolution as a function of track inclination from horizontal for both data and MC (a). The TPC reconstructed momentum resolution as a function of track momentum from a MC study (b), the design resolution required by T2K is superimposed [72].



(a)



(b)

Figure 2.17: Distribution of the energy loss as a function of momentum for reconstructed TPC tracks recorded during the Run 1 data taking period. Tracks with a negative and positive reconstructed charge are shown in (a) and (b) respectively. The expected energy loss curves based on MC for muon, electrons, protons and pions are overlaid [72].

2.3.6 The Electromagnetic Calorimeters (ECals)

The ND280 ECal surrounds the P0D, TPCs and FGDs and acts as a sampling electromagnetic calorimeter providing near-hermetic coverage for particles created by neutrino interactions in the inner detector region. The ECal will complement the inner detectors through the detection of photons[9] and charged particles and by providing the relevant information to identify them, i.e., electron-muon-pion separation.

Fig. 2.9 shows the arrangement of the 13 individual ECal modules: the 6 Barrel-ECal and 6 P0D-ECal modules surround the tracker region and P0D detector (parallel to the beam direction) while the downstream module (Ds-ECal) covers the end of the tracker region (perpendicular to the beam). The Ds-ECal is mounted inside the basket containing the other inner sub-detectors whereas the P0D- and Barrel-ECal modules are mounted to the inside of the UA1 magnet.

The scintillator bars for the ECal have a cross sectional area of 4.0 cm $\times$ 1.0 cm and are instrumented in the same way as the other scintillator detectors, with a WLS fibre running through the whole length of the bar which is then read out using an MPPC and the TFB electronics described in § 2.3.2. The Ds-ECal consists of 34 layers of 2.04 m long bars. In each layer, 50 bars are arranged in parallel with the short edges side by side to give a layer thickness of 1 cm. Alternate layers are rotated 90° to each other and separated by 1.75 mm lead sheets giving a total thickness of 10.6 X_0. For the Ds-ECal, the bars are read out at both ends. Because of limitations imposed by the available space in the UA1 magnet, the Barrel-ECal modules are only 31 layers thick and are also interleaved with 1.75 mm thick lead sheets giving a slightly reduced total thickness of 9.7 X_0. The bars running parallel to the beam in the z-direction are 3.84 m long and are read out at both ends whereas the short bars running perpendicular to the beam in the x- and y-directions, with lengths 1.52 m and 2.36 m respectively, are only read out on a single end with the other painted with a reflective coating to maximise light collection. The P0D-ECals are made up of six active scintillator layers separated by 5 layers of 4 mm thick lead (total thickness 3.6 X_0). Unlike the other modules, the bars in all layers are oriented along the z-direction. All bars are 2.34 m long and read out at a single end. Unlike the other ECal modules, the P0D-ECal is not intended to fully reconstruct π^0s as this will be done inside the P0D itself—its job is rather to detect photons that either do not convert in the P0D or that produce showers that are not contained in the P0D. In addition it can also identify MIPs and act as a veto for external backgrounds.

The Ds-ECal was constructed in 2008. It then underwent a series of tests at the CERN T9 beam allowing characterisation of its response to electrons, muons, pions and protons. After this it was installed and commissioned in ND280 in time for Run 1 data taking. Construction of the Barrel- and P0D-ECal modules was completed in late 2009 and these were installed and commissioned by October 2010 ready for Run 2 data taking.

[9] A key requirement of the ECals is that are able to reconstruct the photons created in the decay of π^0s so that they can be used to study NC-π^0 production in the tracker region.

2.3.7 The Side Muon Range Detector

The SMRD consists of 440 scintillator modules inserted into the 1.7 cm air gaps between the 4.8 cm thick steel plates that make up the UA1 return yokes. A number of modules fill each gap. Horizontal gaps are filled with four scintillation modules of dimensions 875 mm $\times$ 167 mm $\times$ 7 mm and the vertical gaps are filled with 5 modules with dimensions 875 mm $\times$ 175 mm $\times$ 7 mm. Each scintillator module is covered in a white diffuse layer, which acts as a reflector to maximise collection of light, and is instrumented with a single WLS which rests in a machined S-shaped groove along the length of the module. Each fibre is read out by a single MPPC and with the same TFB electronics as the ECal and P0D detectors. Fig. 2.18 show a single scintillation module prior to assembly.

Each yoke has 15 air gaps in the radial direction but only the inner air gaps are instrumented as the main purpose is to reconstruct particles entering the SMRD from inside the magnet. The SMRD is used to reconstruct the momentum and direction of muons escaping the inner detector at high angles with respect to the beam direction. The SMRD also acts as a trigger for cosmic events and can be used to provide a veto for beam-related activity from interactions in the surrounding cavity and walls.

As already mentioned we will leave discussion of the various reconstruction algorithms pertinent to the selections presented in § 3 to § 2.6. This is also where we will outline the ND280 software and analysis chain.

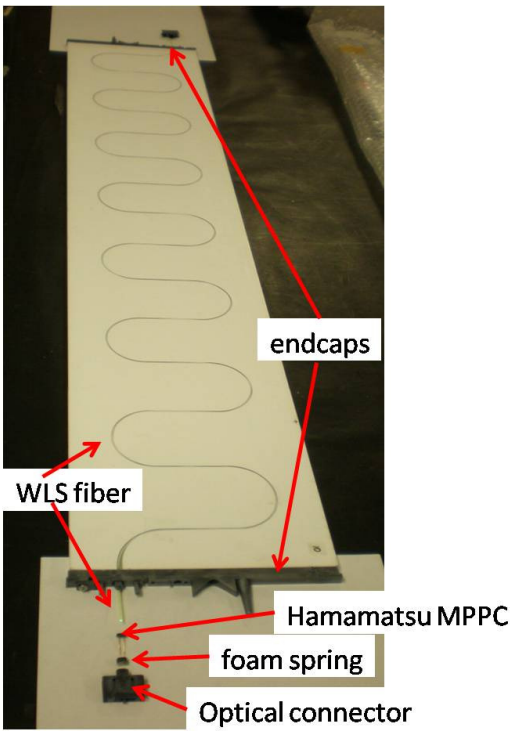


Figure 2.18: A single SMRD scintillation module prior to assembly, showing its instrumentation with the WLS running down the S-shaped groove and readout using a single MPPC [72].

2.4 The Far Detector: Super-Kamiokande

Located 295 km west of J-PARC, the Super-Kamiokande [73] water Cherenkov detector acts as the far detector for T2K. Built 1 km deep in the heart of Mt. Ikenoyama, Super-Kamiokande is a cylindrical cavern filled with 50 ktons of pure water[10]. Using $\sim$ 13,000 photomultiplier tubes it is able to image neutrino interactions by measuring the Cherenkov light given off by energetic charged particles produced when a neutrino interacts with the water. It samples the beam's flavour composition after its journey across Japan and, through comparison of this to the beam's initial composition, allows T2K to search for $\nu_\mu \rightarrow \nu_e$ appearance and ν_μ disappearance.

The water is contained within a steel cylinder 39 m in diameter and 41 m high, which is separated into two main regions: the inner detector (ID) and the outer detector (OD). A schematic of Super-Kamiokande is shown in Fig. 2.19. The ID is a cylinder 33.8 m in diameter and 36.2 m in height. Its inner wall is instrumented with 11,129 50 cm diameter photo multiplier tubes (PMTs) facing inwards to image the inner detector volume. They provide a surface coverage of 40% and have a combined quantum and collection efficiency of approximately 20%. The high surface coverage is needed to extract the necessary physical quantities from the neutrino interaction. The OD is a shell of width 2 m (both radially and on each axis end) which surrounds the ID and is instrumented with 1,885 outward-facing 20 cm diameter PMTs. The inner and outer regions are separated by a 50 cm-wide steel structure covered with plastic sheets which provides both optical separation between the two regions and a structure to attach the instrumentation. This results in a 50 cm wide 'dead space' between the two regions.

In order to successfully characterise the beam flavour composition such that it can be compared to the measured spectrum at the production point, both the type of neutrino interaction (ν_μ and ν_e) and the energy of the incident neutrino need to be measured. To achieve this, Super-Kamiokande counts the number of charged current quasi-elastic (CCQE) interactions taking place within the fiducial volume of the ID. CCQE interactions,

$$\nu_l + n \rightarrow p + l^-,$$

where $l = e, \mu, \tau$, allow both the flavour of the neutrino, based on the flavour of the outgoing lepton, as well as the incoming neutrino energy because of the simplicity of the final state, to be determined.

Super-Kamiokande measures the beam flavour composition by looking for the outgoing leptons produced in charged current quasi-elastic (CCQE) neutrino interactions. The ID acts as the fiducial volume for neutrino interactions. Above a certain energy threshold, charged particles from the neutrino interactions will produce a cone of Cherenkov light which results in a ring-shaped hit pattern on the PMTs. The shape of this ring pattern will depend on the type of charged particle which created it. Because of their large mass

[10] At 50 ktons Super-Kamiokande is the world's largest land-based water Cherenkov detector.

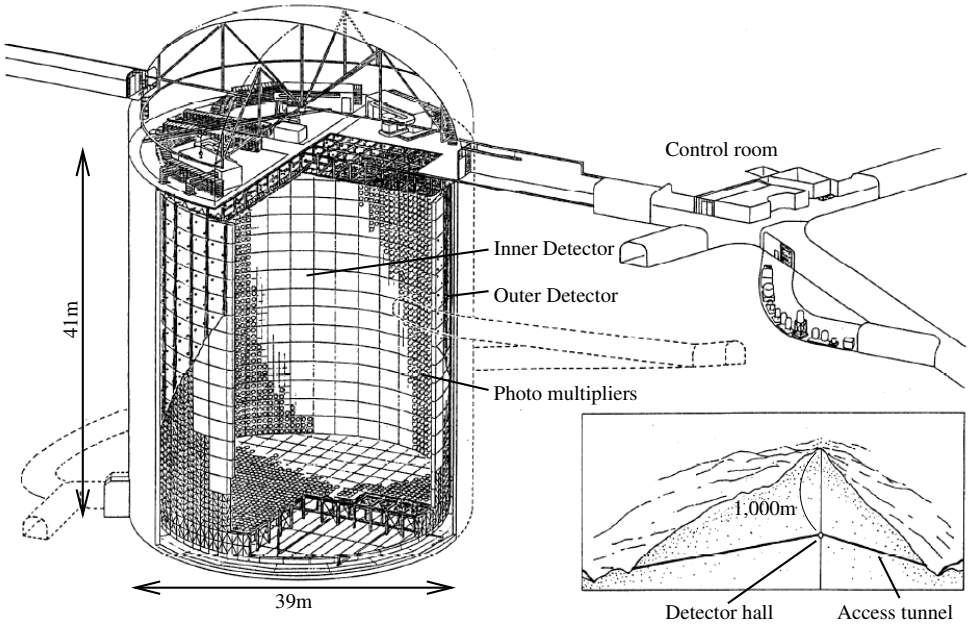


Figure 2.19: A schematic of the Super-Kamiokande detector showing the inner and outer detector regions as well as its location within Mt. Ikenoyama [57].

muons do not re-scatter easily and are resilient to changes in their momentum which leads to a well defined Cherenkov cone of light resulting in a sharp and clear hit pattern. In contrast, electrons, easily deflected because of their small mass and often giving rise to electromagnetic showers, will produce many overlapping Cherenkov rings resulting in a 'fuzzy' ring pattern. The Super-Kamiokande event reconstruction software uses this difference to identify whether rings seen in the ID are electron-like or muon-like and by identifying whether the outgoing lepton is a muon or an electron, it is possible to count the number of ν_μ and ν_e interactions respectively.

The OD serves mainly as an active veto for through-going cosmic rays and other backgrounds. Its walls are lined with highly reflective material and despite the fact that, compared to the ID, it is sparsely instrumented it still achieves a cosmic ray muon background rejection of almost 100%.

Super-Kamiokande has been running since 1996 and has already had a distinguished career producing a wide range of physics results. Because of its long running operation, both the calibration of the absolute energy scale and the modelling of the detector response are known to be accurate at the sub-percent level. In the following two sections, the calibration and event reconstruction methods will be discussed.

2.4.1 Data Reduction and Event Reconstruction

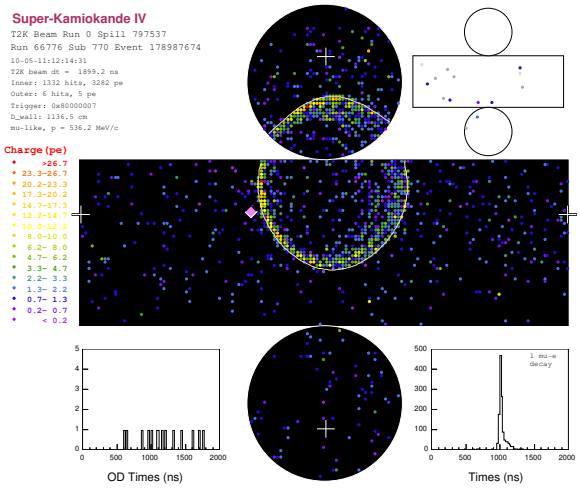
Before the full reconstruction and particle identification algorithms are run on the T2K beam data an event reduction step is performed. Events which are tagged with a T2K

beam coincidence trigger are selected and categorised into three mutually exclusive samples. Those for which the total charge is above a certain threshold and which appear to only contain particle tracks which start and then stop within the ID are categorised as 'fully contained' (FC) events. To determine if the particle tracks were contained within the ID a cut on the total number of OD hits is used, for FC events it is required that there be less than 15 hits in the largest OD hit cluster. Events passing the total charge cut but failing the OD hits cut are classified as 'outer detector' or (OD) events. Finally, events with total charge less than the threshold but which still appear to have hit patterns indicative of neutrino interactions are classified as 'low energy' (LE) events. After this categorisation the reduction software then applies a set of cuts, unique to each sample, to remove backgrounds. Both the FC and OD events then undergo a series of reconstruction steps in order to classify their origin and properties.

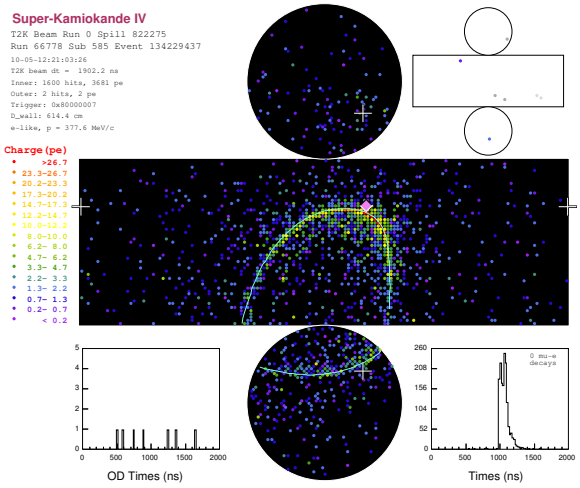
First the vertex position of an event is determined by searching for the point which best fits the distribution of PMT hit times (taking into account the propagation time for the Cherenkov light) and an initial track direction is calculated using this and a well defined edge in the PMT charge pattern. Using the initial vertex position as a seed an iterative technique based on a Hough transform is then used to determine the number of Cherenkov rings in an event and their directions: Secondary rings are searched for, choosing possible ring directions based on an initial Hough map, and a likelihood method is used to determine whether adding a second ring is more consistent with data than having just a single ring. This procedure is repeated, each time fixing the rings found in the previous iteration, until no further rings are necessary to fit the data (for practical purposes a limit of 5 rings is imposed).

Next a particle identification algorithm which exploits the systematic differences in the observed pattern of charge from the Cherenkov rings is applied. Muons, which do not scatter much and are more resilient to changes in momentum due to their relatively large mass, tend to produce a well defined cone of Cherenkov radiation resulting in a clear and sharp ring of PMT hits. On the other hand, electrons, which scatter more easily because of their smaller mass and induce electromagnetic showers, produce a less defined Cherenkov cone leading to a 'fuzzy' ring pattern seen by the PMTs. The algorithm compares the hit pattern to that expected for muons and electrons and produces a likelihood for each and a cut is applied based on the difference between these two likelihoods to decide whether a ring is muon-like or electron-like. Fig. 2.20 shows a muon-like and an electron-like event reconstructed from T2K beam data.

In the final step the momentum for each particle is determined. All the Cherenkov rings are re-fit to take into account the light pattern expected given the particle identification (in the case of a single ring a specialised event fitter is used) and the total charge of the event is apportioned between all of the rings. The charge associated with each ring is then used to infer the reconstructed momentum using a relationship derived from Monte Carlo simulations and detector calibrations from a number of sources. Full details of the event reconstruction and particle identification can be found in [74].



(a) muon-like event



(b) electron-like event

Figure 2.20: Example of reconstructed T2K events in Super-Kamiokande for (a) a muon-like ring and (b) an electron-like ring. Both figures show a projection of the cylindrical detector unrolled onto a plane. The PMT hits are indicated by coloured markers and the reconstructed ring is shown as a white line. The white crosses indicate the vertex location and the diamond marker shows where a line drawn from the vertex to the beam production point would intersect the detector wall. The small hit maps in the upper right show the OD hits [57].

2.4.2 Charged Current Quasi-Elastic Event Selections

We now describe the standard event selection cuts applied to the T2K beam data collected at Super-Kamiokande. The purpose of these is to provide the ν_e and ν_μ charged current quasi-elastic (CC-QEL) events samples needed to allow reconstruction of the neutrino energy. Due to the simple 2-body scattering of CC-QEL events the neutrino energy can be reconstructed using only the reconstructed momentum and direction of the outgoing lepton[11].

For both the ν_e and ν_μ selections cuts are applied to select single-ring fully contained fiducial volume events:

- No activity in the 100 μs before the beam trigger time.
- Number of PMT hits in highest charge outer detector cluster is < 16. This is to select fully contained events where all energy has been deposited in the inner detector. This is a requirement when reconstructing the momentum of the event.
- The reconstructed vertex is required to be at least 2 m from the inner detector wall. This is optimised to give the highest sensitivity whilst avoiding systematic edge effects.
- The total visible energy E_{vis} in the ID is required to be greater than 30 MeV.
- Only a single reconstructed ring is found.

Then for the ν_e selection the following cuts are used to identify the electron:

- The ring is e-like based on the particle identification.
- The total visible energy is greater than 100 MeV.
- There is no delayed electron signal.
- The reconstructed invariant mass given by forcing the ring finding algorithm to find a second ring is inconsistent with that of a π^0.
- The final reconstructed neutrino energy is required to be less than 1250 MeV.

Whereas for the ν_μ selection the following are used to identify the muon and reject events where there is a Michel electron indicating an additional charged pion:

- The ring is μ-like based on the particle identification.
- The total reconstructed momentum of the ring in the muon hypothesis is greater than 200 MeV.
- There is at most 1 delayed electron signal.

[11] This neglects the Fermi momentum of the hit nucleon and is only an approximation.

These selections result in samples with high signal efficiency for selecting ν_e and ν_μ CC-QEL interactions whilst also providing effective rejection of background non-CC-QEL interactions. The signal efficiency for the ν_e appearance selection is estimated from MC to be 66% with rejection of the $\nu_\mu + \bar{\nu}_\mu$ CC, intrinsic beam ν_e and NC backgrounds at $> 99\%$, 77% and 99% respectively [30]. The ν_μ cuts select ν_μ CC-QEL events with a final purity of 57% (82%) for the predicted flux with (without) oscillations applied[12]. As we will see in the next section the dominant background for the ν_μ selection comes from ν_μ CC interactions where a single π^+ is produced in the final state [75].

2.5 Motivation for ν_μ-CC-1π^+ Measurement at ND280

As described in § 1 T2K has recently published its first physics result based on the Run 1 and 2 data showing indications of ν_e appearance that hint at a non-zero value of θ_{13}. In addition to this a measurement of ν_μ disappearance was also performed, which consisted of two independent analyses with different fitting techniques but using the same set of inputs. The core of this thesis is the measurement of neutrino induced charged current charged pion production at ND280. This is done both as a preliminary check of the data-MC agreement between current flux and GENIE predictions for the neutrino interaction rate and as a demonstrator for future studies that can be performed using ND280 to reduce systematic uncertainties affecting ν_μ disappearance analyses due to uncertainties in the simulation of the neutrino interaction physics. We now present some results from one of these disappearance analyses [75] and discuss the motivation for studying charged current charged pion production at ND280.

The general analysis technique was to compare the predicted and measured spectrum of events passing the ν_μ cuts described in the previous section. Using a likelihood-ratio method as a function of the oscillation parameters Δm^2_{23} and $\sin^2 2\theta_{23}$, the best fit points were determined and the 68%, 90% and 95% contours extracted using the Feldman-Cousins method [76]. An inclusive measurement of the rate of charged current neutrino interactions in the ND280 tracker region was used to normalise the overall rate prediction at Super-Kamiokande. Needless to say this analysis was the culmination of much work involving inputs from many different aspects of T2K including systematics for Super-Kamiokande and ND280 detector uncertainties, flux simulation shape and normalisation uncertainties and neutrino generator uncertainties. The final contours and best fit point are shown in Fig. 2.21 where they are consistent with the most recent results from the MINOS experiment. We can see that even though the combined Run 1+2 data set of 1.431×10^{20} POT represents less than 2% of the full 5 year data set (equivalent to approximately 8×10^{21} POT) the sensitivity is still comparable with the current best measurements by other experiments. This highlights the effectiveness of T2K's high intensity narrow-band

[12] This is based on MC simulations for the expected T2K neutrino flux including three-flavour oscillations for $\sin^2 2\theta_{13} = 0.1$ and $\delta_{CP} = 0$ as well as the current best fit values for the atmospheric mixing parameters.

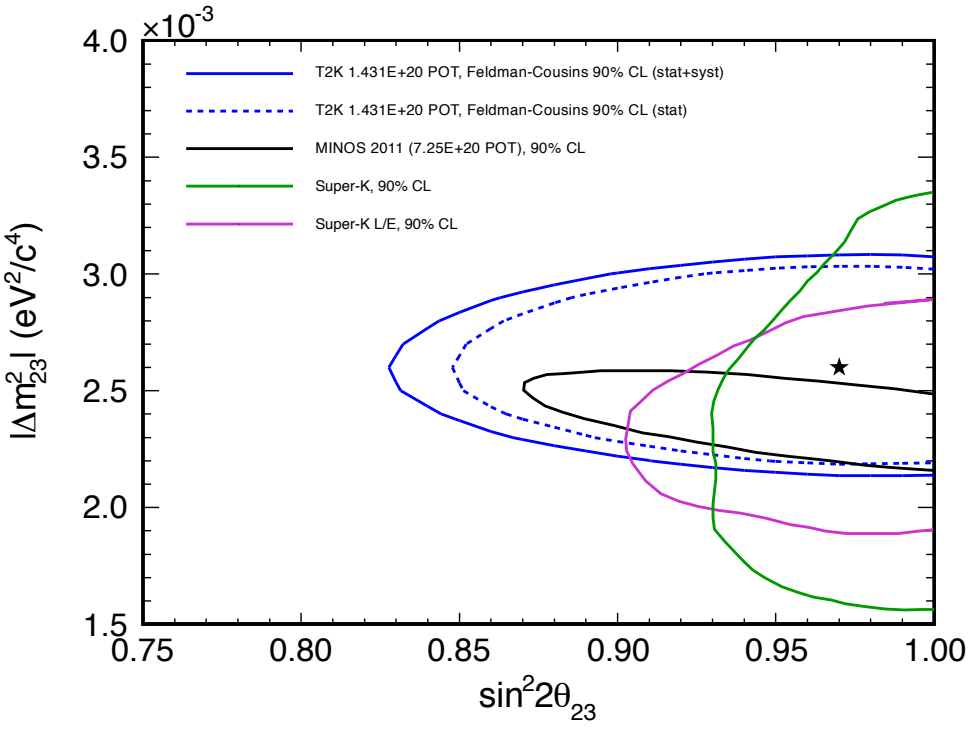


Figure 2.21: Run 1 and 2 ν_μ disappearance analysis result showing the best fit point (star) and the 90% contours (with and without systematics) in the $\sin^2 2\theta_{23}$ Δm^2_{23} plane. Also shown for comparison are the latest MINOS and Super-K results. Taken from [75].

neutrino flux tuned to give oscillations at Super-Kamiokande. This analysis was still very much statistics-limited as can be seen by the small difference between 90% contours with and without the inclusion of systematics.

With the full 5 year data set, T2K aims to make precision measurements of the disappearance parameters to uncertainties within $\delta(\Delta m^2_{23}) \sim 10^{-4}\text{eV}^2$ and $\delta(\sin^2 2\theta_{23}) \sim 0.01$. At this level of precision, systematic uncertainties coming from the uncertainties in the simulation of neutrino interactions become important. In Fig. 2.22 we can see a breakdown of the events passing the ν_μ selection cuts used in the disappearance analysis. The number of reconstructed events per 50 MeV bin is shown for both the true and reconstructed neutrino energy, where the reconstructed energy is calculated based on the assumption of quasi-elastic scattering. The events are separated into 5 categories based on the true neutrino interaction type:

- The ν_μ-CC-QEL signal events making up approximately 57% of the selected total.
- The ν_μ-CC-1π events for which there is a single $\pi^\pm$ in the final state make up approximately 24%.
- The ν_μ-CC-Other events which are dominated by DIS/multi-π final states also make up approximately 6% of the total.
- The $\nu_{\mu(\tau)}$-NC events make up approximately 6% of the total.

- A final category includes all other ν_e and $\bar{\nu}$ interactions.

From this we can see that the dominant background comes from ν_μ-CC-1π final states. Any non-QEL background event will have its energy systematically mis-reconstructed to lower values. This is because the energy reconstruction assumes quasi-elastic scattering and misses the component of the neutrinos energy and momentum carried away by the additional outgoing particles present. The ν_μ-CC-1π above the oscillation dip in Fig. 2.22(a) can be seen to shift to lower values and fill it in when looking at reconstructed energy in Fig 2.22(b). Systematic uncertainties associated with the modelling of these neutrino interaction modes result in a systematic uncertainty on the measurement of Δm^2_{23} and $\sin^2 2\theta_{23}$. The effect of these systematics were calculated as part of the disappearance analysis [77]. The combined effect on the error associated with the best fit point of Δm^2_{23} and $\sin^2 2\theta_{23}$ due to uncertainties in the relative rates of ν_μ-CC-1π/ν_μ-CC-QEL and ν_μ-CC-Other/ν_μ-CC-QEL and the uncertainties associated with the final state interactions of the hadrons before they escape the nucleus were found to be:

- $\frac{\delta(\Delta m^2_{23})}{\Delta m^2_{23}} \sim 0.4\%$
- $\frac{\delta(\sin^2 2\theta_{23})}{\sin^2 2\theta_{23}} \sim 1.6\%\%$

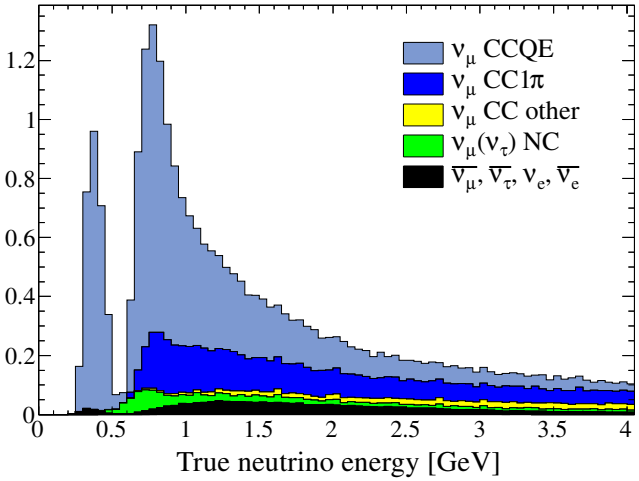
This shows that the systematic uncertainties on the simulation of charged current neutrino interactions with charged pions in the final state, especially those producing a single π^+, will have a significant impact on the final sensitivity that T2K will be able to achieve in the precision measurement of the atmospheric oscillation parameters Δm^2_{23} and $\sin^2 2\theta_{23}$—the ability of T2K to make measurements of neutrino induced π^+ production at the near detector in order to reduce these uncertainties associated and to resolve ambiguities where the model is known to be incorrect will be necessary for T2K achieve its physics goals. This motivates the main topic of this thesis which is a measurement of the rate of neutrino induced charged current charged π production using the tracker and ECal components of ND280.

2.6 Reconstruction and Particle Identification in the Off-Axis Near Detector

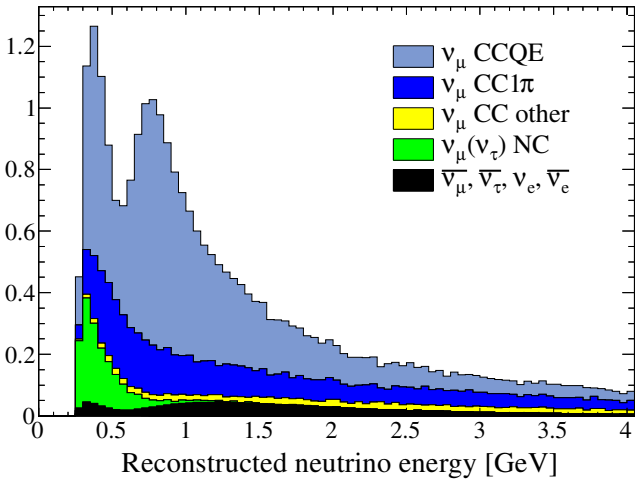
Before developing the selections necessary to study charged current charged pion production at ND280 we will detail some of the reconstruction and particle identification algorithms that drove particular analysis choices as well as describing general aspects of ND280 pertinent to an analysis based on the tracker- and ECal sub-detectors.

2.6.1 The ND280 Analysis Chain

The ND280 offline software suite handles processing of both data and MC. Fig 2.23 shows a schematic of the most important elements. The data is first unpacked from its raw format



(a)



(b)

Figure 2.22: Composition of events passing the Super-Kamiokande ν_μ selection used in the disappearance analysis. The expected number of reconstructed events for 3.23×10^{19} POT are shown as a function of true (a) and reconstructed (b) neutrino energy. A flux prediction with oscillations assuming $\Delta m^2_{23} = 2.32 \times 10^{-3} \text{eV}^2$ and $\sin^2 2\theta_{23} = 1.0$ is used. The contributions are broken down into various categories based on the neutrino interaction process. The ν_μ-CC-1π selected events form the dominant background and, by comparing their distribution in (a) and (b), can be seen to systematically mis-reconstruct to lower neutrino energies and fill the deficit at the oscillation maximum [75].

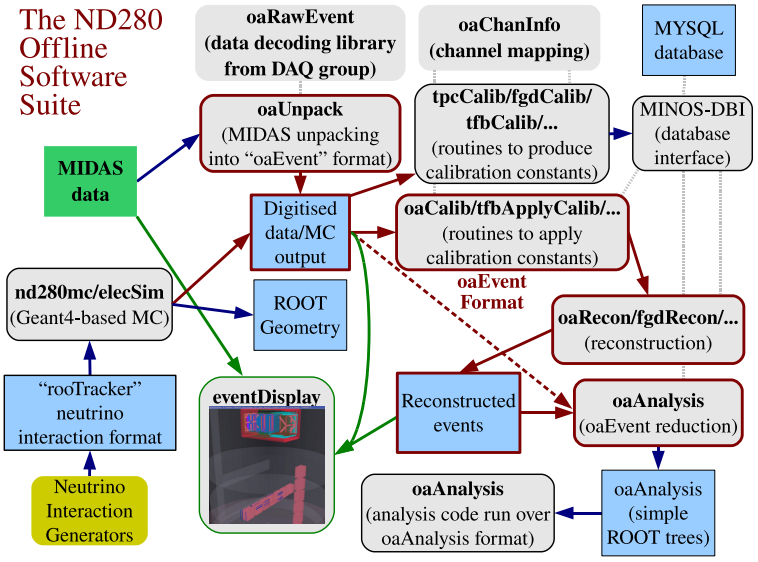


Figure 2.23: Schematic of the ND280 offline software suite showing the analysis chain for both MC and raw data through to the final reduced output analysis files. Taken from [30].

and converted into the C++ based oaEvent format which, with the exception of the pure ROOT final analysis files, forms the common base for all input/output in the offline software. It is then processed by the relevant sub- and pan-detector calibration packages which apply the relevant calibration based on constants generated using dedicated data processing. These are stored in a MYSQL database allowing the storage and retrieval of constants with specific validity in time.

The MC chain starts off with the external simulation of the flux at ND280 using the JNUBEAM simulation described in § 2.1.4. This is then used as input, along with a realistic ROOT based description of the ND280 geometry, to the neutrino interaction generators. At present both the NEUT [78] and GENIE[13] neutrino interaction generators are fully integrated with the ND280 analysis chain. They output the distribution of neutrino vertices along with the equivalent flux POT in a ROOT-based tree which is then read into the detector simulation packages nd280mc and elecSim. nd280mc uses the GEANT4-libraries[14] to simulate the energy deposits and trajectories of the particles produced by the neutrino interactions and elecSim then simulates the response of the detector to this. At this point the output is in the same format as the unpacked data, meaning that from this stage on both the MC data and the real data are treated in the same way. The MC data is also then processed by the calibration packages although only a subset of the calibrations, corresponding to those for which the corresponding effect has been simulated in elecSim, are applied.

Both data and MC are then processed by the relevant sub-detector and global reconstruction packages. The final stage in the software chain is oaAnalysis which takes the

[13] As described in § 1.3.
[14] Note that the ND280 software uses a more recent version of GEANT than that used by the beam group.

detailed and heavy output of the reconstruction stage and reduces and distils it into a format useful for individual users performing analyses. oaAnalysis has to interface with many different areas of the software and handles the addition of extra information, such as the beam summary data and the ND280 data quality flags necessary to successfully analyse data. It also handles the addition of detailed flux and generator MC information necessary for the event reweighting that will be described in § 4.1. During my time on T2K I acted as the package manager for oaAnalysis and was responsible for ensuring that it produced reliable, up to date, and useful information for people doing ND280-based analyses.

The processing of the full data and MC sets through the whole software chain is a significant effort in terms of both manpower and CPU time. Official productions take place to provide stable data sets to develop analyses from and to enable full exploitation of the parallel processing and data distribution resources available through the GRID. These are organised by the ND280 computing group. In this thesis we use the output of the ND280 production 4 for both MC and data. This is the most recent large scale production of both the Run 1 and 2 data sets and included a high statistics MC sample of about 10 times the POT equivalent of the data. For a given production a number of re-spins are expected as bugs are found and to include updated calibration constants. In particular here we use the re-spin 4C for MC and 4D for data. There were no changes in the physics models or reconstruction algorithms between these two re-spins. Production 4 was based on the ND280 software versions v9r7p9 for MC and v9r9p1 for data. MC samples for both NEUT, using version v5.1.1, and GENIE, using version v2.6.2, were generated and these used the flux simulation output from JNUBEAM version 11a. The ND280 software matches up information regarding the ND280 detector status (provided by the ND280 data quality group) and the database of beam status parameters (provided by the Beam group) with the ND280 data runs so that the data POT and quality cut flags are available to the ND280 user when performing an analysis. This is vital to ensure a proper comparison between data and MC.

2.6.2 ND280 Reconstruction Algorithms

The ND280 reconstruction starts with the individual sub-detector reconstruction algorithms. These are responsible[15] for taking the output of the calibration stage and applying the appropriate reconstruction algorithms specific to the particular sub-detector. There are separate software packages for each sub-detector and also a tracker reconstruction package which handles combination of the TPC and FGD information. This is done before the global reconstruction. It uses special techniques to incrementally match the TPC tracks back to hits in the FGD and then tries to combine these with other matched tracks in the tracker. Any leftover FGD hits not associated with a TPC track are then passed to specialised FGD-only reconstruction algorithms. Each reconstruction package is

[15] With exception of the FGD reconstruction which only tries to reconstruct hits left over from the TPC-FGD matching stage.

managed and developed by the relevant experts and full details of the ND280 sub-detector reconstruction algorithms are described elsewhere [79]. Here we focus on the global reconstruction and global vertexing algorithms[16].

The job of the global reconstruction is to combine the various sub-detector reconstruction outputs. It uses a series of standardised fitting, propagation and matching routines provided by the external RECPACK [80] software package. The final output from the sub-detector reconstruction is a list of either track-like or shower-like objects. With the exception of the TPC reconstructed output the track-like objects are then re-fit using the RECPACK Kalman filter. This takes into account the momentum of the tracks and the expected energy loss based on a simplified model of the ND280 geometry. Next the global reconstruction tries to match together these outputs to form global tracks:

- First an attempt is made to match each of the tracker objects with each object in the adjacent detectors. These are the P0D, the P0D-ECal and the Barrel- and Ds-ECal. The reconstructed state of one object is extrapolated to the matching plane of the other object using RECPACK matching functions. A matching χ^2 is computed for the two objects. This is based on their positions and relative direction only and not on the reconstructed momentum. If the matching χ^2 is less than 100 (200 for matching to P0D or SMRD objects) and the two objects are within 300 ns of each other then they are matched together. The matching χ^2 were chosen to optimise performance and the large values reflect the fact that in the version of the reconstruction used for this analysis the covariances associated with the various sub-detector tracks were being underestimated. Matched objects are then refitted using the RECPACK Kalman filter which takes into account the reconstructed momentum and expected energy loss of the tracks. The matching is only performed on a single pair of objects at a time. An iterative process of matching is then followed until no more objects can be combined together.

- This is then repeated but starting with the so far unmatched objects in non-tracker sub-detectors, first for those in P0D and then for those in the ECal.

The output of the matching stage is a set of globally reconstructed tracks. These also store the associated PID information of the various sub-detector objects that they are made up of. For example, a global track with a contribution from a TPC sub-track would have the TPC pull PID variables attached to it for later use. Fig. 2.24 shows the same Run 2 beam spill as in Fig. 2.10 but now with the output of the global reconstruction overlaid. It shows that the global reconstruction has successfully matched the output from the various sub-detector and tracker-reconstruction stages.

[16]The following is a summary of the full description given in [79].

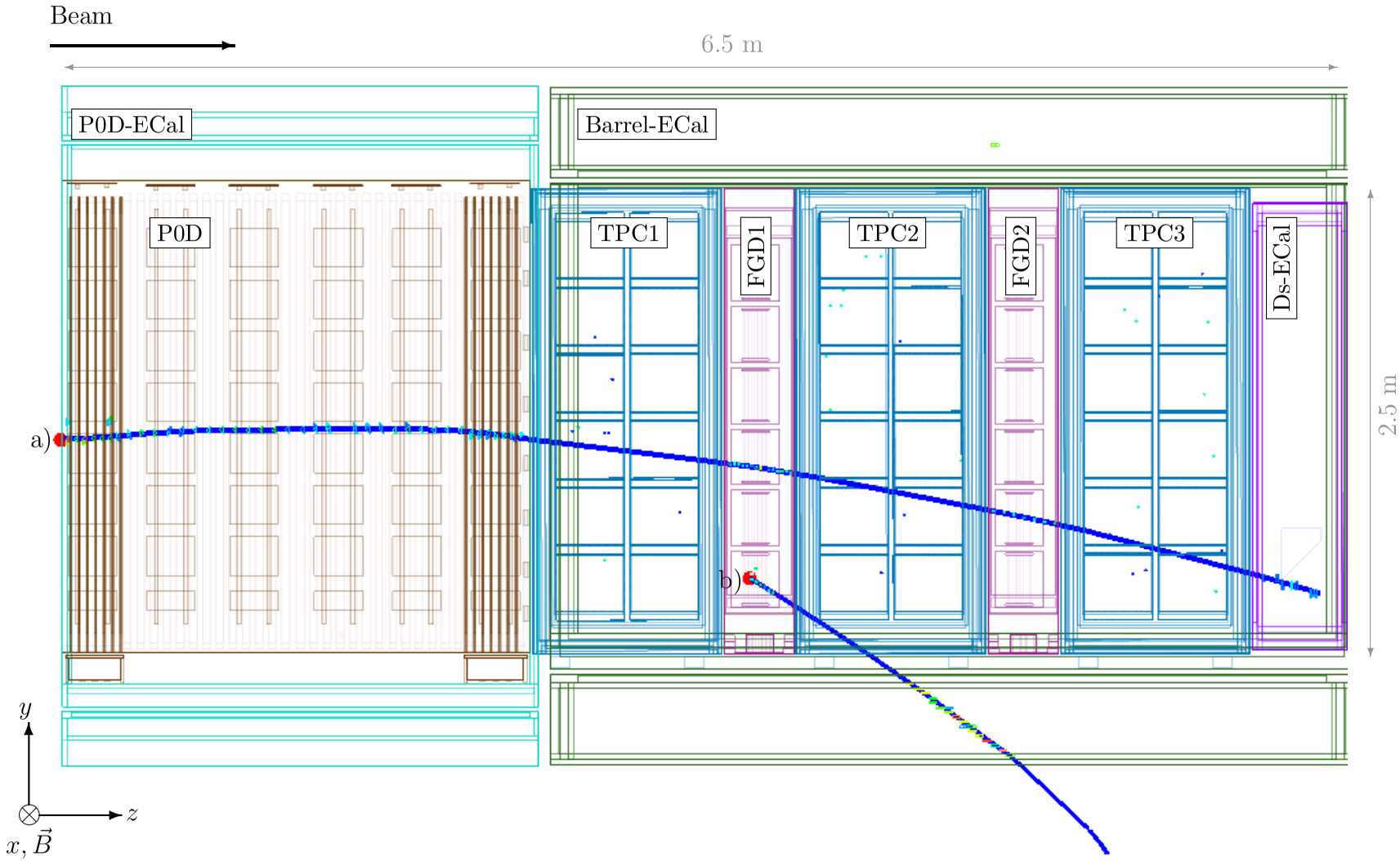


Figure 2.24: ND280 event display showing the same beam spill as in Fig. 2.10 but now with the output of the global reconstruction overlaid. The blue lines indicate a global track with a negative reconstructed charge and the red circles indicate the global reconstructed vertex position, which for these single track cases is simply the upstream end of each track. We can see that for both tracks the global reconstruction has successfully matched the contributors from different sub-detectors. At present the charge information is only trusted if the track has a TPC sub-track as one of its constituents.

2.6.3 Global Vertexing

The output of the global reconstruction is then passed on to the global vertex finder. Here we focus on a description of the global vertexing as we discuss its performance and validation in § 3.3.

The global vertexing uses a Kalman filter approach to decide which tracks are associated with a common neutrino vertex [81]:

- A preliminary track clustering stage provides the list of potential vertices to be filtered. This is done in an iterative procedure that uses the closest point of approach in the XZ plane, in which the tracks are unperturbed by the magnetic field, to decide if they are associated. If a track is not matched to any other then it is stored as a single-track vertex.

- For each cluster of tracks a Kalman filtering technique is then used to decide which of the tracks are actually associated with the vertex. Each track is extrapolated to the average cluster position and its associated covariance is updated and used to calculate an updated χ^2 for matching to the vertex.

- If more than one track survives this first stage then they are refit using an inverse Kalman filter. This uses the final estimate for the vertex position along with the modified track covariances to test the effect of removing tracks associated with the vertex. From this the individual contribution to the χ^2 from each track is determined and used to remove tracks.

- In the case of single-track vertices no global vertexing algorithm is performed and the start of the highest momentum track in the bunch is taken as the vertex location.

A detail important to this thesis is that the version of the global vertexing used in the production 4 processing runs not on the output of the global reconstruction but instead on the output of the tracker reconstruction. The list of tracks associated with each reconstructed vertex is matched up to the corresponding global reconstructed track at a later stage. This has two implications for the analysis presented later: Firstly, the start point of the single-track vertices is not always identical to the start point of the associated global track as it will have been refitted using the Kalman filter at the global reconstruction stage. Secondly, only tracks with a tracker component will be matched to a vertex. For example, a track in the ECal which pointed back towards a global vertex but was not matched to any tracker objects would not be associated with that vertex.

2.6.4 Measurement Strategy

We now discuss how the reconstruction and particle identification details drove the choices made when developing the selections in § 3. The overall aim was to produce a robust selection with minimal detector-based systematics that will allow first data MC comparison to be made and also demonstrate the use of future analysis techniques.

Fig. 2.25 shows the results of a truth study[17] where the paths of the μ^- and π^+, from ν_μ-CC interactions in an FGD with only a single $\pi+$ in the final state, are categorised based on which sub-detector active volumes they pass through. It shows that 57.2% of the time the π^+ passes through only the FGD active volume (blue row) indicating that approximately half of the charged pions produced do not make it out of the FGD. The green box shows the topologies that could in principle[18] be selected if the TPC is used to identify both the μ^- and the π^+ and the yellow box shows the fraction of events which could be recovered if the ECal was used to identify the π^+. In addition the overlap of the blue row and red column shows that for 13.7% of interactions both the μ^- and the π^+ go through only the FGD active volume. The results of this study do not include the effect of reconstruction but serve to demonstrate the complex geometrical and kinematical acceptance of the detector.

We chose to perform an inclusive analysis, where we define the signal as any neutrino interaction with at least one charged pion in the final state, as opposed to trying to select those with only one pion in the final state. This was to avoid the use of the PID information associated with FGD-only tracks for which, as discussed in § 2.3.4, there are currently not-yet-understood detector systematics. As approximately half of the charged pions produced in an FGD do not travel through another sub-detector trying to select interactions with only one pion in the final state would be difficult as without the FGD PID information it is not possible to reject higher multiplicity interactions where the additional pions stop in the FGD. By using such an inclusive definition of the signal is we are less sensitive to the complex geometrical acceptance of the ND280 detector.

To avoid dependence on the momentum scale we choose to perform a measurement of the integrated rate of neutrino interactions rather than a differential measurement with respect to the reconstructed kinematic quantities of the event, such as the track momentum and angle. This also removes dependency on a purely software-based bug which caused incorrect momenta for any global tracks that were matched to the SMRD sub-detector.

With these considerations in mind we decided to develop two types of selections. The first aims to select all ν_μ-induced charged current interactions and the second only selects those with at least one charged pion in the final state.

The following signal and background definitions were used when developing and evaluating the selections described in the next section. For the ν_μ-CC-Inclusive selection we define various truth categories based on the neutrino interaction type:

Non-FGD: Neutrino vertices from outside the FGD FV that are mistakenly reconstructed as inside the FV.

ν_μ-CC: Charged current ν_μ neutrino interaction in the FGD FV.

NC-All: All neutral current neutrino interactions in the FGD FV.

[17] Using GENIE 2.6.2 MC and the current nd280 geometry.

[18] Note that no minimum track length through an active volume is required so some of these topologies would not be reconstructed.

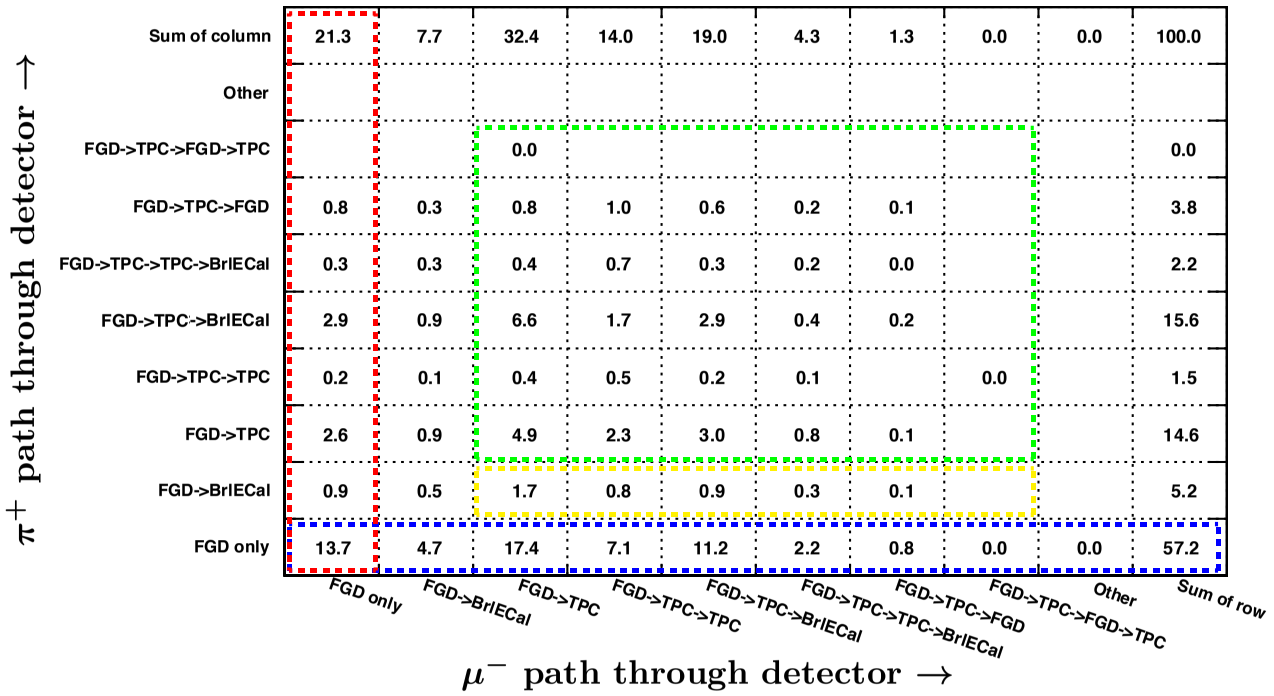


Figure 2.25: Detector trajectory topologies for ν_μ-CC-1π^+ FGD interactions showing various categories for the true path of the outgoing μ^- and π^+ based on which active detector volumes they pass through (unlike elsewhere, here the detector category BrlECal is inclusive of both the Barrel- and Ds-ECal). All entries are in percent and the total number of ν_μ-CC-1π^+ FGD interactions predicted by GENIE for the Run 2 POT is $\sim$ 2300 (where we use the FGD fiducial volumes described in § 3.3.1). The column characterises the path of the μ^- and the row the path of the π^+. The cases where the particle travels through only the FGD and no other active volume are highlighted in blue and red for the μ^- and π^+ trajectories respectively. We see that this happens for 57.2% of interactions when considering the path of the π^+ and 21.3% of interactions when considering the μ^-. For 13.7% of interactions both the π^+ and μ^- travel through only the FGD. The green box shows those topologies that are in principle possible to select when using a TPC-based PID for both the μ^- and the π^+ (28.4% of interactions) and the yellow box shows those topologies which could in principle be recovered using an ECal-based PID to identify the π^+ candidate (an additional 3.4%).

$\bar{\nu}_\mu$: Any $\bar{\nu}_\mu$ interaction in the FGD FV.

$\nu_e/\bar{\nu}_e$: Any ν_e or $\bar{\nu}_e$ interaction in the FGD FV.

$\eta_{match} < 0.5$: Reconstructed vertices where the truth matching failed to successfully find a corresponding truth vertex.

It is also useful to separate by neutrino interaction scattering process type (here the Non-FGD and $\eta_{match} < 0.5$ are same as before):

ν_μ-CC-QEL: All charged current (CC) quasi-elastic scattering.

ν_μ-CC-RES: All CC resonance production.

ν_μ-CC-DIS: All CC deep inelastic scattering.

ν_μ-CC-COH: All CC coherent pion production.

Non-ν_μ-CC: All other interactions (mainly NC, $\bar{\nu}$ and ν_e).

Finally we also categorise based on the final state pion topology of the neutrino interaction. This defines what the detector could resolve in theory (Non-FGD, $\eta_{match} < 0.5$ and Non-ν_μ-CC are same as before):

ν_μ-CC-$\pi^\pm$FS: Any CC interaction with at least one charged pion in the final state.

ν_μ-CC-0πFS: Any CC interaction with no pions or other hadrons, excluding ejected nucleons, in the final state.

ν_μ-CC-π^0FS: Any CC interaction with no charged pion in final state but with at least one neutral pion.

ν_μ-CC-OtherFS: All other final states including those with other charged mesons in the final state.

We also define a final **ν_μ-CC-π^+FS** category where the final states with at least one π^+ are singled out rather than those with either a π^+ or a π^-.

3 Near Detector Event Selections

In this chapter we describe the ν_μ-induced charged current inclusive (ν_μ-CC-Inclusive) and CC charged pion (ν_μ-CC-$\pi^\pm$) event selections which have been developed to study neutrino induced charged current $\pi^\pm$ production rates using the ND280 tracker and ECals. When developing these selections an emphasis was placed on demonstrating the use of new tools and techniques for use in future analyses—these include the use of a novel global vertexing algorithm and a newly developed ECal particle identification variable. The structure of this chapter is as follows: After an overview of the general analysis strategy the data sets used in this analysis and the validation checks of the global vertexing algorithm are shown. Next we describe the event selections and their performance and then go on to evaluate the effect of detector and reconstruction systematics.

3.1 Analysis Context and Strategy

The ND280 is a relatively new detector whose capabilities are still being explored. Throughout its commissioning and during the Run 1 and 2 data-taking periods most of the ND280 analysis effort focussed on providing a robust input to be used in the first T2K ν_e-appearance and ν_μ-disappearance physics results. This consisted of a tracker-based measurement of the inclusive rate of ν_μ-CC interactions, occurring in the FGDs, which was used to constrain the overall expected rate of neutrino interactions at Super-Kamiokande. This approach was appropriate for an analysis based on Run 1 and 2 data, where, given the relatively low statistics at the far detector the emphasis was on making sure that the result was not sensitive to any differences in the overall rate of neutrino interactions due to uncertainties in the simulation of the neutrino beam and the modelling of neutrino interaction physics. However, for T2K to reach its design sensitivity it will be necessary to reduce the uncertainty in both the shape and normalisation of the prediction of the neutrino events seen at Super-Kamiokande. ND280 is expected to play a key role in this by providing the spectral measurements of exclusive neutrino interaction channels that are necessary to disentangle data and MC discrepancies coming from the flux and neutrino interaction simulations.

Within the ND280 physics working groups there is now a concerted effort to achieve this. In particular the NuMu physics group is close to providing a spectral measurement, in terms of outgoing μ^- momentum and angle, of both the total rate of ν_μ-CC interactions occurring in FGD1 and that of a sub sample of CCQE-like events (where only a μ^- is produced, in addition to any ejection nucleons). This work is at the cutting edge of our

understanding of ND280 and is driving a lot of the studies of detector and reconstruction systematics associated with ND280-based selections. Where possible, the work described here will draw on that of the NuMu group in order to estimate the effect of these systematics on the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ selections. It should be emphasised that the NuMu group work is very much ongoing, and that in places preliminary results will be used as conservative estimates.

The selections described here have been developed to allow comparisons between data and MC for the ν_μ-CC-$\pi^\pm$ production rate as well as the total rate of ν_μ-CC-Inclusive production and put in the context of current uncertainties in neutrino interaction and flux simulation. The motivation is to provide a robust comparison of data and MC for a semi-exclusive production process and, as mentioned previously, to demonstrate the use of new techniques which can be used in future analyses. This includes the introduction of a new ECal-based PID, using the deposited charge per unit length, to tag charged pions and demonstrate how using the ECal can increase the geometrical acceptance for ν_μ-CC-$\pi^\pm$ production. Given the relatively low statistics expected for the ν_μ-CC-$\pi^\pm$ selection and the additional uncertainties introduced by using new analysis techniques, we have chosen to restrict the measurement to the comparison of relative, rather than absolute, quantities and to the total integrated rate, as opposed to a differential measurement with respect to μ^- momentum and angle.

3.2 Data Samples and Quality Cuts

This analysis will use ND280 data collected between Nov. 2010 and March 2011 during the Run 2 data-taking period. We do not use the Run 1 data as one of the objectives of this analysis is to demonstrate how using the complete ECal can increase the geometrical acceptance of the selection and for the Run 1 period the Barrel ECal was not yet installed. Details of the run periods were given in § 2.1.5.

Two types of pre-selection cuts are applied to the data. The first is a good beam spill cut to remove any ND280 events triggered on beam spills which do not pass the quality cuts of the beam monitoring system. The second is the ND280 data quality cut which requires that all the sub-detectors and subsystems, such as the magnet, were operating within their nominal conditions for that ND280 event. The total collected integrated POT for the Run 2 data-taking period before and after these pre-selection cuts is shown in Table 3.1. In general the good spill cut removes a very small fraction of the total POT because of the high stability of the beam during the Run 2 period. The drop in efficiency of almost 25% when applying the ND280 data quality cut is mainly due to a hardware problem in one of the MicroMegas on TPC3 during the early stages of the Run 2 period.

For the MC data sample we use a full spill simulation where neutrino interactions are simulated over the entire ND280 geometry including the Magnet volume and with the correct number of interactions per spill for the Run 2 beam intensity (this depends on the measured beam intensity and the predicted rate of events from the neutrino interaction

Cut	Integrated POT cut
None	10.55×10^{19}
Good spill cut	10.53×10^{19}
ND280 DQ cut	7.83×10^{19}

Table 3.1: Run 2 integrated POT used for analysis before and after beam spill and ND280 data quality cuts.

generators). The model of the detector geometry used to both simulate the neutrino interactions and the response of the detector to them was an accurate representation of the physical state of the detector during the Run 2 period and included the Barrel-ECal modules. It is important to point out that no interactions were simulated in the detector cavern and material (mainly sand) between the beam production point and ND280—This has implications on the study of backgrounds coming from neutrino interactions outside of ND280 which will be discussed in § 3.6.4. As mentioned before T2K is fortunate enough to have two independent neutrino interaction generators, both of which are fully integrated with the detector simulation and reconstruction software. For this analysis only the stream based on the GENIE generator output is used where the version of GENIE used as input to the production 4 processing is v2.6.2, details of the physics models implemented in this version were given in § 1.3. As input to the neutrino interaction generator the 11a flux simulation was used with a tuning based on external hadron production data, more details on this are given in § 2.1.4 and § 4.1.2.

For both data and MC samples we use the output of the 4D and 4C productions respectively, details of the software versions used for these was given in § 2.6.

3.3 Global Vertexing in ND280

One of the main ways in which this analysis differs from most other ν_μ-based analyses is that it uses the output of a novel Kalman-filtering global vertexing algorithm as the starting point when trying to select neutrino interactions in the ND280 tracker region. The global vertexing algorithm was developed at the University of Geneva in response to the need for a more precise and generic vertexing technique for the ND280 detector. As described in § 2.6.3 it uses a basic clustering stage to group together tracks from the same beam bunch into potential vertices and then a Kalman filter to iteratively decide whether these tracks are associated with the same vertex and to use the correlations between the final set of tracks associated with the vertex to fit a more precise vertex position. In contrast, the simple vertexing technique used in many other analyses centres around using the start point of the most energetic track[1] as the vertex location and then associating any tracks within a certain radius of this start point with that vertex. In the case of a global vertex where no preliminary clusters were found, or there was only a single track present,

[1] Typically it is actually the start point of the most energetic negative track that is used but this is more of an analysis choice to reflect the assumption that a μ^- will be produced.

the Kalman filtering stage is not performed and instead the simple technique based on the start point of the most energetic track is used. For multi-track vertices the global vertexing has a number of advantages over the simple approach:

- The final fitted vertex resolution taking into account the correlations between the different tracks has a superior resolution to that given by taking the start point of most energetic track.
- By creating potential vertices using only the spacial and directional orientation of tracks it provides a more inclusive selection of vertices leaving analysis-specific choices based on things like the charge or momentum of the constituent tracks to a later stage.
- Treating the vertexing in a generic way means it can easily be extended to vertices in subdetectors other than the tracker region.

It should be noted that many of the parameters which control the behaviour of the global vertexing have only been tuned approximately during its development [81]. However, even with this first set of parameters it does perform well and, as will be shown in § 3.4, yields similar results to analyses based on the simple vertexing technique. Because of this and the fact that it is a new tool which has not been used by many other analyses the focus of this section is on validation of the global vertexing and not optimisation of its performance.

3.3.1 Bunch Timing and Fiducial Volume Cuts

Bunch timing and FGD fiducial volume (FGD FV) cuts were applied before carrying out the global vertexing checks. Fig. 3.1 shows the vertex times for data and MC with the bunch timing cut times indicated. For this a relaxed set of timing cuts was used, consisting of eight equally-spaced bunch windows aligned with the start of the beam trigger and with the expected bunch periodicity. The double peak structure present in the data for each bunch is due to a known shift in the arrival time of the beam during Run 2 data taking. In future analyses it may be preferable to tighten the cuts to more closely surround the expected arrival time of the bunches, and hence more effectively reject background, but for this analysis the choice was made to use the relaxed set to ensure the selection is robust to both changes in the arrival time of the bunches and the differences in shape between data and MC.

Three considerations drove the choice in the FGD1 and FGD2 fiducial volume cuts:

1. Maximising the total number of selected signal events.
2. Reducing contamination of the ν_μ-CC selected events from backgrounds, especially the out-of-FGD background (Non-FGD background) where an interaction occurring outside the FGD FV is mis-reconstructed inside the FV.
3. Ensuring the selection was robust to any possible shifts between the FGD detector locations between the actual detector and the MC simulation of the detector geom-

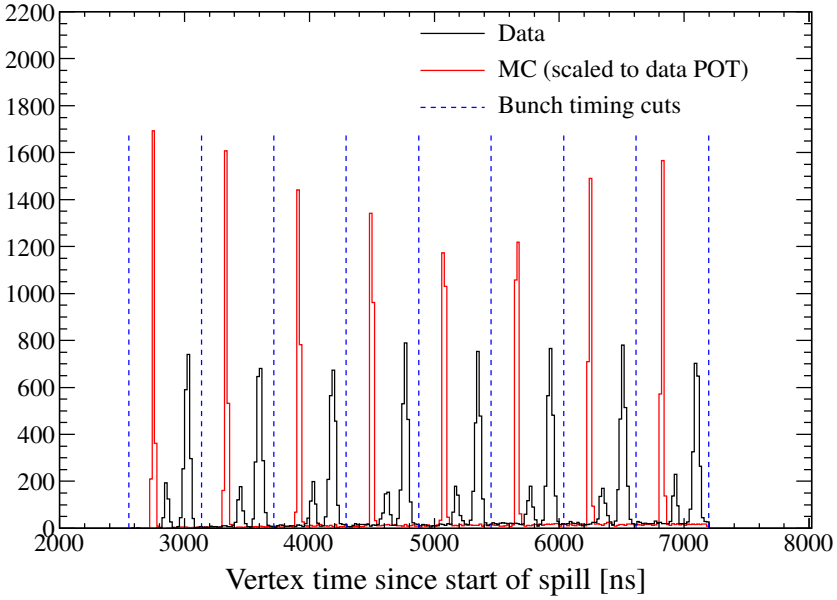


Figure 3.1: Time since start of spill for global vertices showing the 8 bunch spill structure. A double peak is visible per bunch for data, which is expected because of a known shift in the arrival time of the beam during the Run 2 data taking. The timing cuts used to separate into bunches are overlaid and show that they are not sensitive to the change in arrival time or the differences in the width of the bunches between data and MC.

etry; Alignment studies [82] indicate that these are of the order 1 mm for x and z, and 4 mm in y.

Fiducial volumes that satisfy the first two of these requirements were chosen by using the largest FV possible whilst not entering a region where the fraction of background events starts to dominate. The third requirement is met by ensuring that the actual FV boundaries align with the gaps between scintillator bars. This is important as the reconstructed vertex positions are heavily quantised at the bar centres in the z direction so any shift in the FV boundary in z should not cause it to migrate over a bar centre[2]. The fiducial volumes chosen are:

FGD1 FV (mm): $-874.5 < x < 874.5$, $-819.5 < y < 929.5$ and $136.9 < z < 447.0$

FGD2 FV (mm): $-874.5 < x < 874.5$, $-819.5 < y < 929.5$ and $1500.0 < z < 1807.1$

For FGD1 these values are based on those suggested by the NuMu group for its official FGD1-only spectrum analysis. The upstream z boundary was chosen to lie between the first and the second x-y modules as this optimised signal to background ratio. The x and y cut values are based on previous studies within the NuMu group which also optimised the signal to background ratio. For FGD2, the x and y values match those of FGD1 and a qualitatively similar choice was made for z.

[2]Even though the quantisation is less pronounced the same procedure is adopted for the x and y directions.

Fig. 3.2 shows the reconstructed vertex x, y and z positions for global vertices passing the bunch timing cut in FGD1. The dashed lines indicate the boundaries of the FV regions. The MC is normalised to the data POT and Figs. 3.2(a) and (b) show the distributions with and without the fiducial volume cuts applied. The peaks due to bar quantisation in the z-direction are clearly visible and correspond to the centre of each of the 28 planes making up 14 x-y modules. Fig 3.2(a) shows that the upstream FV boundary lies between these peaks after the first x-y module. The MC is broken down by true neutrino interaction type with the signal ν_μ-CC interactions shown in red and the dominant Non-FGD background in blue. The cuts can be seen to successfully remove the outer regions where the fraction of Non-FGD background events increases. There is a data excess in these outer regions, which is most likely from un-simulated sources of Non-FGD backgrounds, resulting in an overall data excess in the FV region when projecting the vertices onto a particular x, y or z axis. This data excess is mostly removed when the FV cuts have been applied as can been seen in Fig. 3.2(b). Similar results are shown in Fig. 3.3 for vertices in FGD2, where the sparser peak structure in the z-direction is caused by the fact that there are only 7 x-y modules interleaved with the water targets. As with FGD1, the upstream z FV boundary was chosen to lie just after the first x-y module as this gave the most effective rejection of background. Overall there is a good agreement between data and MC. There is a slight data excess which could be due to un-simulated Non-FGD background but, as will be shown in § 4, this is well within the uncertainties from flux and cross section errors.

As a check of the stability of the beam and the effectiveness of the ND280 data quality cuts, the event rate, after bunch and FV cuts, was plotted throughout the Run 2 period. This is shown in Fig. 3.4 where the run period has been divided into bins of 10^{18} POT that are contiguous in time. A Kolmogorov-Smirnov probability of 0.88 shows consistency with a constant event rate throughout the run.

3.3.2 Data-MC Comparisons

This analysis is one of the first to use the global vertexing tool. Because of this it was important to perform a number of checks to show that there is no obvious bias introduced by the global vertexing algorithm responding differently between data and MC. This approach has the inherent problem that there is no reason to assume that the input quantities to the global vertexing are the same in data and MC, so in general, we restrict these checks to looking for agreement at the level to which we expect the underlying distributions may vary.

As described in § 2.6, an important feature of the global vertexing is that for multi-track vertices it will decide whether a track should be associated with the vertex based on the effect that including that track has on the overall χ^2 of the Kalman filter. If including an extra track increases the χ^2 by a critical value, it is then not associated with the vertex. Fig. 3.5 shows a comparison between data and MC for the final χ^2 returned from the vertex for single-track vertices on the left and multi-track vertices on the right. In the

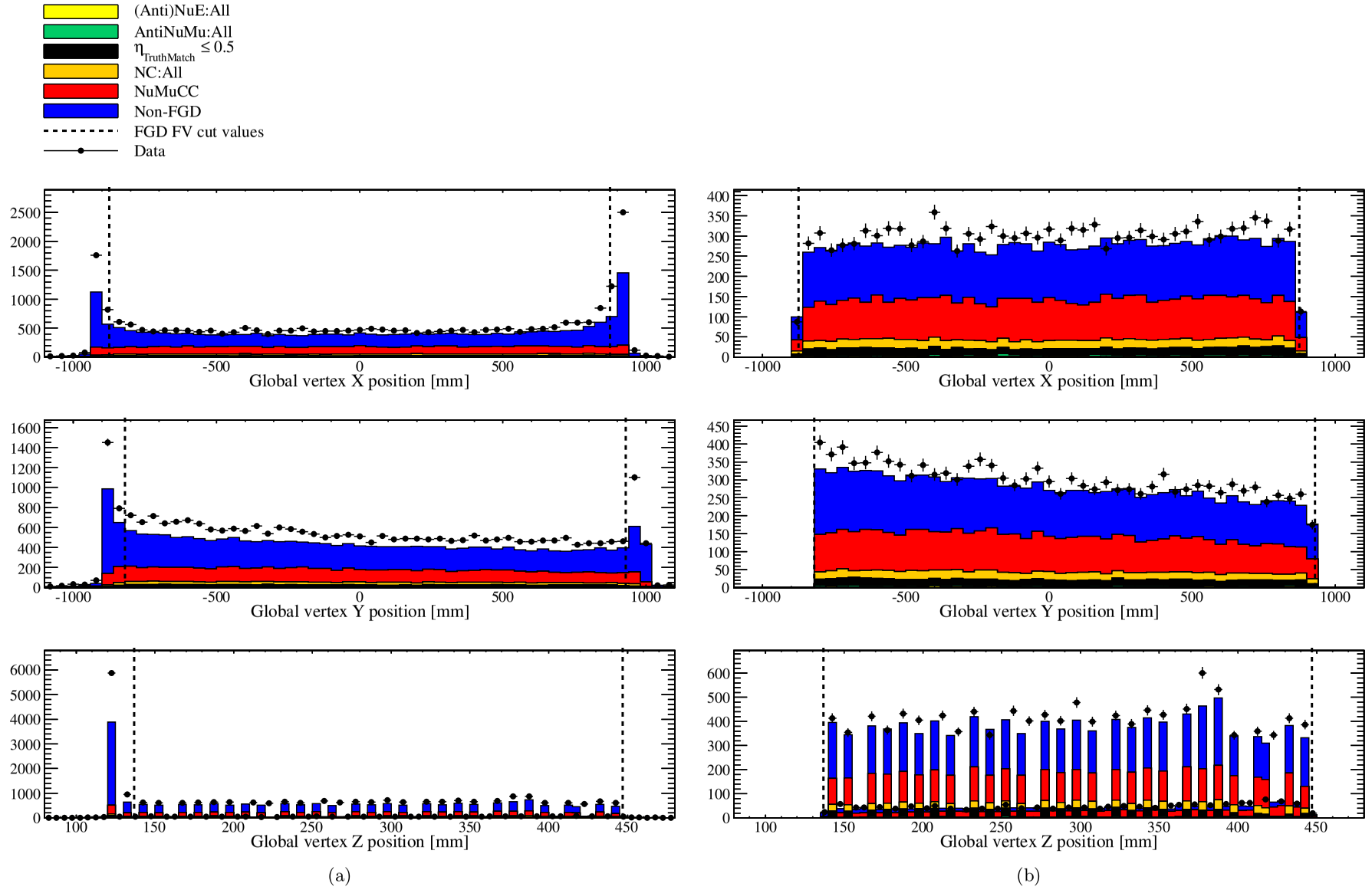


Figure 3.2: Comparison of reconstructed global vertex positions in FGD1 for data and MC both (a) with and (b) without the FGD fiducial volume cuts applied. Statistical errors are shown only for data, with those for MC being approximately 2.3 times smaller. The MC is normalised to the integrated Run 2 data (7.83×10^{19} POT).

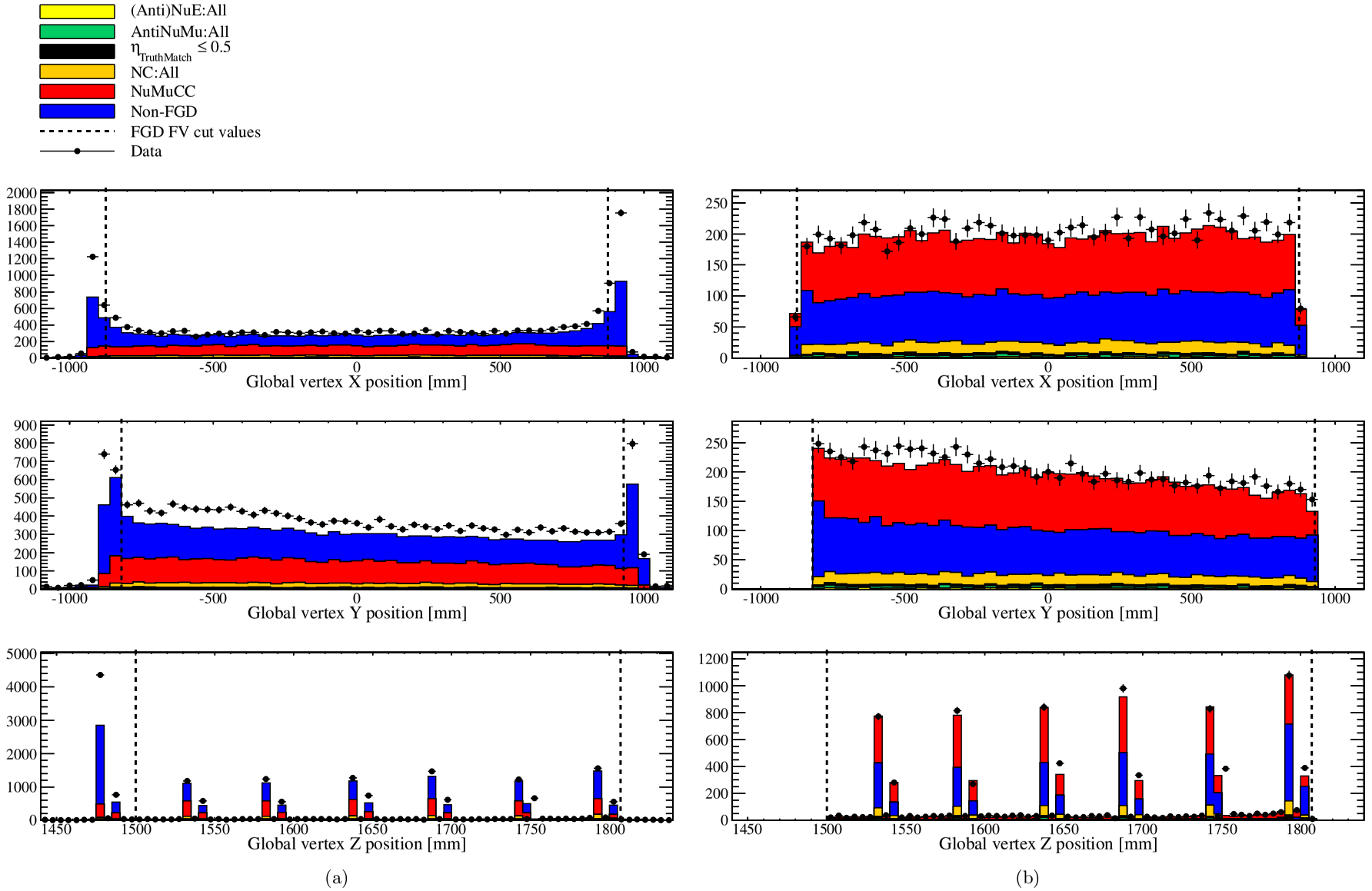


Figure 3.3: Comparison of reconstructed global vertex positions in FGD2 for data and MC both (a) with and (b) without the FGD fiducial volume cuts applied. Statistical errors are shown only for data, with those for MC being approximately 2.3 times smaller. The MC is normalised to the integrated Run 2 data (7.83×10^{19} POT).

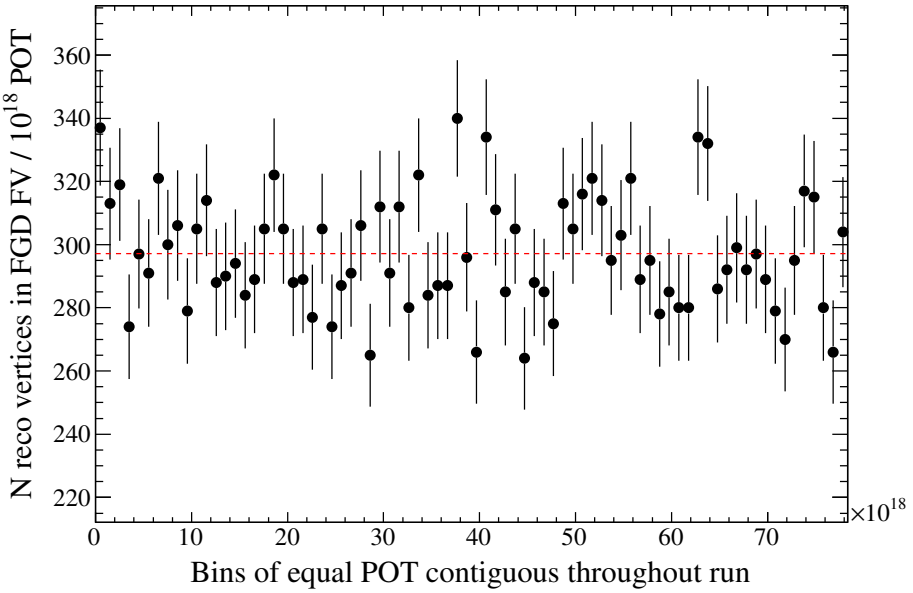


Figure 3.4: Number of reconstructed vertices after data quality and FGD FV cuts for 10^{18} POT regions throughout the Run 2 data taking period. The Kolmogorov-Smirnov probability for a constant rate is 0.88.

case of single-track vertices the χ^2 is simply that of the fitted track, and overall this shows good agreement between data and MC. There is a slight MC excess at low χ^2 which is not understood but as the 1-track vertexing algorithm does not make any choices based on the track χ^2 this is unlikely to have an impact on its performance. For the multi-track vertices there is good agreement of the χ^2/dof between data and MC. This indicates that the global vertexing responds in a similar way to data and MC and also that the input distributions are similar for data and MC.

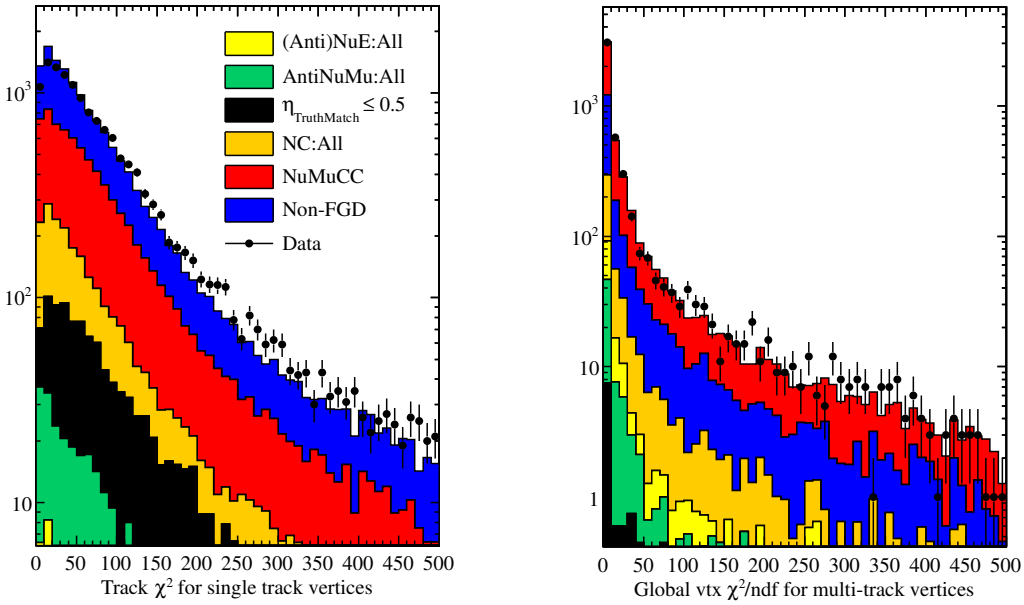
Fig. 3.5(a) shows the MC broken down by the neutrino interaction type. There seems to be no correlation between high values of χ^2 and larger fractions of background interactions, such as the Non-FGD background, indicating that there is no advantage in applying a quality cut on the vertex χ^2 before using it in the selection. For multi-track vertices such a correlation might be expected due to the fact that the vertexing algorithm is effectively creating a false vertex from two or more tracks originating from vertices outside of the FV. In Fig. 3.5(b) the MC is broken into truth categories based on the fraction of reconstructed tracks which match back, using MC truth, to the original vertex. For example the *all tracks from primary* category in green indicates that all the reconstructed tracks came from the same truth vertex, whereas vertices where at least one, two or more of the tracks did not originate from the truth vertex are shown in increasing shades of red. There seems to be no correlation between high χ^2 values and the fraction of unmatched tracks, indicating that the uncertainty on the track directions and positions dominate over any topological differences in the arrangement of tracks from the same vertex and the coincidence of tracks

from multiple vertices[3]. As with the breakdown by interaction type, there seems to be no advantage in applying a vertex quality cut to remove vertices where tracks have been incorrectly matched.

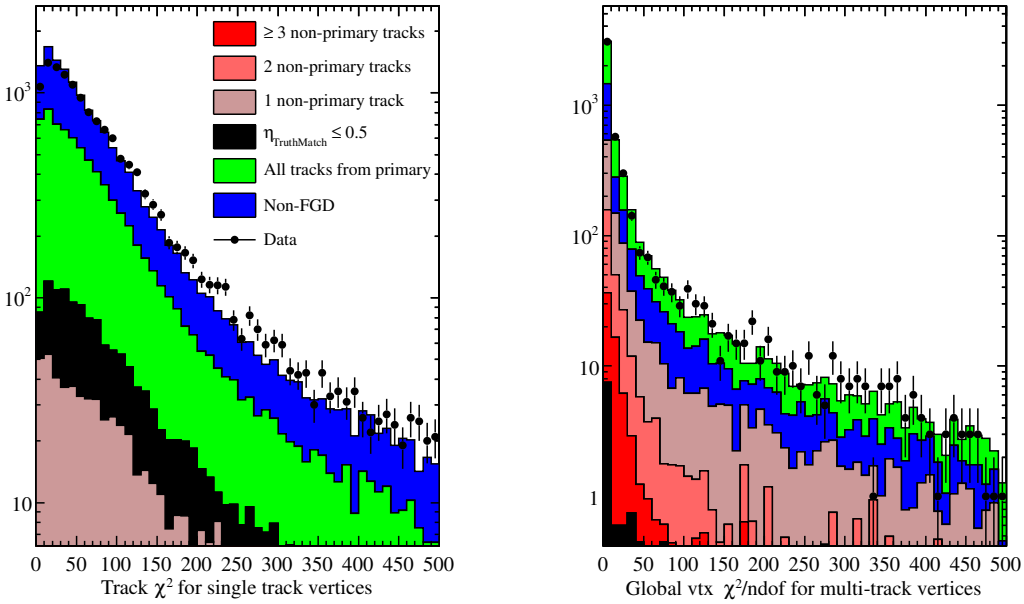
In the case of the simple vertexing algorithm, any track within a certain distance of the vertex, which is defined as the start point of the highest momentum track, is associated with that vertex. This will depend only on the distribution of tracks around the vertex and not on their direction or momentum. This is not the case for the global vertexing which, as discussed above, may remove tracks which seem inconsistent with coming from the vertex regardless of how close they are. The global vertexing will therefore be more sensitive to any changes in momentum or direction of the tracks being used to make the vertices. Fig. 3.6 compares the distance between the closest end of a track and the vertex to which it is associated for data and MC. Single-track vertices are shown on the left, and multi-track vertices on the right. Vertices in FGD1 and FGD2 are shown in Fig. 3.6(a) and 3.6(b) respectively. The MC is separated by the particle type which made that track indicating that there is no strong correlation between the type of particle which made the track and the typical distance of the start point of that track to the vertex. There is good agreement between data and MC for both the single and multi-track vertices and, as with the χ^2 comparisons, this indicates both that the vertexing algorithm responds in the same way to data and MC and that the input distributions are similar.

The final data-MC comparison check relates to the vertex clustering efficiency. Given a neutrino interaction which produces a number of tracks which are successfully reconstructed, we ask what the efficiency is for the global vertexing algorithm to correctly match them back to the same vertex. If the global vertexing efficiency were different for data and MC then this could easily introduce a bias into the analysis by causing migration between the different vertex samples defined by the number of tracks associated with the vertex. Calculating this vertexing efficiency exactly is difficult as it would require disentangling effects caused by differences in the input distributions and those due to the vertexing efficiency, and for data it is impossible to say whether a track actually originated from a vertex or came from some other activity in the detector. A basic check can be performed, assuming similar input distributions for both data and MC, by comparing the number of tracks which the global vertexing algorithm associates with a given vertex to the total number of tracks within a given distance of that vertex. This is an effective vertexing efficiency, as it does not require that the track actually originated from the vertex, but it should still be sensitive to any differences in the true vertexing efficiency for data and MC. Table 3.2 shows the result of a study using the entire Run 2 data and a statistically equivalent set of MC. The columns indicate the total number of tracks associated with the vertex, for MC and data side by side, and the rows indicate how many tracks were within 100 mm and 400 mm of the original global vertex position for the upper and lower six rows respectively. The large values in the diagonal columns (bold) show that, as expected,

[3]It may well be the case that the initial clustering stage of the vertexing, combined with the removal of tracks causing a significant increase in χ^2, effectively removes cases where an effect would be noticeable.

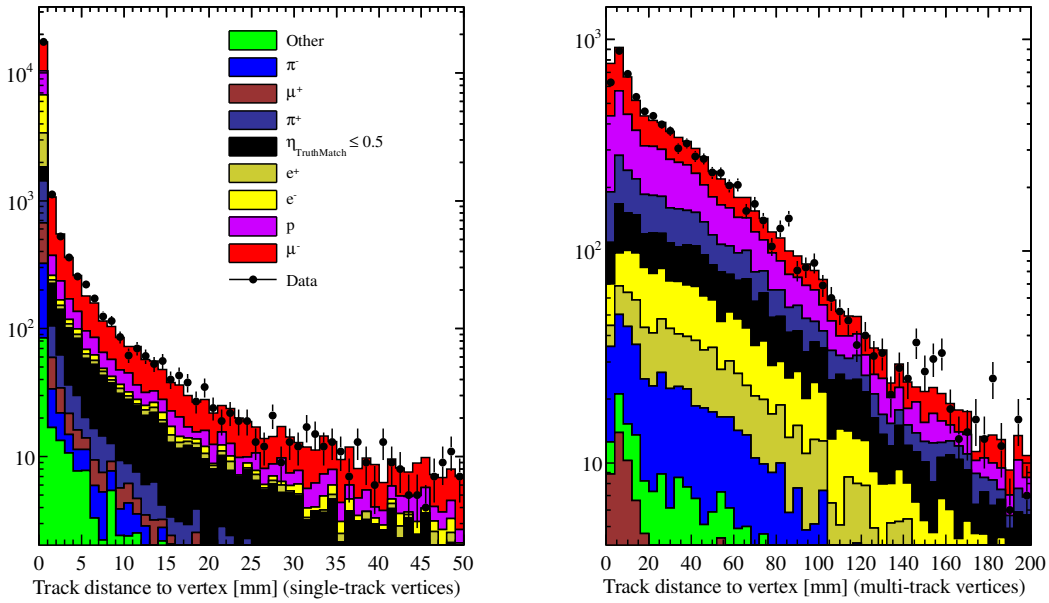


(a) MC categorised by CC-inclusive interaction type.

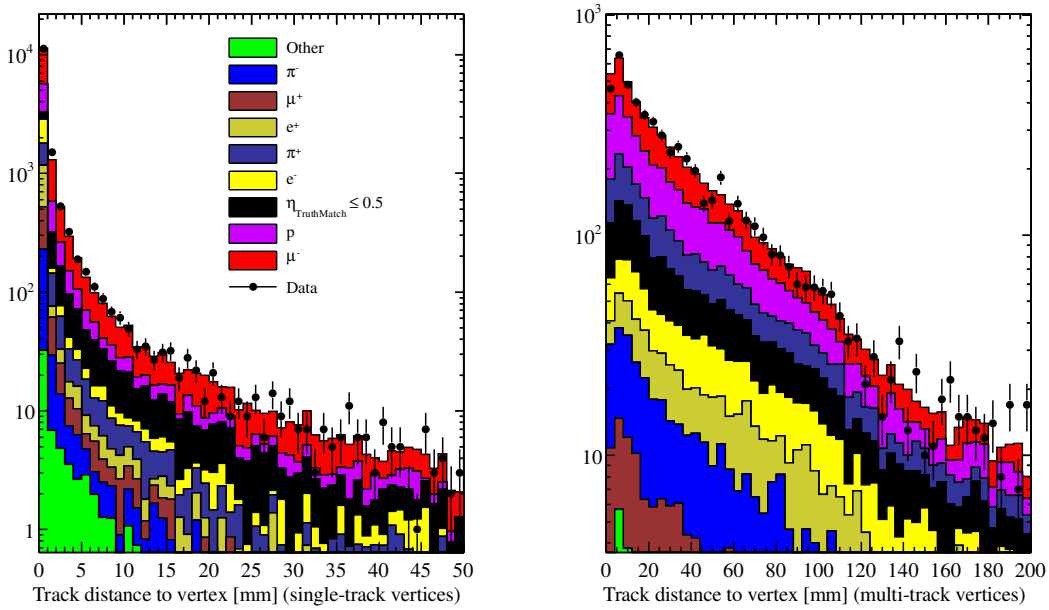


(b) MC categorised by truth match to primary vertex type.

Figure 3.5: Data and MC comparison of the χ^2 from the global vertexing algorithm for single-track (where χ^2 is that from the global fit of the contributing track) and multi-track vertices (where χ^2 is that returned by the final iteration of the Kalman filter). The MC is normalised to the total number of vertices in data for the one- and multi-track samples and the statistical errors are about 4 times smaller than those shown for data. In (a) MC is broken down into truth categories by neutrino interaction type and in (b) by the number of tracks clustered by the vertexing algorithm that do not originate from the primary neutrino interaction (based on MC truth matching). There is no obvious correlation between larger values of χ^2 and the amount of background contamination or the number of incorrectly clustered tracks.



(a) Tracks associated with vertices in FGD1 FV.



(b) Tracks associated with vertices in FGD2 FV.

Figure 3.6: The distance from the closest end of a reconstructed track and the vertex to which it is associated is shown for single and multi-track vertices on the left and right respectively and separately for (a) FGD1 and (b) FGD2. The MC is normalised to the total number of tracks for single and multi-track vertices for the Run 2 data set and the statistical errors are about 4 times smaller than those shown for data. The MC is broken down by the MC true particle type which created the track. Overall good agreement between data and MC for single and multi-track vertices is shown for both FGD1 and FGD2.

		Number of tracks associated with the vertex							
		1 tracks		2 tracks		3 tracks		≥4 tracks	
		Data	MC	Data	MC	Data	MC	Data	MC
$N_{Trks} < 100$mm	0 tracks	$0.1\% \pm 0.1\%$	$0.2\% \pm 0.1\%$	$1.1\% \pm 0.3\%$	$0.4\% \pm 0.2\%$	$0.0\% \pm 0.0\%$	$0.4\% \pm 0.4\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$
	1 track	**$89.7\% \pm 2.4\%$**	**$91.2\% \pm 2.1\%$**	$8.2\% \pm 0.8\%$	$7.5\% \pm 0.7\%$	$1.6\% \pm 0.8\%$	$0.4\% \pm 0.4\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$
	2 tracks	$6.3\% \pm 0.5\%$	$5.4\% \pm 0.4\%$	**$81.9\% \pm 3.3\%$**	**$84.7\% \pm 3.2\%$**	$10.9\% \pm 2.2\%$	$13.3\% \pm 2.3\%$	$6.5\% \pm 3.0\%$	$3.8\% \pm 2.3\%$
	3 tracks	$2.2\% \pm 0.3\%$	$2.1\% \pm 0.2\%$	$6.4\% \pm 0.7\%$	$6.0\% \pm 0.6\%$	**$66.9\% \pm 6.7\%$**	**$75.5\% \pm 6.9\%$**	$19.5\% \pm 5.5\%$	$24.4\% \pm 6.2\%$
	4 tracks	$1.1\% \pm 0.2\%$	$0.7\% \pm 0.1\%$	$1.7\% \pm 0.4\%$	$1.0\% \pm 0.3\%$	$16.1\% \pm 2.7\%$	$9.0\% \pm 1.9\%$	**$42.9\% \pm 8.9\%$**	**$52.6\% \pm 10.1\%$**
	≥ 5 tracks	$0.6\% \pm 0.1\%$	$0.4\% \pm 0.1\%$	$0.7\% \pm 0.2\%$	$0.3\% \pm 0.1\%$	$4.4\% \pm 1.4\%$	$1.4\% \pm 0.7\%$	$31.2\% \pm 7.3\%$	$19.2\% \pm 5.4\%$
$N_{Trks} < 400$mm	0 tracks	$0.0\% \pm 0.0\%$	$0.2\% \pm 0.1\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$
	1 track	**$81.9\% \pm 2.2\%$**	**$84.5\% \pm 2.0\%$**	$1.2\% \pm 0.3\%$	$0.6\% \pm 0.2\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$	$0.0\% \pm 0.0\%$
	2 tracks	$10.9\% \pm 0.6\%$	$9.2\% \pm 0.5\%$	**$77.0\% \pm 3.2\%$**	**$78.9\% \pm 3.0\%$**	$4.0\% \pm 1.3\%$	$5.4\% \pm 1.4\%$	$1.3\% \pm 1.3\%$	$0.0\% \pm 0.0\%$
	3 tracks	$3.5\% \pm 0.3\%$	$3.8\% \pm 0.3\%$	$14.3\% \pm 1.1\%$	$14.5\% \pm 1.0\%$	**$59.3\% \pm 6.2\%$**	**$70.1\% \pm 6.6\%$**	$15.6\% \pm 4.8\%$	$14.1\% \pm 4.5\%$
	4 tracks	$2.3\% \pm 0.3\%$	$1.3\% \pm 0.2\%$	$4.4\% \pm 0.6\%$	$3.1\% \pm 0.4\%$	$23.8\% \pm 3.4\%$	$17.3\% \pm 2.7\%$	**$33.8\% \pm 7.7\%$**	**$52.6\% \pm 10.1\%$**
	≥ 5 tracks	$1.5\% \pm 0.2\%$	$1.2\% \pm 0.2\%$	$3.1\% \pm 0.5\%$	$2.9\% \pm 0.4\%$	$12.9\% \pm 2.4\%$	$7.2\% \pm 1.7\%$	$49.4\% \pm 9.8\%$	$33.3\% \pm 7.5\%$

Table 3.2: Global vertexing efficiencies for data and MC. The columns indicate how many tracks were associated or matched back to the vertex by the Kalman filter, and the rows indicate how many tracks there were with their closest end within either 100mm or 400mm of the vertex. For example, the first entry that is highlighted with a border shows that for data 81.9% of vertices composed of two tracks have only two tracks within 100mm of the vertex, whereas the second highlighted entry shows that 6% of MC vertices composed of 2 tracks actually have 3 tracks within 100mm. Overall there is good agreement between data and MC at the level of the statistical uncertainty of the samples, indicating that the efficiency of the global vertexing for associating tracks with a common vertex is similar between data and MC.

a vertex composed of n tracks will typically only have n tracks within 100 or 400 mm. For example, the upper left highlighted box shows that, for data, 81.9% of two-track vertices only have two tracks within 100 mm of the vertex. The lower left off-diagonal elements indicate when the global vertexing algorithm has not matched all the tracks within 100 mm or 400 mm, and the upper right off-diagonal elements show how often it has associated tracks from outside 100 mm or 400 mm. There is good agreement between data and MC indicating that the global vertexing efficiency will not introduce a significant bias into the analysis.

3.3.3 Vertexing Performance

As mentioned before, one of the advantages of the global vertexing is that it provides an inclusive selection of neutrino interactions without making any analysis-specific physics choices on the types of particles making up the interaction. This means it can be used as a general tool in many different types of analyses which then apply further cuts specific to their needs. Table 3.3 shows the efficiency for selecting different types of neutrino interactions using the output of the global vertexing with only FGD FV and bunch timing cuts applied. The inclusiveness of the global vertexing is shown by the high efficiencies, around 80%, for all CC interaction modes. The fractions reflect the relative abundance of the various type of interaction expected from the T2K neutrino beam. The dominant background at this stage is the Non-FGD background.

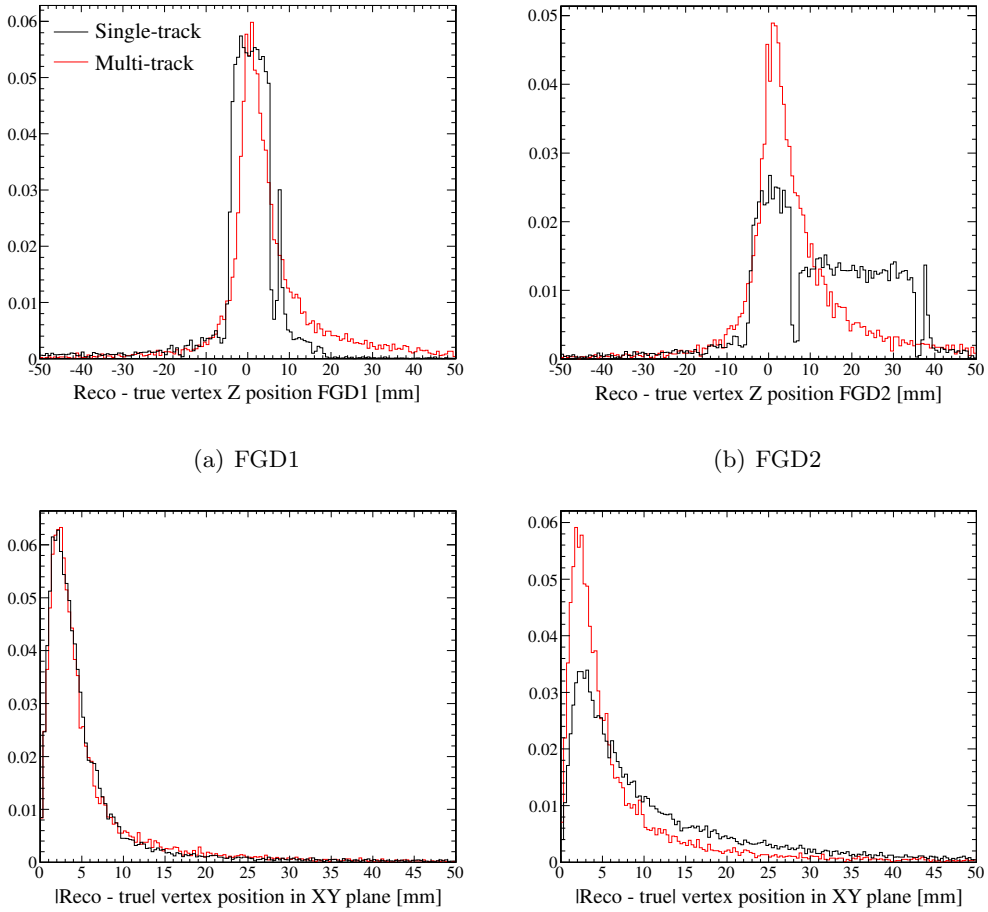
Interaction type	Fraction	Efficiency
ν_μ-CC	39.2%	79.8%
NC-All	7.8%	36.6%
$\bar{\nu}_\mu$	1.2%	88.3%
$\nu_e/\bar{\nu}_e$	0.7%	80.5%
Non-FGD	48.6%	—
$\eta_{match} < 0.5$	2.4%	—

Table 3.3: Global vertexing efficiencies with only FGD FV and bunch timing cuts applied. For each interaction type the fraction and efficiency are defined as the number of selected interactions over the total selected and true number in the FGD FV respectively. The efficiency for the last two rows is 100% by construction as for the Non-FGD and truth failure modes there is no meaningful way to calculate the true number in the FGD FV.

In the case of multi-track vertices the global Kalman filter employed by the global vertexing is expected to provide superior vertex resolution to the single-track case[4]. Figs. 3.7(a) and 3.7(b) show the difference between reconstructed and true vertex z position for FGD1 and FGD2 respectively. As expected the multi-track vertices exhibit a narrower peak implying a higher resolution. For single-track vertices in FGD1 we see the expected ~ 10 mm step-like function due to the finite quantisation of reconstructed positions in the FGD

[4]Or equivalently simple vertexing techniques using the start point of the highest momentum track.

bars. For FGD2 we see a much wider distribution for the single-track vertices with a second step-like function reflecting the fact that in FGD2 there are ~ 40 mm water gaps between each XY-module which increase the quantisation of reconstructed positions. In Fig. 3.7(b) the narrower peak for multi-track vertices shows the power of the global vertexing to recover the positions of neutrino interactions occurring in the water modules. For the multi-track vertices both plots show a slight bias towards reconstructing the z position of the vertex downstream of its true position. If this bias is not the same for data and MC then it would introduce a systematic by causing events to migrate into and out of the FGD FV. Before the global vertexing can be used for official results this would need to be checked.



(a) FGD1 (b) FGD2

(c) FGD1 (d) FGD2

Figure 3.7: Global vertexing resolution for single- and multi-track vertices passing the ν_μ-CC-Inclusive cuts described in § 3.4 and shown separately for FGD1 and FGD2. The distance between reconstructed and true vertex z position is shown in (a) and (b) and the magnitude of the distance between reconstructed and true vertex x-y position is shown in (c) and (d). To allow comparison between single- and multi-track vertices all plots are normalised to unity.

The absolute distance between reconstructed and true vertex positions in the x-y plane

is shown in Figs. 3.7(c) and 3.7(d). For vertices in FGD2 there is a clear improvement in vertexing resolution in the multi-track case whilst for those in FGD1 we see a similar resolution for both single- and multi-track vertices. As before, this demonstrates the power of the global vertexing to reconstruct the position of neutrino interactions occurring in the water modules.

In conclusion, the global vertexing seems to perform well and, as no significant differences in its response to data and MC were found, we will use it as the starting point in the following analysis. In the longer term, and before it is used for any official results, it will need to be optimised and more detailed studies of possible biases between data and MC explored.

3.4 ν_μ-CC-Inclusive Selection

In general, the approach for selecting ν_μ-CC interactions occurring in the FGDs is to tag the outgoing μ^- produced by the interaction. From the reconstruction point of view this translates to looking for a negative MIP-like track originating from a global vertex inside an FGD that is in time with an expected neutrino bunch. The selection described here relies heavily on the use of TPC information to identify the track as MIP-like and to ascertain the charge, allowing rejection of $\nu_e/\bar{\nu}_e$ and $\bar{\nu}_\mu$ interactions producing an outgoing e or μ^+ respectively[5]. As described in § 2.3.5 the TPC PID is based on comparing the measured and expected energy loss per unit length given a particular particle hypothesis. For each TPC track with reconstructed momentum P a pull is calculated for a given particle hypothesis α using:

$$\text{Pull}_\alpha = \frac{(dE/dx)^{meas} - (dE/dx)^{exp}_\alpha}{\sigma_{(dE/dx)^{exp}_\alpha}}$$

where $(dE/dx)^{meas}$ is the measured energy loss, and $(dE/dx)^{exp}_\alpha$ and $\sigma_{(dE/dx)^{exp}_\alpha}$ are the expected energy loss and the expected width of the energy loss respectively for a particle α of momentum P. For a given particle hypothesis a small value for the pull indicates consistency with that hypothesis.

As described in § 2.6.2 the relative timing information of tracks traversing both FGDs is used by the reconstruction to flip the track direction to be backwards-going. Prior to this, tracks are assumed to be forward-going. By requiring tracks to have a direction consistent with a track exiting the FGD, the contamination from neutrino interactions occurring outside of the FGD but which mimic a track starting inside the FGD is reduced.

To identify potential vertices with at least one negative MIP-like track, the selection uses the output of the global vertexing algorithm and the list of associated global fitted tracks on which to apply vertex level cuts, such as requiring a vertex in the FGD FV, and

[5]NC background is also reduced: The MIP-like cut gives rejection of NC-elastic events, producing only nucleons, and the charge cut provides rejection of NC-inelastic events producing any positive MIP-like hadron, typically a π^+, which would otherwise be mistaken for the muon candidate.

then successive track-level cuts, such as requiring the TPC pull to be consistent with the muon-hypothesis. The track-level cuts are applied in a cumulative fashion: in order for a track to pass a given track-level cut it must also have passed the previous track-level cut. Therefore, requiring at least one track to pass the nth track-level cut is equivalent to requiring that at least one track has passed all n track-level cuts. A more complete description of the detector components and reconstruction outputs used in this selection was given in §2.6 and here they will only be discussed with regards to their use in, and impact on, the selection. The full set of ν_μ-CC inclusive cuts are:

1) Vertex in FGD FV and bunch time: Require that the vertex position is within an FGD fiducial volume and that the vertex time falls within an expected bunch window. The cut values used are those described in § 3.3 and will not be discussed further here.

2) $\geq$ 1 track with a TPC constituent: Require there to be at least one global reconstructed track, with at least one constituent TPC sub-track, associated with the vertex.

3) $\geq$ 1 track with a good TPC quality: Of the tracks passing cut 2, require that at least one passes the TPC track quality cut requiring the TPC sub-track to be made up of at least 19 hits.

4) $\geq$ 1 track with a TPC μ-like-pull: Of the tracks passing cut 3, require that at least one has a TPC PID value consistent with the muon hypothesis, i.e., that the TPC sub-track closest to the vertex has $|\text{Pull}_\mu| < 2.0$.

5) $\geq$ 1 track without a TPC e-like-pull: Of the tracks passing cut 4, require that at least one has a TPC PID value which is inconsistent with the electron hypothesis, i.e., that the constituent TPC sub-track closest to the vertex has $|\text{Pull}_e| > 2.0$.

6) $\geq$ 1 track has negative charge: Of the tracks passing cut 5, require that at least one has a negative reconstructed charge.

7) $\geq$ 1 track which starts in FGD FV: Of the tracks passing cut 6, require that at least one starts, as reckoned by the global reconstructed direction of the track, in the same FGD FV as the global vertex.

Before using these cuts in the analysis, a number of data-MC checks were performed to ensure that no major biases were being introduced by the any of the cuts or cut values. In general these cuts have not been optimised for performance and instead are based on sensible first guesses which were made with an emphasis on robustness. Unless otherwise stated, the following comparisons use the output of the global vertexing with only FGD FV and bunch timing cuts applied.

Cuts 2 to 6 use the TPC information to select reconstructed vertices with at least one track which looks like a μ^-. As described in § 2.3.5, the TPC PID is based on comparing

the amount of charge deposited by a particle traversing the TPC to that expected based on the reconstructed momentum and a given particle hypothesis. The result of this is a pull value, in units of 1 sigma deviation, for each particle hypothesis. A small (large) absolute value of a pull for a given particle hypothesis indicates that the energy loss was consistent (inconsistent) with that type of particle. Requiring a small pull for the μ-hypothesis and a large pull for the e-hypothesis results in a high purity sample of vertices containing at least one MIP-like particle. The charge cut helps remove background from $\bar{\nu}_\mu$-CC interactions producing a μ^+, and from inelastic ν-NC interactions producing MIP-like hadrons.

Both the PID and the charge cuts depend on the reconstructed momentum of the track. By requiring that there be at least 19 hits in the TPC track, the quality cut ensures that there is enough information to properly reconstruct the momentum. Studies within the NuMu group have shown that a value of ≥ 19 eliminates low quality tracks whilst only rejecting a small fraction of the overall number of selected events. Fig. 3.8 shows the effect of varying the value of the TPC quality cut for events passing the full CC-Inclusive selection. The purity, efficiency, significance and data MC ratio are shown as a function of the minimum number of TPC hits required to pass the TPC quality cut. There is very little dependence of the data MC ratio on the cut value indicating that the systematic introduced by this cut is small. The significance is approximately flat for cut values of smaller than 20.

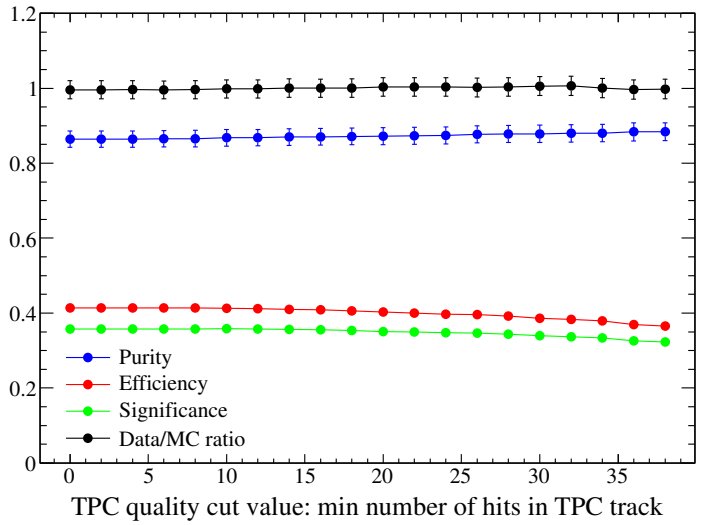


Figure 3.8: The effect of varying the TPC quality cut. The purity, efficiency, significance (purity $\times$ efficiency) and data MC ratio of the number of selected events are shown as a function of TPC track quality cut for events passing the CC-Inclusive selection. The signal definition is for a ν_μ-CC interaction in an FGD FV. The data MC ratio is normalised to 1.0 and shows almost no dependence on the cut value, indicating a small systematic effect.

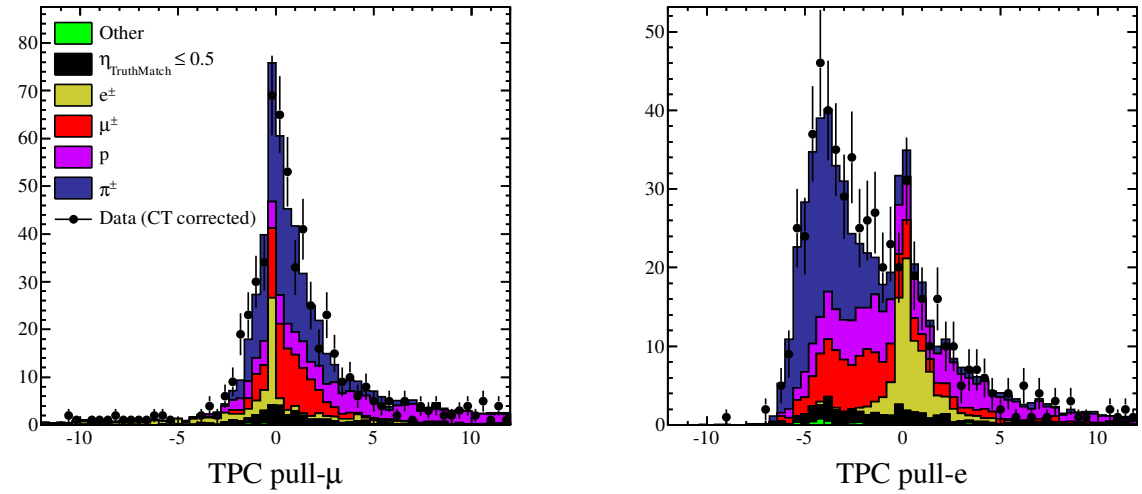
Fig. 3.9 shows data-MC comparisons for both pull-μ and pull-e distributions for tracks in TPC1, TPC2 and TPC3 separately. Generally there is a good agreement, indicating that these are reliable quantities to cut on. There is slightly more of a data excess for

the pull-e distributions and in particular for TPC3. This is not necessarily due to a systematic uncertainty in the way the TPC pulls respond to data and MC and could simply be caused by an excess of e-like tracks in data. Further discussion on estimating the systematics associated with these cuts will be given in § 3.6.

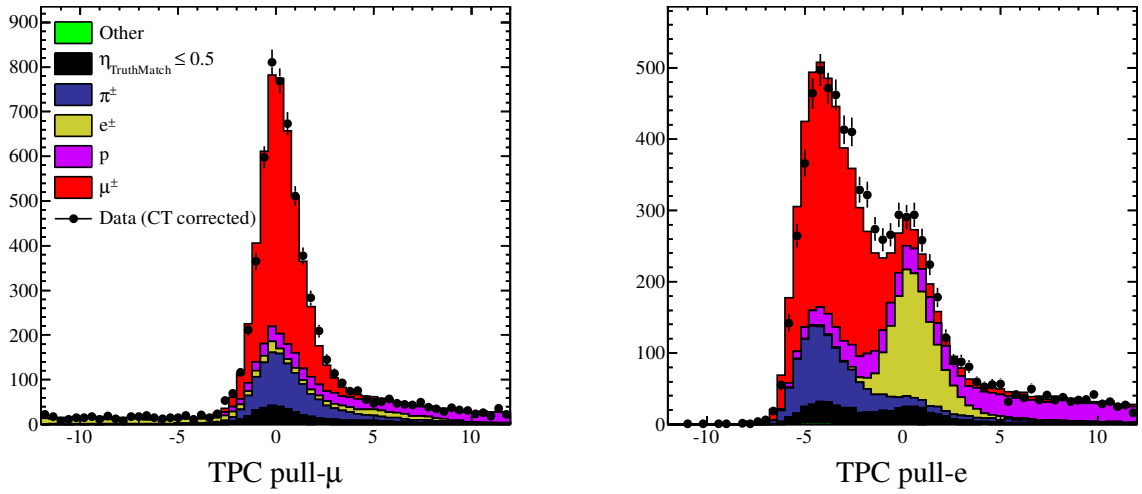
Fig. 3.10 shows the effect that varying the TPC pull cuts has on the purity, efficiency and data MC ratio for events passing the ν_μ-CC-Inclusive cuts. Fig. 3.10(a) shows the effect of varying the pull-μ cut value for the ν_μ-CC-Inclusive cuts up to and including cut three. As we would expect for the cut on $|\text{Pull}_\mu|$, the efficiency increases with a larger cut value and the purity decreases. There is no clear maximum for the significance and any cut value between approximately 2.0 to 5.0 gives similar performance. The data MC ratio is stable in the region of the cut value of 2.0. Fig. 3.10(b) shows the effect of varying the pull-e cut value after applying the ν_μ-CC-Inclusive cuts up to and including cut three, i.e., not applying the pull-μ cut. As expected for a cut which excludes e-like tracks, the efficiency and purity behave in the opposite manner to those of cut four as the cut value is varied. The maximum significance (purity $\times$ efficiency) occurs in the region of the cut value of 2.0 chosen and the data MC ratio is stable around this point. The conclusion is that the initial cut values of 2.0 are a reasonable starting point.

In addition to the actual cut quantities themselves, a number of other reconstructed quantities were checked for general agreement between data and MC. Fig. 3.11 shows good agreement for the track occupancy of the ND280 TPC and ECal sub-detectors for global vertices within the FGD FVs. This tells us that to first order the reconstruction efficiency in each of the sub-detectors is simulated correctly. The low number of tracks in TPC1 compared to the other TPCs is expected based on the direction of the beam and the angular distribution of outgoing particles. Fig. 3.12(a) shows the reconstructed charge times momentum for tracks passing the TPC quality cut and shows relatively good agreement between the relative abundance of positive and negative particles as well as for the reconstructed momentum. Fig. 3.12(b) shows the angular distribution of all tracks associated to the global vertices and, as with the other comparisons, this shows a good agreement in the shape between data and MC.

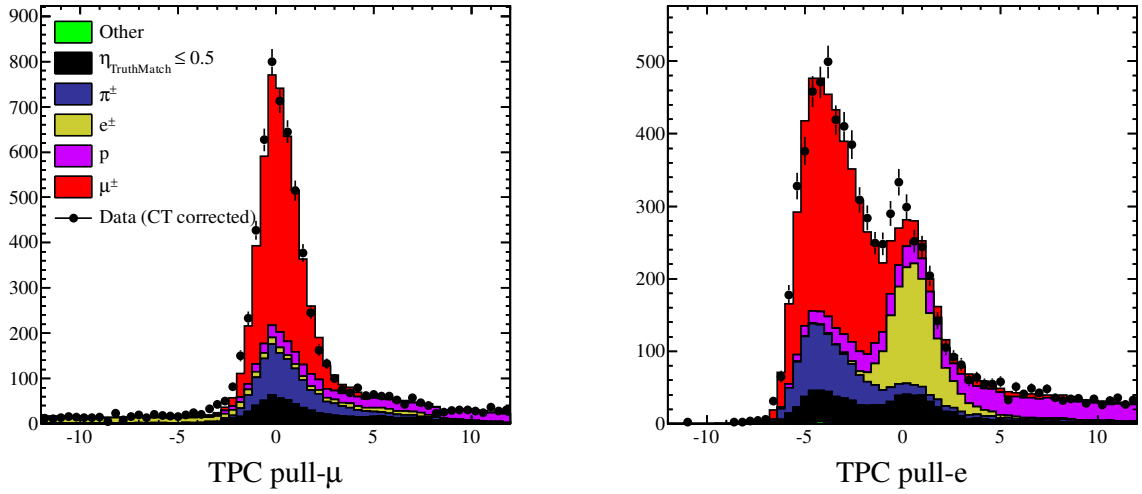
In this section we have shown that, in general, there is good agreement between data and MC both for the quantities that are being cut on and for a number of reconstructed track-level quantities. This, combined with the validation of the global vertexing algorithm in § 3.3 indicates that there are no major biases being introduced by the use of these in the analysis. The systematics introduced by these cuts and those due to the global and sub-detector reconstruction will be discussed further in § 3.6.



(a)

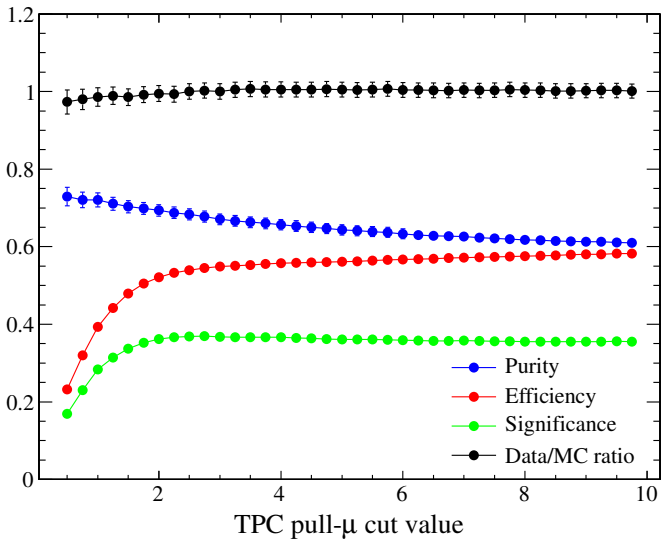


(b)

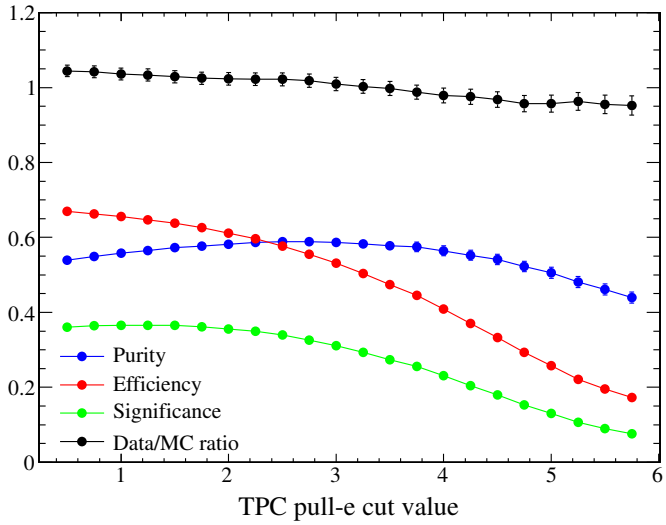


(c)

Figure 3.9: Data-MC comparisons for the TPC pull distributions for pull-μ and pull-e for all tracks with TPC information from vertices passing the FGD-FV and bunch timing cuts. Shown separately for TPC1, TPC2 and TPC3 in (a), (b) and (c) respectively. The MC is normalised to the total number of tracks in data for each sample.

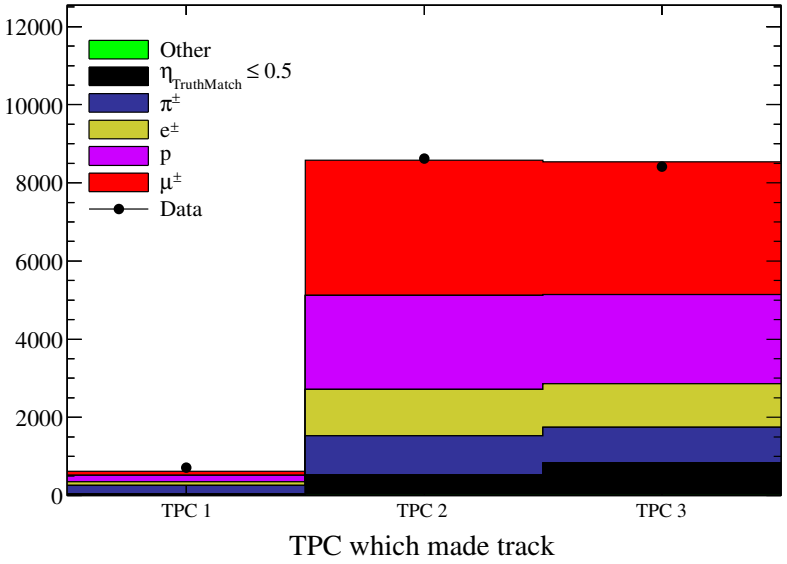


(a)

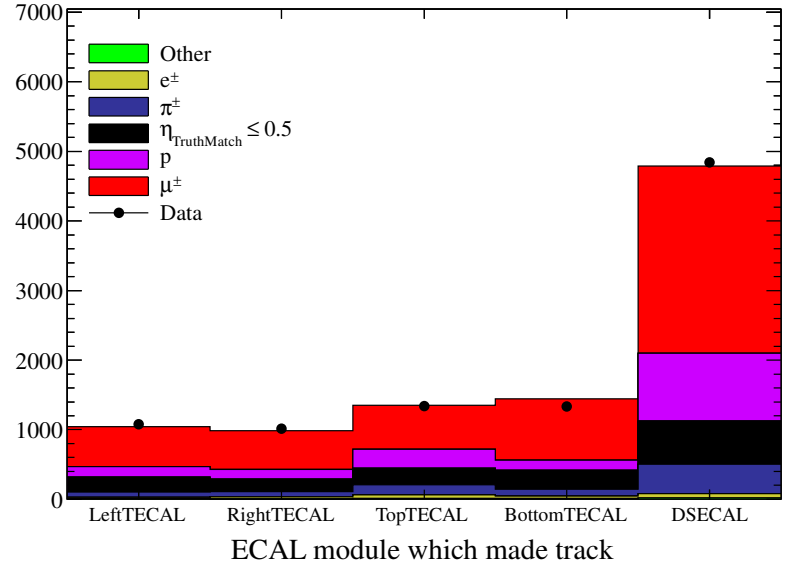


(b)

Figure 3.10: The effect of varying the TPC pull cuts for the CC-Inclusive selection. The purity, efficiency, significance (purity $\times$ efficiency) and data MC ratio of the number of events passing the cuts are shown as a function of TPC pull cut value for events passing the CC-Inclusive selection. The signal definition is a ν_μ-CC interaction in an FGD FV. The effect of varying the cut on $|\text{Pull}_\mu|$ which excludes non-μ-like tracks, is shown in (a) and the effect of varying the $|\text{Pull}_e|$ cut to exclude e-like tracks, is shown in (b). To avoid potential bias when tuning the cut value the data MC ratio was scaled so that the average for all points was 1.0.

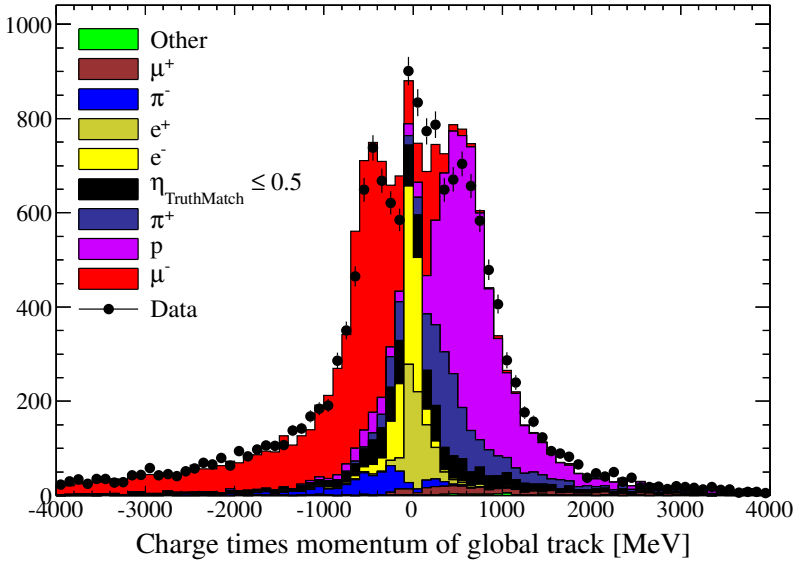


(a)

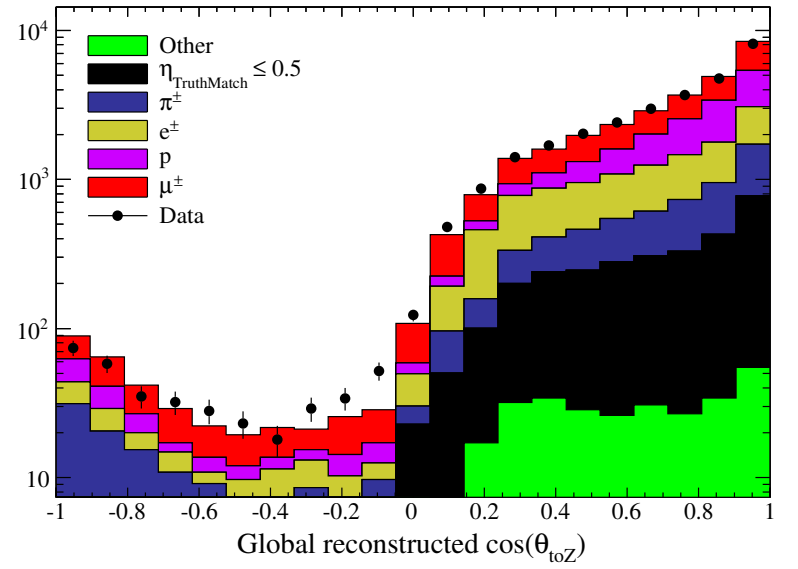


(b)

Figure 3.11: Detector occupancies for tracks associated with global vertices passing the FGD FV and bunch timing cuts. The MC is normalised to the total number of tracks in data for each sample.



(a)



(b)

Figure 3.12: General data-MC comparison for reconstructed quantities of tracks associated with global vertices passing the FGD FV and bunch timing cuts. The MC is normalised to the total number of tracks in data for each sample.

3.4.1 Performance of ν_μ-CC-Inclusive Cuts

The performance of the ν_μ-CC-Inclusive cuts based on their application to the full MC sample is shown in Table 3.4. Here the purity and efficiency for selecting different neutrino interaction types is shown as a function of cut number. Here, and unless otherwise stated, we use an absolute efficiency defined as the ratio of selected true interactions of a particular type to the number of true interactions of that type with a true vertex in the FGD FV. Also shown, in the left two columns, are the total number of selected interactions based on the MC sample size and scaled to the expected POT of the data set. The purities and efficiencies for selecting ν_μ-CC signal events are highlighted with double vertical lines and after application of all seven cuts we achieve a purity of 88.1% and an efficiency of 40.1%. These numbers are comparable to those obtained by other analyses within the NuMu group which do not use the global vertexing.

The dominant background is from Non-FGD interactions which make up 8.2% of the 11.9% loss in purity. Cuts 2 to 5 which require at least one μ-like particle reduce this Non-FGD background contamination from 48.6% to 18.9% which implies that a large fraction of it comes from electron- or proton-like particles. The remaining Non-FGD background represents a range of failure modes including:

- Single track vertices where a μ-like particle from an interaction outside the FGD stops in an FGD FV, but is mistakenly reconstructed as a track starting in the FV and travelling out. Unless relative timing information from the FGD can be used to determine the direction this mode is hard to reject.
- Failures in reconstruction causing a through-going μ-like particle from an interaction outside the FGD which crosses and does not stop in the FGD, to be broken into multiple-tracks, mimicking a particle starting in the FGD.
- Hard scattering of through-going particles which can result in two reconstructed tracks with a reconstructed vertex at the kink.

The final two cuts, based on the charge and the direction of the track respectively, reduce the remaining 18.9% down to its final value of 8.2%. It would be possible to reduce this further through the introduction of vetoes on tracks in various sub-detectors, such as TPC1 and the P0D, and through cuts on event kinematics, such as removing events containing two tracks with an opening angle consistent with hard scattering of a single particle. However, as one of the main purposes of this analysis is to study how using the ECal can increase the angular acceptance of secondary tracks it was decided not to apply any of these so as not to remove potentially interesting events.

The other main contributor to the background contamination is from neutral current interactions (NC-All). Cuts 2 to 6, by requiring at least one negative μ-like track, reduce this down from 7.8% to 2.9%. The remaining background is likely due to inelastic ν-NC-interactions producing one or more negative MIP-like particles, such as a π^-, or a positive MIP-like particle, such as a π^+ whose charge is then mis-reconstructed.

			ν_μ-CC		NC-All		$\bar{\nu}_\mu$		$\nu_e/\bar{\nu}_e$		Non-FGD		$\eta_{match} < 0.5$	
	N_{Total}	N_{Total}^{Scaled}	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ
1) Vtx in FV and bunch	312904	21522	39.2%	79.8%	7.8%	36.6%	1.2%	88.3%	0.7%	80.5%	48.6%	100.0%	2.4%	100.0%
2) ≥ 1 trk with TPC	216264	14875	49.6%	69.8%	8.2%	26.6%	1.6%	83.7%	1.0%	74.4%	39.0%	55.4%	0.6%	16.6%
3) & with good TPC	198443	13649	52.1%	67.2%	8.5%	25.3%	1.7%	81.4%	1.0%	72.1%	36.3%	47.3%	0.4%	10.9%
4) & TPC-μ-Pull	112753	7755	69.7%	51.1%	4.8%	8.1%	2.7%	73.6%	0.6%	22.5%	22.1%	16.4%	0.1%	1.9%
5) & not TPC-e-Pull	98450	6771	72.9%	46.7%	4.9%	7.2%	2.8%	65.6%	0.4%	14.9%	18.9%	12.2%	0.1%	1.3%
6) & is negative	74918	5153	84.4%	41.1%	2.9%	3.3%	0.4%	7.5%	0.1%	3.7%	12.1%	5.9%	0.1%	0.8%
7) & start in FGD FV	69988	4813	88.1%	40.1%	3.0%	3.1%	0.4%	7.3%	0.1%	3.6%	8.2%	3.8%	0.1%	0.7%

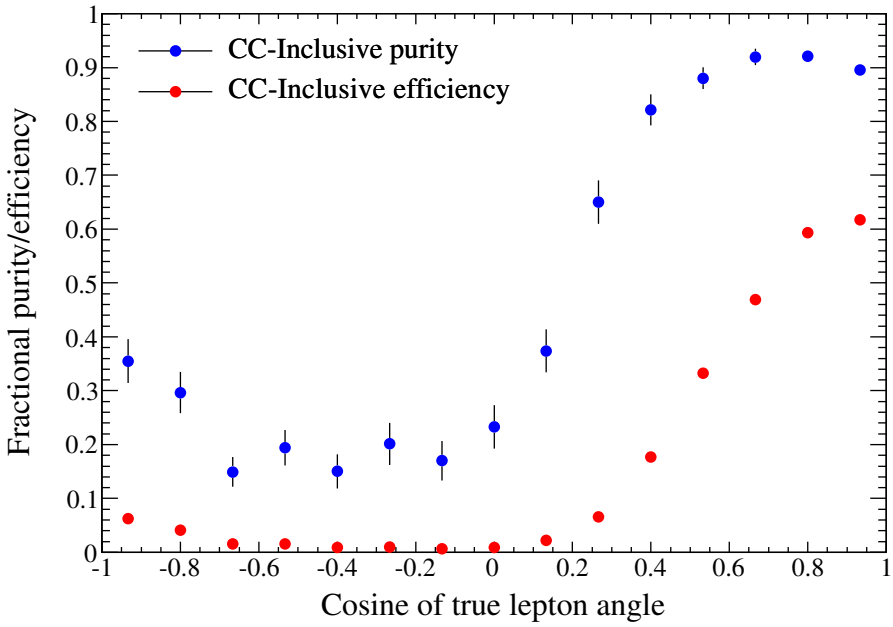
Table 3.4: ν_μ-CC-Inclusive purity and efficiency vs cut number as a function of CC–Inclusive process type. The efficiencies shown are absolute not relative efficiencies. The total number of MC events N_{Total} has also been scaled by a factor of 0.069 to give N_{Total}^{Scaled} which is normalised to the Run 2 data POT of 7.83×10^{19}. The truth category $\eta_{match} < 0.5$ is those events for which there was a less than 50% reconstruction to MC truth match. It is not possible to calculate absolute efficiencies for the Non-FGD and $\eta_{match} < 0.5$ categories and so by definition they are set to be 100% after the first cut.

In general the cuts behave as expected. For example, the efficiency for selecting the $\bar{\nu}_\mu$ background interactions, which if they are CC will produce a μ^+, drops from 65.6% to 7.5% when the negative charge cut is applied, and the background from $\nu_e/\bar{\nu}_e$ interactions is reduced from 72.1% to 14.9% when requiring that the track is μ-like and not e-like in cuts four and five.

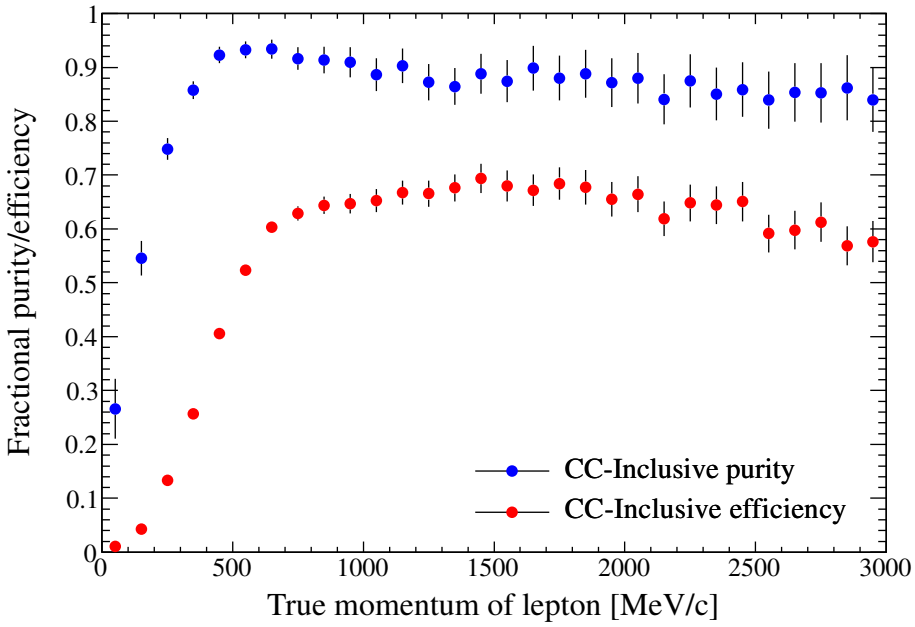
Fig. 3.13 shows how the purity and efficiency for selecting ν_μ-CC signal events varies as a function of the true outgoing muon angle and momentum. Looking at the efficiency curve in Fig. 3.13(a) we see that the selection has an acceptance which is strongly peaked in the forward direction. This is a key feature of all tracker-based CC selections requiring a TPC track to identify the lepton candidate. The dominant effect comes from the geometrical layout of the tracker region where at its most extreme we cannot reconstruct any events where the lepton travels perpendicular to the z direction, see the region $-0.2 < \cos(\theta) < 0.2$. For interactions with leptons travelling in the forward direction we see a steady decrease in efficiency between $\cos(\theta) = 0.8$ and 0.2. This is consistent with the degradation in performance of the TPC reconstruction as a function of track angle, see Fig. 2.16(a), as well as the fact that high angle tracks are less likely to be long enough to fulfil the TPC track quality cut of containing at least 19 hits. Looking at values of $\cos(\theta) < 0.5$ we can see that we are able to reconstruct some interactions where the lepton is travelling backwards but that the absolute efficiency is an order of magnitude less than in the forward direction. This reflects the a priori assumption in the current reconstruction algorithms that all tracks are travelling in the forward direction unless timing information for tracks passing through both FGDs indicates otherwise. Understanding and reducing this forward-peaked acceptance in tracker-based selections is important for T2K—the measurements at ND280 are used to predict the expected interaction rates at the far detector but the difference of the 4π angular acceptance of Super-Kamiokande and the forward-peaked acceptance of the current tracker-based analyses mean that the constraint from the near detector data covers only a subset of the full interaction kinematical phase space seen at the far detector.

Fig. 3.13(b) shows the efficiency for reconstructing ν_μ-CC signal events as a function of the true lepton momentum. We see a significant drop in efficiency below $\sim$ 750 MeV which is expected based on the angular acceptance discussed previously and the tendency of lower momentum muons to correspond to those which scatter through larger angles. For very low momenta the efficiency tends to zero due to the reduced path length for muons below about $\sim$ 200 MeV meaning that they will be contained in the FGD volume and not be able to make it into a TPC. At higher momenta the efficiency plateaus and then above $\sim$ 1500 MeV there is a slight decreases as a function of momentum, which is consistent with increased reconstructed momentum and charge identification uncertainty for higher momentum and hence straighter tracks.

In Fig. 3.14 we can see how the purity and efficiency for selecting ν_μ-CC signal events depends on the neutrino interaction kinematics. In Fig. 3.14(a) the efficiency can be seen to increase as a function of neutrino energy. This is consistent with the more forward-peaked angular distributions of outgoing muons at higher energies combined with the narrow



(a)



(b)

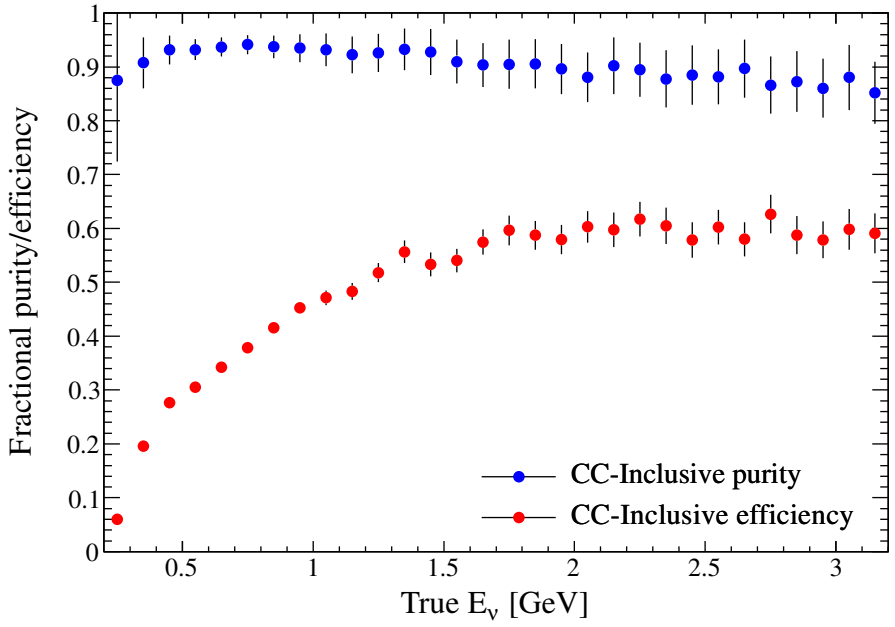
Figure 3.13: Performance of ν_μ-CC-Inclusive selection. Efficiency and purity are shown versus the cosine of true outgoing lepton's angle to the beam direction in (a) and verses the true momentum in (b). Care should be taken when interpreting the behaviour of the purity as a function of lepton kinematics as when calculating the purity the denominator includes both NC events, where the outgoing lepton is a ν, and Non-FGD background events, where the true vertex is outside of the FGD FV: For example, the low purity seen for events where the lepton is backwards-going, $\cos\theta < 0$, is dominated by Non-FGD interactions outside the FGD for which lepton is travelling backwards but the track is likely to have been reconstructed as forward-going.

angular acceptance of this selection. Fig. 3.14(b) shows that the purity and efficiency drop for higher values of Q^2, the square of the four-momentum transfer. This is consistent with the larger outgoing lepton scattering angles associated with high Q^2 events as well as the increased background from inelastic neutral current processes which turn on at higher Q^2 values.

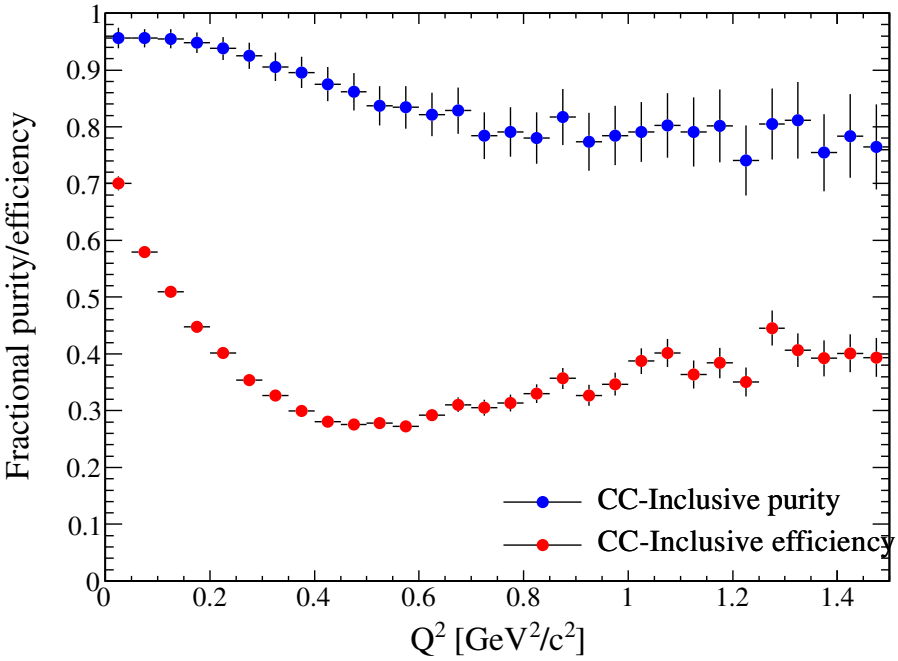
Fig. 3.15 shows a data-MC comparison for the reconstructed momentum and angle for the highest momentum μ-candidate track in events passing all cuts. The MC is normalised to the number of tracks in data. Although in this analysis we will only compare data and MC for the absolute number of events, it is still important to check quantities such as the momentum and angular distributions to demonstrate that the reconstruction is working as expected for both data and MC.

An example of a single-track Run 2 data event passing the ν_μ-CC-Inclusive cuts can be seen in Fig. 3.16. Here we see a muon-candidate starting in FGD1 and leaving a clear track of hits through all sub-detectors before exiting. The reconstruction has successfully matched the whole track which is typical of events where the muon is so forward-going. In Figs. 3.17 and 3.18 we can see examples of multi-track Run 2 data events passing the ν_μ-CC-Inclusive cuts, only tracks which are associated with the vertex based on the output of the Kalman filter are shown—in these cases the global vertexing seems to be working well. Fig. 3.17 is an example of a high multiplicity event in which three tracks pass through a significant length of the active TPC-volume. Fig. 3.18 shows a 2-track event—although only one track passes through a TPC the global reconstruction is able to successfully associate the ECal track with an FGD-only track and hence with the global vertex. We discuss these multi-track events more in the next section.

In this section we described the ν_μ-CC-Inclusive cuts and the checks performed to ensure that no major biases were being introduced by any of the inputs. Although not fully optimised, they still perform well with a purity of 88.1% and an efficiency of 40.1% for selecting ν_μ-CC interactions in the FGD FV. We now turn our attention to the second set of cuts which build upon these to make a more exclusive selection sensitive to neutrino interaction processes controlling π^+ production.

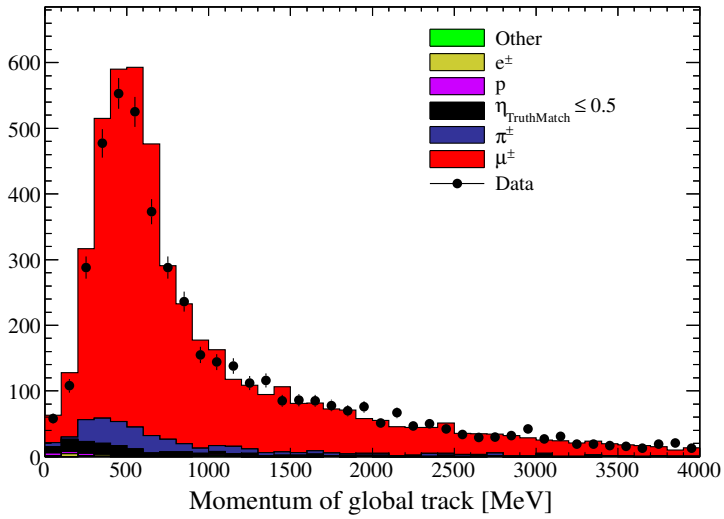


(a)

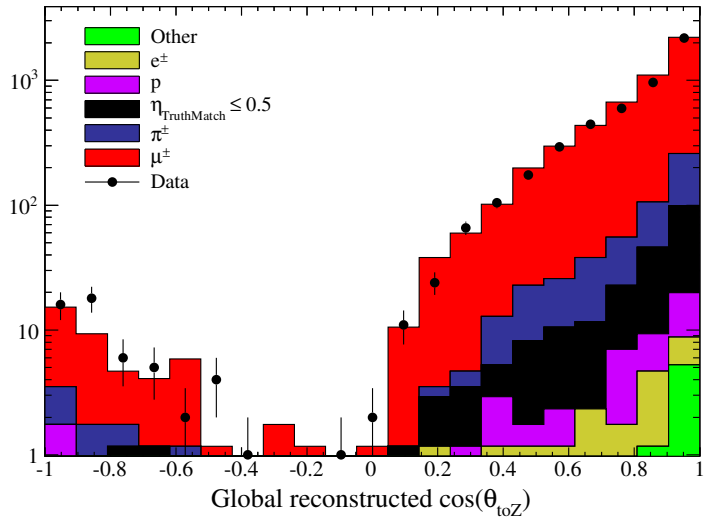


(b)

Figure 3.14: Purity and efficiency of the ν_μ-CC-Inclusive selection vs neutrino energy in (a) and the four-momentum transfer Q^2 in (b).

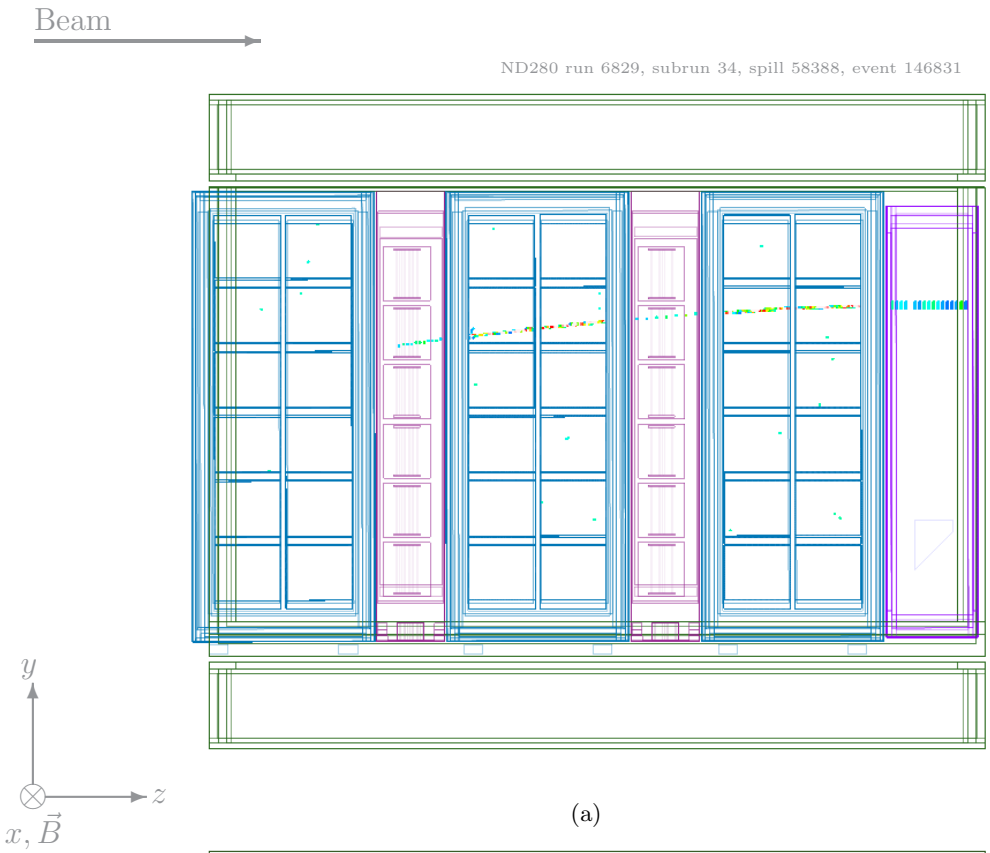


(a)

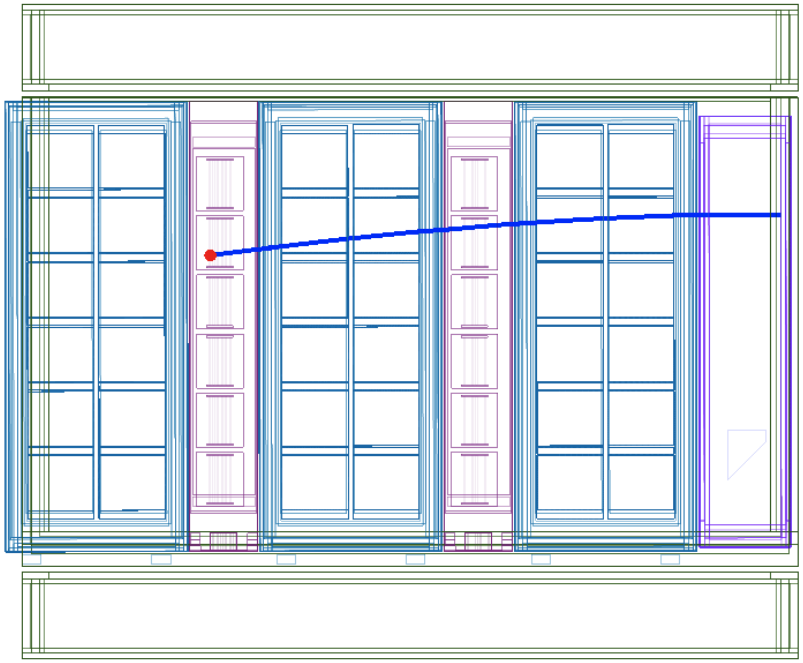


(b)

Figure 3.15: Data-MC comparison for momentum and angle of the highest-momentum μ^--candidate of events passing the ν_μ-CC-Inclusive cuts. The MC is normalised to the total number of tracks in data for each sample.

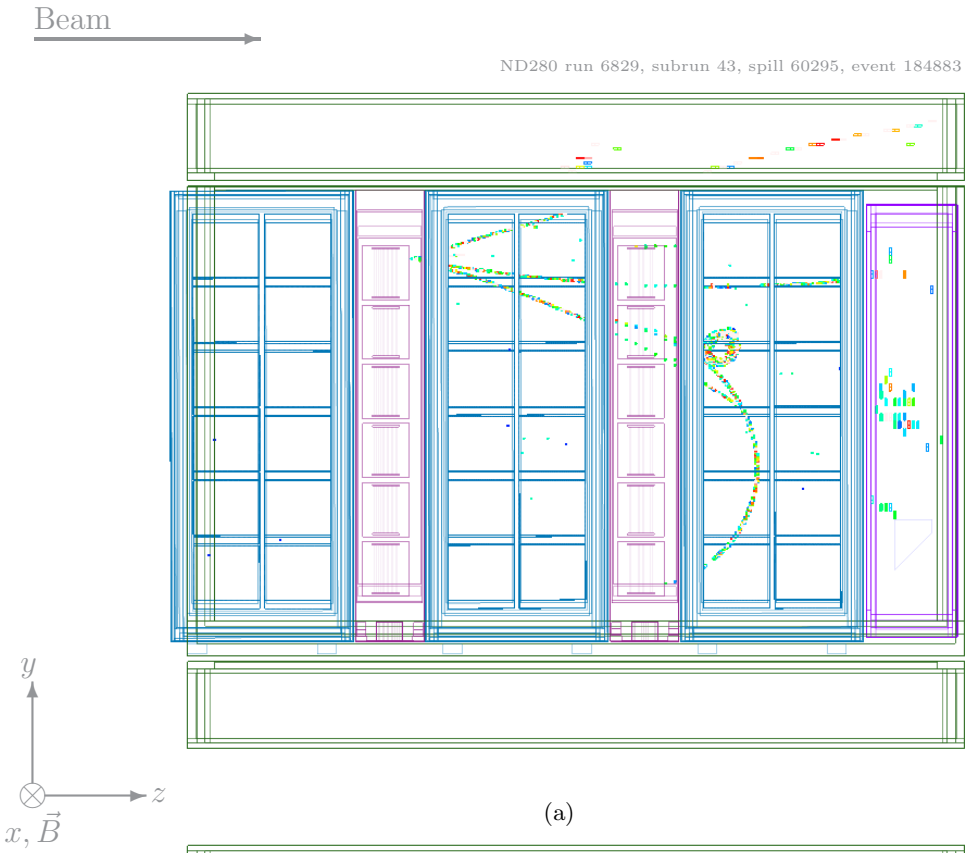


(a)

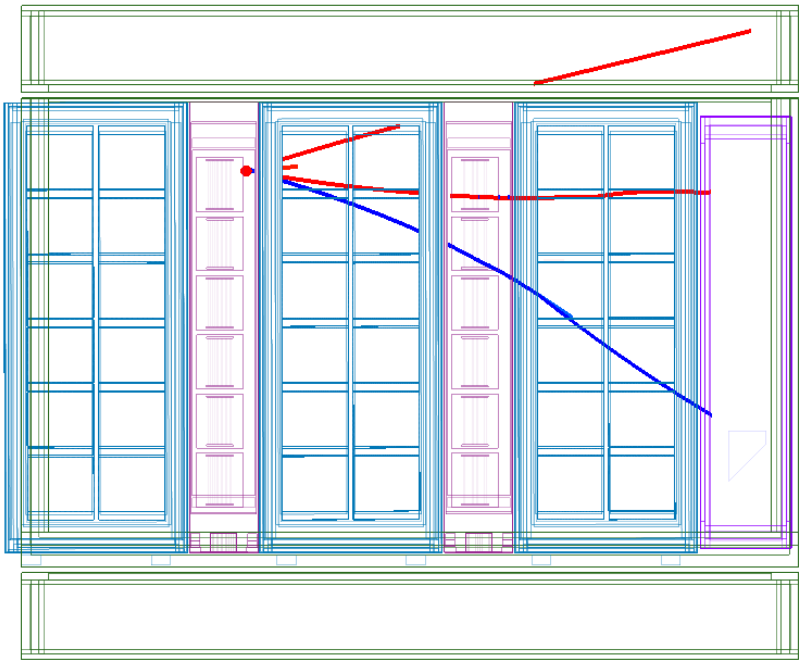


(b)

Figure 3.16: Event display showing a candidate single track FGD1 vertex from a Run 2 data event passing the ν_μ-CC-Inclusive cuts. The calibrated hits are shown in (a) and the global reconstructed track and vertex (red circle) are shown in (b). The global reconstruction is shown to have successfully matched a FGD1→TPC2→FGD2→TPC3→Ds-ECal track. This μ-candidate has a global reconstructed momentum of 1530 ± 100 MeV/c and a negative reconstructed charge (indicated by blue). It is forward-going and has a reconstructed opening angle of 12° corresponding to $\cos(\theta) = 0.98$. In addition to passing the TPC-based cuts used to identify it as a μ^--candidate the track also passes the ECal dQ/dL PID cuts developed to identify charged pions.

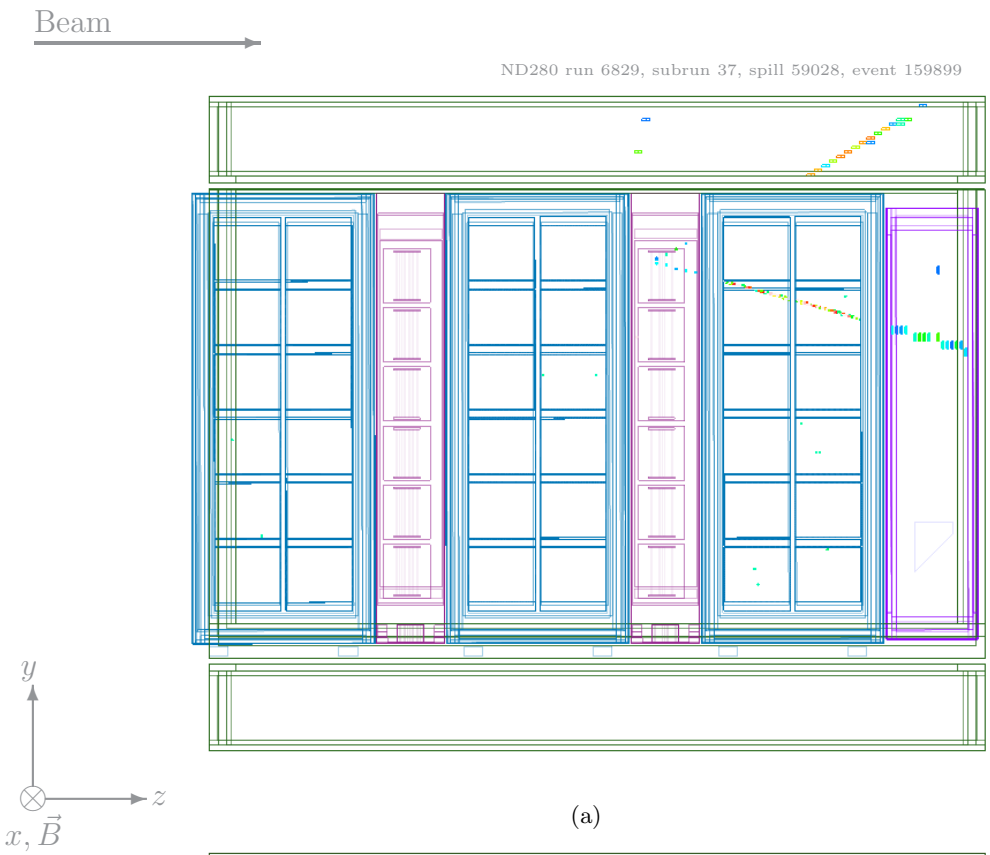


(a)

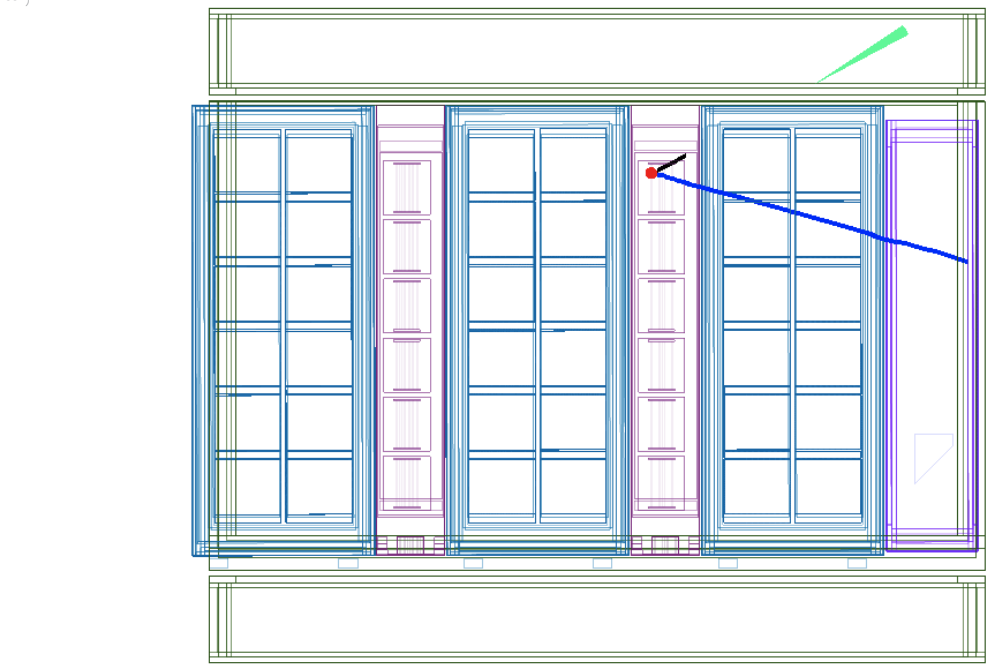


(b)

Figure 3.17: Event display showing a multi-track Run 2 data event passing the TPC based ν_μ-CC-$\pi^\pm$ cuts. The calibrated hits (a) show multiple tracks originating from a common vertex in FGD1 which is then successfully reconstructed in (b). We see that one negative (blue) and three positive (red) tracks are associated with the vertex (red circle), the former is the μ-candidate and has a global reconstructed momentum of 670 ± 40 MeV/c and an opening angle of 24°, corresponding to $\cos(\theta) = 0.93$. Both the long positive tracks pass the TPC based $\pi^\pm$ cuts indicating that this event was caused by an interaction with at least two π^+ in the final state. Due to a problem when displaying global tracks on the event display we only show the output of the individual tracker and sub-detector reconstruction.



(a)



(b)

Figure 3.18: Event display for Run 2 data event passing the ECal based ν_μ-CC-$\pi^\pm$ cuts. Calibrated hits are shown in (a) and the global reconstructed vertex and tracks in (b). The global reconstruction successfully matched the ECal (green) and FGD-only (black) sub-tracks and associated them with the negative μ-candidate (blue) when reconstructing the global vertex in FGD2 (red circle). The μ-candidate has a reconstructed momentum of 4680 ± 880 MeV/c and an opening angle of 16° ($\cos(\theta) = 0.96$). The FGD→Barrel-ECal $\pi^\pm$-candidate has a reconstructed outgoing angle of 29.3° ($\cos(\theta) = 0.87$) and passed the ECal *dQ/dLayer* cuts. This event is a good example of how the ECal based PID can be used to recover information for sideways-going tracks that miss the TPC active volume.

3.5 ν_μ-CC-$\pi^\pm$ Selections

The purpose of the ν_μ-CC-$\pi^\pm$ selections are to provide samples which are sensitive to neutrino interaction processes controlling π^+ production. As discussed in § 2.6.4, we aim for a relatively inclusive measurement where the signal is defined as any ν_μ-CC interaction with at least one $\pi^\pm$ in the final state. This inclusive approach should provide enough sensitivity to the neutrino interaction processes controlling π^+ production for interesting first data-MC comparisons to be made whilst also being robust against any as yet not understood detector systematics.

Here we describe two sets of cuts, which are applied to the output of the ν_μ-CC-Inclusive selection to tag at least one additional $\pi^\pm$ in the tracks associated with the vertex. The first set of cuts make use of the TPC PID, as in the previous section, to identify the $\pi^\pm$, and the second set uses a new ECal-based PID, using the deposited charge per unit length to identify the $\pi^\pm$. This amounts to selecting a vertex with at least two MIP-like tracks, the first being the highest-momentum negative μ^--candidate selected by the ν_μ-CC-Inclusive selection and the second being any other MIP-like track as identified by either the TPC or ECal and with no requirement on the charge. For these selections we only aim to tag the $\pi^\pm$s which behave as MIPs—this results in a loss in efficiency in cases where they decay or shower which, although not dealt with here, may be recoverable in future analyses. In the following description of the TPC- and ECal-based[6] cuts we label the tracks from events passing the ν_μ-CC-Inclusive but excluding the μ^--candidate as the secondary tracks.

3.5.1 TPC-Based Selection Cuts

To identify a $\pi^\pm$, the TPC PID is used in a similar way as in the ν_μ-CC-Inclusive cuts. The only change is that now we require that the secondary $\pi^\pm$-candidate has a pull that is inconsistent with that of a proton, rather than that of an electron, because for the secondary tracks associated with events passing the ν_μ-CC-Inclusive cuts the dominant background to secondary $\pi^\pm$ are protons. As with the ν_μ-CC-Inclusive selection these cuts are applied in a cumulative fashion:

8) $\geq$ 2 tracks: Require there to be least two tracks associated with the reconstructed vertex.

9a) $\geq$ 1 secondary track with a TPC constituent: Require that at least one of the secondary tracks has a constituent TPC sub-track.

10a) $\geq$ 1 secondary track with a good TPC quality: Of the tracks passing cut 9a, require that at least one of these has a constituent TPC sub-track with at least 19 TPC hits.

[6] It should be highlighted that although these are labelled as TPC- or ECal-based this refers only to the source of PID information and not to the sub-detectors contributing to the reconstructed tracks. For example, all of the tracks with ECal information will have been matched to either a TPC-track or an FGD-track.

11a) $\geq$ 1 secondary track with a TPC μ-like-pull: Of the tracks passing cut 10a, require that at least one has a TPC PID value consistent with the muon hypothesis, that the constituent TPC sub-track closest to the vertex has $|\text{Pull}_\mu| < 2.0$.

12a) $\geq$ 1 secondary track without a TPC p-like-pull: Of the tracks passing cut 11a, require that at least one has a TPC PID value which is inconsistent with the proton hypothesis, that the constituent TPC sub-track closest to the vertex has $|\text{Pull}_\text{p}| > 2.0$.

13a) $\geq$ 1 secondary track with positive charge: Of the tracks passing cut 12a, require that at least one has a positive reconstructed charge. This cut was an optional cut to see the how requiring a positive MIP affects the sensitivity to different cross section processes.

A number of checks were performed to ensure that no major biases were being introduced by any of the cuts or cut values used. Unless otherwise stated these checks are performed only for events passing the ν_μ-CC-Inclusive and only for the secondary tracks, where secondary is defined as all but the highest-momentum μ-candidate passing the ν_μ-CC-Inclusive track-level cuts.

Fig. 3.19 shows the TPC pull-μ and pull-p separately for each TPC. There is good agreement between data and MC for both TPC2 and TPC3. The statistics for TPC1 are very low but are consistent with the MC prediction. Apart from the overall number of tracks, the main difference between these and those shown in Fig. 3.9 is that having removed the μ-candidate we now see distributions dominated by $\pi^\pm$ and p tracks. The pull cut values chosen are the same as those used for the CC-Inclusive cuts.

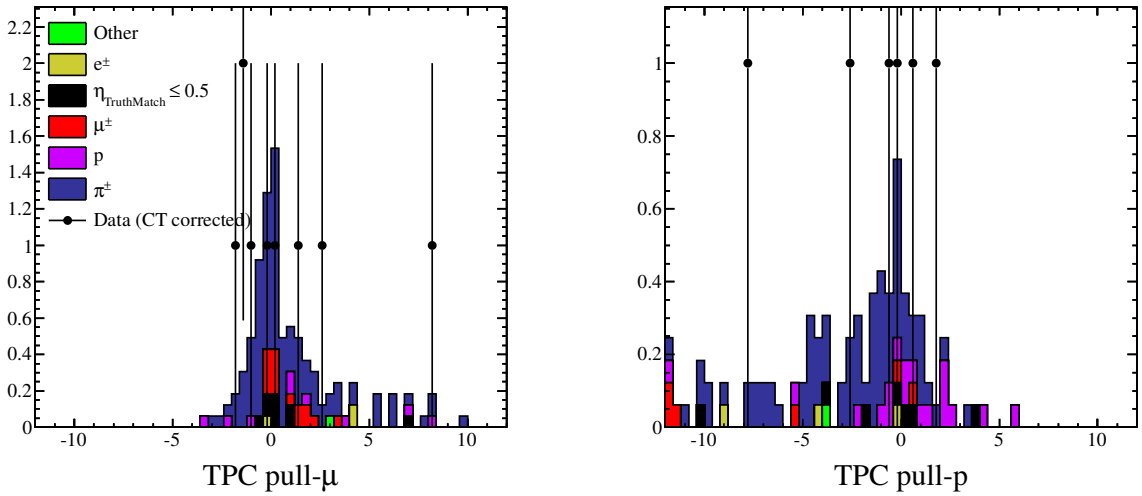
The reconstructed charge times momentum and track angle is shown for secondary tracks with a constituent TPC sub-track, passing data quality cuts, in Fig. 3.20. These are normalised to the total number of tracks in data and show good agreement between data and MC, indicating that these secondary tracks are still well modelled by the simulation.

3.5.2 ECal-Based Selection Cuts

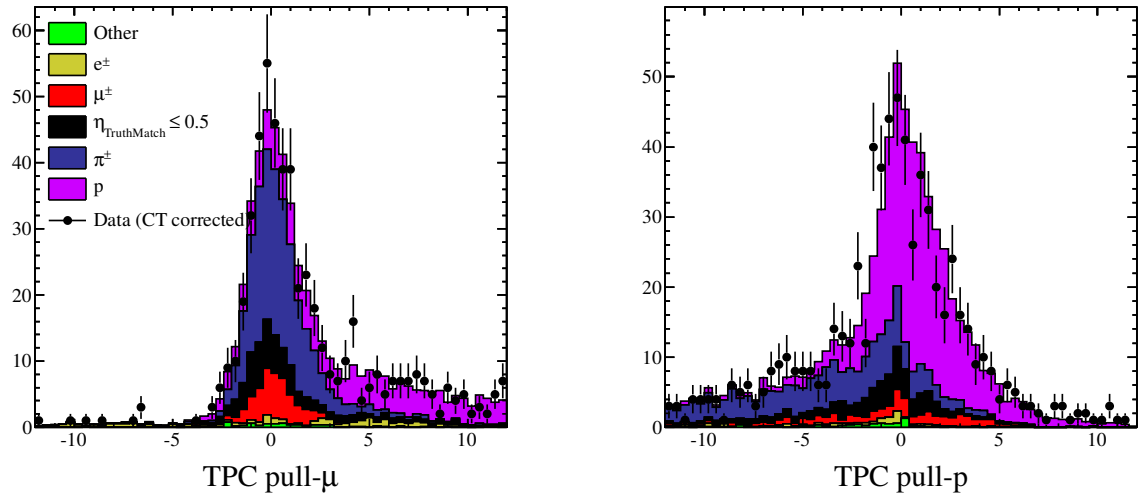
The methodology for applying the ECal-based cuts is the same as for the TPC-based ones—to identify a $\pi^\pm$-candidate amongst the secondary tracks of events passing the ν_μ-CC-Inclusive selection. Two ECal PID variables are used to identify the $\pi^\pm$-candidates:

- **TrShVal:** An existing PID which is the discriminator from a neutral network trained to separate tracks and showers in the ECal. We use it to select track-like objects[7].
- **dQ/dL:** A new variable based on the deposited charge per unit length of the track-like object. The sum of the charge associated with the ECal track is divided by the straight-line track length, calculated using the track incidence angle and the depth

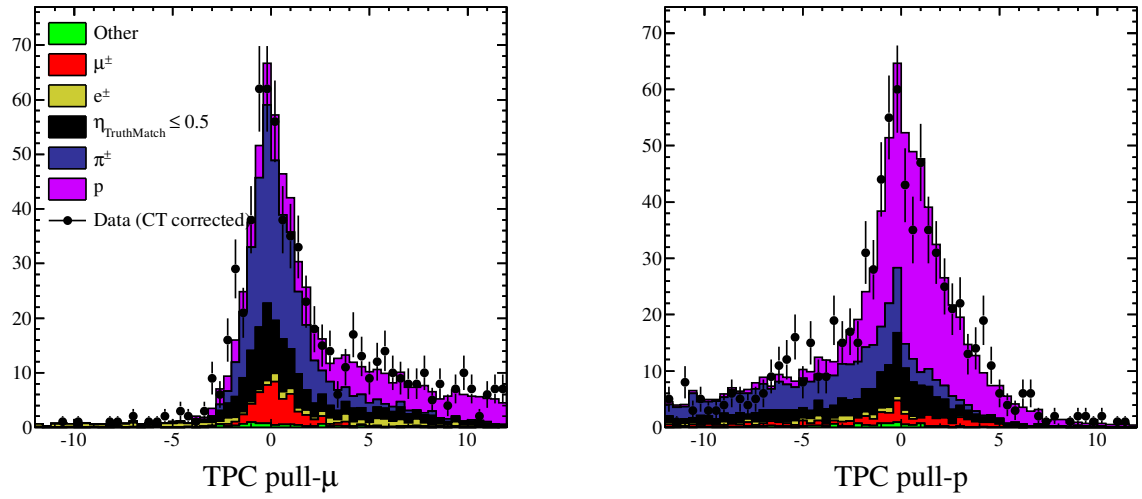
[7] As mentioned at the start of this section, we do not try to recover $\pi^\pm$s which decay or shower within the ECal.



(a)

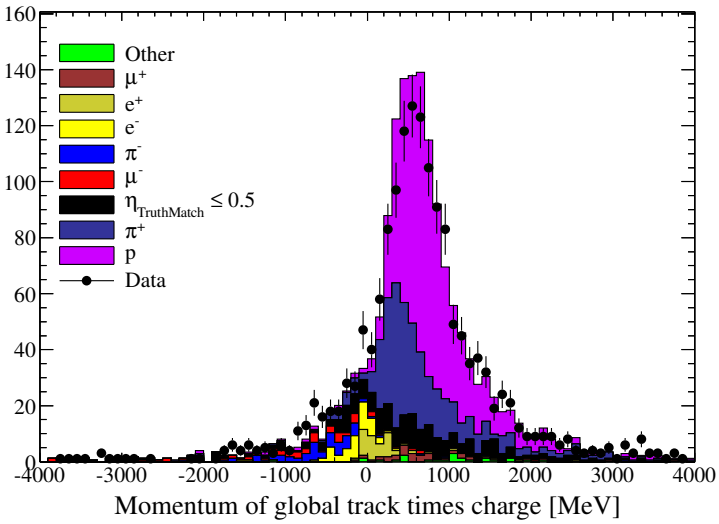


(b)

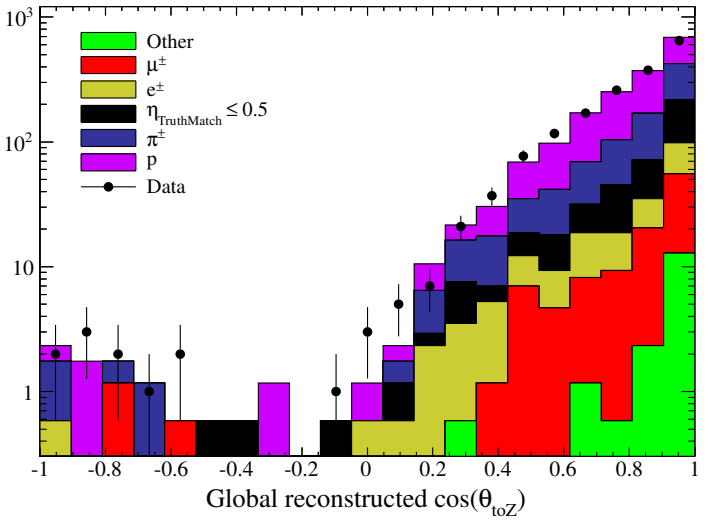


(c)

Figure 3.19: Data-MC comparisons for the TPC pull distributions for pull-μ and pull-p for all tracks except the highest-momentum μ-candidate of events passing the CC-Inclusive cuts. These distributions contain the tracks to which the ν_μ-CC-$\pi^\pm$ cuts will be applied. The pulls are shown separately for TPC1, TPC2 and TPC3 in (a), (b) and (c) respectively. The MC is normalised to the total number of tracks in data for each sample.



(a) Track momentum times charge



(b) Cosine of track angle

Figure 3.20: Data-MC comparison for reconstructed quantities of secondary tracks with TPC information from events passing the ν_μ-CC-Inclusive cuts. The MC is normalised to the total number of tracks in data for each sample.

of the track as worked out from the number of ECal layers hit. Placing a cut on the maximum deposited charge per unit length allows selection of MIP-like tracks whilst removing highly ionising tracks such as those created by protons. The final output of the ECal calibration is a charge in units of MEU where 1 MEU is defined as the charge deposited by a MIP-like particle with a hypothetical position next to the sensor and is extracted from fits to cosmic ray muon data.

We now present the full set of cuts which, as with the TPC-based cuts, are applied to only those events passing the ν_μ-CC-Inclusive selection and in a cumulative fashion:

8) $\geq$ 2 tracks: Require there to be least two tracks associated with the reconstructed vertex.

9b) $\geq$ 1 secondary track with ECal information: Require that at least one of the secondary tracks has a constituent ECal sub-track.

10b) $\geq$ 1 secondary track passing ECal TrShVal cut: Of the tracks passing cut 9b, require that at least one of these has an ECal TrShVal consistent with a track: TrShVal > 0.6.

11b) $\geq$ 1 secondary track passing ECal dQ/dL cut: Of the tracks passing cut 10b, require that at least one of these has an ECal charge deposit per unit length consistent with a MIP: $dQ/dL < 0.22$ [MEU/mm].

As with the TPC-based selection a number of checks were made to ensure that no major biases were introduced by the cuts or cut values used. These were performed separately for both the Ds- and the Barrel-ECal and with a particular emphasis on checking the new dQ/dL PID variable.

Figs. 3.21(a) and 3.21(b) show the TrShVal for secondary tracks of events passing the ν_μ-CC-Inclusive selection, where the MC is normalised to the number of tracks in data. In general, there is reasonably good agreement but there is a slight data excess for TrShVal < 0.5, in particular for the Barrel-ECal, indicating an excess of shower-like objects. This excess can be seen to be correlated with tracks with a shallow angle of incidence to the ECal detector surface in Figs. 3.21(c) and 3.21(d), where the cosine of the incidence angle of a track to the ECal detector face is shown. As expected the Ds-ECal sees a more forward-peaked distribution of incidence angles, because of its position downstream of the FGD FVs with its surface orientated perpendicular to the beam direction. In contrast, the Barrel-ECal sees a distribution peaked at around 45°, reflecting the fact that the Barrel-ECals surfaces lie approximately in line with the beam direction and to the side of the FGD FVs. Figs. 3.21(c) and 3.21(d) show that applying a TrShVal > 0.6 cut removes much of the data excess seen in the incidence angle distributions. For the remaining distributions we show in this section we will impose this TrShVal cut before making comparisons.

To calculate the dQ/dL PID quantity, we divide the total deposited charge for a track by its length. We use an extrapolated track[8] length based on the incidence angle of the track, θ, and the depth of the track obtained from the number of layers spanned multiplied by the layer thickness of 10 millimetres:

$$dQ/dL = \frac{Q_{\mathrm{Sum}}[\mathrm{MEU}]}{N_{\mathrm{Layers}} \times 10[\mathrm{mm}]/\cos\theta}.$$

Fig. 3.22 shows data-MC comparisons for the dQ/dL PID and the quantities which are used to construct it. Figs. 3.22(a) and 3.22(b) show a good agreement for the depth of the ECal track. These are given in terms of the number of ECal layers spanned, and the peaks correspond to the total number of layers in the Ds- and Barrel-ECal of 34 and 31 respectively. These combined with the angular distributions shown in Figs. 3.21(e) and 3.21(f) show that there is good agreement for the quantities used to extrapolate the track length. The total charge deposited by the track is shown for the Ds-ECal in Fig. 3.22(c). There is good agreement between data and MC for both the shape and the peak position of the distribution. The peak position will be affected by the overall energy scale of the ECal module and this agreement reflects the fact that, for the Ds-ECal, the MC energy scale was tuned directly to match that of the data. Fig. 3.22(d) shows the same but for tracks in the Barrel-ECal. Here there is a fairly large data-MC difference in the peak position of the order 20%. This is most likely due to the fact that for the production 4 MC the energy scale of each of the Barrel-ECal modules were not tuned to their individual responses as measured in data, but were given the same value as that used for the Ds-ECal. The width of the charge response for the Barrel-ECal is wider than that for the Ds-ECal because of the wider range of incidence angles, and hence the wider range of path-lengths through the scintillator, for particles originating in the FGD and entering the Barrel-ECal[9]. As will be discussed in § 3.6, a reasonably large systematic will have to be included to account for the data-MC discrepancies seen for the Barrel-ECal.

Data-MC comparisons for the actual dQ/dL PID quantity for the Ds- and Barrel-ECal are shown in Figs. 3.22(e) and 3.22(f) respectively. There is good qualitative agreement in the shape of both distributions, but for the Barrel-ECal the charge scale difference can still be seen in a misalignment of the peaks. The MC is broken down by true particle type where the more highly-ionising proton tracks, shown in purple, can be seen to dominate at values of $dQ/dL > 0.3$ [MEU/mm]. In order select a high purity sample of $\pi^{\pm}$, a cut value of $dQ/dL < 0.22$ [MEU/mm] was chosen. The purity, efficiency and significance for selecting ν_{μ}-CC-$\pi^{\pm}$ as well as the data MC ratio for the number of selected events are shown as a function of cut value in Fig. 3.23. A cut value was chosen which both gave a reasonably high significance whilst also lying in a region where the data MC ratio was not

[8] This was to avoid the use of the stored sub-track information which was filled incorrectly for the current processing.

[9] There is also a spread introduced by the fact that for the Barrel-ECal the short single-ended readout bars have a different response to the double-ended readout bars which is currently not corrected for in either data or MC.

changing rapidly. It is important to note that in order to avoid potential bias that could be introduced by choosing a particular cut value that gives a data MC ratio of 1.0, the data MC ratio was scaled so that the average value for all of the points was 1.0.

3.5.3 Performance of ν_μ-CC-$\pi^\pm$ Cuts

The performance of both the TPC- and ECal-based ν_μ-CC-$\pi^\pm$ cuts, based on their application to the full MC sample, is shown broken down by neutrino interaction final state pion topology in Tables 3.5 and 3.6 and by neutrino interaction process type in Table 3.7. These MC truth categorisations follow those set out in § 2.6.4. In each table, the purity and efficiency for selecting each different category is shown as a function of cut number. The rows of each table are broken into three distinct sections to make clear which cuts are common to both analyses, and which are specific to the TPC- or ECal based selections. Also shown, in the left two columns, are the total number of selected interactions based on the MC sample size and scaled to the expected POT of the data set.

First we look at the purity and efficiency for selecting events as categorised by final state pion topology. The final state topology defines what, in principle[10], the detector is able to see and is a better starting point for evaluating the performance. Table 3.5 shows that application of the TPC-based cuts 7 to 12a results in a final purity of 82.7% and efficiency of 6.9% for selecting the signal ν_μ-CC-$\pi^\pm$FS. In Fig. 3.17 we see an event display for a Run 2 data event that passed the TPC-based ν_μ-CC-$\pi^\pm$ cuts. In this event there was TPC information for three tracks—all three passed the MIP-like cuts indicating a multi-pion final state.

As we would expect, the most effective cut at selecting these multi-track events is simply to require that there is more than one track associated with the vertex. Requiring more than one track also reduces the Non-FGD contamination from 8.2% down to 4.4%, which reflects the dominance of the single track Non-FGD failure modes. Cuts 9a to 12a can be seen to reduce the contamination from final states where no secondary MIP was produced. In particular the background contamination from the ν_μ-CC-0πFS drops from 24.6% down to 1.5%. The dominant background now comes from the Non-ν_μ-CC-BG category. This category encompasses NC, $\bar{\nu}_\mu$ and $\nu_e/\bar{\nu}_e$ interactions which can all fake the ν_μ-CC-$\pi^\pm$ signal through inelastic processes producing high-multiplicity final states with at least two MIP-like particles in them[11].

After applying the ECal-based cuts 7 to 11b, a final purity of 69.7% and efficiency of 4.2% for selecting ν_μ-CC-$\pi^\pm$FS is reached. As with the TPC-based cuts, just requiring more than one track is the most effective cut for rejecting the non-ν_μ-CC-$\pi^\pm$ final states. Aside from this, the main increases in purity come from cuts 9b and 11b, where cut 9b does not discriminate based on the qualities of the track and reflects the tendency that MIP-

[10] This assumes a perfect detector and in reality only fairly long-lived particles above a certain momentum threshold can be resolved.

[11] This is even more true of charged current $\bar{\nu}_\mu$ events which in addition to the outgoing μ^+ would only require an additional negative MIP-like particle to be produced.

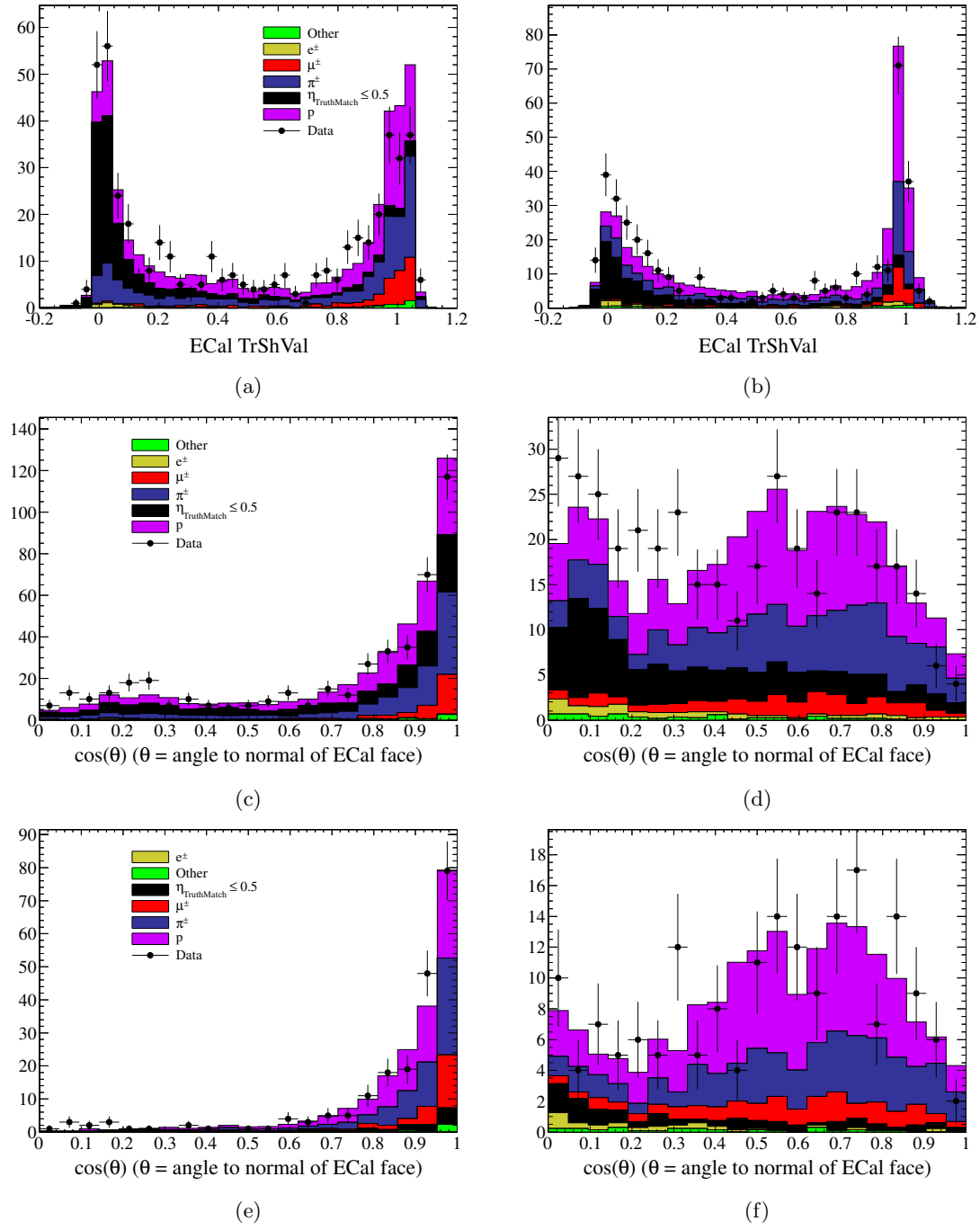


Figure 3.21: Data-MC checks for Ds-ECal (left) and Barrel-ECal (right) showing the ECal TrShVal PID and the cosine of the reconstructed angle to the normal of the ECal face. Only secondary tracks (not including the track identified as the muon candidate) from events passing the passing the ν_μ-CC-Inclusive selection are drawn. The MC is normalised to the total number of tracks in data for each sample. The TrShVal PID variable is shown in (a) and (b). The incidence angle is shown in (c) and (d) for all tracks, and just for tracks with a TrShVal > 0.6 in (e) and (f). Values of cosine close to 1 and 0 indicate tracks entering the ECal normal to and parallel with the inner surface respectively.

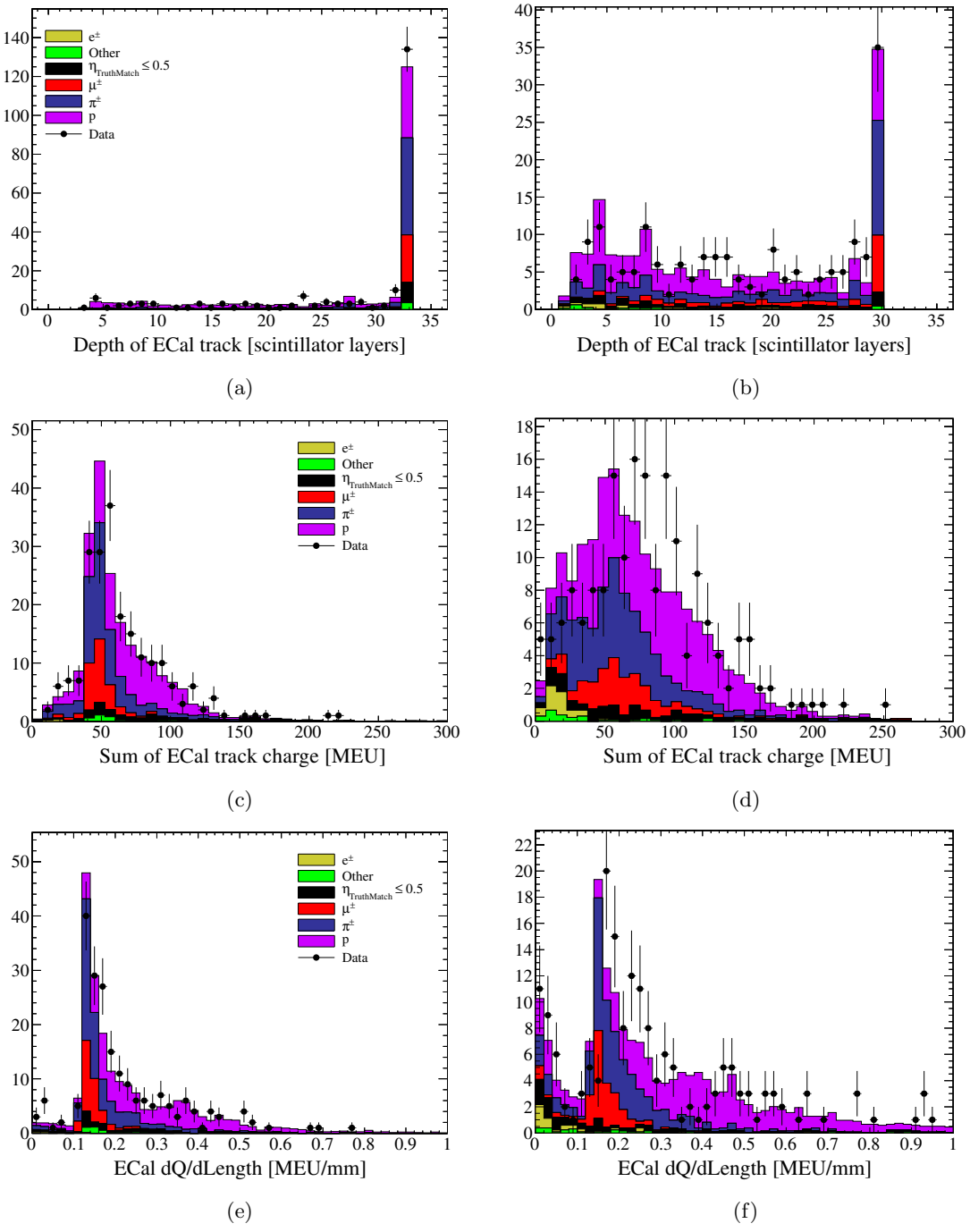


Figure 3.22: Data-MC checks of the dQ/dL ECal PID for the Ds-ECal (left) and Barrel-ECal (right). Only secondary tracks (not including the track identified as the muon candidate) from events passing the passing the ν_μ-CC-Inclusive selection are drawn. The MC is normalised to the total number of tracks in data for each sample. The depth of the ECal track, in terms of number of ECal layers traversed, is shown in (a) and (b) and the total charge deposited in the ECal for that track is shown in (c) and (d). These are then used, along with the track incidence angle, to divide the charge by the extrapolated track length to give the dQ/dL PID quantity shown in (e) and (f).

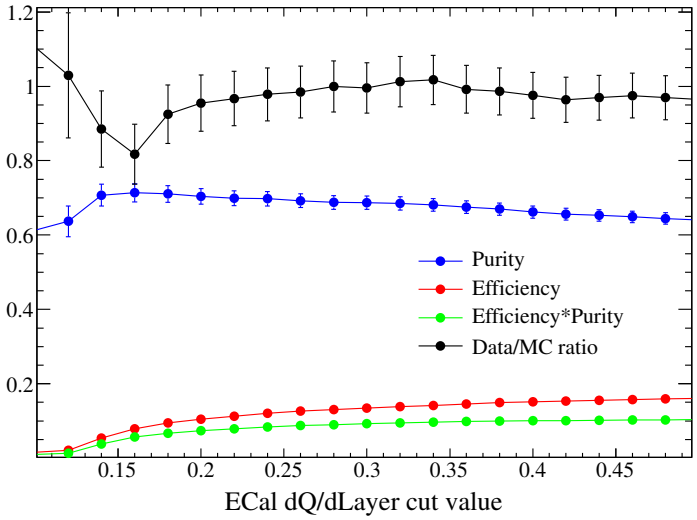


Figure 3.23: The effect of varying the ECal dQ/dL cut values for the ECal-based ν_μ-CC-$\pi^\pm$ selection. The purity, efficiency, significance (purity $\times$ efficiency) and data MC ratio of the number of selected events are shown as a function of cut value for the cut which excludes higher values of dQ/dL. The signal definition is a ν_μ-CC-$\pi^\pm$ final state interaction in an FGD FV. To avoid potential bias when tuning the cut value the data MC ratio was scaled so that the average for all points was 1.0.

like particles are more likely to make it to the ECal. By rejecting more highly-ionising tracks, cut 11b reduces the contamination from ν_μ-CC-0πFS events, where the second track is likely to be a proton. Although the TrkShVal cut does not improve the purity, it is important from the point of view of removing the not-yet-understood data-MC excess at low incidence angle, as discussed in the previous section.

Table 3.6 shows the quantities in Table 3.5 broken down as a function of final state π^+ topology. Here the signal is defined as any ν_μ-CC final state with at least one π^+, as opposed to requiring either a π^+ or a π^-. The fact that the final purities change so little, from 82.7%, 84.5% and 69.7% to 81.3%, 84.2% and 67.9% for the selected events passing cuts 12a, 13a and 11b respectively, shows that the $\pi^\pm$ category is dominated by final states with at least one π^+.

Table 3.7 shows the performance of the cuts broken down by the underlying neutrino interaction process. We see that the effect of the cuts is to reduce the contribution from ν_μ-CC-QEL processes from 48.2% (see first row) down to 1.6% and 5.5% for the TPC- and ECal-based cuts respectively. The final set of selected events for both the TPC- and ECal-based cuts are dominated by the ν_μ-CC-DIS (about 50%), ν_μ-CC-RES (30%), ν_μ-CC-COH (3%) and Non-ν_μ-CC (9%) interaction categories. The effect of applying the additional TPC-based charge cut 13a is to increase the relative fraction of selected events coming from ν_μ-CC-RES interactions, which is consistent with the dominance of $1\pi^+$ final states for ν_μ-CC-RES interactions.

Figs. 3.24 and 3.25 show how the purity and efficiency for selecting a ν_μ-CC-$\pi^\pm$ final

			ν_μ-CC-$\pi^\pm$FS		ν_μ-CC-0πFS		ν_μ-CC-π^0FS		ν_μ-CC-OtherFS		Non-FGD		Non-ν_μ-CC-BG	
	N_{Total}	N_{Total}^{Scaled}	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ
7) CC–Inclusive output	69988	4813	28.4%	38.5%	53.5%	41.3%	5.9%	37.5%	0.3%	43.0%	8.2%	3.8%	3.6%	3.1%
8) ≥ 2 tracks	23945	1647	50.6%	23.5%	29.0%	7.7%	8.9%	19.4%	0.7%	31.2%	4.4%	0.7%	6.4%	1.9%
TPC-based $\pi^\pm/\pi^+$ cuts														
9_a) & 2nd trk with TPC	18631	1281	53.8%	19.4%	24.6%	5.1%	9.6%	16.3%	0.9%	29.0%	3.6%	0.4%	7.5%	1.7%
10_a) & with good TPC	17777	1222	54.6%	18.8%	23.8%	4.7%	9.6%	15.5%	0.9%	28.6%	3.5%	0.4%	7.6%	1.7%
11_a) & TPC-μ-Pull	7775	534	75.6%	11.4%	4.4%	0.4%	5.0%	3.5%	1.2%	17.1%	3.7%	0.2%	10.1%	1.0%
12_a) & not TPC-p-Pull	5056	347	82.7%	8.1%	1.5%	0.1%	2.4%	1.1%	1.0%	8.7%	3.3%	0.1%	9.1%	0.6%
13_a) & is positive	4203	289	84.5%	6.9%	1.1%	0.1%	1.8%	0.7%	0.8%	6.4%	2.1%	0.1%	9.6%	0.5%
ECal-based $\pi^\pm$ cuts														
9_b) & 2nd trk ECal	10282	707	60.9%	12.1%	16.5%	1.9%	9.9%	9.3%	1.1%	20.1%	2.9%	0.2%	8.7%	1.1%
10_b) & 2nd trk TrkShVal	5525	380	61.1%	6.5%	18.4%	1.1%	7.7%	3.9%	1.2%	12.1%	2.5%	0.1%	9.0%	0.6%
11_b) & 2nd trk dQ/dL	3128	215	69.7%	4.2%	8.5%	0.3%	6.7%	1.9%	1.4%	7.7%	2.6%	0.1%	11.1%	0.4%

Table 3.5: ν_μ-CC-$\pi^\pm$ selection purity and efficiency versus cut number as function of neutrino interaction final state $\pi^\pm$ topology (MC truth) for the various TPC- and ECal-based π selections. The efficiencies shown are absolute not relative efficiencies. The total number of events has been scaled by a factor of 0.069 to match the data POT of 7.83×10^{19}. Cuts 9 and greater do not consider the track already selected as the muon candidate by the ν_μ-CC-Inclusive cuts, if more than one muon candidate survives cut 7 then only the highest momentum one is excluded.

			ν_μ-CC-π^+FS		ν_μ-CC-0πFS		ν_μ-CC-OtherπFS		ν_μ-CC-OtherFS		Non-FGD		Non-ν_μ-CC-BG	
	N_{Total}	N_{Total}^{Scaled}	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ
7) CC–Inclusive output	69988	4813	28.0%	38.4%	53.5%	41.3%	6.3%	38.0%	0.3%	43.0%	8.2%	3.8%	3.6%	3.1%
GDFV 8) ≥ 2 tracks	23945	1647	49.7%	23.4%	29.0%	7.7%	9.7%	20.0%	0.7%	31.2%	4.4%	0.7%	6.4%	1.9%
TPC-based $\pi^\pm/\pi^+$ cuts														
9_a) & 2nd trk with TPC	18631	1281	52.9%	19.4%	24.6%	5.1%	10.5%	16.8%	0.9%	29.0%	3.6%	0.4%	7.5%	1.7%
10_a) & with good TPCC	17777	1222	53.6%	18.7%	23.8%	4.7%	10.6%	16.1%	0.9%	28.6%	3.5%	0.4%	7.6%	1.7%
11_a) & TPC-μ-Pull	7775	534	74.3%	11.3%	4.4%	0.4%	6.3%	4.2%	1.2%	17.1%	3.7%	0.2%	10.1%	1.0%
12_a) & not TPC-p-Pull	5056	347	81.3%	8.1%	1.5%	0.1%	3.8%	1.7%	1.0%	8.7%	3.3%	0.1%	9.1%	0.6%
13_a) & is positive	4203	289	84.2%	7.0%	1.1%	0.1%	2.2%	0.8%	0.8%	6.4%	2.1%	0.1%	9.6%	0.5%
ECal-based $\pi^\pm$ cuts														
9_b) & 2nd trk ECal	10282	707	59.7%	12.0%	16.5%	1.9%	11.1%	9.8%	1.1%	20.1%	2.9%	0.2%	8.7%	1.1%
10_b) & 2nd trk TrkShVal	5525	380	59.7%	6.5%	18.4%	1.1%	9.1%	4.3%	1.2%	12.1%	2.5%	0.1%	9.0%	0.6%
11_b) & 2nd trk dQ/dL	3128	215	67.9%	4.2%	8.5%	0.3%	8.6%	2.3%	1.4%	7.7%	2.6%	0.1%	11.1%	0.4%

Table 3.6: ν_μ-CC-$\pi^\pm$ selection purity and efficiency versus cut number as function of neutrino interaction final state π^+ topology (as opposed to $\pi^\pm$ topology). The efficiencies shown are absolute not relative efficiencies. The total number of events has been scaled by a factor of 0.069 to match the data POT of 7.83×10^{19}. The purities for ν_μ-CC-π^+ final states for the 12a, 13a and 11b cuts are 81.3%, 84.2% and 67.9% respectively. These are very close to the equivalent purities for the ν_μ-CC-$\pi^\pm$ final states shown in Table 3.5. This indicates that despite the looser definition the ν_μ-CC-$\pi^\pm$ final state is made up mostly of final states with at least one π^+ in them.

			ν_μ-CC-QEL		ν_μ-CC-RES		ν_μ-CC-DIS		ν_μ-CC-COH		Non-FGD		Non-ν_μ-CC	
	N_{Total}	N_{Total}^{Scaled}	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ	η	ϵ
7) CC–Inclusive output	69988	4813	48.2%	41.6%	22.7%	36.3%	16.3%	41.1%	0.9%	57.7%	8.2%	3.8%	3.6%	3.1%
8) ≥ 2 tracks	23945	1647	23.0%	6.8%	34.2%	18.7%	30.5%	26.3%	1.6%	35.2%	4.4%	0.7%	6.4%	1.9%
TPC based $\pi^\pm/\pi^+$ cuts														
9_a) & 2nd trk with TPC	18631	1281	18.8%	4.3%	33.7%	14.3%	34.7%	23.3%	1.7%	29.1%	3.6%	0.4%	7.5%	1.7%
10_a) & with good TPC	17777	1222	17.9%	3.9%	33.9%	13.7%	35.4%	22.7%	1.7%	28.2%	3.5%	0.4%	7.6%	1.7%
11_a) & TPC-μ-Pull	7775	534	3.4%	0.3%	26.7%	4.7%	53.3%	14.9%	2.8%	20.7%	3.7%	0.2%	10.1%	1.0%
12_a) & not TPC-p-Pull	5056	347	1.6%	0.1%	31.8%	3.7%	50.8%	9.2%	3.5%	16.5%	3.3%	0.1%	9.1%	0.6%
13_a) & is positive	4203	289	1.2%	0.1%	35.1%	3.4%	47.9%	7.3%	4.0%	16.1%	2.1%	0.1%	9.6%	0.5%
ECal based $\pi^\pm$ cuts														
8_b) & 2nd trk ECal	10282	707	12.2%	1.5%	31.6%	7.4%	43.0%	15.9%	1.6%	15.4%	2.9%	0.2%	8.7%	1.1%
10_b) & 2nd trk TrkShVal	5525	380	13.8%	0.9%	32.2%	4.1%	40.9%	8.1%	1.6%	8.3%	2.5%	0.1%	9.0%	0.6%
11_b) & 2nd trk dQ/dL	3128	215	5.5%	0.2%	26.3%	1.9%	52.4%	5.9%	2.0%	5.9%	2.6%	0.1%	11.1%	0.4%

Table 3.7: Breakdown by neutrino interaction process type (MC truth) of events as a function of cut number for the various TPC and ECal based π selections. The total number of events has been scaled by a factor of 0.069 to match the data POT of 7.83×10^{19}. Cuts 9 and greater do not consider the track already selected as the muon candidate by the ν_μ-CC-Inclusive cuts, if more than one muon candidate survives cut 7 then only the highest momentum one is excluded.

state vary as a function of the neutrino interaction kinematics, for both the TPC- and ECal-based cuts separately and also for a combined selection that requires either of the TPC- or ECal-based cuts to be passed.

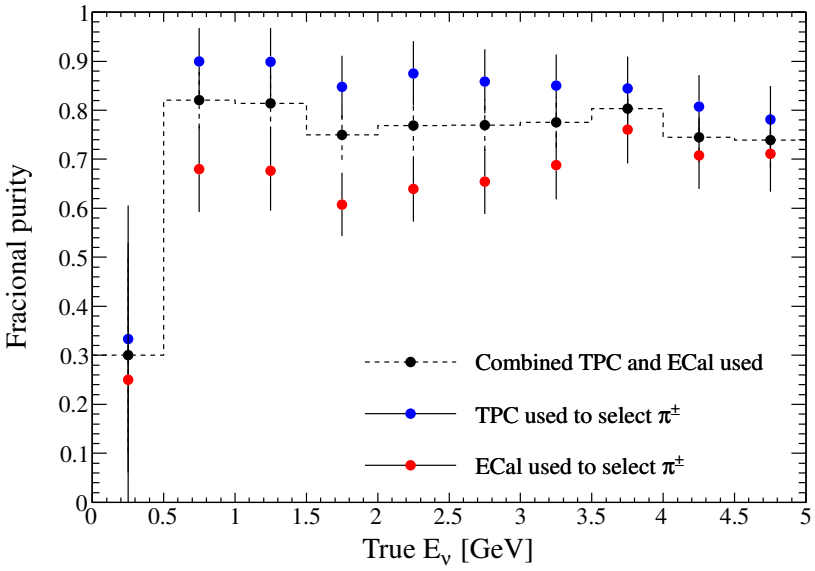
In Fig. 3.24(a) the purity is shown versus neutrino energy. As with the CC-Inclusive selection, the purity here is approximately independent of neutrino energy with the exception that for energies of less than about 0.5 GeV the relative fraction of ν_μ-CC-$\pi^\pm$ final states produced is very low, meaning that the purity tends to much lower values. As we would expect the purity for the combined sample lies between that for the two independent selections. Fig. 3.24(b) shows that the efficiency increases as a function of neutrino energy. The combined selection yields a significantly higher efficiency for selecting the ν_μ-CC-$\pi^\pm$ final states which is important as it demonstrates that the ECal-based cuts are not just selecting a subset of the events passing the TPC-based cuts.

The behaviour of the purity and efficiency as a function of the four-momentum transfer Q^2 is shown in Figs. 3.25(a) and 3.25(b) respectively. The purity is slightly anti-correlated with Q^2, this is most likely due to the higher fraction of background Non-ν_μ-CC events expected for higher momentum transfer. The efficiency increases as a function of Q^2, this is the opposite behaviour to that shown for the ν_μ-CC-Inclusive selection in Fig. 3.14(b): In the case of the ν_μ-CC-Inclusive selection, all that is required is that the μ^- track is identified by the TPC, which means that low-Q^2 interactions, where the lepton keeps most of its momentum and is more likely to pass through a TPC volume, have a higher efficiency. The converse is true in the case of the ν_μ-CC-$\pi^\pm$ cuts where both the μ^- and a secondary $\pi^\pm$ need to be identified, meaning that enough momentum has to be transferred to the $\pi^\pm$ for it to make it into a TPC or the ECal, resulting in an increase[12] in efficiency with Q^2.

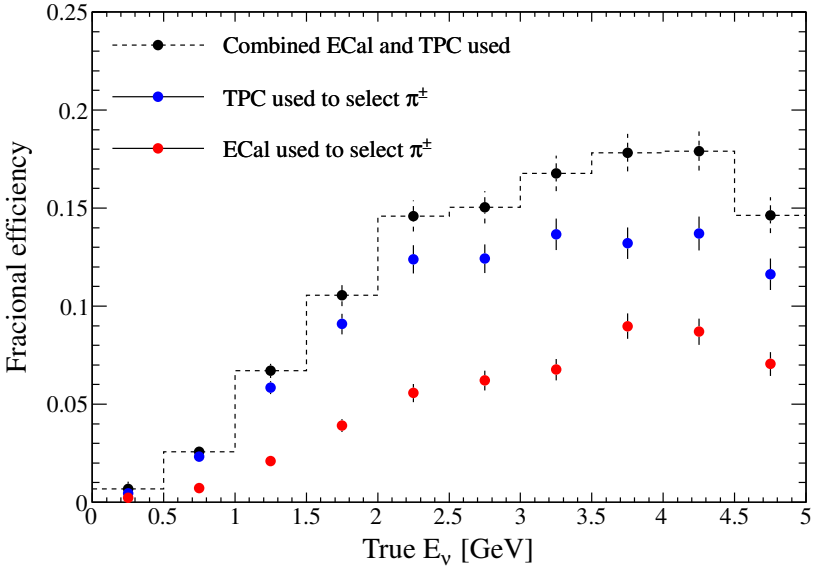
One of the motivations for developing the ECal-based $\pi^\pm$ selection was that the Barrel- and Ds-ECal together provide a much larger geometrical acceptance for particles originating from the FGD than that the TPC alone. Fig. 3.26(a) shows the cosine of the reconstructed angle to the z-direction, θ, for the $\pi^\pm$ candidates of events passing the various selections. The TPC-based selection reconstructs many more π candidates in the forward- and backwards-going regions but the ECal-based cuts can be seen to reconstruct slightly more sideways-going $\pi^\pm$-candidates with $0 < \cos\theta < 0.4$.

Fig. 3.18 shows a Run 2 data event where the ECal-cuts have allowed selection of a high-angle pion. The effect of this on the types of $\pi^\pm$ interactions selected can be seen in Fig. 3.26(b). Here the efficiency for selecting ν_μ-CC-$\pi^\pm$ final states is plotted versus the cosine of the most collinear outgoing $\pi^\pm$, based on MC truth information from the neutrino interaction. A value of $\cos\theta = 0$ means that all the $\pi^\pm$s produced by the neutrino interaction have travelled out at right angles to the beam direction, whereas a value of $1/\sqrt{2}$ would indicate that at least one $\pi^\pm$ was produced with an outgoing angle within 45°. For the TPC-based cuts, the efficiency is very forward-peaked and, as we would

[12]The actual dependence on Q^2 is evidently more complicated than this, as illustrated by a second minimum in Fig. 3.25(b).

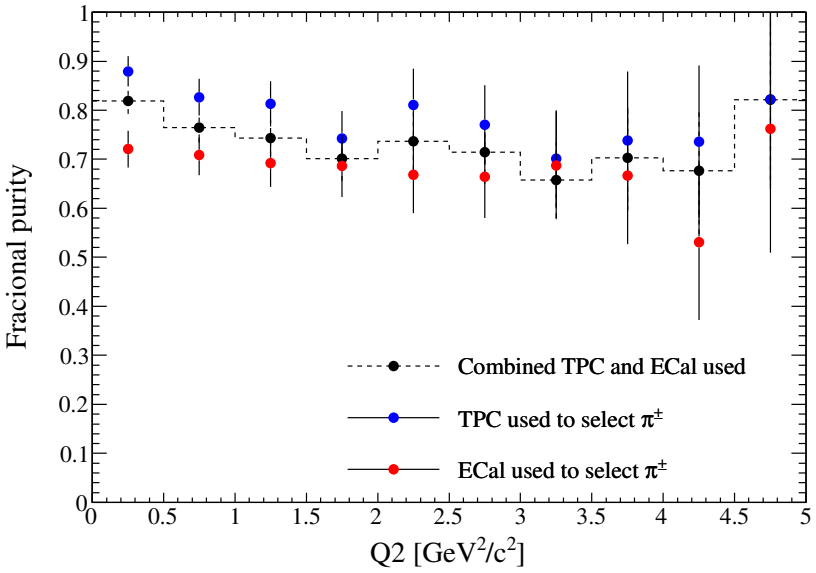


(a)

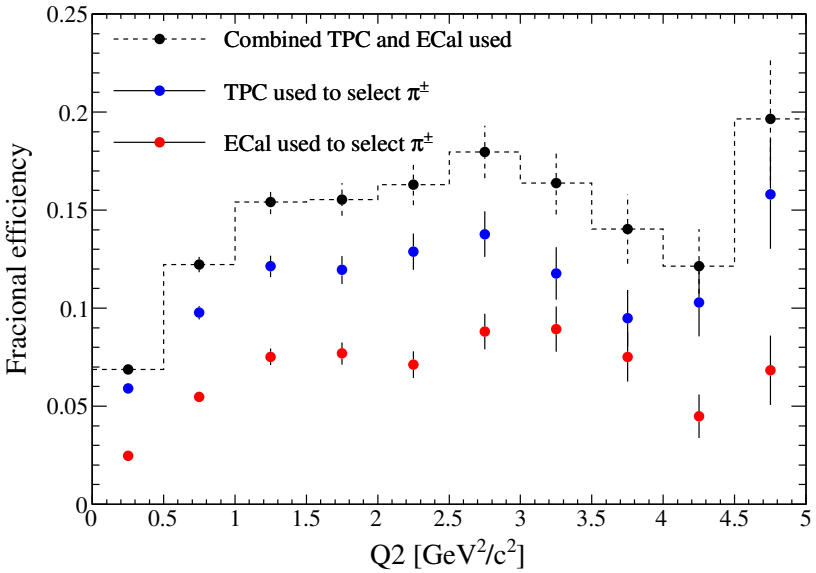


(b)

Figure 3.24: Performance of ν_μ-CC-$\pi^\pm$ selection versus neutrino energy. Purity (a) and efficiency (b) for selecting ν_μ-CC-$\pi^\pm$ events is shown for the TPC- and ECal-based selection separately as well as for a combined selection using both TPC and ECal information to tag the $\pi^\pm$.

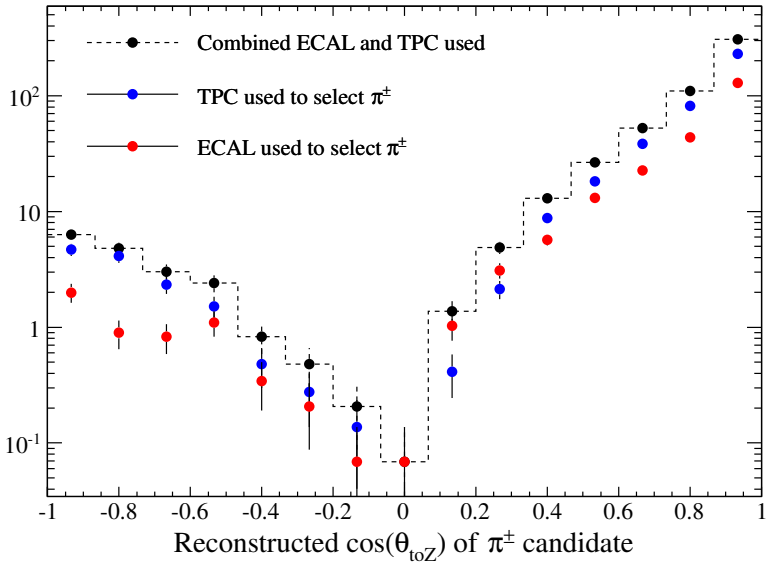


(a)

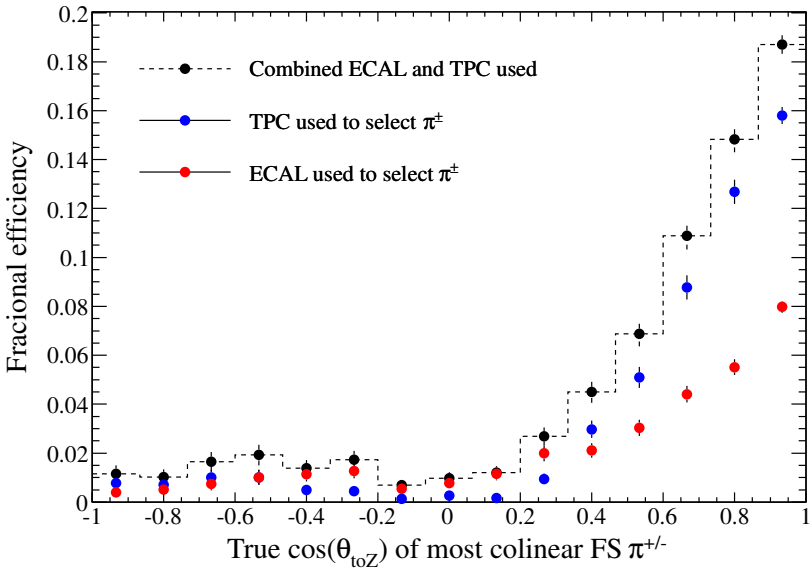


(b)

Figure 3.25: Performance of ν_μ-CC-$\pi^\pm$ selection versus the 4-momentum transfer Q^2 of the interaction. Purity (a) and efficiency (b) for selecting ν_μ-CC-$\pi^\pm$ events is shown for the TPC- and ECal-based selection separately as well as for a combined selection using both TPC and ECal information to tag the $\pi^\pm$.



(a)



(b)

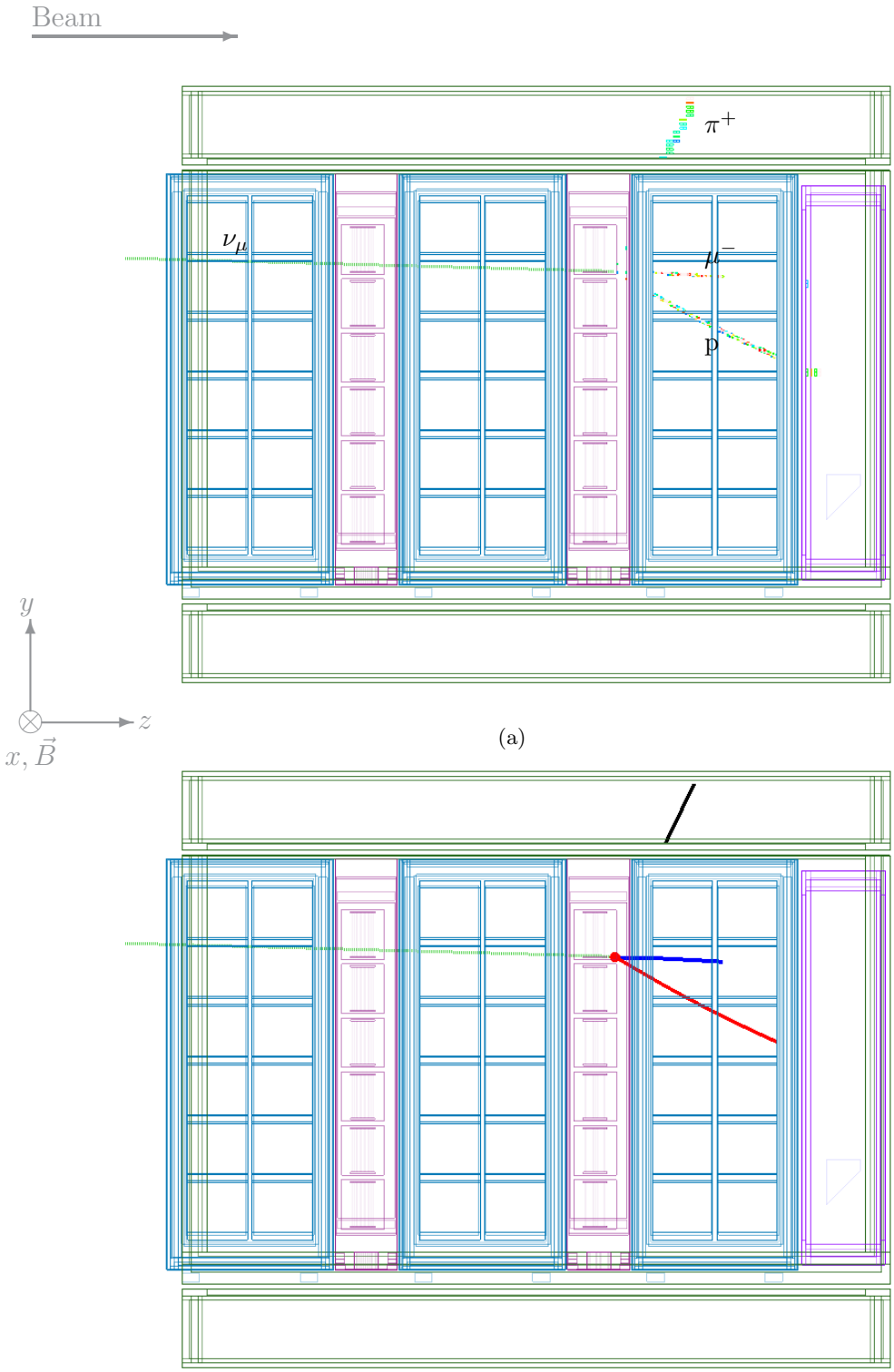
Figure 3.26: Number of selected events and efficiency for ν_μ-CC-$\pi^\pm$ selections as a function of reconstructed and true $\pi^\pm$ kinematics. Shown for the TPC- and ECal-based selections separately, as well as for a combined selection using both TPC and ECal information to tag the $\pi^\pm$. In (a) the angle used is not the reconstructed angle based on the global track direction, but rather the angle made by joining a straight line from the track end closest to the vertex to the end furthest from the vertex.

expect for neutrino interactions where all the $\pi^{\pm}$s were produced at right-angles to the z-axis, tends to 0 as $\cos\theta \to 0$. The behaviour for the ECal-based cuts, although with a lower overall efficiency, is less forward-peaked and has non-zero values in the region of $\cos\theta = 0$. This indicates that by using the ECal-based cuts it is possible to recover some of the interactions with sideways-going $\pi^{\pm}$s. In principle this efficiency could be greatly increased as the current reconstruction is not optimised for matching ECal to FGD tracks—an event display for such a matching failure is shown in Fig. 3.27. This shows a three track MC event for a ν_μ-CC-1π^+ interaction in FGD2. The μ and p both travel through a TPC active volume and are identified as coming from a common vertex by the Kalman filter. The ECal track from the π^+ is not associated with the vertex and so from the point of view of the selection this now looks like a 2-track CC-QEL interaction with just a μ and a p in the final state. This shows how increasing the acceptance of the selection not only leads to an increase in the efficiency for selecting particular final states but also reduces background coming from events which would be classified incorrectly due to missed tracks. Given the Run 2 statistics and the low absolute efficiencies shown here, this extra geometrical acceptance is unlikely to have much of an effect on the selection and serves more as a demonstration, that in principle, the ECal can be used to open up the acceptance.

In conclusion, the TPC- and ECal-based cuts perform well and are able to select ν_μ-CC-$\pi^{\pm}$ interactions with a purity of 82.7% and 69.7% respectively, providing two samples which are sensitive to the neutrino interaction processes controlling $\pi^{\pm}$ production and which, for the identification of the $\pi^{\pm}$, are sensitive to different detector systematics. Although in this analysis the ECal-based cuts will be used to actively include MIP-like tracks, they could also be used in other analyses to veto events with additional MIPs in the final state, such as for those trying to isolate the ν_μ-CC-0π final state characterised by a single μ^- plus nucleons.

3.6 Detector and Reconstruction Systematics

In this section we discuss the detector and reconstruction systematics for the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^{\pm}$ selections just described. As mentioned previously, much of the work towards understanding and quantifying these systematics, and in particular of those relevant to this analysis, is taking part within the official NuMu physics working group. Where possible we try to draw from the output of these studies but it should be emphasised that, as many of them are still ongoing, we make use of preliminary results and, in some cases, resort to using conservative estimates. First we will discuss the methods and values used to estimate the systematics considered, with a separate heading for each, and then in § 3.6.6 we present the results of these as applied to the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^{\pm}$ selections.



(a)

(b)

Figure 3.27: Event display for a true FGD2 ν_μ-CC-1π^+ interaction that did not pass the ECal based ν_μ-CC-$\pi^\pm$ cuts. In (a) the calibrated hits show the expected 3-pronged hit distribution from the μ, π^+ and p (see truth labels). A dashed line indicates the path of the incoming ν_μ. In (b) we see that the TPC successfully identifies the μ (blue) and rejects the p (red) as MIP-like. The Barrel-ECal track was identified as MIP-like by the ECal-PID but was not associated with the global vertex. The large opening angle of the outgoing π^+ (60° or $\cos(\theta) = 0.5$) is typical for these types of FGD-ECal matching failures—improving the reconstruction such that these can be matched to tracks or vertices in the FGD will be important to increase the angular acceptance for the tracker-based selections.

3.6.1 Sub-detector and Global Reconstruction Efficiencies

As described in § 2.6, the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ selections are based on the output of the global reconstruction, which itself uses the output of the tracker- and various sub-detector reconstructions. Any differences between the response to data and MC at any of these levels can introduce a systematic uncertainty to the final selected number of events.

TPC-tracking-eff

First we consider the effect of the TPC sub-detector tracking efficiency. We take directly the output of studies [83] within the NuMu group which have measured a TPC tracking efficiency for both data and MC for all three TPCs. For each TPC, a control sample of through-going muon-like tracks is used to measure the success rate for reconstructing a track with at least 18 hits. For a given TPC, a though-going track is defined as one for which there are at least two reference tracks in the other TPCs (or TPC2 and the P0D when measuring the TPC1 tracking efficiency) which both extrapolate as a contained[13] track into the current TPC, track *a* in Fig. 2.10 is an example of such a track. The reference tracks are also required to contain at least 60 hits and have an energy deposit consistent with the muon hypothesis ($|\text{Pull}_\mu| < 2.5$). Using this method, tracking efficiencies consistent with 99.7% for both data and MC were measured for all TPCs and taking into account the statistical uncertainties of the measurement, they agree to within $\pm 0.5\%$[14]. To evaluate the systematic we run a set of toy MC experiments where the efficiency for reconstructing a TPC track is changed by $\pm 0.5\%$ to see the effect on the total number of selected events. The toy experiment consists of simulating a reduction in the tracking efficiency by randomly discarding 0.5% of tracks with TPC information. This is repeated many times using different random seeds and the mean deviation on the number of expected events for all toy experiments is taken as the systematic. It is not possible to simulate an increase in the efficiency using this technique but at the level of this study it is reasonable to assume that the effect is symmetrical.

TPC-FGD-matching-eff

We also consider the TPC-FGD matching efficiency for the tracker reconstruction described in § 2.6. There are two failure modes for TPC-FGD matching:

- TPC tracks that fail to match to any of the hits in the FGD will lead to an overall decrease in the track finding efficiency for tracks starting in the FGD and travelling through a TPC. As all the selections we have developed rely on a matched TPC-FGD track to identify a μ-candidate starting in the FGD FV, this translates directly to a systematic on the total number of events passing the selection.

[13] A contained track is defined as one which enters and exits through the front and back x-y surface of the TPC.

[14] Because of limited statistics a single number is used for all values of momentum and track angles.

- TPC tracks that fail to match to the complete set of the hits in the FGD can cause the track start point in the FGD to migrate into or out of the FGD FV. This can increase the Non-FGD background component. For example, if a track passed through an FGD but the TPC failed to match the hits in the first few upstream layers then the start point of the track would appear to be a few layers into the FGD and within the FV.

For now we look at the first of these failure modes and will come back to the effect due to the second in § 3.6.4. Any difference between the FGD-TPC matching efficiency in data and MC will lead to a systematic. We take the results of a study [83] which used through-going muons (TPC1-FGD1-TPC2-FGD2-TPC3) produced upstream of the tracker and high angle FGD-triggered[15] cosmic muons traversing a single TPC and FGD to measure the TPC-FGD tracking efficiency as a function of momentum and angle for data and MC. In general the agreement was at the sub-percent level and we use a single value of $\pm 0.4\%$ for all momenta and angles. It is important to note that this study was performed for FGD1 only and that we use the same number for FGD2. The effect of this on the final selections is calculated using toy experiments in a similar manner to that for TPC-efficiency systematic except that now we throw away 0.4% of FGD-TPC matched tracks. As before, the systematic is assumed to be symmetric.

TPC(FGD)-ECal-matching-eff

For the ECal-based $\pi^{\pm}$ selection, we also consider a systematic introduced due to differences in the TPC-ECal matching efficiency performed at the global reconstruction level. For this we use the results from a preliminary study performed within the UK ECal group [84] where the combined efficiency for reconstructing and matching a track in the Ds-ECal, given a μ-like track in the TPC that extrapolates into the Ds-ECal, was measured as a function of track momentum for data and MC and shown to agree to within $\pm 15\%$ and $\pm 4\%$ for tracks with reconstructed momentum below and above 150 MeV respectively, where we have taken an average result for the relevant momentum bins. Unlike the previous toy experiments to estimate the effect of this on the selection output we do not completely discard the track when simulating the decrease in efficiency but instead just remove the ECal-PID information from the event. Because of this this systematic will only affect the ECal-based selection of the secondary $\pi^{\pm}$-candidates. It is important to note that although this study was only performed for the Ds-ECal here we assign the same systematic for tracks entering the Barrel-ECal, this would need to be addressed before this analysis could be used in any official results.

[15] The FGD trigger requires hits within the FGD and this, combined with the extrapolation of the TPC track, leads to an expectation of a reconstructed TPC-FGD track.

Global-matching-eff

Failures at the tracker- and global-reconstruction level where tracks fail to be successfully matched between sub-detectors can cause lost tracks which, if there is a difference in the rate of failures between data and MC, can cause a systematic that affects the overall selection efficiency. Preliminary results from studies within the NuMu group [85], which use a similar technique (based on through-going spill muons) as the other matching systematic studies, have shown that the efficiency between data and MC for the global matching efficiency is the same to within ±1%. Unfortunately these studies were halted at a preliminary stage because the official NuMu spectral analysis stopped using the output of the global reconstruction directly. As we are using the output of the global vertexing we do not have this choice and make use of these preliminary numbers. Typically these matching failures occur at the Kalman filter refit of the matched tracks and not at the initial matching stage. Unfortunately, a bug was found for global tracks containing an SMRD component which effectively increases this data MC difference to ±10% for any global tracks containing an SMRD constituent. The effect of these on the selections is evaluated in the same manner as for the previous systematics except that here we throw away 10% of global tracks with an SMRD constituent and 1% for all others.

3.6.2 Charge Confusion and Track Directionality

When identifying the μ-candidate track in the ν_μ-CC-Inclusive selection, both the charge and direction of the global track are used to select negative MIPs starting in the FGD FV. The version of the global reconstruction used in this analysis assumes a priori that all tracks travel in the downstream direction unless timing information for tracks traversing both FGDs indicates otherwise, as described in § 2.3.4 if $\Delta t_{\mathrm{FGD}} < 3$ns where $\Delta t_{\mathrm{FGD}} = t_{\mathrm{FGD2}} - t_{\mathrm{FGD1}}$ then the track direction is flipped (reversed). There is a finite probability for wrong track inversion where a true forward-going particle is flipped so that its direction is backwards.

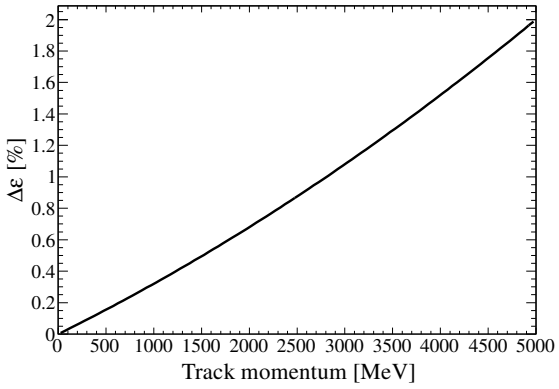
Track-direction

Studies within the NuMu group [86] using through-going spill muons from neutrino interactions upstream of the tracker have measured the fraction of wrong inversions R for both data and MC to be $R_{Data} = 0.82\% \pm 0.2\%$ and $R_{MC} = 0.49\% \pm 0.1\%$. Using the difference between these we extract a conservative estimate on the uncertainty on the fraction of wrong inversions in MC of $\Delta R_{MC} = \pm|R_{Data} - R_{MC}| = \pm 0.33\%$. To simulate the effect of this data MC difference, we run a set of toy MC experiments where we randomly correct the direction of tracks which were wrongly inverted 67% ($= 0.33/0.49 \approx \Delta R_{MC}/R_{MC}$) of the time. This corresponds to reducing the fraction of wrong track inversions, and we assume symmetric errors to account for the case of increasing the fraction of wrong inversions. In addition to this, and as a conservative estimate, we also correct the direction for the same fraction (67%) of backwards-going tracks that were incorrectly reconstructed

as forward-going, i.e., they did not have their direction flipped. This is performed in the same manner as for forward-going tracks.

Charge-confusion

The likelihood of charge confusion causing the charge of a track to be incorrectly reconstructed depends on the track momentum, the track length in the TPC and on the track direction inversion described above. Studies within the NuMu group [87] using samples of through-going muons have shown that for long tracks (> 40 TPC hits) with momentum around 1 GeV/c, the rate of charge confusion ϵ is about 2% for both data and MC. These studies also provided the uncertainty in the agreement between data and MC, $\Delta\epsilon$, as a function of track momentum. We use a parametrisation of this uncertainty, shown in Fig. 3.28, to vary the rate of charge confusion in MC for a set of toy experiments. Here we increase the rate of charge confusion by randomly flipping the charge of the tracks[16] with a probability equal to $\Delta\epsilon(P)$. As with the other systematics we assume the effect on the final selections to be symmetric.



(a)

Figure 3.28: Parametrisation of uncertainty in the fraction of charge confusion as a function of track momentum. As expected the rate of charge confusion is larger for high momentum, and hence low curvature, tracks.

3.6.3 PID Cuts

TPC-Pull-uncertainties

The TPC-pull cuts are central to the selections we have developed. As described in § 2.3.5 the energy loss measurement used to calculate the TPC pulls is based on a truncated charge mean (CT) per unit track length. During data taking the gains of the TPCs are monitored and used to apply a correction to the CT energy loss measurements. By varying the CT corrections according to the uncertainty on their measured values we can quantify

[16] We only flip the charge of tracks that have the correct charge to begin with so as not to cancel with the original charge confusion.

the systematic associated with the TPC pulls. We perform many toy MC experiments where the CT correction for each TPC is varied according to a random number drawn from a Gaussian distribution with the width of the uncertainty on the CT correction. This essentially corresponds to varying the response of the TPCs randomly and by an amount consistent with the uncertainty on their measured values. The analysis is then re-run using TPC-pull values that are recalculated to account for the change in the CT correction. For each selection we perform 50 such toy experiments and plot the distribution of the total number of events passing all cuts. The RMS of this distribution is taken as the size of the systematic error where, to account for the possibility of asymmetric errors, we use the mean of the distribution μ, which can differ to the nominal value, and take the upper and lower bound of the systematic error to be $\mu \pm \text{RMS}$.

ECal-charge-scale

For the ECal-based PIDs we only consider the systematic introduced by the dQ/dL ECal-PID cut as the systematic introduced by the TrShVal cut is expected to be small in comparison. Based on the level of agreement between data and MC for the sum of the ECal-track charge shown in Figs. 3.22(c) and 3.22(d) we introduce shifts in the charge scale of $\pm 5\%$ for the Ds-ECal and $+15-5\%$ for the Barrel-ECal. The selections are then re-run but using the modified dQ/dL to see how introducing these shifts in the underlying charge distribution affect the total number of events passing each selection.

3.6.4 Additional Systematics

Non-FGD-background

Non-FGD background is defined as any reconstructed vertex within the FGD FV where the actual interaction vertex occurred outside the FV. As described in § 3.4.1, there are many possible failure modes which can cause this to happen and in general this is known to be a hard systematic to evaluate. Preliminary studies [83] into the effect of reconstruction and detector systematics on the Non-FGD background indicate that changes of the order $\pm 10\%$ can be introduced to the total rate of Non-FGD events. As a conservative estimate we apply variations of $\pm 20\%$ to the rates of Non-FGD events[17].

An additional problem we face in evaluating the systematic associated with Non-FGD events is that for the current MC data set, neutrino interactions were only simulated within the Magnet geometry and not for the surrounding cavern or material upstream of the detector, where, as the J-PARC research complex is situated on the coast, this is mostly sand. This means that at present there is no way to estimate the fraction of muons from interactions upstream of ND280 in the cavern and sand (known as sand-muons) that contribute to the Non-FGD background. Within the ND280 working groups an effort is currently under way to simulate these interactions but due to computational

[17] This is only applied to Non-FGD events where the vertex is actually outside of the FGD and not for those where it is outside the FGD FV but still inside the active FGD.

difficulties associated with simulation over such a large volume this is still ongoing. We do not consider this systematic in the analysis and it would need to be addressed before this analysis could be used for any official results.

Target-mass-uncertainty

The final systematic we consider is that due to the uncertainty in the FGD target masses. Any difference between the measured and actual target mass will manifest as a systematic in the absolute rate of neutrino interactions seen. A detailed study [88] that calculated the expected weights of the FGD scintillator modules and then compared them to measured values of both the final machined modules and the raw scintillator bars showed consistency to within $\pm 0.67\%$. We use this directly as a systematic on the event rate from true vertices occurring in either FGD1 or FGD2.

3.6.5 Systematics not Considered

We now list some of the sources of systematic that are not considered in this analysis:

- No systematic is calculated to account for possible biases caused by the global vertexing algorithm. However, these are thought to be small based on the good data-agreement shown in § 3.3. The agreement shown in Table 3.2 indicates that these biases would be at the percent-level.

- No systematic is assigned for the TPC quality cut as previous studies [89] have shown that it is likely to be at the sub-percent level and because the data MC ratio on the number of selected events shown in § 3.4 is very stable as a function of cut values within the region we use.

- No systematic was assigned based on the difference between the simulated and measured magnetic field. Because we are essentially performing a counting experiment to first order we are not sensitive to overall shifts in the magnetic field. However, both the extrapolation during global track matching and refitting and TPC-pulls will be affected by both overall scale changes and local variations in the magnetic field strength. Before this analysis could be used for any official results this would need to be addressed.

- No systematic was assigned for the inter-detector timing calibration. This is justified because, with the exception of the track direction reversal based on the FGD times, we have not used any outputs that would be sensitive to the level of timing differences seen between various sub-detectors.

- Similarly, no systematic has been assigned based on differences between the measured and actual geometry alignment. These are expected to be small and to some extent have already been folded into the tracking efficiency studies.

3.6.6 Effect on Selections

Table 3.8 shows how the systematics discussed in this section affect the total number of events passing all cuts for each of the selections we have developed. Each row shows the upper and lower fractional change introduced by a given systematic for all four selections. The final row shows the total error calculated by adding these in quadrature and this is the value that will be used when comparing data to MC for the selections.

The dominant terms for the ν_μ-CC-Inclusive selection come from the global matching efficiency, track-direction, TPC-pull and Non-FGD background resulting in a total systematic of $+3.09-3.08\%$. For the TPC-based ν_μ-CC-$\pi^\pm$ selections, we see that the contributions from the TPC related efficiency and pull systematics are relatively high as they have an effect on the identification of both the μ- and $\pi^\pm$-candidate, giving a final systematic on the total number of selected events of $+4.20-4.26\%$ and $+4.47-4.52\%$ for the $\pi^\pm$- and π^+ selections respectively. The ECal-based ν_μ-CC-$\pi^\pm$ can be seen to be dominated by the ECal and global matching efficiencies and by the ECal charge scale uncertainty used by the dQ/dL PID cut, resulting in a total systematic uncertainty of $+7.50-10.16\%$. The contribution from the Non-FGD background is reduced for both the TPC- and ECal-based ν_μ-CC-$\pi^\pm$ selections and this reflects the dominance of single-track failure modes contributing to Non-FGD events.

	$\delta N^{exp}/N^{exp}$ for selection							
Systematic	ν_μ-CC-Inclusive		ν_μ-CC-$\pi^\pm$ (TPC)		ν_μ-CC-π^+ (TPC)		ν_μ-CC-$\pi^\pm$ (ECal)	
TPC-tracking-eff	+0.50%	−0.50%	+0.95%	−0.95%	+1.01%	−1.01%	+0.79%	−0.79%
TPC-FGD-matching-eff	+0.40%	−0.40%	+0.76%	−0.76%	+0.81%	−0.81%	+0.60%	−0.60%
TPC(FGD)-ECal-matching-eff	+0.00%	−0.00%	+0.00%	−0.00%	+0.00%	−0.00%	+4.18%	−4.18%
Global-matching-eff	+2.00%	−2.00%	+3.59%	−3.59%	+3.65%	−3.65%	+3.95%	−3.95%
Track-direction	+1.31%	−1.31%	+0.11%	−0.11%	+0.04%	−0.04%	+0.30%	−0.30%
Charge-confusion	+0.37%	−0.37%	+0.45%	−0.45%	+0.86%	−0.86%	+0.27%	−0.27%
TPC-Pull-uncertainties	+1.16%	−1.14%	+1.54%	−1.69%	+1.92%	−2.04%	+1.00%	−0.87%
ECal-charge-scale	+0.00%	−0.00%	+0.00%	−0.00%	+0.00%	−0.00%	+4.52%	−8.23%
Non-FGD-background	+1.26%	−1.26%	+0.51%	−0.51%	+0.27%	−0.27%	+0.33%	−0.33%
Target-mass-uncertainty	+0.62%	−0.62%	+0.63%	−0.63%	+0.64%	−0.64%	+0.64%	−0.64%
Total	+3.09%	−3.08%	+4.20%	−4.26%	+4.47%	−4.52%	+7.50%	−10.16%

Table 3.8: Effect of detector systematics on the expected number of events passing each selection. The fractional change due to each detector systematic is shown for each of the selections in the four columns and the final row shows the sum in quadrature of all systematics.

4 Results

In this chapter we present and discuss the final data-MC comparisons of the event rates for the ν_μ-CC-Inclusive and the ν_μ-CC-$\pi^\pm$ selections developed in the previous section. Before doing this we consider the systematics introduced by the neutrino interaction and flux simulation.

4.1 Physics Simulation Systematics

Event reweighting methods are used to propagate the uncertainties in the input parameters for both the interaction and flux simulation to see the effect on the final observables, in our case the total rate of selected neutrino interactions. This is done via the T2KReWeight software framework which we will now briefly describe.

4.1.1 T2KReWeight

The motivation for using an event reweighting scheme to propagate systematics for the neutrino interaction and flux simulation is that it allows the effect of changing the input physics parameters to be studied without having to re-run the full MC chain for each change in input parameter. Consider the full MC chain for simulating neutrino interactions at ND280:

1) Flux simulation (JNUBEAM).

2) Neutrino interaction generators (GENIE or NEUT).

3) GEANT4-based detector simulation.

4) Sub-detector and global reconstruction (full production output of $\sim$ 6 TB).

5) Analysis (total oaAnalysis reduced output $\sim$ 350 GB).

Changing the physics parameters at the first or second stage would mean re-running the full simulation chain, which for a typical ND280 production is expected to take $\sim$ 2.5 CPU years[1] with a final analysis output file size of about 350 GB[2]. Thus CPU time and storage requirements become prohibitive if needing to test the effect of many input parameters and values—it would certainly not be feasible for a single user to do this on a per analysis

[1] Based on local processing time for 2.50 GHz Intel Xeon processor.

[2] This is the total analysis file size for the MC data used in this analysis and represents approximately 17×10^6 neutrino interactions over the full Magnet volume.

basis. Event reweighting provides a solution to this. For each event i, a weight w_i is generated to reflect a change in the set of input physics parameters $\vec{a} \rightarrow \vec{a}'$:

$$w_i = \frac{P_i'}{P_i} = \frac{P_i(\vec{a}')}{P_i(\vec{a})}$$

where P_i is the probability for that event. The weight is then applied to the output of the final analysis stage and takes only the time to recalculate the modified probability. This reduces the total production time to evaluate a single change in physics parameters down to about 10 minutes. Because each reweighting scheme needs to be able to re-evaluate the probability for a given event, they are generally tied to the generators themselves. This reduces the need for code duplication and reduces the chance of problems occurring which could lead to different values of P from those used when generating the original event, and consequently to incorrect weights. This means that each generator has its own set of reweighting routines. For T2K, these currently exist for the GENIE and NEUT neutrino interaction generators and the JNUBEAM flux simulation and are in the form of external reweighting libraries.

T2KReWeight is a global analysis tool designed to unify the treatment of neutrino interaction and flux simulation systematics for the GENIE, NEUT and JNUBEAM reweighting routines. It provides a common interface to each of these external libraries as well as acting as a repository for common tools useful when generating and applying weights. Originally developed only as an interface for the GENIE reweighting libraries (see [90] for details of development and validation) it was then generalised for use with the NEUT and JNUBEAM reweighting libraries. It is written in C++ and uses object-oriented design principles to define the required interface which each of the external libraries then has to implement. The successful use of T2KReWeight in this analysis was an important demonstration of its usability in ND280-based analyses.

4.1.2 Neutrino Flux Systematics

The beam simulation was described in § 2.1.4. This analysis is based on the 11a flux simulation which was tuned to both the NA61 thin target data as well as to external hadron production cross section data. A default tuning is provided via the JNUBEAM reweighting libraries and accessed through T2KReWeight. This default tuning has been applied to all results presented so far. Fig. 4.1 shows the true neutrino energy spectrum prediction at ND280 before and after the default tuning was applied.

Using the reweighting libraries for JNUBEAM interfaced via T2KReWeight, we are able to consider uncertainties in the flux simulation from a number of sources:

- The proton beam parameters beam position (x and y), direction (x' and y') and width (σ_x and σ_y). As well as the divergence of the beam as characterised by the emittance (ϵ_x and ϵ_y) and the Twiss parameter (α_x and α_y). The values of these quantities used in the beam simulation are based on measurements made by the beam

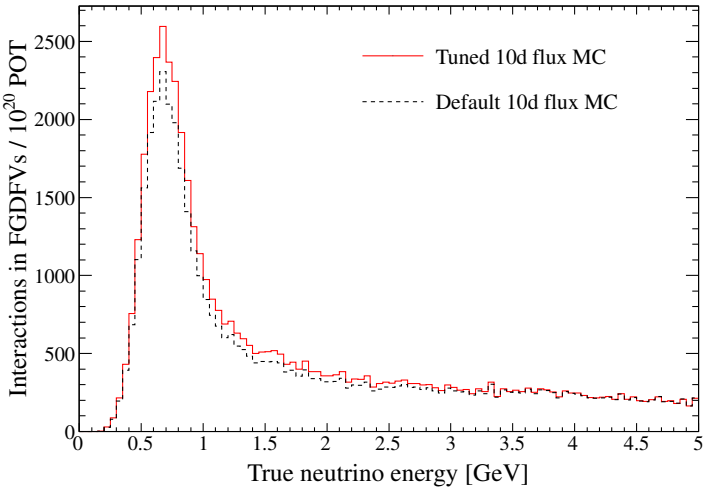


Figure 4.1: The true neutrino energy spectrum before and after applying the 11a flux tune.

monitoring system and uncertainties on these measured values lead to a systematic on the flux prediction.

- The uncertainty on the differential π-production multiplicity and rate based on the errors from the fits to the NA61 thin target data. The reweighting library for JNUBEAM contains a set of pre-calculated parameter variations which are used to make throws of sets of parameters representing the correlated uncertainties in the fits to NA61 data. We make a 100 throws of these correlated parameter sets and use the RMS deviation from the nominal as the systematic due to π-production uncertainties.

- The uncertainty on the K-production multiplicity and rate based on the errors from tuning to external data. The BMPT parametrisation [91] was used to fit to the external data. It provides an analytical formulae describing secondary particle yields in p-A interactions. The uncertainties on these fit parameters are stored within the JNUBEAM reweighting library. We vary each of the BMPT parameters separately within their 1σ error to see the effect on the total number of events passing each selection. These are then summed in quadrature to give the total error due to K-production. It is important to note that the BMPT parameterisation was not actually used in the default flux tuning so that these errors are not exactly correct for the current version of the flux simulation but rather give an indication of the size of the errors on kaon production expected.

An important caveat for the results we are about to present is that the current implementation of the JNUBEAM reweighting libraries do not produce the same results as the official flux uncertainty produced by the T2K Beam group. In addition it is not possible to evaluate further beam related systematics such as the horn and target alignment and

the horn current uncertainties. Therefore these results should not be taken as the final flux uncertainties.

The effect of these systematics on each of the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ selections is shown in Table 4.1. Here we see that the dominant uncertainty for all selections come from the K-production systematics—this is expected as the current beam simulation (with 11d tuning) does not include the NA61 K-production measurements, meaning that the prediction is subject to the large uncertainties associated with existing K-production data. In contrast the smaller uncertainties associated with the π-production systematics reflect the fact that the NA61 π-production multiplicity and rate measurements have been used when tuning the current flux simulation. For the ν_μ-CC-Inclusive selection, the combined uncertainty from K-production is $+8.24$ -7.43%, whereas for the ν_μ-CC-$\pi^\pm$ selection it is around $\pm 20\%$. This difference is expected—as Fig. 4.2 shows the ν_μ-CC-$\pi^\pm$ cuts typically select interactions coming from higher energy neutrinos for which, as was shown in Fig. 2.4(a), K-production is the dominant process. The converse is true for the NA61-based π-production systematics where they have a larger relative impact on the output of ν_μ-CC-Inclusive selection as this samples the low energy neutrino peak where π production is dominant.

The last row of Table 4.1 shows the total systematic uncertainty due to uncertainties in the flux simulation. This was calculated by summing the individual contributions in quadrature. We note that the values for the uncertainties of about $\pm 10\%$ and $\pm 20\%$ for the ν_μ-CC-Inclusive and the ν_μ-CC-$\pi^\pm$ respectively are significantly larger than the corresponding contributions from detector and reconstruction-based systematics that are shown in § 3.6.6. This highlights the feasibility of using the ND280 data in fits to constrain the flux uncertainty, although ideally, from the point of view of making dedicated cross section measurements it would be better if these came solely from the external data and measurements from other T2K near detectors such as INGRID.

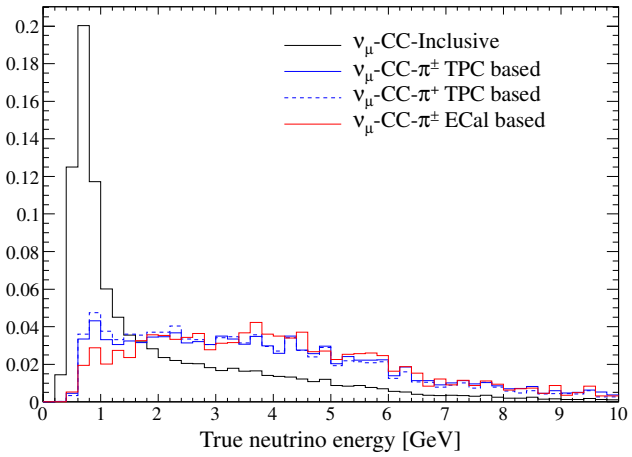


Figure 4.2: True neutrino energy distributions for events passing the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ selections. To allow comparison all are normalised to unity.

	$\delta N^{\text{exp}}/N^{\text{exp}}$ for selection							
Systematic	ν_μ-CC-Inclusive		ν_μ-CC-$\pi^\pm$ (TPC)		ν_μ-CC-π^+ (TPC)		ν_μ-CC-$\pi^\pm$ (ECal)	
Beam x	+0.19%	−0.21%	+0.25%	−0.27%	+0.14%	−0.18%	+0.21%	−0.25%
Beam y	+1.08%	−1.10%	+0.84%	−0.89%	+0.75%	−0.76%	+0.86%	−0.89%
Beam x'	+0.08%	−0.09%	+0.11%	−0.11%	+0.06%	−0.07%	+0.09%	−0.09%
Beam y'	+1.40%	−1.43%	+1.12%	−1.19%	+1.02%	−1.01%	+1.08%	−1.17%
Beam σ_x	+0.09%	−0.15%	+0.14%	−0.21%	+0.24%	−0.33%	+0.21%	−0.31%
Beam σ_y	+0.02%	−0.63%	+0.32%	−1.01%	+0.00%	−0.61%	+0.08%	−0.89%
Beam ϵ_x	+0.00%	−0.00%	+0.00%	−0.00%	+0.00%	−0.01%	+0.00%	−0.01%
Beam ϵ_y	+0.00%	−0.00%	+0.00%	−0.00%	+0.00%	−0.00%	+0.00%	−0.00%
Beam α_x	+0.10%	−0.17%	+0.13%	−0.20%	+0.23%	−0.32%	+0.21%	−0.32%
Beam α_y	+0.03%	−1.40%	+0.42%	−1.71%	+0.00%	−1.12%	+0.25%	−1.58%
Total beam	+1.78%	−2.39%	+1.53%	−2.51%	+1.32%	−1.86%	+1.46%	−2.39%
Kaon BMPT A	+4.35%	−4.35%	+10.15%	−10.15%	+9.58%	−9.58%	+10.77%	−10.77%
Kaon BMPT B	+2.98%	−2.98%	+7.72%	−7.72%	+7.03%	−7.03%	+7.83%	−7.83%
Kaon BMPT α	+4.11%	−3.27%	+10.67%	−8.37%	+9.72%	−7.67%	+10.78%	−8.57%
Kaon BMPT β	+3.20%	−2.76%	+6.93%	−6.02%	+6.72%	−5.82%	+7.53%	−6.53%
Kaon BMPT a1	+0.54%	−0.53%	+1.41%	−1.37%	+1.27%	−1.24%	+1.39%	−1.35%
Kaon BMPT γ	+2.35%	−2.20%	+5.53%	−5.16%	+5.18%	−4.84%	+5.74%	−5.38%
Kaon BMPT δ	+2.68%	−2.01%	+6.03%	−4.60%	+5.75%	−4.36%	+6.41%	−4.87%
Total BMPT	+8.24%	−7.43%	+19.83%	−17.85%	+18.50%	−16.67%	+20.64%	−18.65%
NA61 Pion	+3.36%	−4.75%	+1.99%	−2.76%	+2.11%	−2.92%	+1.74%	−2.52%
Total all	+9.08%	−9.14%	+19.99%	−18.24%	+18.67%	−17.03%	+20.76%	−18.97%

Table 4.1: Flux systematics for fractional change in expected number of events passing each selection. The sum in quadrature of the individual contributions is shown for the beam and K parameters separately and the final row shows the combined systematic for all flux related systematics.

4.1.3 Neutrino Interaction Systematics

The physics models used by the GENIE neutrino interaction generator were described in § 1.3. Using the external reweighting libraries for GENIE [34], we are able investigate the effect of uncertainties on many of the input physics parameters controlling three areas of the neutrino interaction simulation:

- Parameters controlling the neutrino cross section calculations.
- Those controlling intranuclear rescattering, whereby hadrons produced in the nuclear environment may re-scatter before escaping the nucleus.
- Those controlling the hadronisation and decay of intermediate particles produced by the primary neutrino interaction.

To evaluate the systematic on the total rate of events passing each of the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ selections, we vary each input parameter for 9 equally spaced values within the $\pm 1\sigma$ uncertainty about its nominal value. For each of these a set of weights is generated and then applied to the events passing the selections in order to calculate the modified rate corresponding to the change in input parameter. We scan over all 9 parameter values and take the maximum and minimum deviation in rate as the upper and lower bound on the systematic effect due to that parameter. For the level of uncertainty on each input parameter we adopt the values used in a recent study [55] into the systematic uncertainty on the T2K oscillation analysis coming from the physics models in GENIE. Here the input uncertainties were estimated based on either the spread of experimental results, the spread of available theoretical calculations or on GENIE comparisons to external data. These uncertainties are also similar to those used by oscillation analyses on other neutrino experiments (e.g. MINOS [16]) that used a neutrino interaction model very similar to that in the default GENIE v2.6.2 used for this analysis.

Table 4.2 shows the complete list of neutrino cross section systematics that were considered. It gives a brief description and shows the assigned fractional uncertainty for each systematic. The dominant sources of systematic include the axial mass for CCQE scattering and the axial and vector masses for both CC and NC resonance neutrino production (M_A^{CCQE}, M_A^{CCRES}, M_V^{CCRES}, M_A^{NCRES} and M_V^{CCRES} respectively). The uncertainties in nuclear effects due to Pauli suppression in CCQE reactions ($S_{\mathrm{Pauli}}^{\mathrm{CCQE}}$) are taken into account by varying the Fermi momentum level k_F in the modified impulse approximation used by GENIE. Uncertainties in the choice of vector form factors (dipole vs BBA2005) used for CCQE scattering are taken into account using the $V_{\mathrm{FF}}^{\mathrm{CCQE}}$ systematic. This systematic is qualitatively different to the others as it is associated with a choice in input physics models rather than a single input parameter value. For this the reweighting scheme allows the user to smoothly switch between the two choices to see the effect on the output of the selection—where the difference is then taken as the systematic. CC and NC coherent pion production uncertainties are taken into account by modifying the corresponding axial mass (M_A^{COH}) and the nuclear size parameter (R_0^{COH}), which controls the pion absorption

factor in the Rein-Sehgal (RS) model. As described in § 1.3, a number of suppression factors are introduced to avoid double counting the non-resonance background of single- and multi-π final states in the transition region between resonance and DIS production. These suppression factors are tuned to fit experimental data, and uncertainties in the level of the non-resonance background are considered for all CC and NC 1π and 2π final states by varying these factors (via the $R_{\rm bkg}$ systematics). Finally the uncertainties associated with CC and NC DIS scattering are considered by varying the most important parameters in the Bodek-Yang (BY) model used in GENIE.

x_P	Description of P	$\delta P/P$
$M_A^{\rm CCQE}$	Axial mass for CC quasi-elastic	$-15\% + 25\%$
$S_{\rm Pauli}^{\rm CCQE}$	CCQE Pauli suppression (via changes in Fermi level k_F)	$\pm 35\%$
$V_{\rm FF}^{\rm CCQE}$	Choice of CCQE vector form factors (BBA05 $\leftrightarrow$ Dipole)	–
$M_A^{\rm CCRES}$	Axial mass for CC resonance neutrino production	$\pm 20\%$
$M_V^{\rm CCRES}$	Vector mass for CC resonance neutrino production	$\pm 10\%$
$M_A^{\rm NCRES}$	Axial mass for NC resonance neutrino production	$\pm 20\%$
$M_V^{\rm NCRES}$	Vector mass for NC resonance neutrino production	$\pm 10\%$
$M_A^{\rm COHpi}$	Axial mass for CC and NC coherent pion production	$\pm 40\%$
$R_0^{\rm COHpi}$	Nuclear size param. controlling π absorption in RS model	$\pm 10\%$
$R_{\rm bkg}^{\nu p,{\rm CC}1\pi}$	Non-resonance bkg in νp CC1π reactions	$\pm 50\%$
$R_{\rm bkg}^{\nu p,{\rm CC}2\pi}$	Non-resonance bkg in νp CC2π reactions	$\pm 50\%$
$R_{\rm bkg}^{\nu n,{\rm CC}1\pi}$	Non-resonance bkg in νn CC1π reactions	$\pm 50\%$
$R_{\rm bkg}^{\nu n,{\rm CC}2\pi}$	Non-resonance bkg in νn CC2π reactions	$\pm 50\%$
$R_{\rm bkg}^{\nu p,{\rm NC}1\pi}$	Non-resonance bkg in νp NC1π reactions	$\pm 50\%$
$R_{\rm bkg}^{\nu p,{\rm NC}2\pi}$	Non-resonance bkg in νp NC2π reactions	$\pm 50\%$
$R_{\rm bkg}^{\nu n,{\rm NC}1\pi}$	Non-resonance bkg in νn NC1π reactions	$\pm 50\%$
$R_{\rm bkg}^{\nu n,{\rm NC}2\pi}$	Non-resonance bkg in νn NC2π reactions	$\pm 50\%$
$A_{HT}^{\rm BY}$	A_{HT} higher-twist param in BY model scaling variable ξ_w	$\pm 25\%$
$B_{HT}^{\rm BY}$	B_{HT} higher-twist param in BY model scaling variable ξ_w	$\pm 25\%$
$C_{V1\mu}^{\rm BY}$	$C_{V1\mu}$ u valence GRV98 PDF correction param in BY model	$\pm 30\%$
$C_{V2\mu}^{\rm BY}$	$C_{V2\mu}$ u valence GRV98 PDF correction param in BY model	$\pm 40\%$

Table 4.2: Neutrino interaction cross section parameters considered. For the systematic uncertainty on each we adopt those used in [55].

Table 4.3 shows the fractional effect due to these neutrino cross section systematics on the total number of events passing each of the selections. It contains three summary rows showing the sum in quadrature separately for the systematics controlling CCQE- and inelastic-scattering as well as for the total for all cross section systematics. For the ν_μ-CC-Inclusive selection we see a large contribution from the $M_A^{\rm CCQE}$, $S_{\rm Pauli}^{\rm CCQE}$ and $M_A^{\rm CCRES}$ systematics with slightly smaller but still significant contributions from the $M_V^{\rm CCRES}$ and $R_{\rm bkg}^{\nu n,{\rm CC}1\pi}$. This is consistent with expectations for an inclusive selection with a final selection purity of about 50% for CCQE processes and 35% for RES- or DIS-precesses (see first column of Table 3.7 for origin of these numbers). For the TPC- and ECal-based π selections we see that the contributions from parameters controlling RES scattering now dominate and that the effect due to parameters controlling CCQE scattering have dropped significantly down to the sub-percent level for the TPC-based selections and to

approximately 2% for the ECal-based selection. This demonstrates that the selections are performing as required—The ν_μ-CC-Inclusive selection is sensitive to the total rate of neutrino interactions whilst the π selections are mostly sensitive to cross section systematics affecting $\pi^\pm$-production. We can see from the final row that the total systematic uncertainty from neutrino cross section systematics is about $\pm 10\%$ for all selections.

Table 4.4 shows the complete list systematics considered for intranuclear hadron transport, hadronisation, and resonance decay uncertainties. As described in § 1.3, the default intranuclear hadron transport model used in GENIE is an effective model where the total rescattering probability for a hadron created inside the nucleus, and the type of rescattering it will experience, are factorised. This means we consider two kinds of uncertainties affecting GENIE INTRANUKE/hA: uncertainties in the total rescattering probability (mean free path) for hadrons within the target nucleus, and uncertainties in the conditional probability of each hadron rescattering mode (elastic, inelastic, charge exchange, π production and absorption / multi-nucleon knock-out), given that a rescattering did occur. These are treated separately for pions and nucleons.

The lower half of the Table 4.4 shows the uncertainties associated with neutrino-induced hadronisation and resonance decays. These include uncertainties in the assigned pion kinematics in $N\pi$ hadronic states produced by the Andreopoulos-Gallagher-Kehayias-Yang (AGKY) GENIE hadronisation model (AGKY$^{pT1\pi}$ and AGKY$^{xF1\pi}$) and uncertainties on the size of formation zone over which a hadron has not fully materialised and is subject to a reduced probability of intranuclear rescattering, which is accounted for by varying the formation zone length using the f_{zone} parameter. Uncertainties in the angular distributions of pions produced by the decay of Δ-resonances are taken into account by switching between the default isotropic decays used by GENIE and a modified anisotropic decay predicted by the Rein-Sehgal model. The final systematics to be considered are those due to the uncertainties on the branching ratios for radiative and single-η resonance decays.

The fractional effect on the total number of selected events for each selection is shown in Table 4.5. The systematics related to intranuclear hadron transport are shown in the upper half of the table. For the ν_μ-CC-Inclusive selection, the dependence on the hadronic system is minimal, as we only try to identify the outgoing μ, and the total uncertainty on the final number of selected events is small ($+0.33-0.38\%$). In contrast the various TPC- and ECal-based π selections are sensitive to the hadronic system, as they try to identify if there was at least one charged π in the final state, and exhibit larger uncertainties of around 3–4%. We see similar behaviour for the hadronisation and resonance decay systematics where there is little effect on the ν_μ-CC-Inclusive selection and the dominant contributions to the π selections come from uncertainties in the angular distributions for the Δ-resonance decay ($\theta_\pi^{\Delta\to\pi N}$) and the formation zone uncertainty (f_{zone}). The final row shows the sum of all neutrino induced systematics added in quadrature and including the cross section systematics shown in Table 4.3.

In this section we have investigated the effect that uncertainties in the neutrino flux and neutrino interaction simulation have on the total number of events passing the selections

	$\delta N^{\text{exp}}/N^{\text{exp}}$ for selection							
Systematic	ν_μ-CC-Inclusive		ν_μ-CC-$\pi^\pm$ (TPC)		ν_μ-CC-π^+ (TPC)		ν_μ-CC-$\pi^\pm$ (ECal)	
M_A^{CCQE}	+8.10%	−5.73%	+0.56%	−0.28%	+0.45%	−0.21%	+2.40%	−1.11%
$S_{\text{Pauli}}^{\text{CCQE}}$	+4.20%	−4.43%	+0.04%	−0.05%	+0.01%	−0.01%	+0.05%	−0.06%
$V_{\text{FF}}^{\text{CCQE}}$	+0.43%	−0.43%	+0.03%	−0.03%	+0.04%	−0.04%	+0.18%	−0.18%
All CCQE	+9.14%	−7.26%	+0.56%	−0.28%	+0.45%	−0.22%	+2.40%	−1.13%
M_A^{CCRES}	+4.48%	−4.67%	+7.94%	−7.48%	+8.62%	−8.20%	+8.08%	−7.08%
M_V^{CCRES}	+2.45%	−2.26%	+4.32%	−3.78%	+4.69%	−4.12%	+4.37%	−3.68%
M_A^{NCRES}	+0.30%	−0.26%	+0.17%	−0.10%	+0.17%	−0.11%	+0.46%	−0.32%
M_V^{NCRES}	+0.08%	−0.07%	+0.04%	−0.03%	+0.04%	−0.03%	+0.12%	−0.10%
M_A^{COHpi}	+0.05%	−0.14%	+0.20%	−0.54%	+0.24%	−0.63%	+0.15%	−0.39%
R_0^{COHpi}	+0.08%	−0.08%	+0.32%	−0.32%	+0.38%	−0.38%	+0.23%	−0.24%
$R_{\text{bkg}}^{\nu p,\text{CC1}\pi}$	+0.51%	−0.51%	+1.02%	−1.02%	+1.14%	−1.14%	+1.46%	−1.46%
$R_{\text{bkg}}^{\nu p,\text{CC2}\pi}$	+0.86%	−0.86%	+2.69%	−2.69%	+3.04%	−3.04%	+2.40%	−2.40%
$R_{\text{bkg}}^{\nu n,\text{CC1}\pi}$	+2.58%	−2.58%	+3.72%	−3.72%	+3.97%	−3.97%	+4.36%	−4.36%
$R_{\text{bkg}}^{\nu n,\text{CC2}\pi}$	+0.78%	−0.78%	+2.72%	−2.72%	+2.45%	−2.45%	+2.68%	−2.68%
$R_{\text{bkg}}^{\nu p,\text{NC1}\pi}$	+0.02%	−0.02%	+0.01%	−0.01%	+0.01%	−0.01%	+0.03%	−0.03%
$R_{\text{bkg}}^{\nu p,\text{NC2}\pi}$	+0.09%	−0.09%	+0.25%	−0.25%	+0.28%	−0.28%	+0.27%	−0.27%
$R_{\text{bkg}}^{\nu n,\text{NC1}\pi}$	+0.25%	−0.25%	+0.11%	−0.11%	+0.07%	−0.07%	+0.21%	−0.21%
$R_{\text{bkg}}^{\nu n,\text{NC2}\pi}$	+0.12%	−0.12%	+0.32%	−0.32%	+0.33%	−0.33%	+0.23%	−0.23%
A_{HT}^{BY}	+0.13%	−0.42%	+0.16%	−1.45%	+0.23%	−1.28%	+0.36%	−1.50%
B_{HT}^{BY}	+0.29%	−0.56%	+0.69%	−1.91%	+0.73%	−1.72%	+0.92%	−1.99%
$C_{V1\mu}^{\text{BY}}$	+0.08%	−0.35%	+0.06%	−1.27%	+0.07%	−1.05%	+0.20%	−1.26%
$C_{V2\mu}^{\text{BY}}$	+0.05%	−0.34%	+0.00%	−1.24%	+0.00%	−1.03%	+0.11%	−1.23%
All inelastic	+5.89%	−6.01%	+10.59%	−10.45%	+11.38%	−11.14%	+10.95%	−10.36%
All cross section	+10.87%	−9.42%	+10.60%	−10.45%	+11.39%	−11.14%	+11.21%	−10.43%

Table 4.3: Neutrino interaction cross section systematics for fractional change in expected number of events passing each selection. The sum in quadrature of the individual contributions is shown separately for the systematics controlling CCQE-scattering and those controlling inelastic scattering. The final row shows the combined systematic based on the sum in quadrature of all neutrino cross section systematics that were considered.

x_P	Description of P	$\delta P/P$
λ^N	Nucleon mean free path (total rescattering probability)	$\pm 20\%$
P^N_{cex}	Nucleon charge exchange probability	$\pm 50\%$
P^N_{elas}	Nucleon elastic reaction probability	$\pm 30\%$
P^N_{inel}	Nucleon inelastic reaction probability	$\pm 40\%$
P^N_{abs}	Nucleon absorption probability	$\pm 20\%$
$P^N_{\pi-\text{prod}}$	Nucleon π production probability	$\pm 20\%$
λ^π	Pion mean free path (total rescattering probability)	$\pm 20\%$
P^π_{cex}	π charge exchange probability	$\pm 50\%$
P^π_{elas}	π elastic reaction probability	$\pm 10\%$
P^π_{inel}	π inelastic reaction probability	$\pm 40\%$
P^π_{abs}	π absorption probability	$\pm 20\%$
$P^\pi_{\pi-\text{prod}}$	π π production probability	$\pm 20\%$
$\text{AGKY}^{pT1\pi}$	Pion transverse momentum (p_T) for $N\pi$ states in AGKY	–
$\text{AGKY}^{xF1\pi}$	Pion Feynman-x (x_F) for $N\pi$ states in AGKY	–
f_{zone}	Hadron formation zone	$\pm 50\%$
$\theta^{\Delta\to\pi N}_\pi$	Pion angular distribution in $\Delta \to \pi N$ ($isotropic \to RS$)	-
$B^{R\to X+1\gamma}$	Branching ratio for radiative resonance decays	$\pm 50\%$
$B^{R\to X+1\eta}$	Branching ratio for single-η resonance decays	$\pm 50\%$

Table 4.4: Neutrino interaction intranuclear hadron transport, hadronisation, and resonance-decay systematics considered for this analysis. For the systematic uncertainty on each we adopt those used in [55].

developed in the previous chapter. We see that, as intended, the two types of selection are sensitive to different aspects of the flux and neutrino cross section models. The K-production uncertainties in the flux simulation dominate the ν_μ-CC-$\pi^\pm$ selections while the systematic uncertainty in the ν_μ-CC-Inclusive selection still has a significant contribution due to π-production uncertainties (in flux simulation). If, for example, there was a systematic bias in the K-production as modelled in the current flux simulation then this should manifest as a larger data-MC difference in the ν_μ-CC-$\pi^\pm$ selected events, allowing a certain level of disentanglement of possible data-MC discrepancies. The ν_μ-CC-Inclusive selected events are sensitive to neutrino interaction systematics affecting the total rate of ν_μ-CC interactions but not to those affecting the hadronic system, while the ν_μ-CC-$\pi^\pm$ selections are not sensitive to cross section systematics related to CCQE-scattering, but show a strong dependence on those controlling π^+ production and on the intranuclear hadron rescattering and hadronisation systematics.

4.2 Final Data and MC Comparisons

We now present the final data-MC comparisons, with detector and reconstruction systematics for each of the selections developed, and put into context with the flux and neutrino interaction systematics evaluated in the previous section. Table 4.6 shows the number of events passing each cut for data, $N^{\text{Selected}}_{\text{Data}}$, for MC, $N^{\text{Selected}}_{\text{MC}}$, and for MC normalised by a factor of 0.0688 ($= 7.83 \times 10^{19}/1.14 \times 10^{21} = N^{\text{POT}}_{\text{Data}}/N^{\text{POT}}_{\text{MC}}$) to the total POT for the Run 2 period, $N^{\text{Selected}}_{\text{ScaledMC}}$. Also shown is the relative survival probability for each cut $P^{\text{Survival}}_{\text{Data}}$

	$\delta N^{\text{exp}}/N^{\text{exp}}$ for selection							
Systematic	ν_μ-CC-Inclusive		ν_μ-CC-$\pi^\pm$ (TPC)		ν_μ-CC-π^+ (TPC)		ν_μ-CC-$\pi^\pm$ (ECal)	
λ^N	+0.02%	−0.02%	+0.19%	−0.14%	+0.27%	−0.19%	+0.61%	−0.64%
P^N_{cex}	+0.00%	−0.00%	+0.00%	−0.00%	+0.01%	−0.01%	+0.02%	−0.02%
P^N_{elas}	+0.12%	−0.13%	+0.14%	−0.14%	+0.25%	−0.25%	+1.80%	−1.80%
P^N_{inel}	+0.01%	−0.02%	+0.02%	−0.03%	+0.05%	−0.05%	+0.04%	−0.04%
P^N_{abs}	+0.08%	−0.09%	+0.24%	−0.24%	+0.29%	−0.29%	+1.16%	−1.16%
$P^N_{\pi-\text{prod}}$	+0.00%	−0.01%	+0.20%	−0.20%	+0.24%	−0.24%	+0.04%	−0.04%
λ^π	+0.12%	−0.14%	+2.30%	−2.74%	+2.58%	−3.06%	+1.52%	−1.79%
P^π_{cex}	+0.02%	−0.04%	+0.09%	−0.13%	+0.07%	−0.09%	+0.11%	−0.12%
P^π_{elas}	+0.02%	−0.04%	+1.39%	−1.40%	+1.62%	−1.63%	+0.56%	−0.59%
P^π_{inel}	+0.22%	−0.25%	+0.25%	−0.33%	+0.00%	−0.04%	+0.42%	−0.38%
P^π_{abs}	+0.14%	−0.18%	+1.93%	−2.06%	+2.04%	−2.13%	+0.58%	−0.63%
$P^\pi_{\pi-\text{prod}}$	+0.00%	−0.01%	+0.08%	−0.11%	+0.09%	−0.11%	+0.04%	−0.07%
All intranuclear	+0.33%	−0.38%	+3.35%	−3.74%	+3.71%	−4.10%	+2.85%	−3.02%
AGKY$^{pT1\pi}$	+0.29%	−0.00%	+0.62%	−0.00%	+0.87%	−0.00%	+0.71%	−0.00%
AGKY$^{xF1\pi}$	+0.85%	−0.22%	+1.10%	−0.13%	+1.40%	−0.00%	+2.33%	−1.22%
f_{zone}	+0.30%	−0.72%	+4.00%	−4.80%	+3.74%	−4.02%	+2.25%	−4.95%
$\theta_\pi^{\Delta\to\pi N}$	+0.09%	−0.09%	+1.29%	−1.29%	+1.51%	−1.51%	+0.77%	−0.77%
$B^{R\to X+1\gamma}$	+0.01%	−0.01%	+0.01%	−0.01%	+0.02%	−0.02%	+0.01%	−0.01%
$B^{R\to X+1\eta}$	+0.22%	−0.22%	+0.53%	−0.53%	+0.67%	−0.67%	+0.33%	−0.33%
All hadronisation	+0.97%	−0.79%	+4.42%	−5.00%	+4.40%	−4.34%	+3.42%	−5.17%
Total all (incl xsec)	+10.92%	−9.46%	+11.97%	−12.17%	+12.76%	−12.64%	+12.06%	−12.02%

Table 4.5: Neutrino interaction intranuclear, hadronisation, and resonance decay systematics for expected number of events passing each selection.

and $P_{\mathrm{MC}}^{\mathrm{Survival}}$, which acts as a final check that the cuts are not responding in a significantly different way to data and MC.

Cut	$N_{\mathrm{Data}}^{\mathrm{Selected}}$	$N_{\mathrm{MC}}^{\mathrm{Selected}}$	$N_{\mathrm{ScaledMC}}^{\mathrm{Selected}}$	$P_{\mathrm{Data}}^{\mathrm{Survival}}$	$P_{\mathrm{MC}}^{\mathrm{Survival}}$
1) Vtx in FV and bunch	23182	312904.5	21522.4	100.0%	100.0%
2) ≥ 1 trk with TPC	15961	216264.6	14875.2	68.9%	69.1%
3) & with good TPC	14492	198443.7	13649.5	90.8%	91.8%
4) & TPC-μ-Pull	8419	112753.7	7755.5	58.1%	56.8%
5) & not TPC-e-Pull	7215	98450.1	6771.6	85.7%	87.3%
6) & is negative	5383	74918.3	5153.1	74.6%	76.1%
7) & start in FGD FV	4917	69988.2	4814.0	91.3%	93.4%
8) ≥ 2 tracks	1796	23945.1	1647.0	36.5%	34.2%
TPC-based $\pi^{\pm}/\pi^{+}$ cuts					
9_a) & 2nd trk with TPC	1362	18631.5	1281.5	75.8%	77.8%
10_a) & with good TPC	1287	17777.7	1222.8	94.5%	95.4%
11_a) & TPC-μ-Pull	564	7775.9	534.8	43.8%	43.7%
12_a) & not TPC-p-Pull	362	5056.8	347.8	64.2%	65.0%
13_a) & is positive	286	4203.7	289.1	79.0%	83.1%
ECal-based $\pi^{\pm}$ cuts					
9_b) & 2nd trk ECal	729	10282.9	707.3	40.6%	42.9%
10_b) & 2nd trk TrkShVal	366	5525.6	380.1	50.2%	53.7%
11_b) & 2nd trk dQ/dL	212	3128.8	215.2	57.9%	56.6%

Table 4.6: Total number of events and relative cut survival probability for all ν_μ-CC-Inclusive and ν_μ-CC-$\pi^{\pm}$ cuts. The number of selected MC events is shown for both the actual simulated MC sample POT (1.14×10^{21}) and for the number of events scaled to the Run 2 data POT (7.83×10^{19}).

In Fig. 4.3 we now summarise the final results. For each of the four selections we show the total number of events selected, for both the Run 2 data set and for the MC expectation for the Run 2 POT, as well as the data/MC ratio with associated statistical and systematic uncertainties. It should be noted that no MC-based corrections are applied to the data so we are comparing the total number of events passing each selection and not the purity and efficiency corrected rate of ν_μ-CC or ν_μ-CC-$\pi^{\pm}$ interactions. It should also be pointed out that whilst the MC numbers shown are normalised to the data POT the associated statistical errors depend on the actual number of simulated events[3].

.

[3] A further complication is that the MC prediction is reweighted to reflect the tuned flux simulation meaning that the statistical errors are slightly larger than that expected for $\sqrt{N_{\mathrm{MC}}^{\mathrm{Selected}}}$.

1) ν_μ-CC-Inclusive selection ($\eta^{\nu_\mu\text{-CC}} = 88.1\%$ and $\epsilon^{\nu_\mu\text{-CC}} = 40.1\%$):

$$N_{\text{Data}}^{\text{CCIncl}} = 4917.0 \pm 70.1(\text{stat})$$

$$N_{\text{ScaledMC}}^{\text{CCIncl}} = 4814.0 \pm 19.3(\text{stat})^{+148.9}_{-148.4}(\text{det})^{+525.5}_{-455.4}(\text{xsec})^{+436.9}_{-439.9}(\text{flux})$$

$$\boxed{N_{\text{Data}}^{\text{CCIncl}}/N_{\text{ScaledMC}}^{\text{CCIncl}} = 1.021 \pm 0.015(\text{stat})^{+0.032}_{-0.031}(\text{det})^{+0.112}_{-0.097}(\text{xsec})^{+0.093}_{-0.093}(\text{flux})}$$

2) ν_μ-CC-$\pi^\pm$ selection TPC-based ($\eta^{\nu_\mu\text{-CC-}\pi^+} = 81.3\%$ and $\epsilon^{\nu_\mu\text{-CC-}\pi^+} = 8.1\%$):

$$N_{\text{Data}}^{\text{CC}\pi^\pm} = 362.0 \pm 19.0(\text{stat})$$

$$N_{\text{ScaledMC}}^{\text{CC}\pi^\pm} = 347.8 \pm 5.1(\text{stat})^{+14.6}_{-14.8}(\text{det})^{+41.6}_{-42.3}(\text{xsec})^{+69.5}_{-63.4}(\text{flux})$$

$$\boxed{N_{\text{Data}}^{\text{CC}\pi^\pm}/N_{\text{ScaledMC}}^{\text{CC}\pi^\pm} = 1.041 \pm 0.057(\text{stat})^{+0.044}_{-0.044}(\text{det})^{+0.125}_{-0.127}(\text{xsec})^{+0.208}_{-0.190}(\text{flux})}$$

3) ν_μ-CC-π^+ selection TPC-based ($\eta^{\nu_\mu\text{-CC-}\pi^+} = 84.5\%$ and $\epsilon^{\nu_\mu\text{-CC-}\pi^+} = 6.9\%$):

$$N_{\text{Data}}^{\text{CC}\pi^+} = 286.0 \pm 16.9(\text{stat})$$

$$N_{\text{ScaledMC}}^{\text{CC}\pi^+} = 289.1 \pm 4.6(\text{stat})^{+12.9}_{-13.1}(\text{det})^{+36.9}_{-36.5}(\text{xsec})^{+54.0}_{-49.2}(\text{flux})$$

$$\boxed{N_{\text{Data}}^{\text{CC}\pi^+}/N_{\text{ScaledMC}}^{\text{CC}\pi^+} = 0.989 \pm 0.061(\text{stat})^{+0.044}_{-0.045}(\text{det})^{+0.126}_{-0.125}(\text{xsec})^{+0.185}_{-0.168}(\text{flux})}$$

4) ν_μ-CC-$\pi^\pm$ selection ECal-based ($\eta^{\nu_\mu\text{-CC-}\pi^+} = 69.7\%$ and $\epsilon^{\nu_\mu\text{-CC-}\pi^+} = 4.2\%$):

$$N_{\text{Data}}^{\text{CC}\pi^\pm} = 212.0 \pm 14.6(\text{stat})$$

$$N_{\text{ScaledMC}}^{\text{CC}\pi^\pm} = 215.2 \pm 4.0(\text{stat})^{+16.1}_{-21.9}(\text{det})^{+26.0}_{-25.9}(\text{xsec})^{+44.7}_{-40.8}(\text{flux})$$

$$\boxed{N_{\text{Data}}^{\text{CC}\pi^\pm}/N_{\text{ScaledMC}}^{\text{CC}\pi^\pm} = 0.985 \pm 0.070(\text{stat})^{+0.074}_{-0.1}(\text{det})^{+0.119}_{-0.118}(\text{xsec})^{+0.205}_{-0.187}(\text{flux})}$$

Figure 4.3: The final data-MC comparisons for the number of events passing 1) the ν_μ-CC-Inclusive cuts, 2) the TPC-based ν_μ-CC-$\pi^\pm$ cuts, 3) the TPC-based ν_μ-CC-π^+ cuts, and finally 4) the ECal-based ν_μ-CC-$\pi^\pm$ cuts. Also shown are the associated statistical, detector and reconstruction, flux simulation, and neutrino interaction systematic uncertainties. The MC prediction is normalised to the Run 2 data POT, 7.83×10^{19}. The comparison is for the total number of events passing each selection, and no MC-based purity corrections have been applied.

4.3 Discussion

In this section we discuss the results of the final data-MC comparisons for the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ selections. We then consider the limitations and areas of improvement for these and other tracker-based selections and conclude with a critical evaluation of the near detector design.

4.3.1 ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ Results

Based on the results shown in Fig. 4.3, we can see that the data/MC ratios for both the ν_μ-CC-Inclusive and π-based selections are 1.0 to within the statistical uncertainty. The level of data-MC agreement is well within the detector, flux and neutrino interaction systematic uncertainties associated with the MC prediction. There is a certain amount of statistical correlation between these results—events passing the π-based cuts are a subset of those passing the ν_μ-CC-Inclusive cuts and the same is true, but to an even greater extent, for the events passing the TPC-based $\pi^\pm$ and π^+ cuts.

As discussed in § 4.1, the ν_μ-CC-Inclusive selection is sensitive to neutrino interaction systematics affecting the total CC rate, whilst the π-based selections are only sensitive to those affecting the production of $\pi^\pm$s. The data-MC agreement for both types of selections provides some indication that the neutrino interaction models in GENIE are correctly reproducing, to within the statistical and large systematic uncertainties of the measurement, both the total CC interaction rate and the rate of production of charged pions. Similarly, the two types of selection are sensitive to different aspects of the flux simulation—the CC-Inclusive selection is sensitive to systematics affecting the beam simulation of both π- and K-production, whilst the π-based selections are predominantly only sensitive to the simulation of K-production. As before, this provides some indication that the total rate of π- and K-production in the beam are correctly simulated. However, because we only measure the total rate of interactions it is hard to disentangle the effects due to biases in the flux simulation with those from the neutrino interaction simulation. For example, it is entirely possible that an under-prediction in the rate of K-production in the flux simulation, producing higher energy neutrinos that are more likely to interact to produce a $\pi^\pm$ in the final state, is cancelling with some aspect of the neutrino interaction simulation that is over-predicting the rate of neutrino interactions leading to a $\pi^\pm$ in the final state.

The ECal-based ν_μ-CC-$\pi^\pm$ selection also shows data-MC agreement to within the statistical and systematic uncertainties. This selection will be sensitive to different detector systematics affecting the identification of the $\pi^\pm$ and provides a useful cross check for these. As shown in § 3.5.3 the ECal-based cuts have the potential to select neutrino interactions where the $\pi^\pm$ has a greater opening angle. Although at present there is not much difference in this geometrical acceptance it is a promising avenue of development for future analyses.

The main conclusion that we draw from these results is that both the overall rate of CC neutrino interactions and the rate of those producing at least one charged pion in the final

state are well reproduced by the full MC simulation chain. To move beyond this conclusion would mean disentangling the possible effects coming from flux and neutrino interaction simulations. For this it would be necessary to move from a measurement of the integrated number of events to that of a differential measurement. For example, a measurement with respect to the reconstructed event kinematics such as the muon momentum and outgoing angle, as well as some quantifier of the total energy in the hadronic system, would allow a ratio measurement to be performed in which the flux systematics would largely cancel.

Another important avenue for development, and one which will require further understanding of the geometrical acceptance of the tracker-region, would be to subdivide the ν_μ-CC-$\pi^\pm$ selected events into more exclusive categories based on the number of π-candidates, for example subdividing into subsamples of $1\pi^\pm$ and $2\pi^\pm$, $\geq 3\pi^\pm$ candidates will provide more constraints for each of the different cross section processes such as RES and DIS scattering contributing to π production. It should be emphasised that the accumulated POT collected by T2K so far represents less than 2% of that expected for the full 5 years of nominal running, the increased statistics that this will bring makes possible such differential measurements of exclusive subsamples.

4.3.2 Limitations of Current Tracker-based Selections

As discussed in § 3.4.1 the ν_μ-CC-Inclusive selection has a highly forward peaked acceptance. This is due the geometrical layout of the tracker-region and the dependence on TPC information to identify the muon candidate as well as the a priori assumption in the reconstruction that all tracks travel in the downstream direction, unless FGD timing information indicates otherwise. Increasing the acceptance for interactions with high angle and backwards-going leptons is important as these type of interaction sample a different part of the kinematical phase space used to model the neutrino interactions. This is particularly relevant as the far detector, with its 4π angular acceptance, is sensitive to these interactions—meaning the near detector measurements used to predict the interaction rates at the far detector will not constrain the models over the full phase space seen at the far detector. For example, to first order the current models describing CC-QEL scattering are parameterised as a function of the four-momentum transfer Q^2. In Fig. 4.4 we see that a tracker-based ν_μ-CC-QEL selection samples predominantly the low Q^2 region in contrast to the far detector, which samples both the low and high Q^2 region well.

The ECal-based ν_μ-CC-$\pi^\pm$ selection developed as part of this thesis demonstrated the use of the Barrel-ECal to recover high-angle tracks to improve the angular acceptance for charged pions, see Fig. 3.18(b) for a successfully recovered high-angle pion. In principle, high angle primary leptons could be selected in a similar manner. This may take the form of a combined TPC and ECal PID, in the case of high-angle tracks which still pass through a TPC, or by the sole use of the ECal PID, for tracks travelling almost perpendicular to the z direction which do not pass through an active TPC volume. In such cases the charge of the lepton candidate will be hard to reconstruct—this may be possible with

future improvements to the global reconstruction and fits for track curvature based on the contributing sub-detector tracks. The ongoing work to improve matching between the ECal and FGD is of high importance and should bring significant improvement for future analyses. Fig. 3.27(b) shows an event display of the type of high-angle track that could be recovered through improved FGD-ECal matching.

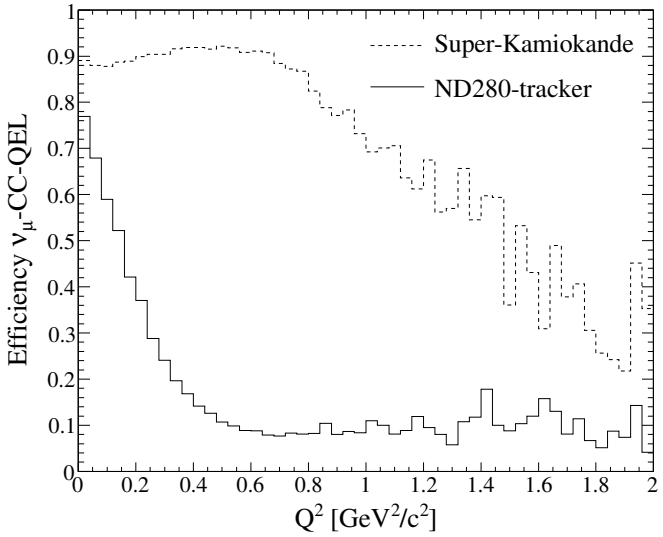


Figure 4.4: The efficiency for selecting ν_μ-CC-QEL interactions as a function of Q^2 is shown for both the tracker-based ν_μ-CC-Inclusive selection developed here (with an additional cut requiring only one reconstructed track) and for the Super-Kamiokande ν_μ-CC-QEL selection described in § 2.4.2 (applied to the unoscillated spectrum at the far detector). It should be noted that these are not official results and are intended only to demonstrate qualitative differences between the acceptance at the near and far detectors.

As we saw in § 2.6.4 by not using the FGD PID information for tracks which stop in an FGD we neglect $\sim 50\%$ of the final state charged pions from neutrino interactions in the FGD. In future analyses it will be important to include these FGD-only tracks as they sample pions with different momentum distribution to those which escape the FGD. Fig. 4.5 shows the result of a truth study of ν_μ-CC-$1\pi^+$ interactions in an FGD showing that, as we would expect, the pions which stop in an FGD have in general much lower momentum. By only sampling the higher momentum pions we are neglecting an important part of the phase space around the Cherenkov threshold for charged pions in water. There is an ongoing effort to improve the FGD-only reconstruction efficiency and much work has gone into validating the FGD-only PID and understanding the associated systematics so that they can be used in future data and MC productions.

In addition to the gain in the acceptance for backwards going tracks, improvements in the ability to tell the sense of direction for tracks not passing through both FGDs will allow increased rejection of the Non-FGD background discussed in § 3.6.4 where the track direction can be used to remove events where the track starts outside of the FGD. At present there is ongoing work on the global reconstruction to allow directionality based on

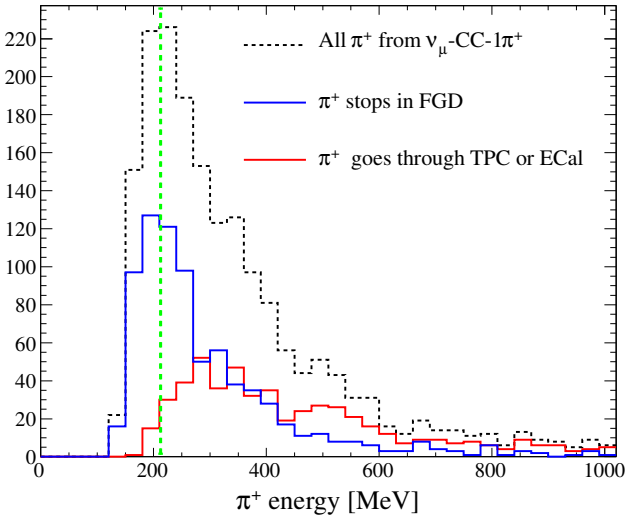


Figure 4.5: True energy distribution for π^+ from ν_μ-CC-1π^+ interactions in an FGD, where the μ^- travels through a TPC active volume. Plotted separately for all pions, those which stop in the FGD and those which escape the FGD and travel through either a TPC or ECal active volume. The vertical dashed line indicates the Cherenkov threshold for charged pions in water (∼210 MeV). Based on results from truth study looking at trajectories of particles through the ND280 geometry [92].

energy deposit as a function of track length as well as development of more sophisticated measurements used for inter-detector timing calibration.

Having considered the limitations in the various tracker-based selections developed in this thesis we next give a critical vision on the detector design and discuss possible alternatives.

4.3.3 Detector Design

To meet the physics requirements of T2K discussed in § 2.3 the near detector needs to reconstruct and make spectral measurements of CC and NC exclusive neutrino interaction channels for both ν_μ and ν_e scattering and then use these to predict the interaction rates at the far detector. Focussing on the measurement of CC interactions we can visualise an idealised detector with the following design and features:

- One or more active high mass targets capable of simultaneously reconstructing multiple tracks from vertices with high multiplicity final states and with the necessary elemental composition to allow extrapolation to interactions on water at Super-Kamiokande.

- A tracking detector that hermetically surrounds the target region and can simultaneously reconstruct the high density of tracks exiting the active region. This needs to be able to measure the momentum and sign for charged particles as well as distinguish between electrons, muons and protons.

- An additional calorimeter surrounding the tracking detector to allow reconstruction of particles exiting the tracker region, in particular particle identification and calorimetric energy measurements for showering electrons, photons, and charged pions.

- The ability to tell the sense of direction for each track in order to determine charge and reject backgrounds entering from outside the target.

Following the description of the off-axis near detector in § 2.3 we can see that with its current design the tracking and surrounding ECal region of ND280 is close to this idealised detector and, with the developments discussed in § 4.3.2, almost all of the requirements can be met. The only area where ND280 significantly deviates from the idealised detector is that the TPC tracking regions do not provide hermetic coverage of the FGD target volumes, which leads to the forward-peaked acceptance discussed in § 4.3.2.

Ideally, and ignoring monetary constraints, each FGD would be boxed in on all sides by separate TPCs in a similar way to that in which the Barrel- and Ds-ECal surround the tracker region. The difficulty is that the magnetic field required for the TPC spectral measurements puts constraints on their possible orientation as their drift directions must be aligned with the magnetic field direction, which for the UA1 magnet is along the x direction. This means that it would not be feasible to add additional TPCs to the left and right side of the tracker region as most of the particles entering would have their trajectories aligned with the drift direction, and hence be subject to the same degradation of reconstruction performance as very high angle tracks in the existing TPCs. However, with the magnetic field along x it would still be advantageous to add additional TPCs to cover the top and bottom region of the tracker. There would be a trade-off with the reduction in vertical height, and hence fiducial mass, of the FGDs which would need to be optimised for physics performance. Although the acceptance of the tracking detectors would still not be 4π having full acceptance in the y-z plane would improve sensitivity to cover the majority of the kinematical phase space.

5 Conclusions

In this thesis we started with a summary of the current state of neutrino physics and an overview of the simulation of neutrino interactions in the few-GeV region by the GENIE MC generator. We then gave an outline of the T2K experiment and discussed the importance of constraining neutrino interaction uncertainties affecting charged current π^+ production and how this constraint is necessary to allow T2K to reach its final physics goals.

The main component of this thesis has been a study of neutrino-induced charged current π^+ production at the T2K near detector (ND280). Using the ND280 tracker and surrounding ECals, both an inclusive selection (ν_μ-CC-Inclusive), sensitive to the total rate of charged current ν_μ interactions, and a set of semi-exclusive selections (ν_μ-CC-$\pi^\pm$), sensitive to the neutrino interaction models controlling π^+ production, were developed. In order to remain robust to not-yet-understood detector systematics the ν_μ-CC-$\pi^\pm$ selection aimed to select any neutrino interaction event where there was at least one $\pi^\pm$ in the final state rather than looking explicitly for one and only pion in the final state. Although this results in a more inclusive spectrum of events than the single π^+ background seen at the far detector, it still provides a valuable first comparison of the data MC agreement for neutrino interaction models controlling π^+ production.

The ν_μ-CC-Inclusive selection gave similar results to other ν_μ-CC-Inclusive selections being developed as part of the official ND280 analysis and demonstrated the use of an existing but novel global vertexing algorithm. It will be important for future ND280 analyses to move away from the simple vertexing algorithms which are based on the start of the highest momentum track. We believe that the demonstration of the global vertexing presented here and the basic data MC checks performed represent a valuable step towards this goal.

Two types of ν_μ-CC-$\pi^\pm$ selections were developed: one identifying the $\pi^\pm$ using TPC information, the other using a new ECal PID based on the charge deposited per unit length. This is the first time ECal information has been used to positively select MIPs. This demonstrated the potential increase in geometrical acceptance possible when using the full Barrel- and Ds-ECal. The same technique could be applied to analyses trying to remove background events with multiple MIP-like particles in the final state, such as those trying to select CC quasi-elastic (CC-QEL) interactions.

Using the Run 2 data collected between November 2010 and March 2011 we presented data MC comparisons for the overall rate of both the ν_μ-CC-Inclusive and ν_μ-CC-$\pi^\pm$ selections. For both, the level of agreement was well within the evaluated uncertainties

associated with the neutrino interaction and flux simulations.

The comparison of the rate of $\pi^{\pm}$ production for both the TPC- and ECal-based selections was a first demonstration of the potential use of ND280 beyond the standard CC-Inclusive and CC-QEL measurements. This highlights how useful, with the collection of more data, ND280 will be in constraining neutrino interaction and flux systematics.

Bibliography

[1] Y. Nir, CERN lectures on neutrino physics (2004): http://www.infocobuild.com/education/audio-video-courses/physics/neutrinos-cern.html.

[2] C. Giunti, C. W. Kim, Fundamentals of Neutrino Physics and Astrophysics, Oxford University Press (2007).

[3] R. Davis, D. S. Harmer, and K. C. Hoffman, Phys. Rev. Lett. **20**, 1205 (1968).

[4] J. Bahcall, W. A. Fowler, J. Iben, I., and R. Sears, Astrophys.J. **137**, 344 (1963).

[5] KAMIOKANDE-II Collaboration, K. Hirata et al., Phys.Rev.Lett. **63**, 16 (1989).

[6] GALLEX Collaboration, P. Anselmann et al., Phys.Lett. **B285**, 376 (1992).

[7] D. Abdurashitov et al., Phys.Lett. **B328**, 234 (1994).

[8] K. Hirata et al., Phys.Rev. **D38**, 448 (1988).

[9] J. N. Bahcall and M. Pinsonneault, Phys.Rev.Lett. **92**, 121301 (2004), astro-ph/0402114.

[10] Kamiokande Collaboration, Y. Fukuda et al., Phys.Lett. **B335**, 237 (1994).

[11] R. Becker-Szendy et al., Nucl. Instrum. Meth. **A324**, 363 (1993).

[12] Super-Kamiokande, Fukuda, Y. and others, Phys. Rev. Lett. **81**, 1562 (1998), hep-ex/9807003.

[13] Soudan-2 Collaboration, W. Allison et al., Phys.Rev. **D72**, 052005 (2005), hep-ex/0507068.

[14] MACRO Collaboration, M. Ambrosio et al., Eur.Phys.J. **C36**, 323 (2004).

[15] K2K, E. Aliu et al., Phys. Rev. Lett. **94**, 081802 (2005), hep-ex/0411038.

[16] MINOS Collaboration, P. Adamson et al., Phys. Rev. Lett. **106**, 181801 (2011).

[17] SNO Collaboration, Q. Ahmad et al., Phys.Rev.Lett. **89**, 011301 (2002), nucl-ex/0204008.

[18] KamLAND Collaboration, T. Araki et al., Phys.Rev.Lett. **94**, 081801 (2005), hep-ex/0406035.

[19] ALEPH Collaboration, D. Decamp et al., Phys.Lett. **B231**, 519 (1989).

[20] Delphi Collaboration, P. Aarnio et al., Phys.Lett. **B231**, 539 (1989).

[21] L3 Collaboration, B. Adeva et al., Phys.Lett. **B231**, 509 (1989).

[22] OPAL Collaboration, M. Akrawy et al., Phys.Lett. **B231**, 530 (1989).

[23] K. Nakamura et al. (Particle Data Group), Journal of Physics G37, 075021 (2010) and 2011 partial update for the 2012 edition.

[24] L. Wolfenstein, Phys.Rev. **D17**, 2369 (1978).

[25] S. P. Mikheev and A. Y. Smirnov, Sov. J. Nucl. Phys. **42**, 913 (1985).

[26] K. Eitel, Nucl.Phys.Proc.Suppl. **143**, 197 (2005).

[27] S. F. King, Contemporary Physics **48**, 195 (2007).

[28] H. Murayama, Summary of neutrino experimental results: http://hitoshi.berkeley.edu/neutrino.

[29] A. de Gouva, A. Friedland, and H. Murayama, Physics Letters B **490**, 125 (2000).

[30] T2K Collaboration, K. Abe et al., Phys.Rev.Lett. **107**, 041801 (2011), 1106.2822.

[31] MINOS, P. Adamson et al., Phys. Rev. Lett. **107**, 181802 (2011), 1108.0015.

[32] G. L. Fogli, E. Lisi, A. Marrone, A. Palazzo, and A. M. Rotunno, Phys. Rev. D **84**, 053007 (2011).

[33] Sanchez, Federico, (ed.) and Sorel, M., (ed.) and Alvarez-Ruso, Luis, (ed.), Neutrino-nucleus interactions in the few-GeV region, Proceedings, 6th International Workshop, NUINT-09.

[34] C. Andreopoulos et al., Nucl. Instrum. Meth. **A614**, 87 (2010), 0905.2517.

[35] E. J. Moniz et al., Phys. Rev. Lett. **26**, 445 (1971).

[36] A. Bodek and J. Ritchie, Phys.Rev. **D24**, 1400 (1981).

[37] R. D. Woods and D. S. Saxon, Phys. Rev. **95**, 577 (1954).

[38] H. De Vries, C. De Jager, and C. De Vries, Atom.Data Nucl.Data Tabl. **36**, 495 (1987).

[39] C. Andreopoulos, GENIE talk given at 45 Karpacz Winter School in Theoretical Physics:
http://wng.ift.uni.wroc.pl/karp45/presentations/Andreopoulos-GENIE.pdf.

[40] C. H. Llewellyn Smith, Phys. Rept. **3**, 261 (1972).

[41] R. Bradford, A. Bodek, H. Budd, and J. Arrington, Nuclear Physics B - Proceedings Supplements **159**, 127 (2006).

[42] D. Rein and L. M. Sehgal, Annals Phys. **133**, 79 (1981).

[43] R. P. Feynman, M. Kislinger, and F. Ravndal, Phys. Rev. D **3**, 2706 (1971).

[44] K Nakamura et al. (Particle Data Group) 2010 J. Phys. G: Nucl. Part. Phys. 37 075021.

[45] D. Rein and L. M. Sehgal, Nucl.Phys. **B223**, 29 (1983).

[46] A. Bodek and U. Yang, J.Phys.G **G29**, 1899 (2003), hep-ex/0210024.

[47] M. Gluck, E. Reya, and A. Vogt, Eur.Phys.J. **C5**, 461 (1998), hep-ph/9806404.

[48] GENIE User and Physics Manual: http://www.genie-mc.org.

[49] T. Yang, C. Andreopoulos, H. Gallagher, and P. Kehayias, AIP Conf.Proc. **967**, 269 (2007).

[50] Z. Koba, H. B. Nielsen, and P. Olesen, Nucl.Phys. **B40**, 317 (1972).

[51] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **0605**, 026 (2006), hep-ph/0603175.

[52] D. S. Baranov, et al., An estimate for the formation length of hadrons in neutrino interactions, PHE 84-04.

[53] S. Dytman and A. Meyer, AIP Conf.Proc. **1405**, 213 (2011).

[54] R. Arndt, W. Briscoe, I. Strakovsky, R. Workman, and M. Pavan, Phys.Rev. **C69**, 035213 (2004), nucl-th/0311089.

[55] C. Andreopoulos et al., T2K internal note 35: Neutrino Generator (GENIE) Uncertainties and 2010a ν_μ-Disappearance / ν_e-Appearance Oscillation Analysis Systematics (2011).

[56] C. Andreopoulos, J. Dobson and P. Guzowski, T2K internal note 06: A GENIE-based Event Generation Driver for T2K and MDC0-era Issues (2009).

[57] T2K Collaboration, K. Abe et al., Nucl.Instrum.Meth. **A659**, 106 (2011), 1106.1238.

[58] Y. Fukuda et al., Nucl.Instrum.Meth. **A501**, 418 (2003).

[59] MINOS, S. E. Kopp et al., Nucl. Instrum. Meth. **A554**, 138 (2005), physics/0507211.

[60] N. Abgrall et al., T2K internal note 38: Neutrino flux prediction (2011).

[61] N. Abgrall et al. (the NA61/SHINE Collaboration), CERN-SPSC-2008-018 (2008).

[62] Full details at: http://www.fluka.org/fluka.php.

[63] GEANT Detector Description and Simulation Tool, Application Software Group, Com- puting and Networks Division, CERN, Geneva (1993).

[64] C. Zeitnitz and T. Gabriel, Proc. of International Conference on Calorimetry in High Energy Physics, Tallahasse, FL, USA, February, (1993).

[65] Official plots provided by the T2K beam group.

[66] T2K, A. Vacheret, Nucl. Instrum. Meth. **A623**, 201 (2010).

[67] F. Retiere et al., Nucl.Instrum.Meth. **A610**, 378 (2009).

[68] A. Vacheret et al., Nucl.Instrum.Meth. **A656**, 69 (2011).

[69] A Vacheret et al., The front end readout system for the T2K-ND280 detectors, IEEE Nuclear Science Symposium Conference Record, Vols 1-11 (2007).

[70] T2K and FGD, F. Retiere, Nucl. Instrum. Meth. **A623**, 495 (2010).

[71] Official plots provided by the FGD group.

[72] N. Abgrall et al., Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **637**, 25 (2011).

[73] Y. Fukuda et al., Nucl.Instrum.Meth. **A501**, 418 (2003).

[74] Super-Kamiokande, Y. Ashie et al., Phys. Rev. **D71**, 112005 (2005), hep-ex/0501064.

[75] C. Andreopoulos et al., T2K internal note 64: T2K 1.431×10^{20}-POT Muon-Neutrino Disappearance Analysis (2011).

[76] G. J. Feldman and R. D. Cousins, Phys.Rev. **D57**, 3873 (1998), physics/9711021.

[77] Private communication with C. Andreopoulos of Rutherford Appleton Laboratory, UK (2011).

[78] Y. Hayato, Acta Phys.Polon. **B40**, 2477 (2009).

[79] A. Hillairet et al., T2K internal note 72: ND280 Reconstruction (2011).

[80] The RECPACK reconstruction package is described here: http://ific.uv.es/recpac.

[81] Gustav Wikstrom et al., Global Vertexing: http://www.t2k.org/docs/technotes/046 (2011).

[82] Kendall Mahn et al., Internal studies of the FGD alignment: http://www.t2k.org/nd280/software/alignment/asbuilt/fgd-geometry-status/fgdfv (2011).

[83] Preliminary results from, and private correspondence with authors of, forthcoming T2K internal note 75: ND280 Tracker Tracking Efficiency (2011).

[84] Based on preliminary results shown at the weekly T2K-UK ECal meetings by Callum Lister of Warwick University (2011).

[85] Based on preliminary results shown at NuMu physics group meetings by Dr Thomas Lindner of TRIUMF national laboratory, Canada (2011).

[86] J. Caravaca and F. Sánchez, Results presented within the NuMu physics group meetings: Forthcoming internal note expected (2011).

[87] J. Caravaca et al., T2K internal note 48: Charge Misidentification in local and global reconstruction.

[88] S. Oser, T2K internal note 91: Elemental composition and masses of FGD XY modules (2010).

[89] L. Monfregola et al., T2K internal note 79: Systematics induced by track quality selection in the 2011 numu CC and CCQE analyses (2011).

[90] J. Dobson and C. Andreopoulos, Acta Phys.Polon. **B40**, 2613 (2009).

[91] M. Bonesini, A. Marchionni, F. Pietropaolo, and T. Tabarelli de Fatis, Eur.Phys.J. **C20**, 13 (2001), hep-ph/0101163.

[92] Plot provided courtesy of Thomas Dealtry of Oxford University (2011).