Minimising Systematic Uncertainties in the T2K Experiment Using Near-Detector and External Data

Carl Vincent Clarence Wret

The Blackett Laboratory,
Imperial College London

A Dissertation Submitted to Imperial College London
for the Degree of Doctor of Philosophy

Minimising Systematic Uncertainties in the T2K Experiment Using Near-Detector and External Data

Abstract

The Tokai to Kamioka (T2K) long baseline neutrino oscillation experiment was designed to make precise measurements of neutrino oscillations. It uses a muon (anti-)neutrino dominated beam produced at the Japan Proton Accelerator Research Complex (J-PARC) on the east coast of Japan, aiming towards the Super-Kamiokande (SK) detector 295 km west. The neutrino beam is sampled by the two near detectors ND280 and INGRID, 280 m downstream of production. These measure the neutrino flux and interaction cross-sections to reduce the impact of systematics for oscillation analyses. The work presented herein details the process of using ND280, INGRID, and external data to best constrain the predicted event rates at SK. The analysis proceeds by using a Markov Chain Monte Carlo method which simultaneously fits ND280 and SK data without assumptions on the underlying posterior probability density function. The two analyses detailed herein reduce event rate uncertainties at SK from 12-14% to 2-4%, enabling world-leading oscillation parameters to be extracted from T2K. Numerous run-by-run and detector-by-detector studies were performed and alternate models investigated, all of which were deemed compatible within uncertainty. This body of work has been included in the official T2K results presented in 2017 and 2018, and its use continues beyond that.

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. This dissertation does not exceed the word limit for the respective Degree Committee.

Carl Vincent Clarence Wret

©The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgements

I have had the great privilege to work, travel, and form friendships with a number of excellent people during my three years on T2K and at Imperial College London. Special mentions to all of those who have shared and discussed lovely ideas at length on never-ending flights, speeding Shinkansens, in coffee shops across the world, and on paradisal islands.

Starting my Ph.D. on neutrino interactions and model selections, I am indebted to Patrick Stowell, Callum Wilkinson, Luke Pickering and Ryan Terri for countless hours of discussions and explanations, culminating in the birth of the NUISANCE project [1]—now in use across many neutrino interaction experiments. We could not have done it without the expertise and advice of my supervisor Morgan Wascko, Kendall Mahn, and Kevin McFarland. I also extend gratitude to Minoo Kabirnezhad for our two months in Japan implementing new single pion production models and eating 7-Eleven goods past the early hours.

Continuing the interaction model studies but now using ND280 data—the topic of this thesis—would have been impossible without the help of the current and past members of the MaCh3 analysis group, spearheaded by Asher Kaboth. I hail Asher for his many wisdoms on the fit mechanics, statistics and Markov Chain Monte Carlo, and Richard Calland for fuelling my interest in high performance computing and GPU programming. I thank the members of the frequentist ND280 fitting group (BANFF), led by Mark Scott with Simon Bienstock and Pierre Lasorak, for many discussions and hard work on studying the physics, statistics and systematics in the fit.

I remain grateful for the time spent and the knowledge transferred through my colleagues in the Imperial T2K group: you all made room 530 (and beyond) a treat and joy. A special thank you to Toby Nonnenmacher for his many hours spent reading this thesis. I thank my year-mates in the Imperial HEP group Jack Wright, Sioni Summers, Thibaud Humair, and Slavomira Stefkova for not always discussing physics at work, and making Imperial a welcome place.

A modern High Energy Physics analysis demands excellent computing facilities. The work in this thesis relied on the Compute Canada Guillimin, Cedar and Graham clusters—made possible by Hiro Tanaka—Imperial College High Energy Physics Computing—maintained by Simon Fayer and Ray Beuselinck, with special thanks

to David Colling for GPU support—and the Imperial College High Performance Computing centre. I also acknowledge the generous hardware sponsorship from NVIDIA.

Finally, I extend my deepest gratitude to my close friends and loved ones, my sister, parents, step-parent, and grandmother who have happily stood by me throughout, the most stable of pillars. Thank you for all the entertaining distractions, buckets of patience, endless support, and much love you've given me. I hope you forever remain in my life.

Contents

Chapter 1

Introduction

Neutrino oscillation physics is entering the precision era with the later stages of T2K [2] and NOvA [3], building towards the high statistics Hyper-Kamiokande [4] and DUNE [5] experiments. Nature will soon reveal if neutrinos oscillate differently to anti-neutrinos, the ordering of the neutrino mass states, and the completeness of the PMNS parametrisation [6–8]. Thanks to excellent beam performance with increasing power, both T2K and NOvA have successfully run in muon neutrino and anti-neutrino dominated modes and see consistent results in the latest analyses, with fresh updates presented in Heidelberg at Neutrino 2018 [9, 10]. As the statistics mount, there is a concerted effort towards percent-level systematics in the community, and a future combination of T2K and NOvA results has been agreed to fully exploit the strengths of each experiment [11, 12]. Reaching the desired uncertainties on neutrino oscillation parameters in the GeV region requires significantly improved interaction and flux modelling over the current constraints [13, 14].

Hand-in-hand with detailed study of ν_e appearance at T2K and NOvA, the channel shares oscillation parameters with short baseline reactor experiments such as Daya-Bay [15], measuring $\bar{\nu}_e \rightarrow \bar{\nu}_e$. Differences in oscillation parameters from the two neutrino sources is a probe of un-modelled physics. Similar is true for the IceCube [16] and Super-Kamiokande [17] neutrino observatories, measuring solar and atmospheric neutrinos.

This work concerns the use of external and internal neutrino and hadron scattering data at T2K to aid precision measurements of neutrino oscillations. By carefully studying systematics present at both ND280 and Super-Kamiokande, the uncertainties on event rates at SK are decreased from 12-14% to 2-4%. Such stringent model constraints enable world-leading oscillation measurements from T2K [18], and opens the door to the first observation of $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$, with strong indications from NOvA presented at Neutrino 2018 [10].

The studies also shine light on degeneracies and weaknesses of the T2K simulation from systematic sources, notably the neutrino flux and interaction modelling. By

focusing effort on highlighted systematics, uncertainties can be driven down further. Narrowing down model selection is also crucial to searches at T2K involving ND280 only, such as sterile neutrinos and rare neutrino interaction searches.

This thesis is organised to first present a historical overview alongside theory in chapter 2. It then introduces T2K on the stage of neutrino oscillation physics in chapter 3, giving an overview of the beamline, the ND280 and INGRID near detectors and the Super-Kamiokande far detector. The fitting procedure using Markov Chain Monte Carlo is introduced in chapter 4, followed by an overview of the T2K oscillation analysis chain and the importance of ND280 data in chapter 5. The selections at ND280 are presented, the parameterisation of systematics is detailed, and a fit to Asimov data and then real data is made, rounded off by a discussion of the impact on T2K oscillation analysis.

The second analysis chapter in chapter 6 concerns updating the fit to ND280 data using new selections and systematics, with an almost twofold increase in statistics for both neutrinos and anti-neutrinos. The model constraints are to be used beyond the 2018 analysis, further reducing uncertainties on expected rates at Super-Kamiokande. The thesis finally closes with a summary and concluding remarks.

The presented results are used in the official T2K 2017 [2] and upcoming 2018 publications, including those presented at Neutrino 2018 [9].

Chapter 2

Neutrino Physics

This chapter introduces the experimental evidence for the neutrino and neutrino oscillations, outlines the theory of neutrino oscillations, and discusses neutrino interactions in the $E_\nu \sim 1 - 10\,\text{GeV}$ range, relevant to long baseline neutrino oscillation experiments.

2.1. The Discoveries of the Neutrinos

Neutrinos were initially proposed as a solution to the apparent violation of the conservation of four- and angular momentum in James Chadwick's measurements of beta decay in 1932 [19,20]. Inspired by Wolfgang Pauli's new elementary particle "the neutron"[a] [21], Enrico Fermi built his theory of β-decay [22], in which the observable process $n \rightarrow p + e^-$ is always accompanied by an invisible four-momentum carrier, the electron anti-neutrino.

The neutrino remained elusive until Reines and Cowan in 1953 devised their experiments using the inverse beta decay (IBD) process, $\bar{\nu}_e + p \rightarrow n + e^+$, near a nuclear reactor [23,24]. The experiment consisted of two tanks of water sandwiched by three liquid scintillator tanks with photo multiplier tubes (PMTs). The water was doped with 40 kg $CdCl_2$, which could detect free neutrons through capture. The electron anti-neutrinos were emitted by the nuclear reactor, interacted with the protons in the water, producing a prompt signal from $e^+ + e^- \rightarrow 2\gamma$. The free neutron was detected $\sim 5\mu$s after the prompt 2γ from $n + {}^{108}\text{Cd} \rightarrow {}^{109m}\text{Cd} \rightarrow {}^{109}\text{Cd} + \gamma$. The experiment also took data from a reactor off period, demonstrating a significant reduction in neutrino event rates. Modern reactor neutrino oscillation experiments such as Daya Bay [15] operate much on the same principle. The experiment was complemented by measurements by R. Davis in 1964 [25], which exposed tanks of ^{37}Cl to reactor electron anti-neutrinos, interacting through $\bar{\nu}_e + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar}$, which

[a] Which had characteristics of what we today call a nucleon and a neutrino

would violate lepton number conservation. The experiment found no excess of ^{37}Ar, and instead set limits on the solar neutrino flux.

The field quickly developed after the first measurements and in 1962 Lederman, Schwartz, Steinberger and others [26] observed another flavour of neutrino, the muon neutrino. They used a beam of protons impinging a target, creating a π dominated beam which decayed following $\pi^+ \to \mu^+ + \nu_\mu$, and looked for subsequent interactions of the ν_μ in a 10 tonne shielded aluminium spark chamber. The experiment was later confirmed by measurements at CERN in 1964 [27, 28].

When the third charged lepton, the τ, was discovered at SLAC's e^+e^- accelerator in 1975 [29], the search for its neutrino partner started. Its existence was already hinted at in τ decays and was discovered at DONUT in 2000 [30]. The discovery of the ν_τ and the three neutrino flavours was largely expected from precise measurements of Z decays at the Large Electron Positron (LEP) and the Stanford Linear Accelerator (SLAC), which found the number of active neutrino flavours, assuming the Standard Model, as $N_\nu = 2.9840 \pm 0.0082$ [31]. This has also been confirmed by cosmological data from Planck and others, $N_{\text{eff}} = 3.04 \pm 0.18$[b] [32].

2.2. Neutrino Oscillations

The discovery of neutrino oscillations, detailed in section 2.4, is a direct consequence of neutrino mass. B. Pontecorvo [6, 7, 33], Z. Maki, M. Nakagawa and S. Sakata [8] developed the PMNS formalism, widely used by the oscillation community today. This section highlights some crucial components of the theory and how it has been applied in the field.

The PMNS formalism starts by introducing a neutrino mass eigenstate $|\nu_i\rangle$, which is a linear superposition of the flavour eigenstates participating in the weak interaction $|\nu_\alpha\rangle$ with n neutrino states,

$$|\nu_i\rangle = \sum_{\alpha}^{n} U_{\alpha i} |\nu_\alpha\rangle \tag{2.1}$$

[b] $N_{\text{eff}} = 3.0 \pm 0.4$ and $\sum m_\nu < 0.22$ eV when varying both N_{eff} and $\sum m_\nu$.

where the unitary matrix U is generally expressed as

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \tag{2.2}$$

This echoes that of quark mixing proposed by Cabbibo [34], Kobayashi and Maskawa [35]. The superposition leads to a probability of observing neutrino flavour change from flavour α to β over distance L for a neutrino with energy E in which the square of two neutrino mass states are separated by $\Delta m^2_{ij} = m^2_i - m^2_j$,

$$P(\nu_\alpha \to \nu_\beta) = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 = \sum_{k,j} U^*_{\alpha k} U_{\beta k} U^*_{\alpha j} U_{\beta j} \exp\left(-i \frac{\Delta m^2_{ij} L}{2E}\right) \tag{2.3}$$

Then using the squared unitarity relation we finally get [36],

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4 \sum_{i>j} Re\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin^2(\Delta m^2_{ij} \frac{L}{4E}) \tag{2.4}$$

$$+ (-) 2 \sum_{i>j} Im\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin(\Delta m^2_{ij} \frac{L}{2E}) \tag{2.5}$$

where the negative sign is picked up for anti-neutrinos.

The PMNS matrix is often parameterised into three separate matrices with their own mixing angles θ_{13}, θ_{23} and θ_{12} and a complex phase δ, where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ [36],

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.6}$$

where the $(1,2)$ parameters are referred to as "solar", $(2,3)$ as "atmospheric", and $(1,3)$ as "reactor". The δ is commonly referred to as the CP violating Dirac phase, δ_{CP}.

Reducing down to two neutrino mass states (where the third mass state is approximately degenerate with another mass state), we obtain a simpler mixing matrix

U,

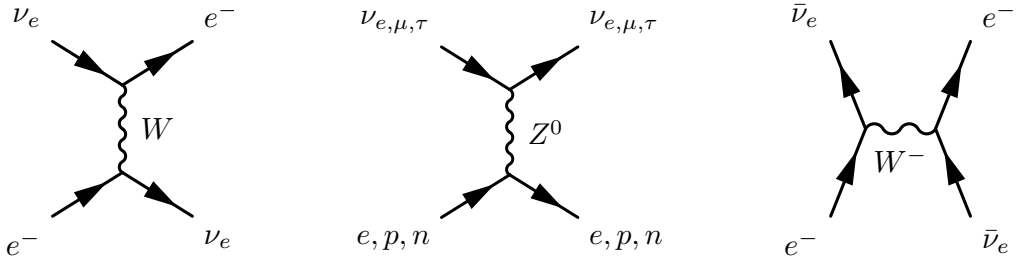
$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{2.7}$$

and an oscillation probability of

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - (+)\sin^2(2\theta)\sin^2\left(\frac{1.267\Delta m^2\left[\mathrm{eV}^2\right] L\,[\mathrm{km}]}{E\,[\mathrm{GeV}]}\right) \tag{2.8}$$

where the positive sign is picked up when $\beta \neq \alpha$. The sinusoidal oscillation of the neutrino flavour states in Equation 2.8 has a period controlled by the parameter Δm^2 and amplitude by the mixing angle θ. The maximum oscillation probability for a fixed mixing angle θ is at $L/E \sim 1.25/\Delta m^2$, which for $\Delta m^2 \sim 2.5 \times 10^{-3}$ eV^2 and $L \sim 300$ km means $E = 0.6$ GeV—placing an experiment like T2K ($L = 295$ km, $E = 0.6$ GeV) near the maximum.

The final ingredient in the oscillation probability is to account for effects from traversing matter rather than vacuum, often referred to as the Mikheyev-Smirnov-Wolfenstein (MSW) effect [37–40]. The effect sets electron neutrinos apart from muon and tau neutrinos, since they have an additional weak interaction with electrons in matter, shown in Figure 2.1.



(a) ν_e on e^- **(b)** $\nu_{e,\mu,\tau}$ on e, p, n **(c)** $\bar{\nu}_e$ on e^-

Figure 2.1.: Interaction diagrams with matter for different neutrino flavours, highlighting (anti-)electron neutrino differences giving rise to different matter-effects for (anti-)electron neutrinos

Electron neutrinos experience a modified Hamiltoninan potential $\Delta V = \pm 2\sqrt{2} G_F E_\nu N_e$, where G_F is the Fermi constant, E_ν is the neutrino energy and N_e is the electron number density of the matter, where the neutrino(anti) picks up the positive(negative) term. The effect modifies the oscillation probability to have stronger dependence

on $\sin\left(\ldots\Delta m^2\ldots\right)$, hence the sign of Δm^2 can be resolved when significant matter interactions occur [41].

2.3. Neutrino Interactions

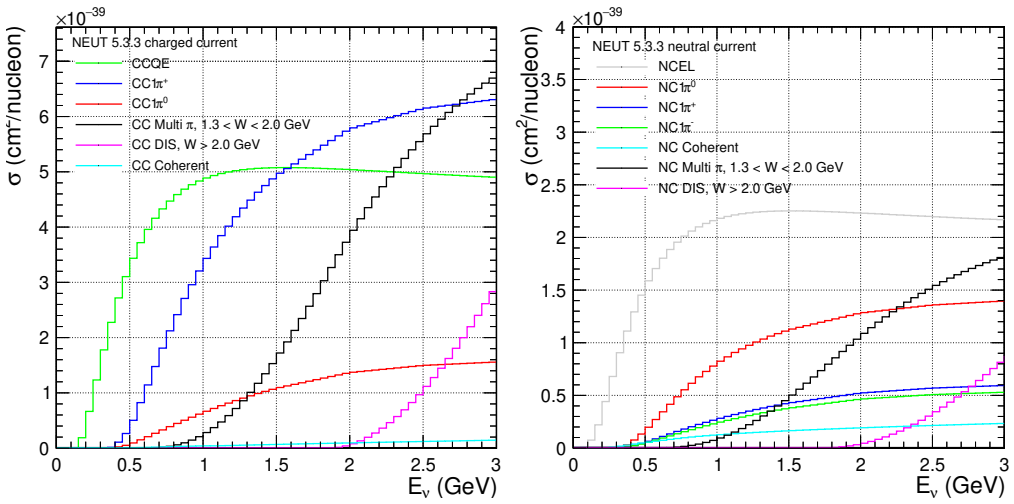
Our limited knowledge of neutrino-matter interactions make them a dominant systematic for current long-baseline, intermediate energy neutrino oscillation experiments. The parametrisation of neutrino interaction systematics for this thesis is given in subsection 5.3.3, with an overview provided here. Detailed discussions can be found in [42–44].

Generally, the $E_\nu \sim 0.5 - 5$ GeV regime is referred to as the "'intermediate energy region". Neutrino-matter interactions in this region involve a range of complex nuclear physics, in which the neutrino can interact with the entire nucleus (e.g. coherent [45] or giant resonance [46] interactions), nucleon pairs, single bound nucleons, and quarks (as in DIS). At the low end of the energy range, additional nuclear effects such as nucleon-nucleon correlations [47, 48], Δ baryon in-medium modifications [49], W-boson self-energy corrections [47], and nucleus spectral functions [50], are important. The high-end of the region is dominated by deep inelastic scattering (DIS), with dependence on parton distribution functions. The transition regions between producing no, single and multiple pions is particularly poorly modelled [42–44]. Neutrino-nucleus and nucleon scattering theory is often inspired by results from electron scattering experiments [51], such as CLAS [52] at JLAB. The electron scattering experiments have the benefit of a very narrow beam energy spectrum, so can study nuclear effects at specific $Q = P_{in} - P_{out}$ to the nucleus—unachievable with current neutrino sources.

In Monte Carlo event generators, the neutrino interaction process is commonly factorised into four parts: 1) A nucleon is simulated from a nuclear model and is used as a target for the neutrino interaction and we boost into its rest-frame, 2) The neutrino interacts with chosen nucleon which is now at rest, equivalent to a neutrino-nucleon interaction, 3) The outgoing particles from the fundamental vertex are propagated through the nucleus with radiative and final state interactions applied, 4) The particles are boosted back into the lab frame.

The total cross-sections in E_ν, $\sigma(E_\nu)$, for the NEUT 5.3.3 [53] generator used by T2K are shown in Figure 2.2. At T2K energies ($E_\nu \sim 0.6$ GeV) the primary interaction mode

is CCQE. Charged pion production becomes important at $E_\nu \sim 1$ GeV, and multi-π and DIS above $E_\nu \sim 2.5$ GeV. Since this analysis aims to minimise systematics for oscillation analyses—which select the charged-current 0π final state at SK—the 0π systematics have the largest impact.



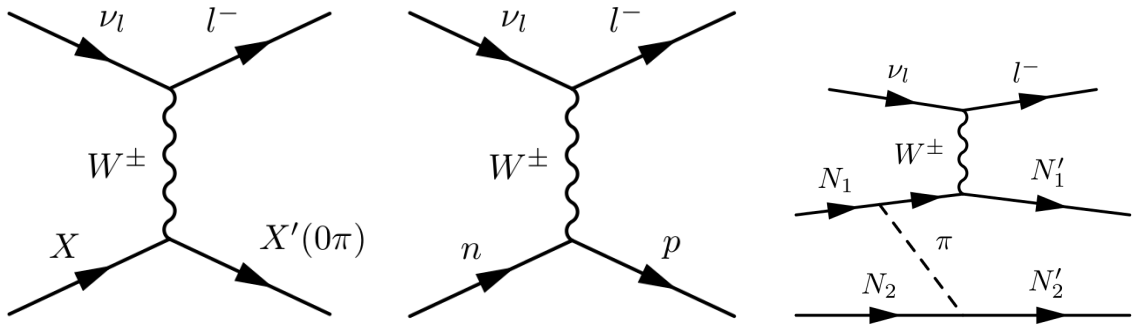
(a) Charged current **(b)** Neutral current

Figure 2.2.: Total cross-sections from the NEUT 5.3.3 [53] neutrino interaction generator

2.3.1. CC0π

Figure 2.3 shows pseudo diagrams for the definition of the 0π, CCQE and one 2p2h process. The CC0π signal definition does not include hadronic information, so the CCQE and 2p2h processes both produce 0π final states. T2K uses the CCQE diagram for neutrino energy reconstruction, assuming a nucleon at rest. If 2p2h events are included in the selection it biases E_ν, and wrongly estimating the bias has a noticeable effect on oscillation parameters. The same holds true for including single pion events due to unreconstructed pions or final state interactions.

Generally, the neutrino-nucleon CCQE interaction is relatively well understood. The current effort in the field is to understand the impact of form factor choices, numerous nuclear effects, and minimising the models' impacts on cross-section and oscillation measurements.



(a) CC0π, defined as having an outgoing charged lepton and no charged or neutral pions

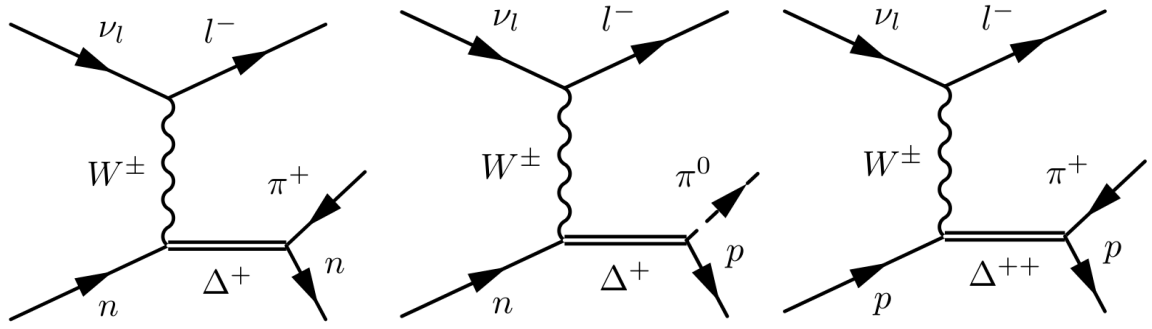
(b) CCQE, defined as having interacted on an initial state neutron, producing an outgoing charged lepton and proton

(c) One of many 2p2h processes, defined as an interaction on two initial state nucleons which are coupled

Figure 2.3.: CC0π, CCQE and 2p2h pseudo-diagrams

2.3.2. Single Pion Production

Figure 2.4 shows the dominant charged-current neutrino-nucleon interactions giving rise to single pion production (SPP). The interaction proceeds by a resonant state here labelled as Δ, which is dominant at T2K energies. SPP makes up $\sim$20% of selected 0π events at SK due to missing pions.



(a) CC1π^+1n **(b)** CC1π^0 **(c)** CC1π^+1p

Figure 2.4.: Charged-current single pion production on a nucleon via a Δ resonance

Contrary to the CCQE interaction, single pion production on free nucleons is poorly modelled. The low hadronic mass (W) regime where the single $I_{3/2}$ $\Delta^{++} \rightarrow p + \pi^+$ interaction dominates can be considered understood, although resonance-resonance and resonance-non-resonance interference present at higher W is not. Additionally, in-

medium effects from resonance fields propagating the nucleus are poorly understood and multi-nucleon couplings often unmodelled [42, 54].

2.3.3. Multi-π and DIS

The transition from SPP to DIS is generally referred to as soft inelastic scattering (SIS). NEUT uses a custom interpolation between $1.3 < W < 2.0$GeV, selecting events with $N_\pi > 1$ to avoid double counting SPP cross-sections. Other generators such as GENIE [55] and NuWro [56] use different implementations. A pseudo diagram is shown in Figure 2.5b, where fragmentation causes the pion emissions.

2.3.4. Subdominant Interactions

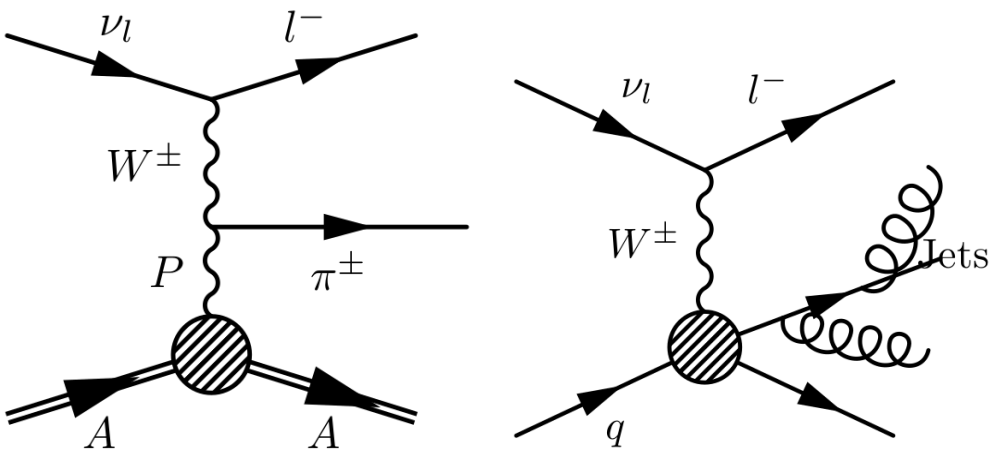
Subdominant interactions with small interaction cross-sections can populate very specific signal regions, e.g. NC1γ and NC1π^0 mimicking $\nu_e + X \rightarrow e^- + X'$, and differences in ν_e vs ν_μ. Historically, there is much less cross-section data on NC interactions than CC and barely any using electron neutrinos, and it is common to trust model extrapolation from CC to NC rather than test against data.

The charged current coherent process shown in Figure 2.5a occupies a very specific phase space: the very forward going, collinear lepton-pion, low Q^2 region. This is important since it has the same final state as single pion production, although not produced through a resonance, making up about 10% of the most forward-going $\cos\theta_\mu$ bin in CC1π selections.

2.3.5. Intranuclear Hadronic Cascades

The nuclear cascade following the initial neutrino-nucleon interaction is handled by a microscopic hadron propagation in NEUT. The hadron interaction probabilities are re-calculated as the hadron travels through the nucleus, seen in Figure 2.6.

Simulations can be considered to agree relatively well with pion scattering data [57], but there are concerns that extrapolating results into the neutrino-nucleus interaction is poorly justified [43]. More sophisticated models exist in the Giessen-Boltzmann-Uehling-Uhlenbeck (GiBUU) [58] generator, although its role as a primary generator in neutrino physics is currently unviable due to computational requirements.



(a) CC coherent **(b)** CC DIS

Figure 2.5.: Coherent and multi-pion/DIS scattering diagrams

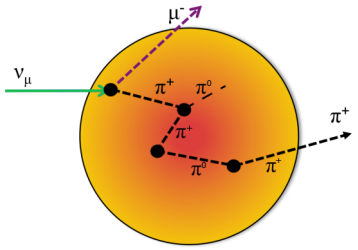


Figure 2.6.: An example of a pion FSI cascade

2.4. Experimental Overview

Neutrino oscillations is now an established physics phenomenon, cemented by awarding the 2015 Nobel Prize in Physics to Kajita-san (SK) and Art McDonald (SNO) for their experiments' measurements of solar and atmospheric neutrino oscillations. This section gives a brief introduction and overview of neutrino oscillation experiments and production mechanisms, summarised in Figure 2.7. Experiments are generally categorised in neutrino energy E and baseline L, and the combination of L/E largely determines which oscillation parameter(s) an experiment has sensitivity too.

2.4.1. Solar Neutrinos

Solar neutrinos emanate from various nuclear fusion products and decays in the sun. Table 2.1 shows the fluxes for various sources, where the pp flux is strongest. However,

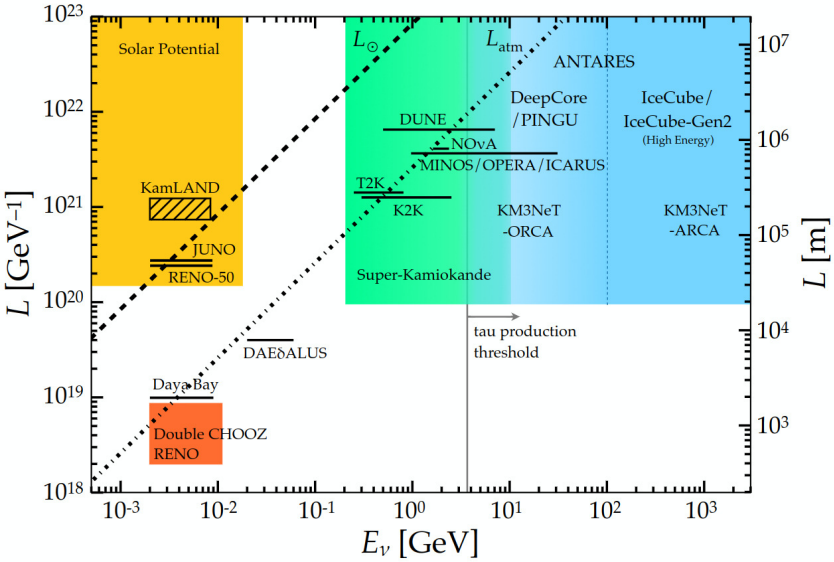


Figure 2.7.: Neutrino oscillation experiments in baseline L and energy E. Figure from [59]

the neutrino energy is often below threshold for the largest contributors to the flux, and most solar neutrino experiments measure the ^{8}B flux, shown in Figure 2.8.

Reaction	Label	Flux (cm^{-2}s^{-1})
$p + p \rightarrow ^2\mathrm{H} + e^+ + \nu_e$	pp	5.95×10^{10}
$p + e^- + p \rightarrow ^2\mathrm{H}\nu_e$	pep	1.40×10^8
$^3\mathrm{He} + p \rightarrow ^4\mathrm{H} + e^+ + \nu_e$	hep	9.3×10^3
$^7\mathrm{Be} + e^- \rightarrow ^7\mathrm{Li} + \nu_e$	^{7}Be	4.77×10^9
$^8\mathrm{B} \rightarrow ^8\mathrm{Be} * + e^+\nu_e$	^{8}B	5.05×10^6

Table 2.1.: Integrated solar neutrino flux from various solar processes in the pp chain. Table replicated from [60]

R. Davis and J. Bachall continued their 1964 measurements [25] of the solar neutrinos from ^{8}B and in 1968 [62] announced a solar ν_e flux a factor of seven below the expected ($\sim 2\sigma$ significance), at the time attributed to inaccurate solar model calculations. This was the birth of the "solar neutrino problem", which Bruno Pontecorvo and Vladimir Gribov in 1969 [33] proposed solving by invoking a $\nu_e \leftrightarrow \nu_\mu$ oscillation similar to $K^0 \leftrightarrow \bar{K}^0$, giving rise to the PMNS paradigm in its current form, first developed around the early 1960s. In 1989, the Kamiokande experiment [63] confirmed the 1968 measurements of Davis and Bachall, measuring a solar neutrino flux from ^{8}B of ~ 0.5 the expected, agreeing with the higher statistics data from Homestake [64]. The solar neutrino deficit was confirmed from the low threshold detectors

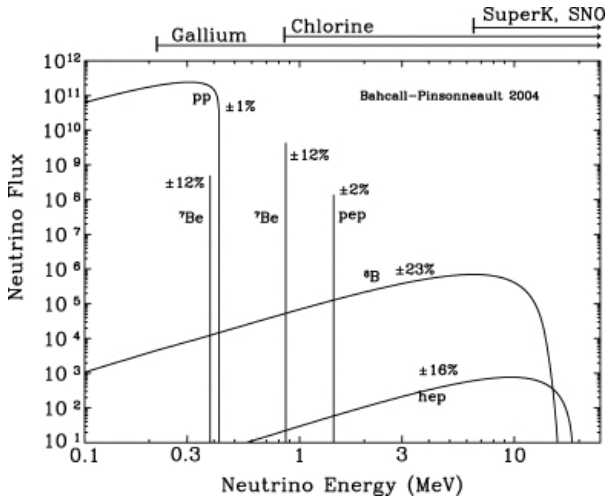


Figure 2.8.: Solar flux from different solar fusion processes, including thresholds of experiments. Figure from [61]

SAGE [65] and GALLEX [66], additionally capable of detecting pp neutrinos using $^{71}\text{Ga} + \nu_e \rightarrow ^{71}\text{Ge} + e^-$.

The Sudbury Neutrino Observatory (SNO) put the nail in the coffin in 2002 [67] by measuring solar neutrinos from ^{8}B in three channels: $\nu_e + d \rightarrow p + p + e^-$ (CC), $\nu_x + d \rightarrow p + n + \nu_x$ (NC) and $\nu_x + e^- \rightarrow \nu_x + e^-$ (ES). The measured neutrino rates had a ν_e component consistent with previous measurements, a strong non-ν_e component 5.3σ above zero, and a NC component consistent with predictions from solar models.

Additionally, the low threshold, low background, Borexino experiment detected solar neutrinos from the ^{8}B, ^{7}Be, pep, and pp processes [68]. The final stages of Borexino aims to measure the CNO cycle and the next-generation SNO+ experiment aims to confirm and improve these measurements, and make detailed measurements of the MSW effect, solar metallicity and luminosity [69].

Although the solar neutrino oscillation parameters Δm^2_{21} and θ_{12} are considered well-constrained, there is $\sim 2\sigma$ tension on Δm^2_{21} between the solar neutrino measurements at SK and SNO (which are internally compatible), and the long baseline reactor anti-electron-neutrino experiment, KamLAND [70], measuring the same oscillation process but with a different neutrino source.

2.4.2. Atmospheric Neutrinos

Atmospheric neutrinos are emitted when cosmic rays interact with nuclei in the earth's atmosphere, producing mesons which decay into neutrinos, amongst other particles. The primary decay is the pion decay,

$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu})$$
$$\mu^{\pm} \rightarrow e^{\pm} + \bar{\nu}_{\mu}(\nu_{\mu}) + \nu_{e}(\bar{\nu}_{e})$$

giving rise to a total of three neutrinos. The atmospheric neutrino flux calculations from Honda [71] is shown in Figure 2.9, which peaks in the 1-10 GeV region, notably higher than the solar neutrinos.

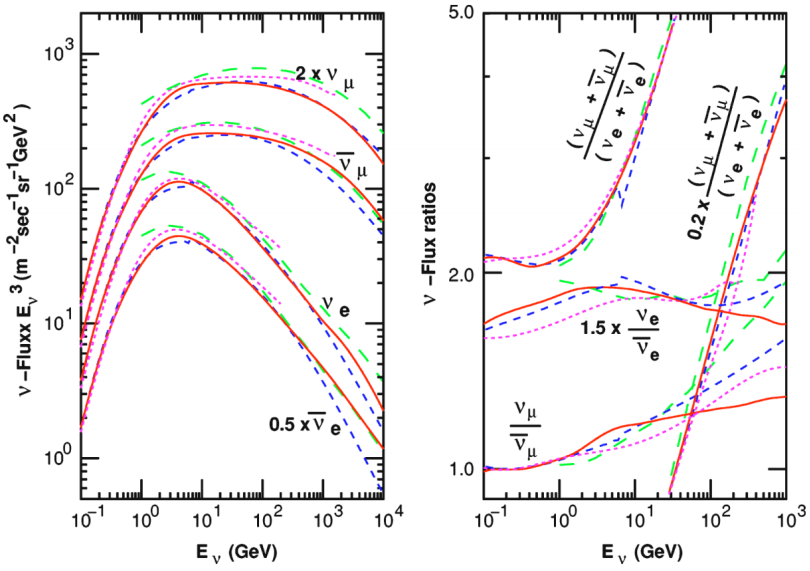


Figure 2.9.: Atmospheric neutrino flux from Honda et al. [71]

In 1965 F. Reines [72] and C.V. Achar [73] first saw hints of atmospheric ν_{μ} disappearance in deep underground laboratories through $\nu_{\mu}(\bar{\nu}_{\mu}) + X \rightarrow \mu^{\pm} + X'$. The Irvine-Michigan-Brookhaven (IMB) experiment observed deficits of ν_{μ} interactions in 1986 [74], and Kamiokande II in 1988 [75] verified this and found muon-like events of $59 \pm 7\%$ the prediction, although good agreement of electron-like single-prong events. The Soudan-2 experiment [76] also saw muon neutrino deficiency with a flavour ratio of $0.72 \pm 0.19^{+0.05}_{-0.07}$ relative the expectation.

When SK in 1998 published [77] their high-statistics[c] ν_μ data, they found $R = (\mu/e)_{\text{Data}} \,/\, (\mu/e)_{\text{MC}} = 0.65 \pm 0.05 \pm 0.08$. They additionally fitted the oscillation parameters, finding the data well described by $\nu_\mu \leftrightarrow \nu_\tau$ rather than $\nu_\mu \leftrightarrow \nu_e$ oscillations. The summary of flavour ratios for atmospheric neutrinos is seen in Figure 2.10, where the majority of the high precision data sits at $R = 0.5 - 0.8$.

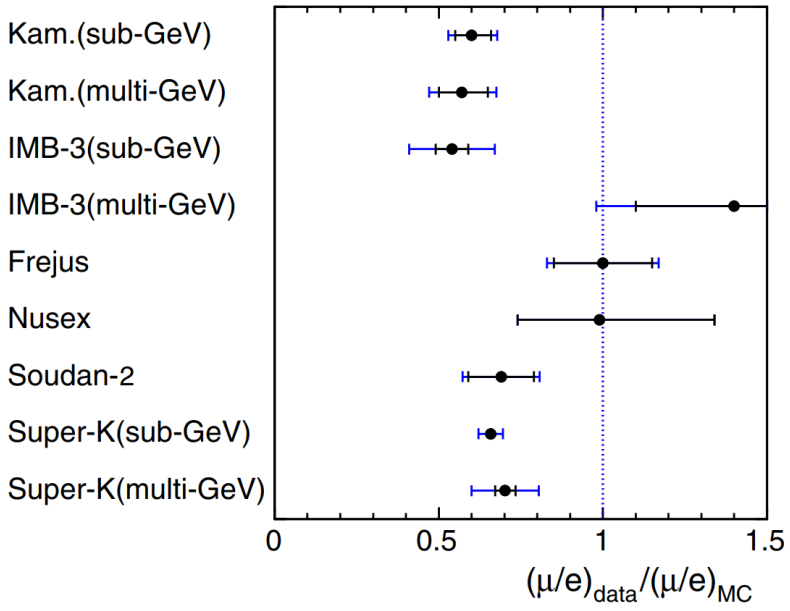


Figure 2.10.: Measured flavour ratios for various historic atmospheric neutrino experiments, showing consistent under-prediction of the μ/e ratio in data and simulation. Figure from [78]

Atmospheric neutrino observatories after the mid 2000s have focussed on measuring $\nu_\mu \to \nu_\mu$ with increasing precision. Furthermore, by isolating regions of specific zenith angle (and therefore baseline L), the extent of the matter effects are also studied, which may resolve the ordering of the mass states. This is largely the focus of the atmospheric neutrino programmes at IceCube [16], ANTARES [79], SNO [80] and SK [17]. The latter has also made attempts at isolating ν_τ events [81], claiming 4.6σ discovery of ν_τ appearance in 2017. A summary of some recent results including complementary long baseline accelerator neutrino experiments can be seen in Figure 2.11.

[c] 4353 fully-contained and 301 partially-contained events

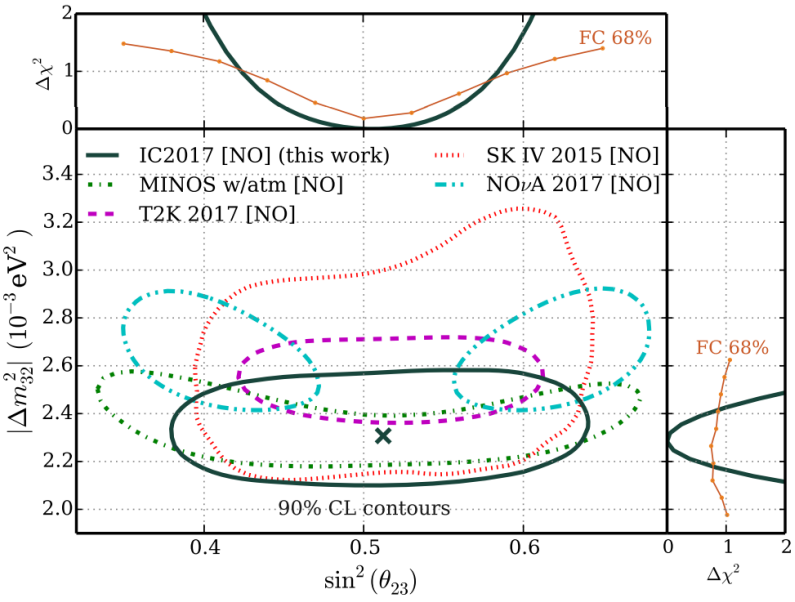


Figure 2.11.: Measured atmospheric oscillation parameters from recent atmospheric and long baseline accelerator neutrino experiments, assuming normal ordering. Figure from [16]

2.4.3. Accelerator Neutrinos

Accelerator neutrino experiments are similar to atmospheric experiments in baseline, neutrino energy, and the neutrino production mechanism. The neutrinos are made by impinging protons from accelerators on targets, producing a flurry of mesons which typically decay into muon flavoured neutrinos, amongst others. Experiments often have the ability to deflect and/or focus mesons after the target, enabling sign selection and thus ν_μ and $\bar{\nu}_\mu$ selection. In contrast to atmospheric neutrinos, accelerator neutrinos are about 90% muon flavoured.

The atmospheric neutrino oscillations, outlined earlier, became the driver behind the accelerator programme in the late 1990s. Since the neutrino energy and baseline can be tuned and chosen in an accelerator experiment, the oscillation dip suggested by atmospheric neutrino experiments can be bombarded with statistics. Furthermore, the dependence on the atmospheric flux simulation is removed [82]. The disadvantage is the reduced total flux at the far-detector, generally forcing the baseline to $L < 1000$ km, which limits the impact of the matter effect and sensitivity to the mass ordering. The majority of long baseline accelerator neutrino experiments include a near detector which samples the beam before any long baseline oscillations have taken place.

The short-baseline ($L \sim 1$ km) accelerator neutrino experiments, such as MiniBooNE [83], MINERνA [84], and the upcoming SBND programme [85], are generally intended to measure neutrino cross-sections and perform short baseline oscillation searches. They may also serve as neutrino beam monitors for other experiments. The interaction measurements are used to inform neutrino event generators [53, 55, 56], aiding in informing systematic uncertainties for neutrino cross-section and oscillation experiments.

The pioneering long-baseline ($L \sim 100 - 1000$ km) experiments MINOS [86] and K2K [87] confirmed the atmospheric neutrino mixing in $\nu_\mu \to \nu_\mu$, finding compatible oscillation parameters. The searches for $\nu_\mu \to \nu_e$ at K2K and MINOS were not statistically significant [88, 89] and were discovered by the next generation experiments T2K [90] and NOvA [91], with evidence of the $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillation presented by NOvA at Neutrino 2018 [10]. The Japanese experiments K2K and T2K have consistently used the 50,000 tonne water Cherenkov detector SK [17] as the far detector, with plastic scintillator based near-detectors and a baseline of $L \sim 250 - 300$ km and $E \sim 0.5 - 2$ GeV. Both MINOS and NOvA, with $L \sim 700$ km and $E \sim 2 - 5$ GeV, use(d) purpose-built functionally identical near and far-detectors, allowing for many uncertainties from systematic parameters to be reduced significantly.

In Europe, the OPERA [92] experiment was designed to look for the dominant $\nu_\mu \to \nu_\tau$ oscillation at $L \sim 700$ km. The ICARUS [93] experiment searched for $\nu_\mu \to \nu_e$ from a fourth "sterile" neutrino, observed by LSND [94] and MiniBooNE [95] which has been questioned by the community [96]. The detection threshold for the charged current interaction $\nu_\tau + X \to \tau + X'$ is $E_\nu \sim 3.5$ GeV, so the neutrino beam from CERN to Gran Sasso (CNGS) [97] was wide-band with $E_\nu = 10 - 25$ GeV. The τ detection requires very fine granularity and OPERA used nuclear emulsions whereas ICARUS pioneered the use of liquid argon TPCs in neutrino physics. OPERA measured ν_τ appearance [98] at 6.1σ, and both OPERA and ICARUS found no evidence of sterile neutrinos [99, 100].

2.4.4. Reactor Anti-Electron Neutrinos

Reactor neutrinos are formed in β decay of fission products in nuclear reactors, e.g. $^{231}\text{Th} \to ^{231}\text{Pa} + e^- + \bar{\nu}_e$ and $^{215}\text{Po} \to ^{211}\text{Pb} + e^- + \bar{\nu}_e$. The neutrino flux depends on the relative fission yields of the products, but generally has a similar energy spectrum

to solar neutrinos, in the 1-10 MeV range. The $\bar{\nu}_e$ flux from a test reactor is shown in Figure 2.12.

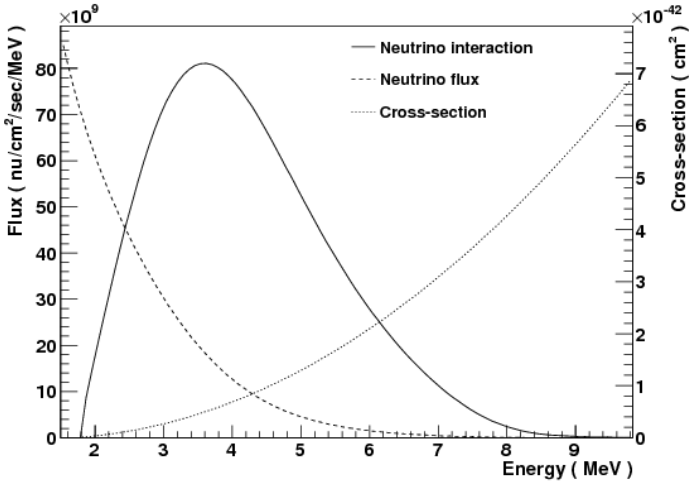


Figure 2.12.: Reactor flux for the Japanese experimental fast reactor, JOYO. Figure from [101]

Reactor anti-electron neutrinos are exclusively detected by the IBD interaction, in which the $e^+ + e^- \rightarrow 2\gamma$ annihilation photons are measured in scintillator. Many experiments additionally dope or surround the scintillator with a high neutron capture element (e.g. ^{6}Li or ^{157}Gd). In the case of Gd doping, the signal consists of the prompt 2γ followed by a $\sim 30\mu$s delayed γ cascade with $E_\gamma^{tot} \sim 8$ MeV from the Gd de-excitation, facilitating signal-background separation [15, 102].

Similar to accelerator neutrino experiments, the reactor experiments can be categorised by baseline. Short baseline experiments with $L \sim 1 - 2$ km perform world-leading measurements of $|\Delta m^2_{13}|$ and $\sin^2\theta_{13}$ and probe parts of the sterile neutrino spectrum. Daya Bay [103], RENO [104] and Double Chooz [105] all measured a relatively large $\sin^2\theta_{13}$, enabling ν_e appearance to be observed at long baseline neutrino experiments such as T2K and NOvA. The short baseline reactor results on $\sin^2\theta_{13}$ are often used in atmospheric and accelerator oscillation analyses for increased sensitivity to the 2,3 parameters, the mass ordering, and δ_{CP}. A summary plot of the measured neutrino oscillation parameters by short baseline reactor and long baseline accelerator neutrino oscillation experiments is shown in Figure 2.13.

The only medium baseline experiment ($L \sim 50$ km) is JUNO [111], currently under construction in China. It aims to measure the neutrino mass ordering by separating the

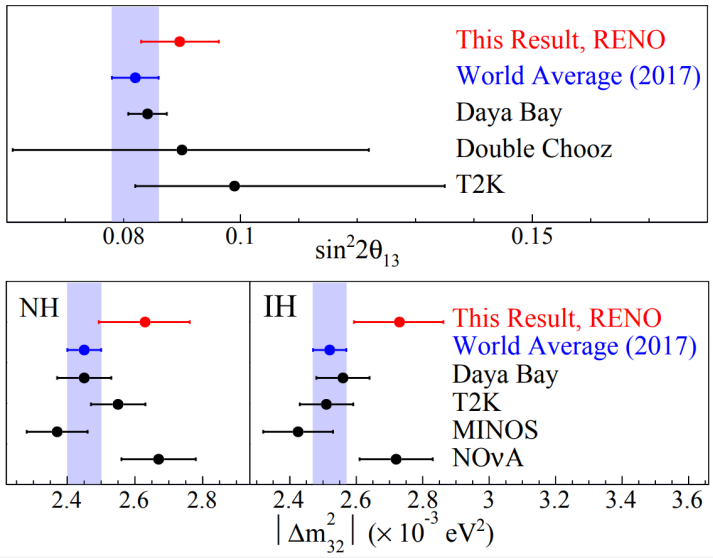


Figure 2.13.: Δm^2_{23} and $\sin^2 2\theta_{13}$ measurements from reactor (Daya Bay [15], RENO [106] and Double Chooz [107]) and accelerator (T2K [108], NOvA [109] and MINOS [110]) neutrinos. Figure from [106]

oscillations into fast and slow parts from Δm^2_{23} and Δm^2_{12}, and improve measurements of $\sin^2\theta_{12}$.

The Japanese KamLAND experiment is the only long baseline ($L \sim 180$ km) reactor anti-neutrino experiment to have run. It measured $\bar{\nu}_e$s from 56 Japanese nuclear power reactors with good sensitivity to Δm^2_{21}. Additionally, combining KamLAND with SNO and SK solar data reduces uncertainties on Δm^2_{21} and $\tan^2\theta_{12}$, as shown in Figure 2.14. These results are used as priors in the T2K oscillation analyses.

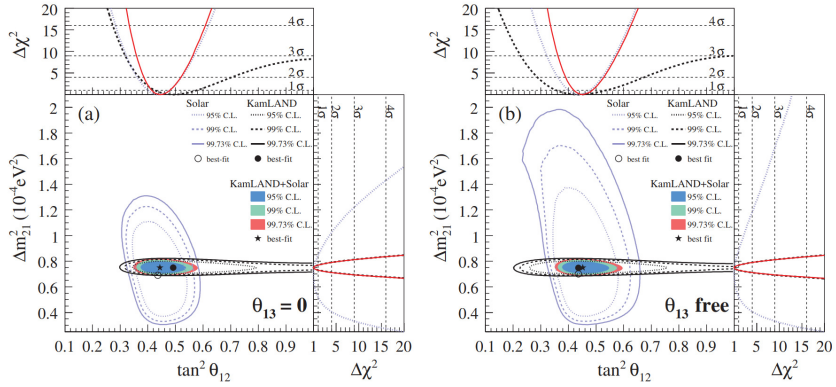


Figure 2.14.: Δm^2_{21} and θ_{21} measurements from KamLAND, SNO and SK (Solar). Figure from [112]

The short baseline reactor experiments at ~ 1 km have consistently measured neutrino event excess at $E_\nu \sim 5$ MeV [15, 102, 105], which is currently unresolved. The culprit is claimed to be either poor neutrino flux modelling or a sterile neutrino [113, 114]. As a result, very short baseline ($L \sim 10 - 20$ m) experiments NEOS [115], DANSS [116], PROSPECT [117], STEREO [118] and SoLiδ [119] have been commissioned to measure $\bar{\nu}_e$ disappearance. None have found evidence of a sterile neutrino, and all have confirmed the 5 MeV excess.

Chapter 3

The Tokai to Kamioka Experiment

The Tokai to Kamioka (T2K) experiment in Japan was proposed and designed in the early to mid 2000s with the intent of observing electron neutrino appearance, alongside precision measurement of muon neutrino disappearance [120, 121]. In 2014 T2K were first to observe electron neutrino appearance, finding a larger mixing angle $\sin^2\theta_{13}$ than measured by the reactors [90], albeit to low significance. NOvA confirmed the appearance measurement in 2016 [91], finding similar values of $\sin^2\theta_{13}$.

The current effort in the $\nu_\mu \rightarrow \nu_e$ channel is reducing the allowed phase space of the CP violating Dirac phase, δ_{CP}, and continuing the measurements of $\sin^2\theta_{13}$. Including the $\nu_\mu \rightarrow \nu_e$ channel(s) also significantly reduces $\sin^2\theta_{23}$ uncertainties [10]. T2K is currently comparing neutrino and anti-neutrino oscillation parameters, seeing weak hints of differences. Furthermore, the preferred mass ordering is normal, although again at weak significance [2].

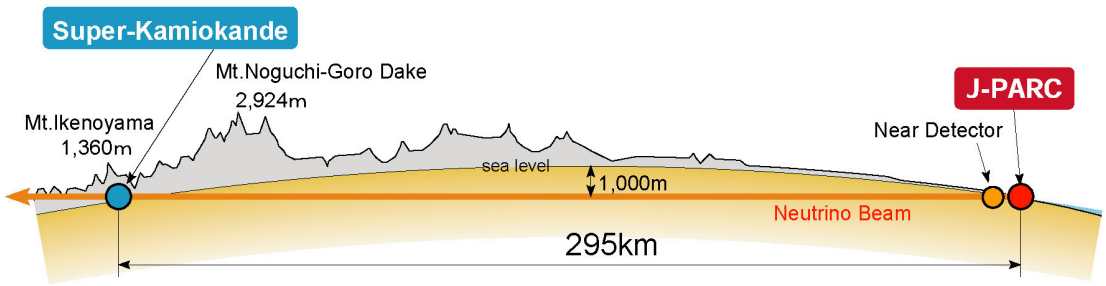


Figure 3.1.: The T2K experiment where neutrinos are created at the J-PARC complex in Tokai. The neutrino beam is characterised at the near-detectors 280 m downstream. 295 km west is the Super-Kamiokande far-detector, measuring the oscillated neutrino spectrum

The neutrinos at T2K are born from particle decay—primarily $\pi^\pm$, $K^\pm$ and $\mu^\pm$—after a proton beam impinges on a target at the J-PARC complex on the Japanese coast, $\sim$120 km north east of Tokyo. The neutrino beam is measured by the ND280 and INGRID detectors $\sim$280 m downstream of the target and a muon monitor station,

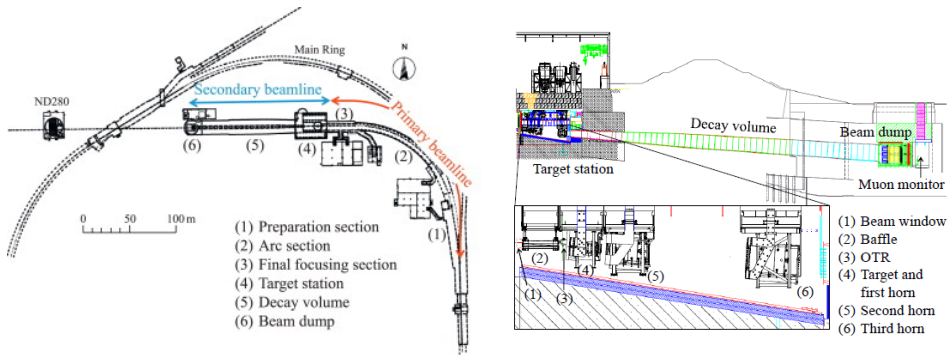
which provide information on the neutrino flux, directionality and interaction cross-section. The far detector, Super-Kamiokande, sits 295 km downstream of the target station and measures the rate of neutrino interactions on 50,000 tonnes of purified water, detailed in section 3.4. A schematic of the neutrinos' travel is shown in Figure 3.1.

Figure 3.2.: The suite of near-detectors at 280 m from the target, showing ND280 and INGRID

A host of other neutrino detectors—not used in this analysis—sit in the same "pit" 280 m from the target station as ND280 and INGRID. Examples include the liquid emulsion NINJA experiment [122], the water target WAGASCI [123] experiment and its magnetic calorimeter Baby-MIND [124].

3.1. Beamline

The J-PARC complex [125] is used to accelerate protons using a linear accelerator (LINAC), a rapid cycling synchrotron (RCS) and a main ring (MR) synchrotron. The MR has the ability to fast-extract into the neutrino beamline with a design power of 750 kW with proton momentum of 30 GeV/c, using $\sim 3 \times 10^{14}$ protons per spill with 8 proton bunches per spill and a spill cycle of ~ 0.5 Hz. The spill width, which opens the trigger window at the neutrino detectors, is $\sim 5\mu$s [126].



(a) Top-view **(b)** Side-view of target-dump section

Figure 3.3.: The neutrino beamline for neutrinos at J-PARC

The neutrino beamline consists of two parts and is shown in Figure 3.3: the primary beamline—which takes fast-extracted protons from the MR, bends them to point towards SK, and impinges them on a graphite target—and the secondary beamline—which directs the mesons from the proton-target interaction through a decay volume, finishing with a beam dump. Shortly after the target in the secondary beamline are three magnetic horns [127] which are used to deflect (focus) wrong-sign (right-sign) mesons to reduce wrong-sign and enhance right-sign neutrinos. Running the magnets at 250 kA (1.7 T) increases the neutrino flux at SK by factor ~ 17 [128]. Taking data in ν_μ dominated mode is referred to as Forward Horn Current (FHC) mode, and in $\bar{\nu}_\mu$ dominated mode as Reverse Horn Current (RHC) mode. After passing the magnetic horns, the focused mesons pass through a ~ 96 m decay volume in which the majority of them decay. Remaining particles strike the beam dump, which stops all hadrons. Surviving high momentum muons ($p_\mu > 5.0$ GeV/c) generally pass through the beam dump, after which they are measured by muon monitors. The MUMON muon monitors (one ionisation chamber and one silicon PIN photodiode) infer the neutrino beam direction to better than 0.25 mrad and the beam intensity better than 3% [129, 130], and are used to inform the beam simulation group.

The neutrinos come primarily from three meson decays

$\pi^+ \to \mu^+ + \nu_\mu$	99.99%	$K^+ \to \mu^+ + \nu_\mu$	63.6%	$K^0_L \to \pi^- + e^+ + \nu_e$	40.6%
$\to e^+ + \nu_e$	10^{-4}%	$\to \pi^0 + e^+ + \nu_e$	5.1%	$\to \pi^- + \mu^+ + \nu_\mu$	27.0%
$\to \mu^+ + \nu_\mu + \gamma$	2×10^{-4}%	$\to \pi^0 + \mu^+ + \nu_\mu$	3.5%		

and one leptonic decay [18]

$$\mu^+ \rightarrow e^+ + \bar{\nu}_\mu + \nu_e \qquad 100\%$$

In Figure 3.4 we trace back each neutrino's parent meson. The π parent is clearly dominant for both the ν_μ and $\bar{\nu}_\mu$ fluxes, although the portion with $E_\nu > 3$ GeV consists of neutrinos whose parents are K mesons. The ν_e and $\bar{\nu}_e$ components of the neutrino beam come primarily from the leptonic μ decay below $E_\nu = 1$ GeV and from K^+ and K^0_L at $E_\nu = 2$ GeV. Tertiary decay products, e.g. a π^- from a K^0_L decay, form large portions of the wrong-sign background.

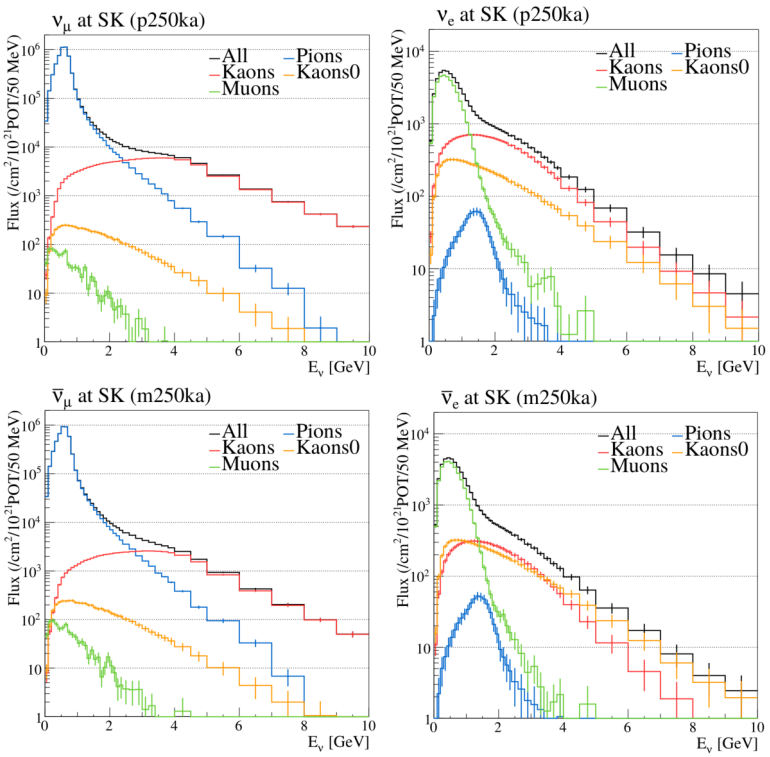


Figure 3.4.: Simulated right-sign neutrino fluxes at SK, showing the neutrino's parent

T2K was the first long baseline accelerator experiment to use the "off-axis technique" in which the far-detector is offset from the neutrino beam center [131]. This has three main effects: 1) it focuses the neutrino energy spectra into a narrower "peak", and 2) it reduces the wrong-sign background for ν_e appearance searches and 3) it lowers the overall rate since most neutrions are forward-going. Assuming the neutrino parents are solely charged pions from $\pi^+ \rightarrow \mu^+\nu_\mu$, we can approximate the neutrino energy E_ν as a function of pion-neutrino angle $\theta_{\pi,\nu}$ (colloquially "off-axis angle"), pion energy

E_π and mass m_π, and the muon mass m_μ,

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

For a chosen $\cos\theta_{\pi,\nu}$ there is a maximum pion energy of $E_\pi^{\max} = p_\pi / \cos\theta_{\pi,\nu}$ giving a maximum neutrino energy of

$$E_\nu^{\max} = \frac{m_\pi^2 - m_\mu^2}{2E_\pi \sin^2\theta_{\pi,\nu}} \tag{3.2}$$

which maximises when π and ν approach collinearity, and as θ increases the allowed neutrino energy spectrum becomes smaller. The calculated flux at SK with the neutrino oscillation probability is shown in Figure 3.5. The off-axis angle is chosen to maximise the flux in the primary oscillation dip at $E_\nu \sim 0.6$ GeV.

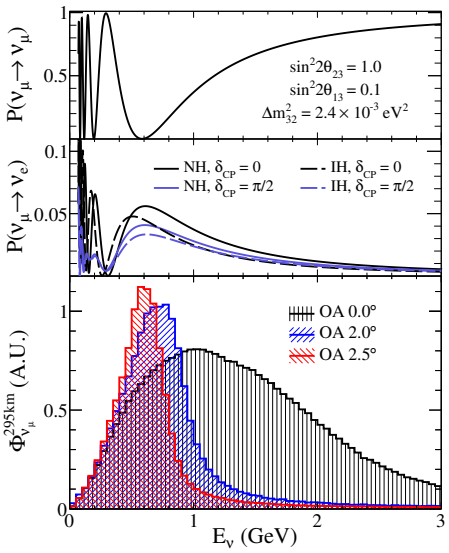


Figure 3.5.: Effect of off-axis (OA) angle on the SK neutrino flux prediction, overlaid with the neutrino oscillation spectrum

The delivered beam power and accumulated POT has been steadily increasing from run 1 in 2010 to run 9 in 2019. Run 9 concluded with $\sim$500 kW beam power, having accumulated a total POT of $\sim 3.16 \times 10^{21}$[a]: 1.51×10^{21} in FHC and 1.65×10^{21} in RHC modes, shown in Figure 3.6.

[a]Or 5.28 mg of protons

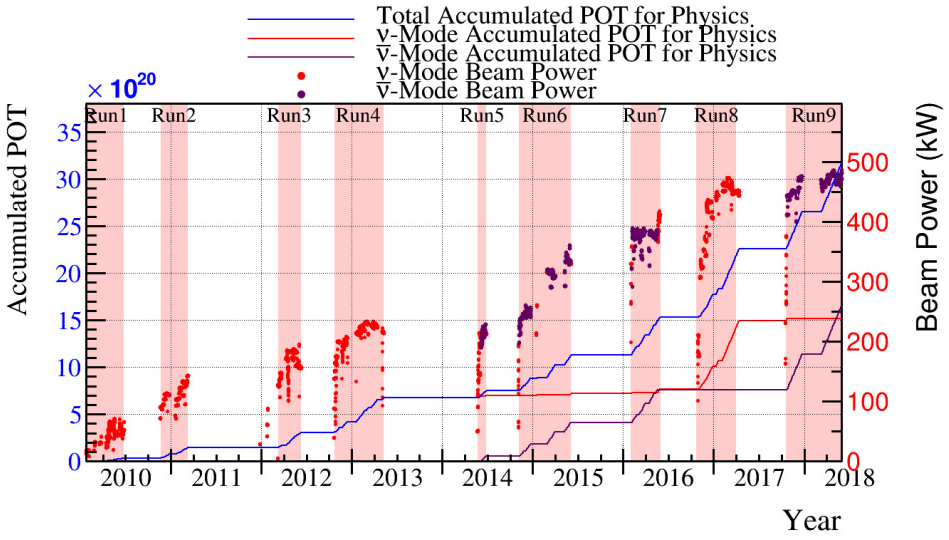
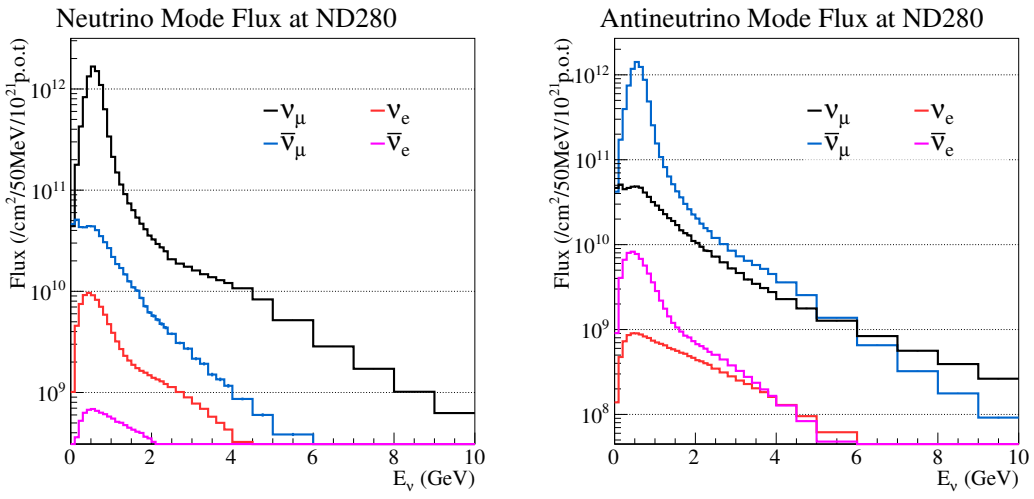


Figure 3.6.: T2K protons on target and beam power for run 1-9

The final simulated neutrino fluxes for run 2 to 8 at ND280 are shown in Figure 3.7. The wrong-sign background in FHC is $\sim$18% at the flux peak which reduces due to the lower anti-neutrino interaction cross-section, and the ν_e component is less than 1% in the flux peak. The $\bar{\nu}_\mu$ flux in RHC is similar to the ν_μ flux in FHC, and the large contamination of ν_μ events in RHC mode comes from the higher ν_μ cross-section rather than the flux.



(a) FHC **(b)** RHC

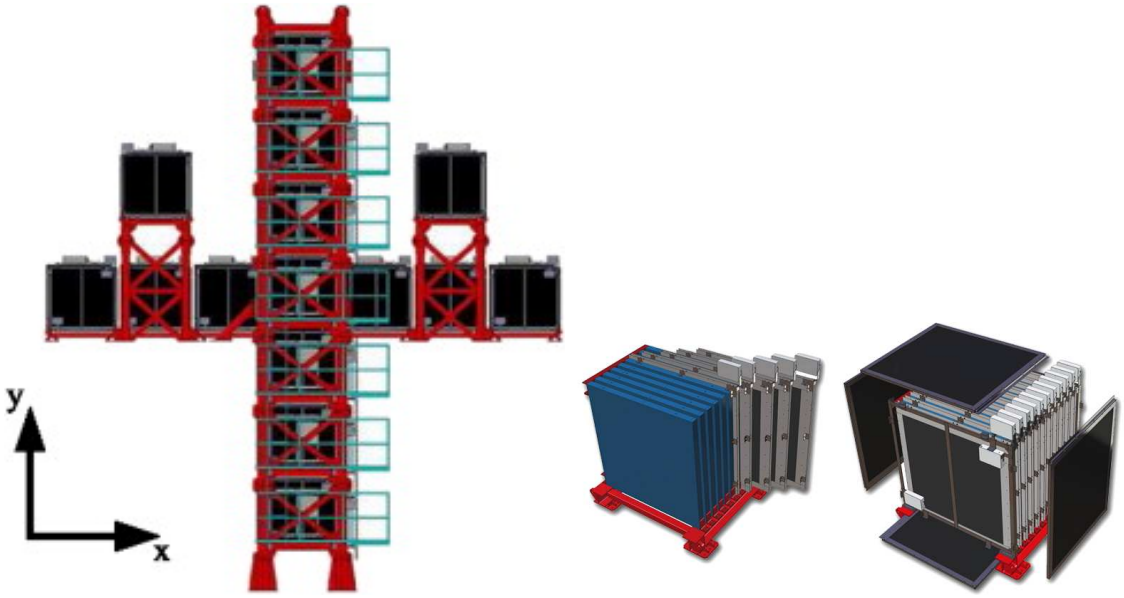
Figure 3.7.: Simulated neutrino fluxes at ND280 in FHC and RHC modes

3.2. The Interactive Neutrino GRID

The off-axis technique introduces the need for precise determination of the neutrino beam direction, as a 1 mrad uncertainty on the beam direction produces a 2-3% uncertainty on the neutrino energy scale [126]. Furthermore, sudden discontinuities in the beam or any of its subcomponents (e.g. magnetic horns or target) directly affect the neutrino flux, so measurements of the beam direction are made on a spill-by-spill basis.

The Interactive Neutrino GRID (INGRID) detector sits on-axis, 280 m downstream of the production target. It was designed to measure the neutrino beam profile to accurately predict the off-axis angle and flux at ND280 and SK, and make inclusive neutrino cross-section measurements [132]. It has a cross shaped geometry and extends 10 m vertically and horizontally, shown in Figure 3.8. The cross consists of 14 identical modules and two off-axis detectors which measure the asymmetry of the neutrino beam. Each INGRID module has nine iron plates and 11 tracking scintillator plates, which are surrounded by veto planes on all sides to reject cosmic backgrounds [133]. Each module has a fiducial iron mass of 7.1 tonnes, and the total cross spans roughly 1σ of the expected beam profile. INGRID also has a second type of module—the proton module—designed to measure neutrino interactions on plastic scintillator. It consists of 34 tracking plates, similar to those in the INGRID modules but with different dimensions, surrounded by the same veto planes [134]. Some T2K runs have also added an INGRID water module to measure neutrino interactions on water.

INGRID has measured the neutrino event rates within 2% of the expected [108], with a precision on the beam directionality of 0.2 mrad—resolving the neutrino beam center to 5 cm. The historical event rate and beam direction from MUMON and INGRID over the full beam period used in this thesis can be seen in Figure 3.9. Generally, INGRID and MUMON agree within 0.2 mrad in both vertical and horizontal directions, and the neutrino event rate at INGRID agrees with expectation from MUMON and the beam simulation.



(a) Downstream beam-view of the INGRID detector

(b) An INGRID module, showing the iron plates and scintillator planes alongside the veto planes

Figure 3.8.: The INGRID experiment, used for neutrino beam monitoring of event rates and direction at T2K

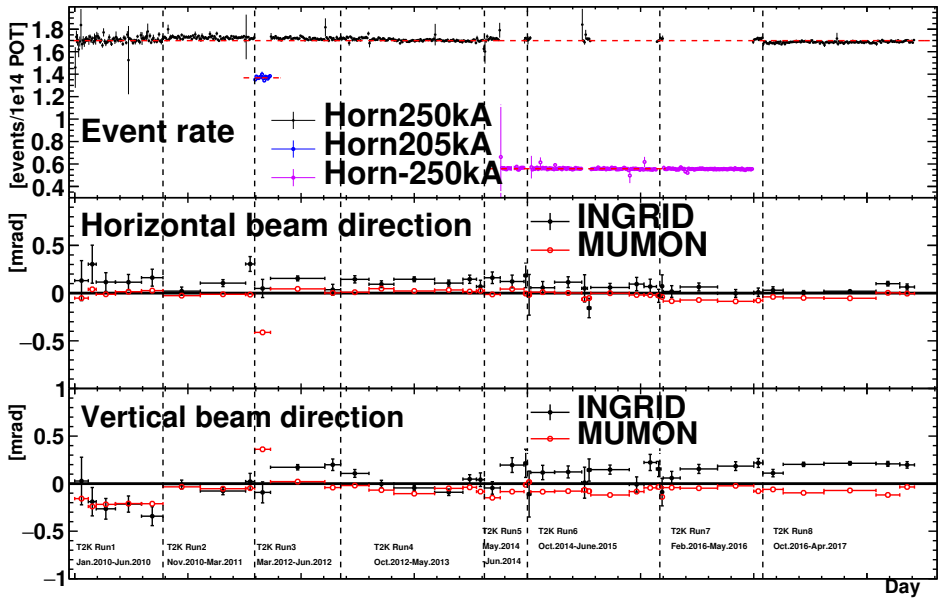


Figure 3.9.: Beam characteristics measured by the INGRID and MUMON detectors over the T2K runs 1 through 8, used in this thesis

3.3. The Near Detector at 280 m

The off-axis near detector for T2K is called ND280, and its neutrino-nucleus interaction data is the subject of this thesis. In contrast to INGRID, ND280 was designed to accurately reconstruct and track particles emanating from a primary neutrino scattering vertex at its central region.

ND280 surrounds its inner target sub-detectors, the two fine grained detectors (FGDs), by three time projection chambers (TPCs). This inner region is referred to as the "tracker" and it is enclosed by a lead-scintillator sampling electromagnetic calorimeter (ECal) on all but the upstream side, at which a dedicated π^0 detector, the P0D, is placed. The whole detector sits in a 0.2 T magnetic field to enable accurate sign selection and momentum measurements with the TPCs. The magnet yoke is interleaved with a side muon range detector (SMRD) made of plastic scintillator strips which enable high angle tracking of μ and provides a cosmic trigger. The exploded detector view is shown in Figure 3.10.

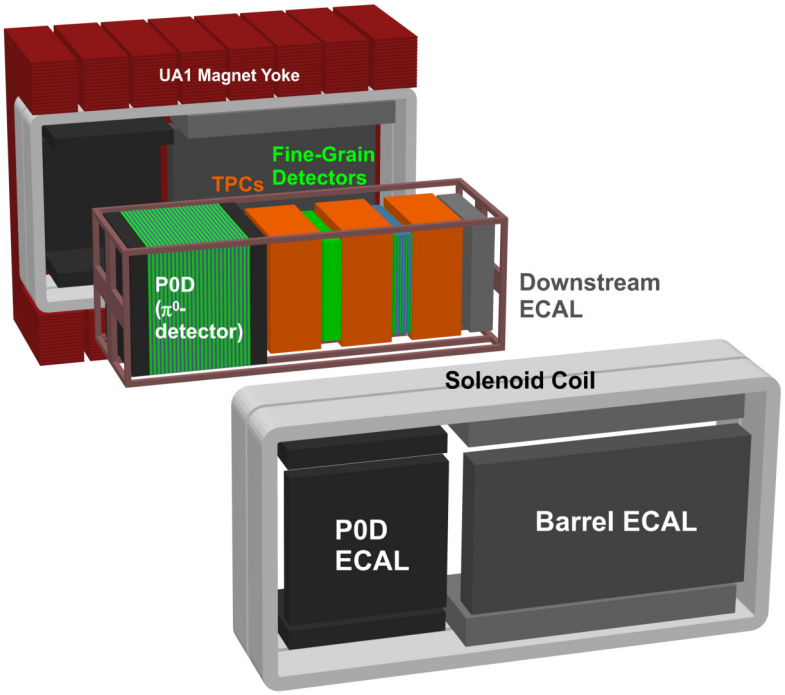


Figure 3.10.: The ND280 detector with its sub-detectors, showing the central tracker region (P0D, TPC, FGD) surrounded by the ECALs and the UA1 magnet with the interleaved SMRD

Since signal at SK is limited to the single ring μ and e selections—vetoing any secondary rings from e.g. pions or high energy protons—the detector is designed for low multiplicity cross-section measurements. ND280 also provides information on the single π^0 cross-sections, a major background for oscillation searches looking for $\nu_\mu \rightarrow \nu_e$.

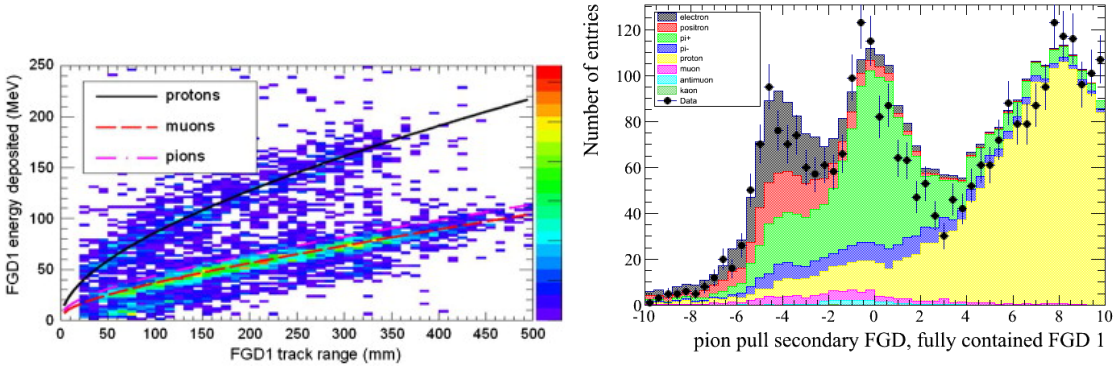
The analyses presented in this thesis only use the FGDs and TPCs for particle detection and ID. Future analyses of ND280 data for oscillation analyses envisage using the ECal for PID, and the P0D for forward-going events going through at least the first TPC.

3.3.1. Fine Grained Detectors

The two fine grained detectors (FGDs) [135] are the central targets for ND280, providing measurements of flux, energy spectrum and electron neutrino contamination at the off-axis angle of SK (2.5°). Each FGD supplies 1.1 tonnes of target material and extends 186.4 cm $\times$ 186.4 cm $\times$ 2.02 cm per scintillator plane. FGD1 sits most upstream of the two and is composed of 15 plastic scintillator XY planes, each plane having 2×192 bars. FGD2 provides a hybrid water-scintillator target, in which seven plastic scintillator XY planes—identical to those in FGD1—are alternated with six 2.54 cm thick layers of water. The two FGDs thus measure interactions on both plastic CH—a common target in external neutrino scattering data—and H_2O—the target in SK.

The FGDs have the ability to reconstruct features of an event independent of the TPCs for contained particles. Isolated tracks are generally of lower momentum and deposit significant energy per unit track length. Summing the total deposited energy of an isolated FGD track provides a means to distinguish protons from minimally ionising particles. Furthermore, stopped pions may give rise to a delayed Michel e, which is searched for in the FGD reconstruction with an efficiency above 90% [136]. Additionally, tracks with hits in the TPC are required to match FGD tracks to determine the interaction vertex. Timing information from the FGD and the magnet is used for tracks to distinguish forward-going positive particles from backward-going negative particles, and vice versa, with a timing resolution of 3 ns.

The measured track length and particle pulls for the pion hypothesis are shown in Figure 3.11. The three dominant distributions (electron/positron, pion and proton) are well separated in the pulls, agreeing relatively well with the data.



(a) Track-length and particle **(b)** Particle pulls

Figure 3.11.: Parts of the FGD particle identification, using the energy deposited per track length

The FGD is used for daily checks of the neutrino event rates and vertex distributions. It also acts as a cosmic trigger for stopping muons, whose selection is used to determine Michel e tagging efficiencies.

3.3.2. Time Projection Chambers

The three time projection chambers (TPCs) [137] provide the majority of the tracking, energy loss, particle identification and momentum measurements in the ND280 detector. Each TPC is composed of an inner and an outer "box". The inner provides the field cage and the outer the ground potential. The inner box measures 1808 mm $\times$ 2230 mm $\times$ 850 mm and the outer box measures 2302 mm $\times$ 2400 mm $\times$ 974 mm. All TPCs use an Ar_{40}:CF_4:C_4H_{10} mixture at 95:3:2, which ionises when charged particles pass through. The ionisation electrons are drifted toward bulk micromegas detectors [138, 139], amplifying the charge. The maximum drift distance from central cathode to Micromegas is 897 mm, and with the nominal cathode voltage at -25 kV and micromegas at -350 kV, the drift field is $\sim$275 V/cm. Figure 3.12 provides a schematic of the TPC design.

The readout provides a 3D image of the tracks from charged particles traversing the TPC, where the bending in the magnetic field determines track momentum. Since two TPCs surround each FGD, they provide good tracking and multiplicity measurements of backward and forwards-going tracks. Furthermore, neutrino interactions produce predominantly forward-going muons and pions which leave tracks in multiple TPCs, giving an improved particle reconstruction.

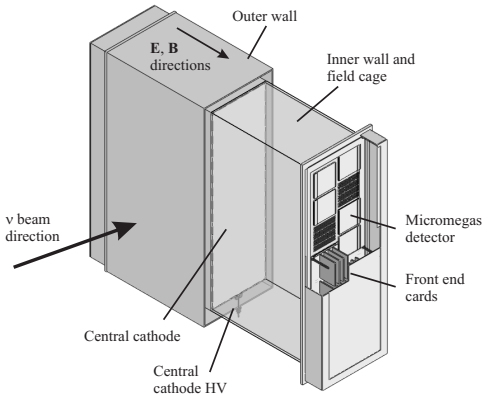
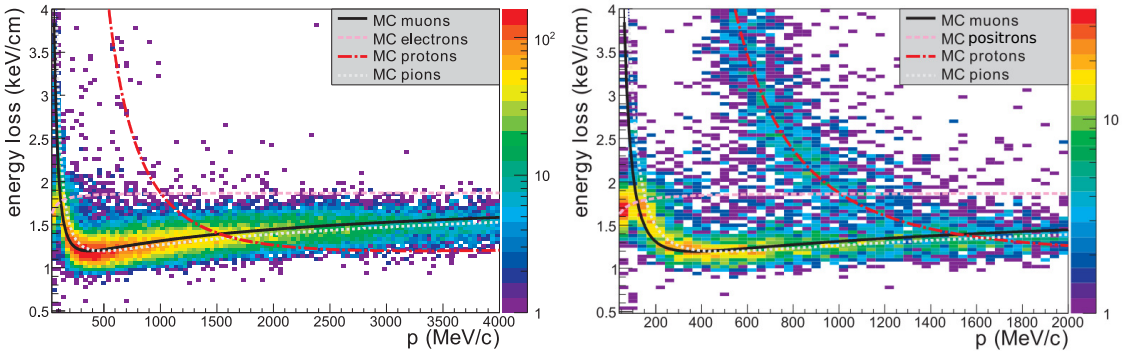


Figure 3.12.: The ND280 TPC design, showing the central cathodes, inner and outer walls, field cages, $\vec{E}$, and $\vec{B}$ fields, and the readout Micromegas and Front End Boards, figure from [137]

The energy loss as a function of momentum in one TPC is shown in Figure 3.13 for negatively and positively charged particles. Muon/electron distinction is achieved, and the MIP resolution is $7.8 \pm 0.2\%$. Using the particle hypotheses outlined in subsection 5.1.1, the probability of assigning a muon to an electron hypothesis is 0.2% for tracks below $p_{true} = 1$ GeV/c. There is also excellent ability to identify proton tracks with $p_{true} < 0.8$ GeV/c [140].



(a) Negative particles **(b)** Positive particles

Figure 3.13.: The energy loss in one of the T2K TPCs as a function of reconstructed momentum

3.3.3. Pi-zero Detector

The pi-zero detector (P0D) [141] was designed to measure the NC1π^0 production cross-section, a large systematic for ν_e appearance.

The design of the P0D is shown in Figure 3.14. The water target consists of an upstream and central section. The upstream (central) water target—consisting of 12 (13) alternating scintillator-water bag (28 mm)-brass sheet planes (1.28 mm)—is surrounded by an upstream and downstream ECal (not to be confused with the P0D-ECal).

The ECals maintain the triangular scintillator bars but have the brass and water layers replaced by a 4.5mm thick lead sheet and is seven planes each. The entire P0D consists of 40 modules, each being two perpendicular arrays of triangular scintillator bars. Each vertical bar (134 per P0D module) is 2200 mm long, and the horizontal bars are 2340 mm long (126 per P0D module)[b]. The active target of the P0D is 2103 mm $\times$ 2239 mm $\times$ 2400 mm and the mass with (without) water is 15.8 (12.9) tonnes.

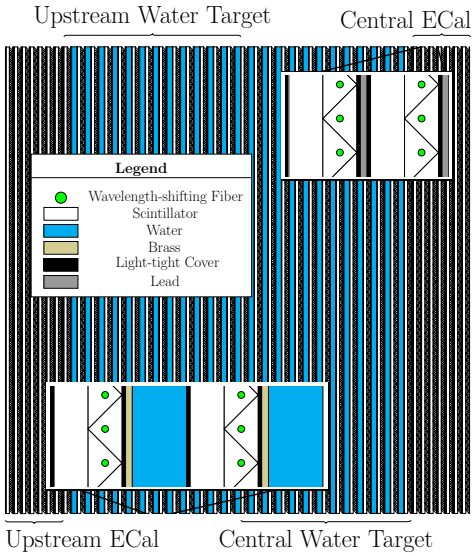


Figure 3.14.: The ND280 P0D side-view

3.3.4. Electromagnetic Calorimeter

The ND280 electromagnetic calorimeter (ECal) [142] is designed to complement the tracker in full event reconstruction due to its near hermetic coverage of the tracker and P0D regions. It measures photon showers' energy and directions, and is used to

[b]Originally designed for the MINERνA experiment [84]

distinguish electrons from muons from pions by shower shape. Its primary purpose is to tag and reconstruct π^0s from the tracker.

The ECal has three main sections: 1) the barrel ECal, surrounding the inner tracking detectors, 2) the downstream ECal, sitting after the last TPC, and 3) the P0D Ecal. The tracker (barrel+downstream) ECal was designed as a tracking calorimeter and reconstructs electromagnetic showers, complementing the TPC where high-angle particles may leave few hits. The ECals are all made of scintillating polystyrene bars 40 mm $\times$ 10 mm which are adhered to 1.75 mm lead sheets. The ECal design was based on good detection efficiency of π^0s emanating from the tracker region. The barrel-ECal has 31 layers and the downstream has 34 layers, corresponding to between 10 and 11 electron radiation lengths. This is required to ensure $\sim$50% of the energy given to photon showers from a π^0 decay is contained.

The downstream ECal is 2300 mm $\times$ 2300 mm $\times$ 500 mm with 1700 2000 mm long scintillator bars (50 per layer). The four barrel ECal top and bottom modules are 4140 mm $\times$ 1676 mm $\times$ 462 mm and the two side barrel ECals are
4140 mm $\times$ 2500 mm $\times$ 462 mm. The top and bottom ECals have 1520 mm bars, and the sides have 2280 mm bars, perpendicular to the beam direction, and 15 3840 mm bars parallel to the beam.

The P0D-ECal is slightly different in design since the P0D performs its own shower reconstruction. The ECal is instead present to tag escaping energy and distinguish photon showers from muon deposits at high angles. It instead has six scintillator layers with thicker 4.0 mm lead sheets. It is 155 mm deep, with the top and bottom modules being 1584 mm wide and the sides being 2898 mm wide, with 2454 mm length.

3.3.5. The UA1/NOMAD Magnet and Side Muon Range Detectors

The entirety of the above detector system (FGD+TPC+ECal+P0D) is placed inside the refurbished UA1/NOMAD dipole magnet. It is operated at 2.7 kA to produce a uniform horizontal magnetic field of 0.2 T. The yoke is split into two sections, each made of eight C-shaped flux return yokes. The inner volume of the magnet is 7.0 m $\times$ 3.5 m $\times$ 3.6 m, placing the main spatial limitations on ND280.

The side muon range detector (SMRD) [143] sits in the innermost gaps of the UA1 magnet return yoke, surrounding the entire ECal, P0D and tracker. It was designed to measure muons which escape the tracker at high angles—punching through the

FGD, ECal and SMRD—leaving few or no TPC hits. The momentum can be inferred from range in the iron and the SMRD, and the tracks can be reconstructed using FGD-ECal-SMRD matching with $\sim$70% efficiency. The SMRD additionally provides a cosmic trigger.

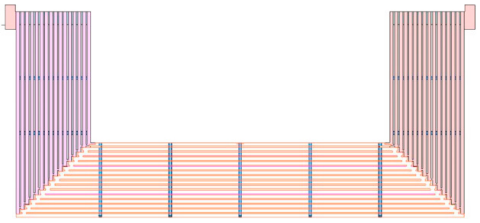


Figure 3.15.: A side-view section of the UA1 magnet yoke showing the interleaved SMRD

Shown in Figure 3.15, there are 16 iron plates in the magnet yoke, each 48 mm thick and which are separated by 17 mm spacers, leaving space for 15 layers of scintillator. The SMRD consists of 192 horizontal and 248 vertical modules in total. The modules measure 9 mm $\times$ 686 mm $\times$ 955 mm and 9 mm $\times$ 892 mm $\times$ 955 mm for horizontal and vertical modules respectively. Each horizontal module has four 7 mm $\times$ 167 mm $\times$ 875 mm scintillation counters and vertical modules have five 7 mm $\times$ 175 mm $\times$ 875 mm, totalling 768 horizontal and 1240 vertical counters. Going downstream, there are three layers of modules in the 1st to 5th yokes, four layers in the 6th downstream yoke, and six layers in the 7th and 8th yokes.

3.4. Super-Kamiokande

The Super-Kamiokande (SK) [144–146] detector has served to constrain proton decay and measure solar and atmospheric neutrino oscillations since 1996 with SK-I. Starting with K2K in 1999 [147], Super-Kamiokande has also been serving as a far detector for long baseline accelerator neutrino oscillation searches in Japan, and continues to do so for T2K with SK-IV and beyond. The detector is placed 295 km from the production target in the Kamioka mine in Ikenoyama, located in Gifu, Japan. The mine provides roughly 1 km of rock overburden—or 2.7 km water equivalent overburden—drastically reducing cosmogenic backgrounds. A sketch of the detector in the mine is shown in Figure 3.16.

The detector consists of 50,000 (25,000 fiducial) tonnes of ultra-pure water in a 41.4 m $\times$ 39.3 m cylindrical tank. It is split into an inner (ID) and outer (OD) detector

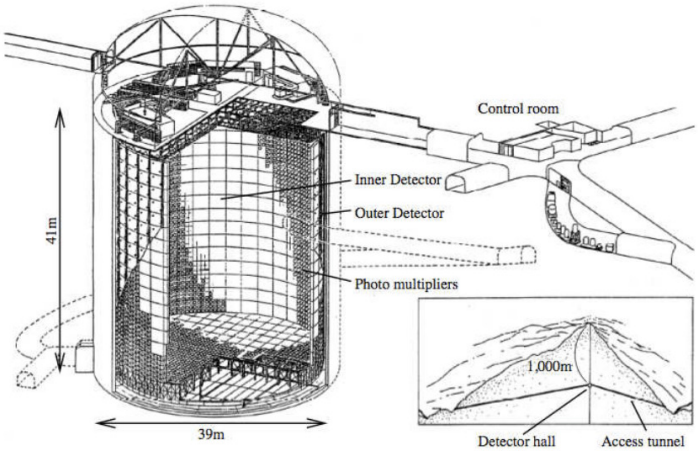


Figure 3.16.: The Super-Kamiokande detector in Ikenoyama

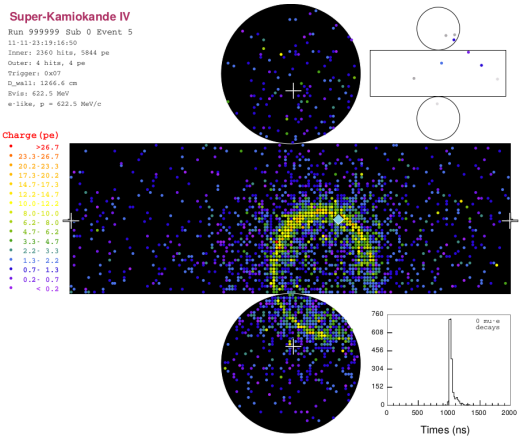
with inner dimension 36.2 m $\times$ 33.8 m, where the OD surrounds the ID. The ID and OD are separated by a Tyvek and "blacksheet" barrier, and the ID has 11,146 inward-facing 20-inch PMTs and the OD has 1,885 8-inch PMTs facing outwards—providing veto and shielding for the inner detector. The ID has approximately 40% photo-coverage and provides excellent μ/e separation, crucial to differentiate muon neutrino disappearance from electron neutrino appearance.

Particle detection in SK happens primarily through production of Cherenkov light. The muon/electron separation is primarily achieved by measuring ring "fuzziness", indicating the amount of scattering of the charged particle emanating from the initial neutrino scattering vertex in the ID. Muons and pions are generally highly penetrating whereas electrons scatter often and give rise to electromagnetic showers at T2K energies. Consequently, the former produces sharp rings whereas the latter produce fuzzy rings. The number of delayed Michel e is also used to identify contained muons and pions, and tracks with kinked trajectories are used to discern rescattered pions from muons.

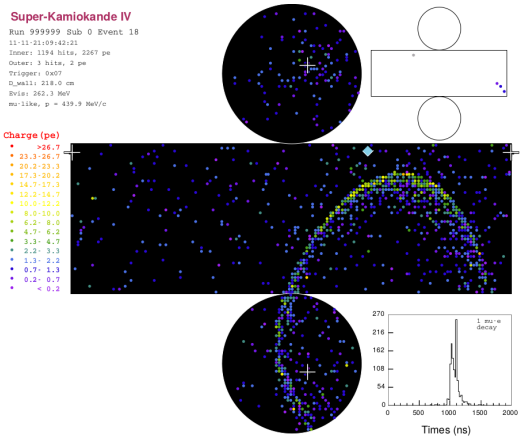
SK is currently closed to prepare it for gadolinium which will add neutron tagging capabilities. Data collection will resume in spring 2019 [148].

3.5. Simulation

The three detectors (INGRID, ND280 and SK) are modelled using the same neutrino interaction Monte-Carlo, namely NEUT [53]. NEUT is an interaction generator written



(a) e-like, $p = 622.5$ MeV



(b) μ-like, $p = 439.9$ MeV

for the (Hyper, Super)-Kamiokande experiments, with large contributions from K2K and T2K collaborators.

ND280 and INGRID are simulated with GEANT4 [126, 149], and the ND280 electronics are simulated with a custom package called ElecSim. The beamline simulation consists of FLUKA2011 [150–152] which simulates hadronic interactions in target and baffle, JNUBEAM (GEANT3-based [153]) which simulates the geometry and handles particle tracking, and GCALOR [154] which simulates hadronic re-interactions and is used as a cross-check for FLUKA [128, 155]. The SK detector is simulated with a custom package, SKDETSIM [144], based on GEANT3 [153].

The systematics treatment from detector effects are detailed in section 5.3.

Chapter 4

Statistical Treatment and Markov Chain Monte Carlo

> There are three kinds of lies: lies, damned lies, and statistics
>
> Mark Twain

The analysis presented in this thesis employs a Bayesian view of statistics, in which the end result is a posterior probability distribution $P(\vec{\theta}|D)$ for the model $\vec{\theta}$ given the observed data D. It is built from the joint probability distribution $P(D|\vec{\theta})$ and the prior information $P(\vec{\theta})$. These quantities are related through Bayes' theorem,

$$P(\vec{\theta}|D) = \frac{P(D|\vec{\theta})P(\vec{\theta})}{\int P(D|\vec{\theta})P(\vec{\theta})d\vec{\theta}} \tag{4.1}$$

which is conditional on the full model space Θ. In this analysis, the $P(D|\vec{\theta})$ is the neutrino event distributions and are modelled with a Poisson probability distribution function for the data n being a fluctuation of the simulation $\lambda(\vec{\theta})$ in each bin. The $P(\vec{\theta})$ is the prior knowledge of the model $\vec{\theta}$ from data not used in this analysis (e.g. external hadron scattering data, cosmic data at ND280) and is modelled using a multivariate Gaussian probability distribution. Each parameter i has a central value μ_i which is varied to X_i during the MCMC sampling and is related to parameter j through the covariance matrix $\mathbf{V}_{ij}$. The joint posterior probability distribution $P(D|\vec{\theta})P(\theta)$ can then formally be expressed as

$$P(D|\vec{\theta})P(\vec{\theta}) = \prod \mathcal{L}_{\text{Total}} = \prod \left(\mathcal{L}_{\text{Bins}} \times \mathcal{L}_{\text{Systematics}}\right) \tag{4.2}$$

or more conveniently in the logarithmic space,

$$-\log \mathcal{L}_{\text{Total}} = \sum_{\text{Bins}} \left[\lambda(\vec{\theta}) - n + n \log \frac{n}{\lambda(\vec{\theta})} \right] + \sum_{\text{Systematics}} \frac{1}{2} \left[(X_i - \mu_i) \left(\boldsymbol{V} \right)^{-1}_{ij} (X_j - \mu_j) \right] \tag{4.3}$$

The posterior is generally not analytically solvable and $\vec{\theta}$ may be of very high dimension—in this thesis $\dim(\vec{\theta}) = 687, 1209, 4369$. It is instead commonplace to sample from the posterior using Monte Carlo methods, producing a density of points proportional to the posterior up to a normalising factor.

Markov Chain Monte Carlo methods have been used extensively in astro-statistics, computational physics and biology, and is central to evaluating the multi-dimensional integrals that occur in Bayesian statistics. The necessary and sufficient conditions for a Markov Chain successfully constructing the posterior density requires

- **Irreducibility:** From any given state, the probability to reach any other state has to always be non-zero
- **Recurrence:** When the stationary distribution has been found, all subsequent steps must sample from that stationary distribution
- **Aperiodicity:** The sequence of steps must not be periodic

A detailed discussion of these criteria can be found in [156–159].

4.1. The Metropolis-Hastings Algorithm

This analysis uses the well established Metropolis-Hastings algorithm [160, 161] to explore the parameter space $\vec{\theta}$ and sample the posterior $P(\vec{\theta}|D)$. It can be considered a random walk which steers towards areas of low $-\ln \mathcal{L}_{\text{Total}}$ in Equation 4.3. The algorithm consists of a set of steps which are repeated N times:

- **Initialisation**: A random starting point of the parameters $\vec{\theta}$ is chosen

- **Step proposal**: A new set of parameters $\vec{\theta}'$ is proposed using a proposal function, symmetric around the previous accepted point

$$\vec{\theta}' = \vec{\theta} + \text{rand}_{\text{Proposal Funct.}}(\vec{\theta}) \tag{4.4}$$

  The proposal function may have tuneable width, which can increase or decrease the acceptance rate but can significantly worsen the parameter space exploration.

- **Acceptance probability**: The total test-statistic in Equation 4.3 is calculated for the current and proposed step, whose likelihood ratio forms the acceptance probability

$$\alpha = \min\left[1, \frac{\ln \mathcal{L}_{\text{Total}}(\vec{\theta})}{\ln \mathcal{L}_{\text{Total}}(\vec{\theta}')}\right] \tag{4.5}$$

- **Accept or reject**: A uniform random number $u \in U[0,1]$ is selected and compared to α, which determines if the new parameter set $\vec{\theta}'$ is accepted or rejected

$$\vec{\theta} = \begin{cases} \vec{\theta}' & \alpha \geq u \\ \vec{\theta} & \alpha < u \end{cases} \tag{4.6}$$

- **Repeat**: Repeat step proposal and continue algorithm

Importantly, the Metropolis-Hastings algorithm will always accept steps in the direction of higher probability, and also accepts steps with lower probabilities. Thus an adequately explored MCMC does not suffer from local minima biases and scales relatively well with increasing dimensionality. However, after N steps there is no guarantee that the posterior has been adequately sampled and the parameters have converged. The general guidance on how big N should be is as large as computationally feasible [157].

As an example, Figure 4.1 shows the evolution of two neutrino-nucleon interaction parameters present in this analysis, detailed in subsection 5.3.3. The MCMC is fitting the simulation to data, and initially starts at $\{M_A^{QE}, \text{BeRPA B}\} \sim \{1.0, 1.1\}$ but this region is disfavoured by the test-statistic. The poor test-statistic almost exclusively comes from the data being better described by a different set of values for the interaction parameters. After only 1,000 steps the BeRPA B parameter roughly reaches stationarity, whereas M_A^{QE} explores the space up to $\sim 20,000$ steps.

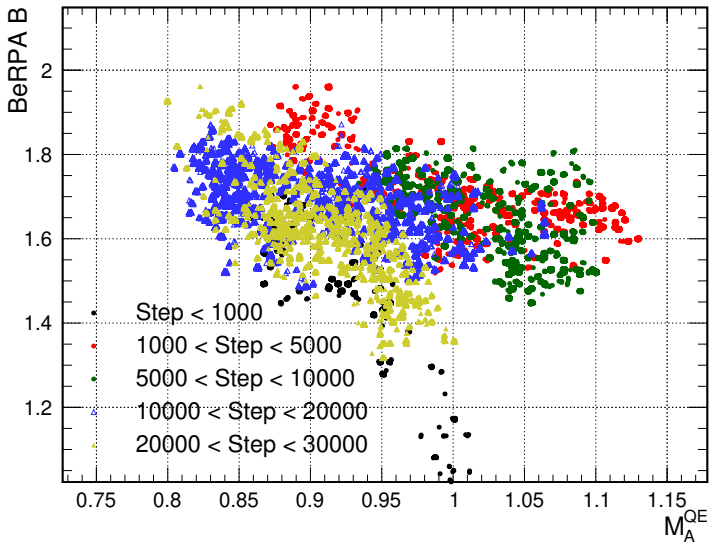


Figure 4.1.: Evolution of two neutrino-nucleon interaction parameters with MCMC step for a fit to data observed at ND280

Since separate MCMC are statistically independent once they've reached stationarity, they can be combined to lower the MC statistical error in the sampling, resulting in better point and interval estimates. Hence MCMC is commonly run in parallel: for ND280-only fits we often use at least three chains, whereas full ND280+SK oscillation fits run several hundred.

4.2. Diagnostics

The main tuneable parameters in a Metropolis-Hastings MCMC algorithm are the step-sizes of the proposal functions and the number of steps N. The MCMC variables are tuned until convergence in the parameters are obtained. For this analysis a set of diagnostics are used for monitoring convergence, as recommended in literature [157]: parameter traces, likelihood traces, batched means and autocorrelations.

The traces are simply the accepted parameter values with each MCMC step, which is a function of the step-size of the proposal function. Too fine step-sizes cause high acceptance probability with poor parameter space exploration and highly correlated steps, with the opposite being true for too coarse step-sizes. The optimal acceptance probability is 0.234 [157, 162] and step-size tuning is done as to roughly coincide with this number.

The batched means are similar to the traces but cuts up the MCMC steps into batches and compares the average parameter values in each batch. This is typically a good indicator of burn-in, as small batches near the start of the chain may have different values to batches near the end of the chain.

The auto-correlation function r_k after lag k for a parameter Y_i at step i is identical to that in signal processing

$$r_k = \frac{\sum_{i=1}^{N-k} (Y_i - \bar{Y})(Y_{i+k} - \bar{Y})}{\sum_{i=1}^{N} (Y_i - \bar{Y})^2} \tag{4.7}$$

and is a measure of how correlated a step in the chain i is with a step k steps ahead. This is an important tool for MCMC to diagnose chain stability, since the optimal acceptance probability of 0.234 can be easily achieved by varying the step-sizes. However, tuning using only this metric may lead to fine step-sizes, which incurs significant step-to-step correlations in the chain, which the auto-correlation assesses. Generally, this analysis aims to have auto-correlations of less than 0.2 after lag $k = 10,000$. Figure 4.2 shows example auto-correlation functions for a chain which did pass this criteria in the flux parameter but not the interaction parameters, so was step-size tuned further.

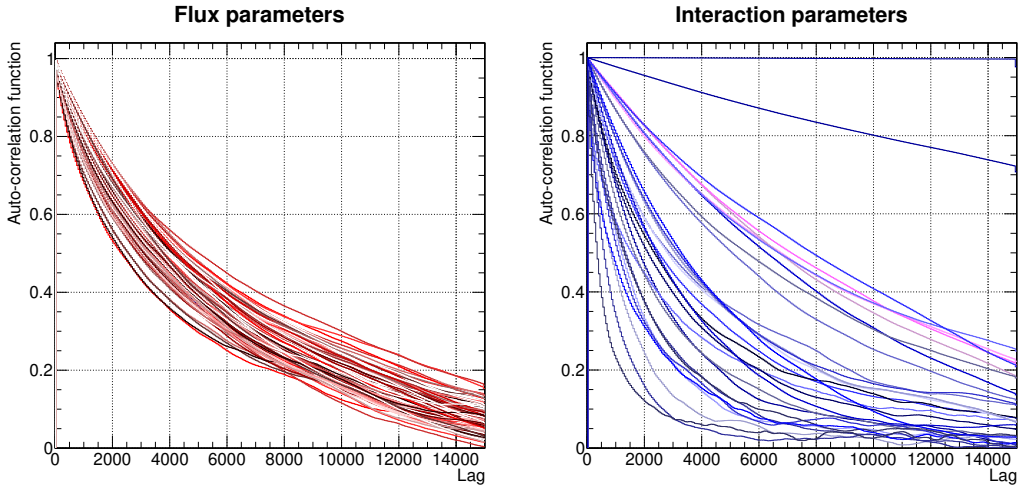


Figure 4.2.: Auto-correlation functions for an example fit to ND280 data

In the case where multiple MCMC have sampled the same posterior we also compare the above diagnostics chain to chain. In the case of suspected non-convergence after significant number of steps, the $\hat{R}$ test [163] is also used. $\hat{R}$ estimates the improve-

ment in estimating the parameter variance that may be achieved by running a chain for longer.

4.3. Burn-in

Since the initial parameters $\vec{\theta}$ aren't necessarily in a region of high probability density, it generally takes an MCMC some time to reach the stationary posterior distribution. This parameter exploration period is normally referred to as the "burn-in" period of a MCMC and is usually discarded since it does not represent the probability density function. Figure 4.3 shows the evolution of a neutrino-nucleon parameter present in the analysis with step for six separate MCMCs, all starting with the parameter value around 0.

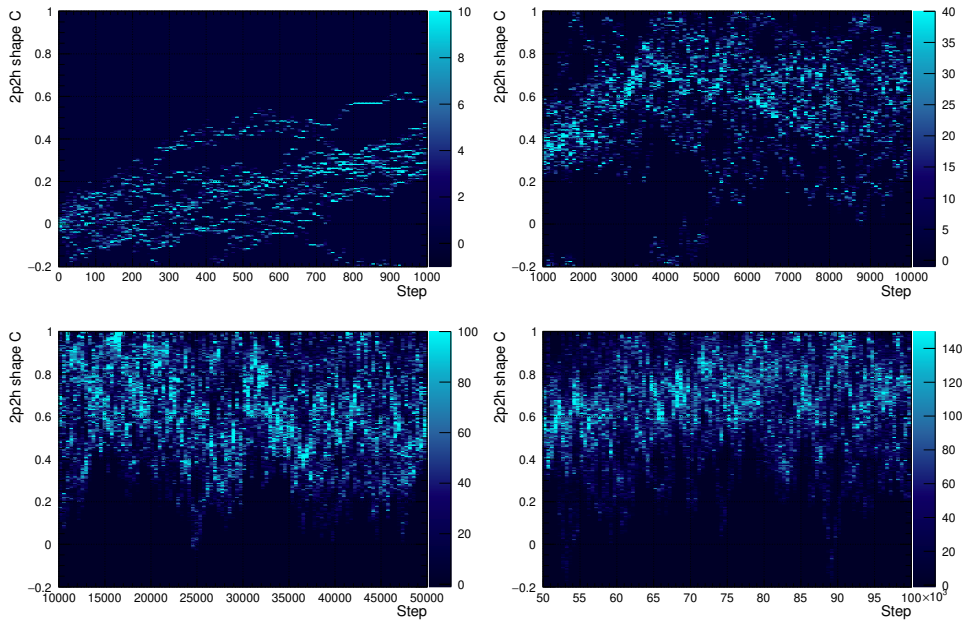


Figure 4.3.: A neutrino-nucleon interaction parameter's evolution with MCMC in six separate chains running on data

The chains appear to converge after about 20,000 steps, which indicates the approximate burn-in. This judgement is largely made by monitoring chain stability and chain-to-chain variation. The auto-correlations are also checked, along with a random subset of systematic parameters' stability with step. The total test-statistic in Equation 4.3 is also checked over steps, and in this case results in Figure 4.4. The

test-statistic initially walks towards the minimum and reaches it after about 800 steps for all the six chains. The chains then globally step out of the minimum and explore the area around it, which appears stable after 15-20,000 steps.

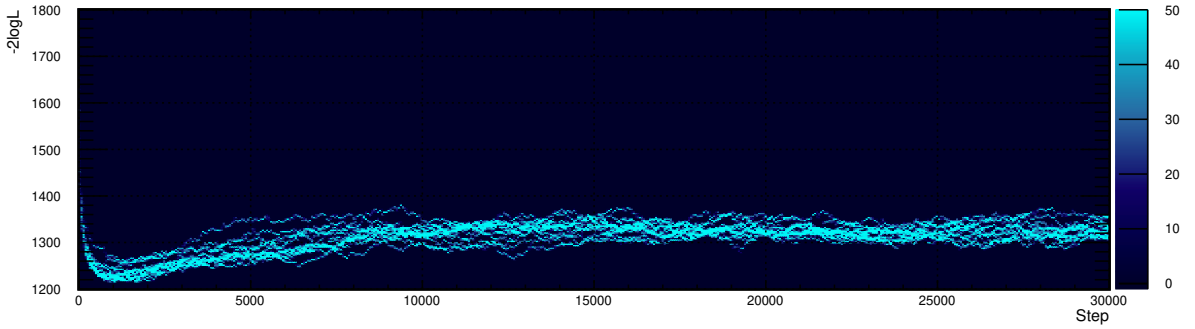


Figure 4.4.: Test-statistic evolution with MCMC steps for six separate chains running on data

The analyses presented in this thesis use a more conservative approach than above since computational power was not an issue. Generally, individual chains are at least 1,000,000 steps long and at least three such chains are run in parallel and the burn-in is always at least a quarter of the total.

4.4. Point Estimates and Uncertainties of Parameters

This analysis propagates a full high-dimensional posterior—rather than central values with uncertainties related through a covariance matrix—so is less concerned with the accuracy of point estimates and their uncertainties. Furthermore, the posterior only contains parameters considered as nuisance parameters in oscillation analyses. This saves significant computational time, since point estimation in high dimensions with correlated parameters requires a significant number of MCMC steps or a better suited algorithm than Metropolis-Hastings.

However, many point estimates with uncertainties are presented in this work, especially when validating against the frequentist "BANFF" framework in Appendix E. Naturally, this process loses information about the full posterior, so is not propagated to the oscillation analysis.

4.4.1. Parameters of Interest and Marginalisation

At minimum, 73 out of 686 parameters are propagated to oscillation analyses in this work. The parameters that aren't propagated are ND280-only parameters—e.g. cross-section parameters on Carbon, ND280 detector parameters, or ND280 flux parameters. However, for an ND280-only analysis, the number of parameters of interest is high and visualising parameter behaviour becomes difficult. Throughout the ND280 fits we project the high-dimensional posterior onto one or two dimensions—the latter being used to form the parameter covariance matrices. The projection uses the marginalisation method, in which we integrate out the posterior's dependency on all but the single "parameter of interest".

Consider the single parameter of interest x and the remaining parameter space is denoted $\vec{\theta}$. The marginalised posterior density $P(x|D)$ of the parameter x over the full parameter space Θ given the data D is given by

$$P(x|D) = \int_{\Theta} P(x|\vec{\theta})P(\vec{\theta}|D)d\vec{\theta} \tag{4.8}$$

This one-dimensional posterior distribution is used to obtain point estimates and uncertainties. We chiefly use four methods for the point estimate and its uncertainty:

- The arithmetic mean and RMS
- The fitted Gaussian's mean and 1σ
- The highest posterior density (HPD) with symmetric or asymmetric errors to cover 68% of the area going outwards from the HPD

The three methods are used to flag when posteriors have non-Gaussian shapes, since in the Gaussian case the above are all equivalent. This can happen for parameters that have hard cut-offs, strong correlations, and for parameters that are switches ("on" or "off"). Figure 4.5 demonstrates the differences between the methods for a non-Gaussian beam parameter.

In Figure 4.5 we note asymmetric errors in the case of the asymmetric HPD method and different estimates of the central value. Importantly, none of the methods are de-facto wrong, and in many cases the one dimensional posterior has to be investigated further for e.g. marginalisation effects with other non-Gaussian parameters. Unless otherwise stated, the arithmetic mean and RMS are used due to their simplicity.

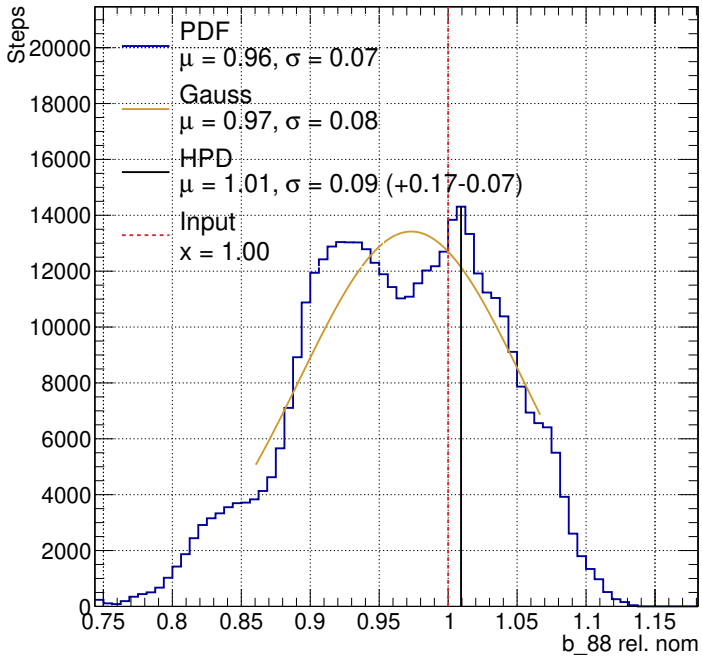


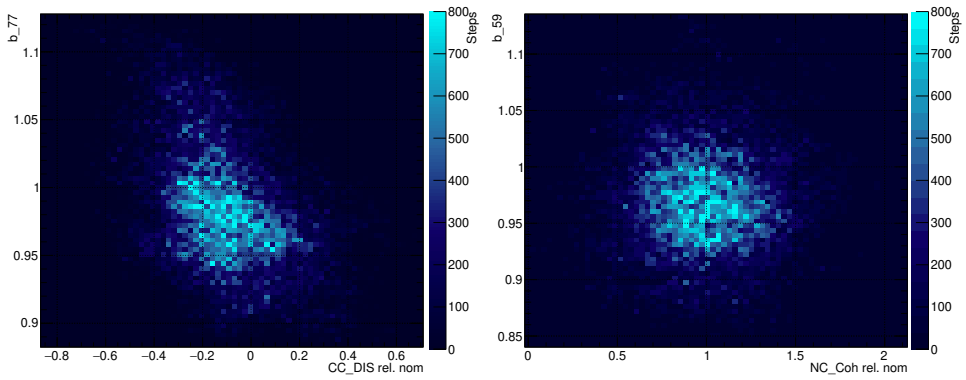
Figure 4.5.: One-dimensional marginalised posterior density for a beam parameter for an MCMC running on data, showing the four methods of point and error estimation. "PDF" denotes the arithmetic method , "Gauss" denotes the Gaussian method, "HPD" denotes the Highest Posterior Density using symmetric or asymmetric errors

4.4.2. Covariance Estimates

We marginalise the high-dimensional posterior onto the two parameters whose covariance we wish to calculate. The marginal posterior is binned and the covariance is calculated arithmetically without assuming a shape of the posterior. Figure 4.6 shows two example two-dimensional marginal posteriors between beam and interaction parameters, one of which results in a strong correlation and the other doesn't.

4.5. Predictive Checks and p-values

A Bayesian analysis conditions on the entire model's probability, so can lead to misleading results when the model space strongly disagrees with the data space. Assessing the model goodness against data is therefore critical in any Bayesian analysis, and this analysis closely follows what is detailed in [164–167].



(a) Strong correlation **(b)** Weak correlation

Figure 4.6.: Two-dimensional marginal posteriors used to calculate parameter covariance

We use two "goodness-of-fit" checks, in which the test-statistic from the sample for observed events n with predicted events λ,

$$-\log \mathcal{L}_{\text{Samples}} = \lambda - n + n \log \frac{n}{\lambda} \tag{4.9}$$

is of central importance. The "predictive spectrum" of a distribution from a model is also crucial, and is defined as having a point estimate $\bar{\lambda}$ equal to the average predicted events over N randomly chosen MCMC steps after burn-in,

$$\bar{\lambda} = \frac{1}{N} \sum_{i}^{N} \lambda_i \tag{4.10}$$

with the error $\Delta\bar{\lambda}$ taken as the root mean square,

$$\Delta\bar{\lambda} = \sqrt{\frac{1}{N} \sum_{i}^{N} \lambda_i^2} \tag{4.11}$$

For this analysis it was checked that the predictive spectrum using the arithmetic mean and RMS were consistent with using a Gaussian fit[a], and also consistent with the mode and 68% central highest posterior density of the bin.

The goodness-of-fit tests involve applying statistical fluctuations to a drawn simulation and calculating the test-statistic in Equation 4.9 of the fluctuation versus some other distribution. Locating the test-statistic of the posterior predictive spectrum ("best-fit") given the data gives a p value. The methods are:

[a] Extracting $\bar{\lambda}$ and $\Delta\bar{\lambda}$ from the fitted parameters

- A one-dimensional plot of the test-statistic between the drawn distribution and the statistical fluctuation of that drawn distribution
- A two-dimensional plot of the test-statistic between the data and the drawn distribution, versus the test-statistic between a fluctuation of the drawn distribution and the drawn distribution

The former case simply locates the realised test-statistic in a distribution, whereas the second informs about the predictive nature of the model after having seen data, were we to observe more data. Formally,

$$P(D'|D_{obs}) = \int P(D'|\vec{\theta})P(\vec{\theta}|D_{obs})d\vec{\theta} \tag{4.12}$$

To construct the p-value in the two-dimensional case we need to compare the result to a reference distribution, which is the statistically fluctuated drawn histogram in this case. In practice, 20,000 random draws of the parameters are performed after burn-in, which samples the posterior. For each draw we

- Reweight the Monte-Carlo using the new parameter set
- Poisson fluctuate each Monte-Carlo bin according to its bin contents
- Calculate the test-statistic between the fluctuated histogram and the drawn histogram, denoted $\chi^2_{\text{Draw, Draw Fluc}}$
- Calculate the test-statistic between the observed data and the drawn histogram, denoted $\chi^2_{\text{Data, Draw}}$.
- Fill a two dimensional histogram of the two test-statistics

The predictive spectrum is formed by taking the mean in each bin as the central number of events, and the error is estimated by using the RMS. Hence the posterior predictive distribution is not observed from one MCMC draw but is rather the distribution representative of the posterior after sampling. The two-dimensional posterior predictive p-value is then finally

$$p = \frac{N\left(\chi^2_{\text{Data, Draw}} < \chi^2_{\text{Draw, Draw Fluc}}\right)}{N\,(\text{Total})} \tag{4.13}$$

where $N\left(\chi^2_{\text{Data, Draw}} < \chi^2_{\text{Draw, Draw Fluc}}\right)$ is the number of times a drawn distribution had a smaller test-statistic against data than it did against a fluctuation of itself, and $N(\text{Total})$ is the total number of draws.

The p-values are also used for statistical closure tests, where the drawn distribution is instead the predictive distribution. Furthermore, the drawn distribution can be from both the prior and posterior densities.

Chapter 5

Constraining Model Parameters at T2K using ND280 Data

> At this point you're re-arranging deck chairs on the Titanic
>
> Senior collaborator at T2K Collaboration Meeting, 2016

The T2K oscillation analyses consider numerous systematics constrained by external data, presented in section 5.3. Since ND280 and SK are in the same neutrino beam the high-statistics neutrino interactions at ND280 are used to constrain the simulation prior to analyses at SK—greatly reducing uncertainties on the extracted oscillation parameters.

The event count uncertainties on the 2015 T2K oscillation analysis without using ND280 data is shown in Table 5.1, with uncertainties ranging from 12-22%. The 1Rμ refers to reconstructing 1 muon-like ring, the 1Re refers to 1 electron-like ring, and the 1Re1de refers to 1 electron-like ring with a delayed Michel electron, indicating the presence of a below-threshold π^+. By fitting ND280 data, the neutrino flux, interaction and ND280 detector model is constrained: the two former being large contributors to the 12-22% error budget. This chapter details that analysis.

SK selection	$\delta N/N(\%)$
1Rμ FHC	12.0
1Re FHC	12.7
1Re1de FHC	21.9
1Rμ RHC	14.5
1Re RHC	12.5

Table 5.1.: Uncertainty on event rates at SK using only prior information without an ND280 fit in the 2015 oscillation analysis [108]

5.0.1. Overview

T2K has two separate groups which use near-detector data with the intent of maximising model likelihood to reduce uncertainty in oscillation analyses. BANFF (Beam And Near detector Flux task Force) and MaCh3 (Markov Chain for 3 flavour oscillation fitting). The two frameworks use identical selections, binning and systematics, outlined in section 5.1, section 5.2 and section 5.3, but different methods of evaluating the model goodness, exploring the parameter space and propagating results to oscillation analyses.

BANFF interfaces to the gradient-descent minimizer MINUIT [168] and MaCh3 uses a custom Markov Chain Monte Carlo to sample the high dimensional parameter space, outlined in chapter 4. Importantly, BANFF attempts to find the global minimum of the test-statistic given the data and the model, whereas MaCh3 explores an area around the minimum test-statistic with the intent of sampling the Bayesian posterior. Therefore, MaCh3 does not necessarily locate a set of "best-fit" parameters with covariances assuming a parabolic minimum: instead it provides a full high-dimensional posterior with arbitrary shape. MaCh3 has the advantage of a near and far detector implementation, meaning a simultaneous fit of data from both detectors can be made [57, 108, 169–171]. This avoids making assumptions on the underlying probability distribution functions of the parameters and the likelihood surface, and benefits from fully correlating the models at both detectors, allowing one to affect the other as the fit proceeds.

The following sections detail the ND280 implementation in the MaCh3 framework. The chapter discusses the selections and systematics, fitting to mock-data to find expected sensitivities and perform closure tests, and real data to constrain the systematics for oscillation analyses. Finally, it evaluates the impact on the predicted neutrino spectra at SK used in the oscillation analyses, ending with compatibility and alternative model studies, and comparisons and validations to the BANFF framework.

5.1. Selections

The goal of the selection is to map interaction modes (e.g. CCQE, CC1π^+, CC DIS) to observable ND280 selections, so that theory parameters receive their largest constraints from a few exclusive selections. Equivalent FGD1 and FGD2 selections are separated

due to differences in systematics and reconstruction: forward-going tracks emanating in FGD1 leaves a track in FGD1, TPC2, FGD2 and TPC3, so has more hits recorded than the FGD2 equivalence, which only passes through FGD2 and TPC3. Furthermore, FGD2 contains plastic scintillator interleaved with passive water layers whereas FGD1 is fully plastic scintillator. Separating FGD1 and FGD2 also allows constraints on water interactions to come strictly from the FGD2 selections.

The events in the analysis are binned in reconstructed muon candidate variables p_μ and $\cos\theta_\mu$. The muon variables are chosen primarily due to excellent detector resolution of muons and for consistency with analyses at SK. There is ongoing effort to include pion variables when such are present (e.g. for CC1π^+ or CCOther selections), and composite variables in the plane transverse to the neutrino, but will not be presented here.

The selections are entirely defined by the observed reconstructed event topology of an event in the detector and there is no attempt at correcting for misidentified particles. There is no attempt to correct for nuclear effects such as final-state-interactions (FSI).

5.1.1. ν_μ in FHC

The different topological selections start by isolating CC-inclusive candidates in FGD1 and FGD2. First, an event is required to contain one reconstructed track of negative charge crossing the TPC downstream of either FGD. The event also needs to fulfil data quality and fiducial volume requirements. The muon is assumed to be the highest momentum negative track (HMNT) found in the event, and it is required that the track is identified as a muon.

The detailed selection criteria for the CC-inclusive sample is:

- **Event quality cut**: The full beam spill has a good global ND280 data quality flag, meaning all ND280 sub-detectors and magnet were operational and reading out data. The event must occur within the bunch time window of the neutrino beam. Event pile-up is mitigated by associating each event to a beam bunch within a beam spill.
- **Quality and fiducial volume cut**: At least one reconstructed track is present in the FGD1 or FGD2 fiducial volumes. The fiducial volume for FGD1 is $|x| <$

874.51 mm, $|y - 55| < 874.51$ mm, $136.875 < z < 446.955$ mm and for FGD2 $|x| < 874.51$ mm, $|y - 55| < 874.51$ mm, $1481.45 < z < 1807.05$ mm[a].

The x and y cuts are designed to accept interactions which have their vertex five bars from the edge of the XY module of each FGD. The z cut excludes the first XY module of each FGD and includes the remaining (14 for FGD1, seven for FGD2). To reject short tracks, for which the TPC reconstruction is unreliable, tracks are required to have more than 18 TPC clusters.

- **Upstream background veto**: If the second highest momentum track starts at least 150 mm upstream of the selected muon candidate the event is rejected. This cut eliminates events in which the muon candidate might be the second part of a broken track which started further upstream. For events with a reconstructed vertex in FGD2 there is the added criterion of having no reconstructed tracks in FGD1.

- **Broken track cut**: The starting position of the muon candidate track needs to be less than 425 mm away from the FGD upstream edge if the event has at least one reconstructed FGD-only track. The cut vetoes events where the reconstruction has cut a muon candidate track into two tracks: one of which is fully contained in the FGD and the other which starts downstream of the fully contained, misplacing the interaction vertex.

- **Muon PID cut**: Once a particle is considered a muon candidate, the particle identification is applied based on the observed dE/dx measurement of the track in the TPC. The measured energy deposit E in the TPC is compared with the expected energy deposit under muon, pion, electron and proton hypotheses and pulls and discrimination functions are applied.

  The pulls δ_i for particle type i are defined as

$$\delta_i = \frac{C_T^{obs} - C_T^{exp}}{\sigma^{exp}} \tag{5.1}$$

  where the expected energy loss C_T^{exp} is parameterised as

$$C_T^{exp} = \frac{53.87 \text{ ADC}}{\beta^{2.283}} \left(5.551 - \beta^{2.283} - \log \left[0.001913 + \frac{1}{(\beta\gamma)^{1.249}} \right] \right) \tag{5.2}$$

[a]The 55mm offset in y reflects the shift in XY modules relative the center of the ND280 coordinate system.

C_T^{obs} is the observed energy loss and σ^{exp} is the deposited energy resolution of the TPC. ADC refers to the analog to digital counts read out from the detector, indicative of the energy deposited by the particle. $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$ are the relativistic variables of the track. The TPC pulls after preselection are shown in Figure 5.3

The likelihoods $\mathcal{L}_i$ are then defined as

$$\mathcal{L}_i = \frac{e^{-\delta_i^2}}{\sum_n e^{-\delta_n^2}} \tag{5.3}$$

where the denominator is summed over the particle types $n = \mu, \pi, e, p$. The likelihoods are shown in Figure 5.2. In the PID algorithm, electrons are rejected by requiring a MIP likelihood

$$\mathcal{L}_{MIP} = \frac{\mathcal{L}_\mu + \mathcal{L}_\pi}{1 - \mathcal{L}_p} > 0.8 \tag{5.4}$$

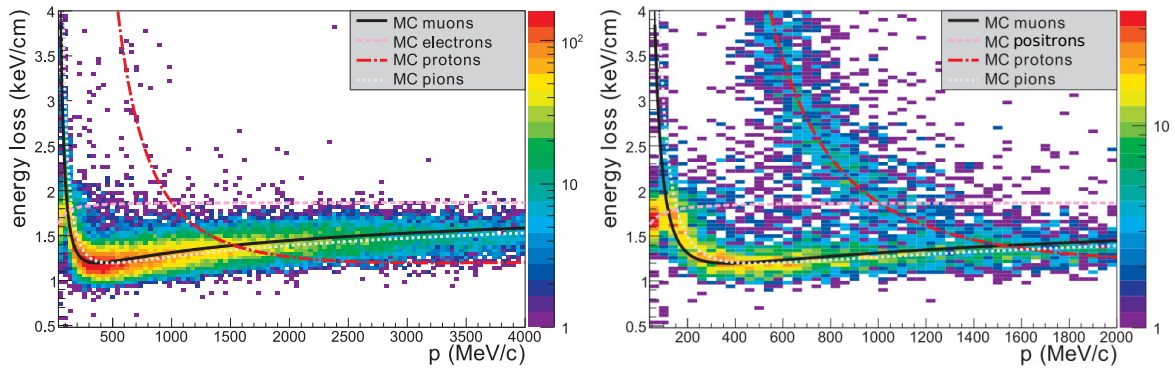
for tracks with $p < 500$ MeV/c. To remove protons and pions, it is also required that

$$\mathcal{L}_\mu > 0.05 \tag{5.5}$$

The constants 0.8, 0.05 and 500 MeV/c are chosen from particle gun studies in the TPC and test-beam data [136, 137, 140]. The energy loss in the TPC from which the pulls are derived are shown in Figure 5.1. Importantly, TPC segments need to pass the TPC track quality cut to contribute to the likelihood: bad quality tracks do not. If a track passes through multiple TPCs all TPC tracks are taken into account.

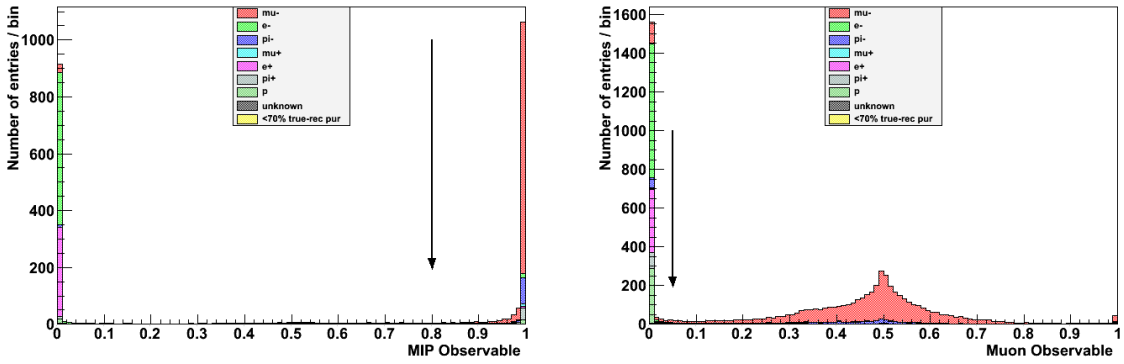
The selection criteria proceeds to split the CC-inclusive sample into the three subsamples: CC0π, CC1π and CCOther. This is based entirely on pion identification in the TPCs and FGDs. To identify pion candidate(s) a number of cuts are applied:

- **Muon candidate**: The track can not be identified as the above muon candidate.
- **Matching beam spill and bunch**: The pion candidate is required to originate from the same beam bunch and spill to the identified muon candidate.
- **Track origin**: The pion candidate is required to start in the same FGD fiducial volume as the muon candidate and enter the downstream TPC for PID purposes.



(a) Negative particles **(b)** Positive particles

Figure 5.1.: The energy loss for particles travelling through the TPC



(a) $\mathcal{L}_{MIP}$ **(b)** $\mathcal{L}_{\mu}$

Figure 5.2.: Likelihood distributions for preselected MC events, showing cuts placed for ν_μ in FHC analysis

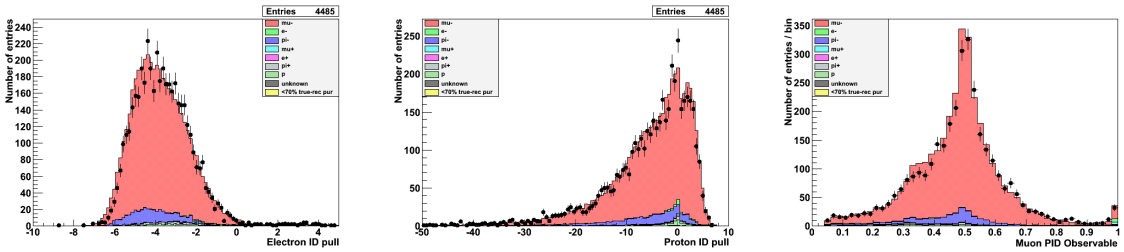
The same FGD and TPC track quality and fiducial volume cut is applied for the pion candidate as for the muon candidate.

- **Pion PID**: For positive tracks in the TPC, pion, positron and proton hypotheses are tested. For negative tracks, pion and electron hypotheses are tested.

  As for the muon candidate, Equation 5.1 and Equation 5.3 define the particle likelihoods. For the pion PID, the MIP likelihood in Equation 5.4 is required and in addition a cut on the pion likelihood is invoked,

$$\mathcal{L}_{\pi} > 0.3 \tag{5.6}$$

When there is no particle track in the TPC, the FGD PID is used to count the number of charged pions. However, it can not be used for neutral pions because



(a) δ_e—the electron ID pull **(b)** δ_p—the proton ID pull **(c)** δ_μ—the muon ID pull

Figure 5.3.: Pull distributions for particle hypotheses after selection showing data and MC for ν_μ analysis

there is currently no electron or positron reconstruction available in the FGD. The FGD pion PID proceeds either by:

- **Michel electron tag**: A search for a Michel electron tag is made for low-momentum tracks that fail to leave enough hits for track reconstruction. It looks for a time-delayed FGD hit cluster out of time with a beam bunch window. A Michel candidate is found if the number of hits in the delayed time bin is greater than six for FGD1 and five for FGD2[b]. Since there is no measurement of the track, the pion has no associated momentum or angle.
- **FGD reconstruction**: Higher momentum pions may leave fully contained tracks in the FGD. If such a track belongs to the same bunch as the muon candidate and there is only one pion track reconstructed in the FGD, it is considered a pion candidate. The pion candidate is required to be upwards or downwards-going by invoking $|\cos\theta_{\pi,\nu}| > 0.3$, which limits the possibility of travelling along the FGD bars. Finally we require a pion pull of $-2 < P_\pi < 2.5$ in the FGD from its track length, shown in Figure 5.4.

Finally, the remaining particles can be identified using the TPC PID:

- For a positive particle, it is tagged according to highest probability. If the most likely particle is a positron but the $p_{reco} > 900$ MeV it's tagged as a proton, otherwise it is a positron.
- For a negative particle, if the probability of a pion is $P_\pi > 0.8$ it is tagged as a negative pion, and if not it is assumed an electron.

[b] Roughly corresponding to 200 photoelectrons, which can't be used as a criteria in FGD2 due to the water layers

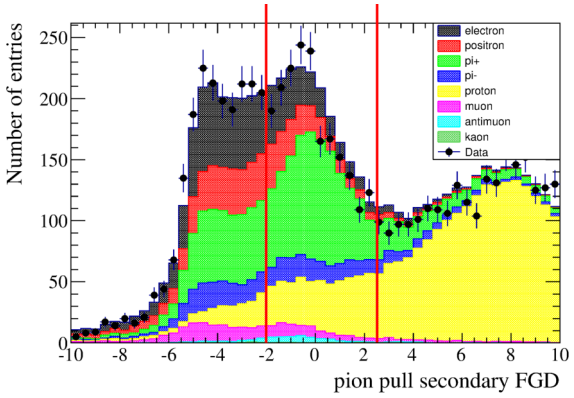
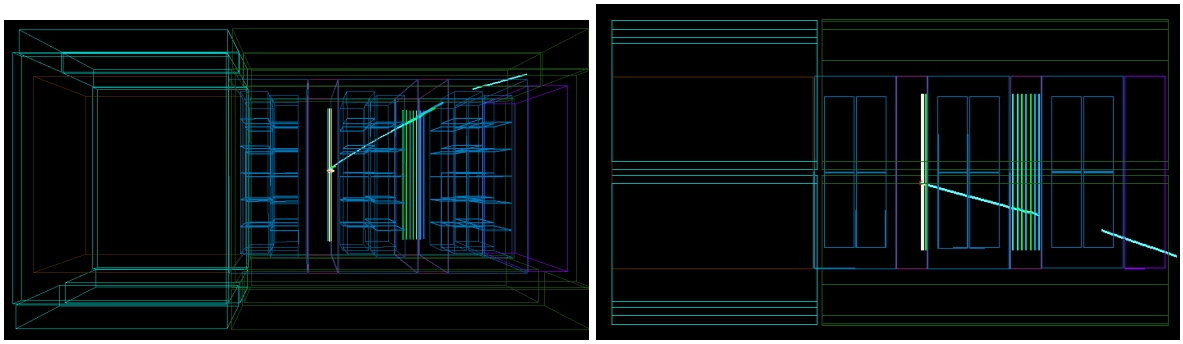


Figure 5.4.: FGD1 pion pulls for a fully contained track

Using information from the TPC PID, FGD Michel electron and FGD PID algorithms, the ν_μ CC-inclusive sample is categorised into the CC0π, CC1π and CCOther samples:

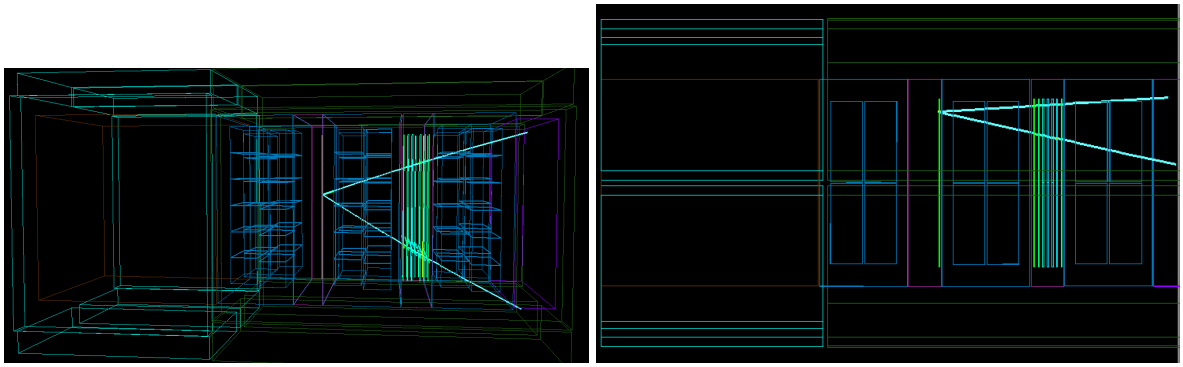
- **CC0π**: Contains events with one negative muon candidate, no identified charged or neutral pions in the TPC or FGD, and no electrons or positrons in the TPC. The selection predominantly contains CCQE and 2p2h events and is the largest sample at ND280. An example event display is shown in Figure 5.5.



(a) Side-view **(b)** Top-view

Figure 5.5.: True FGD1 CC0π event display in ND280

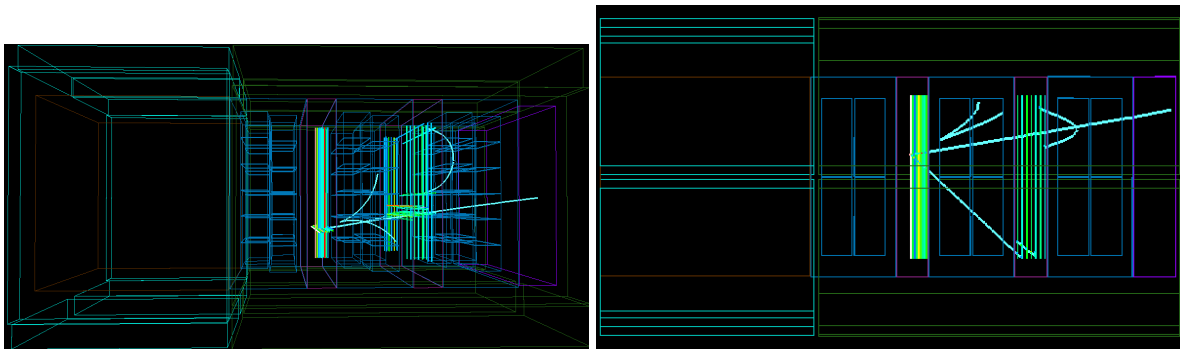
- **CC1π**: Contains events with one negative muon candidate and one positive pion candidate. The sum of the number of positive pions found in the TPC and the number of Michel electrons is one and if there are no Michel electrons the sum of positive pions in the TPC and fully contained in the FGD is one. If there is a negative pion, electron or positron reconstructed in the TPC it is rejected. The selection contains mostly CC1π^+ and multi-π^+ events from resonant and DIS interactions. An example event display is shown in Figure 5.6.



(a) Side-view **(b)** Top-view

Figure 5.6.: True FGD1 CC1π event display in ND280

- **CCOther**: All events with one negative muon candidate which are not classified as CC0π or CC1π fall into this sample. Events with one or more reconstructed negative pion(s), or one or more neutral pion(s) reconstructed as electron or positron candidates in the TPC, are thereby selected. Event with more than one positive pion based on the TPC and FGD pion counting criteria are also accepted. The selection contains mostly multi-π, DIS and CC1π^0 interactions. An example event display is shown in Figure 5.7.



(a) Side-view **(b)** Top-view

Figure 5.7.: True FGD1 CCOther event display in ND280

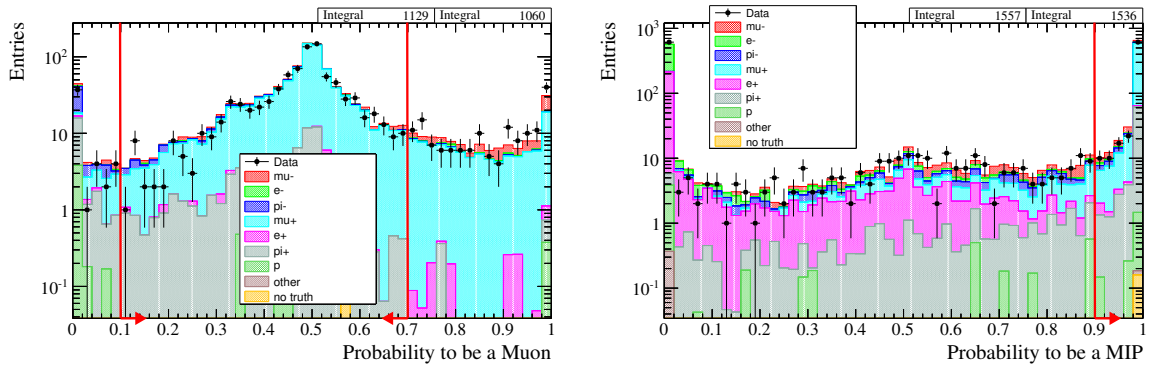
5.1.2. $\bar{\nu}_\mu$ in RHC

The anti-neutrino CC1Track and CCNTrack selections have the same event quality and fiducial volume cut as the neutrino selection, and the muon candidate track is required to pass through the TPC downstream of the struck FGD. The highest momentum track is required to be the highest momentum positive track for its muon PID. The selections

generally have a larger background of "wrong-sign" events: ν_x interactions producing x^- which are identified as μ^+, and ν_x interactions producing a π^+ which may be identified as the lepton candidate. Hence, the selection cuts proceed differently:

- **Positive multiplicity**: The muon candidate track charge is required to be a highest momentum positive track, which removes a large amount of wrong-sign interactions
- **TPC veto**: Veto backwards-going events starting in the FGD and events coming from the P0D and the magnet by utilising the upstream TPCs. If the upstream TPC of an FGD has hits the event is rejected
- **Positive muon identification**: The TPC PID outlined for the ν_μ selections are used to select the positive muon candidate, with the cuts optimised for μ^+.

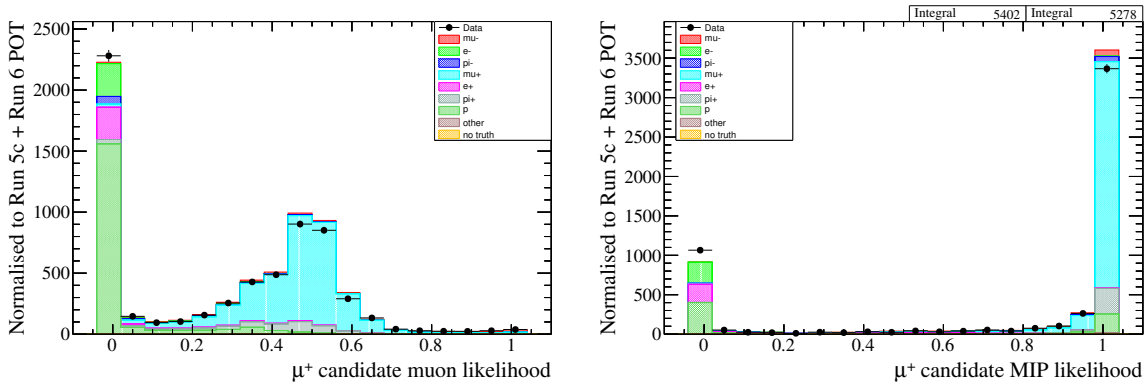
  $\mathcal{L}_{MIP}$ is defined identically to Equation 5.4 although the cut is now placed at 0.9, and still applies only to particles with $p < 500$ MeV. The muon likelihood cut is modified to $0.1 < \mathcal{L}_\mu < 0.7$ which removes protons and positive pions from the ν_μ background. The upper bound at 0.7 is present to reject low energy wrong-sign muons, which may be misidentified as positive tracks. The likelihood distributions and impact of these cuts are shown in Figure 5.8 and Figure 5.9 for the selected lepton candidate. Figure 5.10 shows the TPC PID pulls for run5+6.



(a) $\mathcal{L}_\mu$ **(b)** $\mathcal{L}_{MIP}$

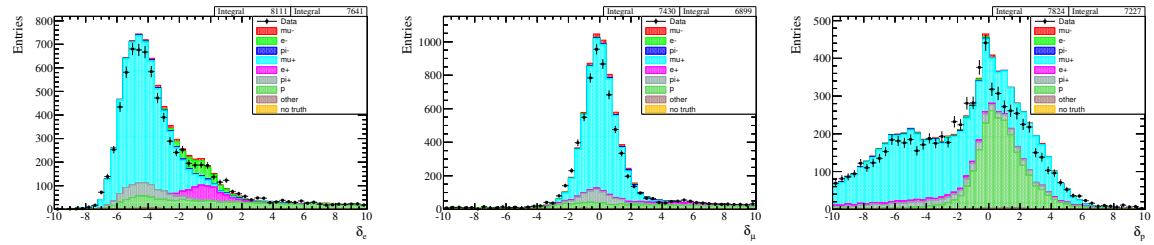
Figure 5.8.: Likelihood distributions for μ and MIP using run5+6 $\bar{\nu}_\mu$ data, used in $\bar{\nu}_\mu$ RHC selections

Once the $\bar{\nu}_\mu$ CC-inclusive selection is run the aforementioned pion reconstruction is applied. The $\bar{\nu}_\mu$ CC1Track selection has one positive muon and does not have any charged or neutral pions in the final state. The $\bar{\nu}_\mu$ CC1Track selection has a higher efficiency in selecting the muon candidate than the ν_μ CC0π selection from the $\bar{\nu}_\mu$



(a) $\mathcal{L}_{\mu}$ **(b)** $\mathcal{L}_{MIP}$

Figure 5.9.: Likelihood distributions for the selected lepton candidate using run5+6 $\bar{\nu}_{\mu}$ data, used in $\bar{\nu}_{\mu}$ RHC selections



(a) δ_e **(b)** δ_{μ} **(c)** δ_p

Figure 5.10.: Pulls (δ) used in the TPC PID used in $\bar{\nu}_{\mu}$ RHC selections

resonant interactions producing a π^-, not a π^+, which can capture on the nucleus and so do not leave a track to be (falsely) identified as a muon candidate. The $\bar{\nu}_{\mu}$ CCNTrack selection contains the remaining particles passing the $\bar{\nu}_{\mu}$ CC-inclusive selection, containing at least one neutral or charged pion.

5.1.3. ν_{μ} in RHC

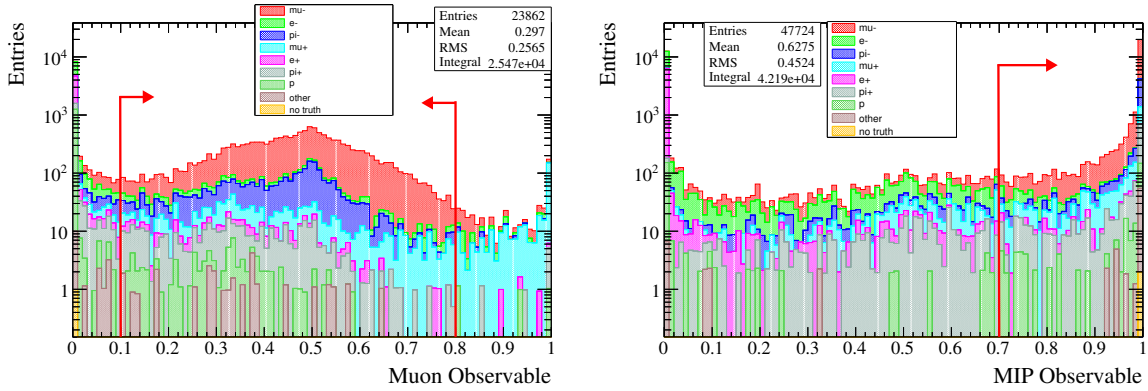
In RHC running there is a large fraction of ν_{μ} interactions, owing mostly to the larger ν_{μ} cross-section. The same pre-selection cuts are applied for the ν_{μ} in RHC selection as for the previous selections.

The CC-inclusive selection proceeds by:

- **Negative multiplicity**: The highest momentum track is required to be the highest momentum negative track, which starts the seeding track. The μ^- identification uses the TPC PID on the highest momentum negative track.

- **TPC PID**: The PID proceeds by the MIP requirement in Equation 5.4 for particles with $p_\mu < 500$ MeV/c, accepting candidate tracks with $\mathcal{L}_{MIP} > 0.7$.

  Similar to subsection 5.1.2, a lower and upper bound is set $0.1 < \mathcal{L}_\mu < 0.8$, which rejects protons and low momentum μ^+. The effect of these cuts can be seen in Figure 5.11.



(a) $\mathcal{L}_\mu$ **(b)** $\mathcal{L}_{MIP}$

Figure 5.11.: Likelihood distributions for μ and MIP using run5+6 $\bar{\nu}_\mu$ data, used in ν_μ in RHC selections

The ν_μ in RHC selection then breaks down the CC-inclusive selection into CC1Track and CCNTrack, based entirely on the number of TPC-FGD matched tracks. Events with one such reconstructed track enters the CC1Track selection, and events with any other number of tracks regardless of PID, enter the CCNTrack selection. Hence the ν_μ RHC selection is analogous to the $\bar{\nu}_\mu$ RHC selection.

A summary of all selections' efficiency and purities is shown in Table 5.2. We note $90+$ % efficiency for CC0π and 1 right-sign 1Track efficiencies, with about 75% purity. As the track multiplicity increases, the efficiencies and purities drop. The NTrack selections in $\bar{\nu}_\mu$ perform the worst at 55% efficiency and 45% purity. For more detail see Appendix A.

5.2. Binning the Selections

We expect largely similar kinematics across the two FGDs so apply the same binning in reconstructed muon momentum, p_μ, and cosine of the average neutrino-muon angle, $\cos\theta_\mu$. The binning for the fit is primarily influenced by MC statistics: we

Selection	Efficiency (%)	Purity (%)
FGD1 CC0π ν_μ	93.8	75.5
FGD2 CC0π ν_μ	93.2	73.5
FGD1 CC1π ν_μ	83.3	58.0
FGD2 CC1π ν_μ	83.1	57.1
FGD1 CCOther ν_μ	73.0	65.3
FGD2 CCOther ν_μ	73.4	64.9
FGD1 CC1Track $\bar{\nu}_\mu$	90.0	76.7
FGD2 CC1Track $\bar{\nu}_\mu$	89.6	76.7
FGD1 CCNTrack $\bar{\nu}_\mu$	54.1	45.1
FGD2 CCNTrack $\bar{\nu}_\mu$	53.8	43.9
FGD1 CC1Track ν_μ in RHC	76.5	52.2
FGD2 CC1Track ν_μ in RHC	74.9	51.8
FGD1 CCNTrack ν_μ in RHC	73.9	60.9
FGD2 CCNTrack ν_μ in RHC	74.2	61.4

Table 5.2.: Efficiency and purity summary for all selections with the range $0 < p_{reco} < 3$ GeV/c

require $\sim$20 raw MC events per bin (roughly equivalent to 1-2 data events). The momentum resolution is $\sim$50 MeV up to 1 GeV and the angular resolution $\sim 2^\circ$, seen in Appendix H.

The binning in p_μ (MeV/c) $\cos\theta_\mu$ for each sample is shown below. The FHC selections all have similar binning and has the highest number of bins. The total number of bins is 1624, of which 902 are FHC (six selections) and 722 are RHC (eight selections).

- FGD1+2 CC0π, CC1π and CCOther ν_μ: 154 bins CC0π, CCOther; 143 bins CC1π
  p_μ: 0, 300, 400, 500, 600, 700, 800, 900, 1000, 1250, 1500, 2000, 3000 (not for CC1π), 5000, 30000
  $\cos\theta_\mu$: -1, 0.6, 0.7, 0.8, 0.85, 0.9, 0.92, 0.94, 0.96, 0.98, 0.99, 1

- FGD1+2 CC1Track $\bar{\nu}_\mu$: 130 bins
  p_μ: 0, 400, 500, 600, 700, 800, 900, 1100, 1400, 2000, 10000
  $\cos\theta_\mu$: -1.0, 0.6, 0.7, 0.8, 0.85, 0.88, 0.91, 0.93, 0.95, 0.96, 0.97, 0.98, 0.99, 1

- FGD1+2 CCNTrack $\bar{\nu}_\mu$: 77 bins
  p_μ: 0, 700, 950, 1200, 1500, 2000, 3000, 10000
  $\cos\theta_\mu$: -1.0, 0.75, 0.85, 0.88, 0.91, 0.93, 0.95, 0.96, 0.97, 0.98, 0.99, 1

- FGD1+2 CC1Trk ν_μ in RHC: 66 bins
  p_μ: 0, 400, 600, 800, 1100, 2000, 10000
  $\cos\theta_\mu$: -1.0, 0.7, 0.8, 0.85, 0.9, 0.93, 0.95, 0.96, 0.97, 0.98, 0.99, 1
- FGD1+2 CCNTrk ν_μ in RHC: 88 bins
  p_μ: 0, 500, 700, 1000, 1250, 1500, 2000, 3000, 10000
  $\cos\theta_\mu$: -1.0, 0.7, 0.8, 0.85, 0.9, 0.93, 0.95, 0.96, 0.97, 0.98, 0.99, 1

5.3. Systematics

> "The garbage of the past often becomes the treasure of the present (and vice versa)"
>
> A. M. Polyakov at Gauge Fields and Strings, London, 1987

The fit aims to constrain systematic parameters for T2K-SK oscillation analyses by using near-detector data. The shared parameters between ND280 and SK are the neutrino flux parameters, and neutrino-nucleus interaction parameters. The nuisance parameters can be considered as the ND280 detector parameters and cross-section parameters that are parametrised as effective on Carbon only. As such, there are many parameters of interest in this analysis, which this section covers.

The systematics enter the fit by changing the predicted event distributions by shape and/or normalisation, which often incur a likelihood penalty for moving parameters away from their prior central values. The penalty takes two forms: either Gaussian or constant. In the case where there is firm reason to believe a parameter is constrained from other sources than ND280 data, the Gaussian penalty is imposed. When external data and/or recent model developments indicate lacking or conflicting knowledge of a parameter, a flat prior is chosen.

For the Gaussian penalty we have

$$-2\log\mathcal{L}_{\text{Penalty}} = (X_i - \mu_i)\,(\mathbf{V})^{-1}_{i,j}\,(X_j - \mu_j) \tag{5.7}$$

for parameter i, with current value X_i, constant priors μ_i and covariance matrix $\mathbf{V}$. For a flat penalty instead

$$-2\log\mathcal{L}_{\text{Penalty}} = C \tag{5.8}$$

where C is a constant.

5.3.1. The Beamline and Neutrino Flux

The flux systematics are split into six categories:

- Hadron interaction uncertainties
- Proton beam profile and off-axis angle
- Horn current and field
- Horn and target alignment
- Materials modelling
- Number of protons on target

The simulations are updated each year to improve the modelling, often taking new data in to account. An example is using the dedicated NA61/SHINE T2K replica target data [172] to tune the hadron production model at the T2K beam target, and including results from the HARP experiment [173].

The fractional errors for the ND280 neutrino flux prediction are shown in Figure 5.12 for FHC running and Figure 5.13 for RHC running. The uncertainties are $\sim$10% in the flux peak region and are dominated by hadron interaction uncertainties, which arise primarily from multiplicity, pion rescattering and interaction length uncertainties. The proton beam profile and off-axis angle become important shortly after the flux peak at about 1 GeV for the right-sign component of the flux.

Importantly, the hadronic interaction uncertainties are reducible by improved modelling and tuning to hadron production data. An example is the black dashed line and the black solid line in Figure 5.12, which shows the reduction in flux uncertainty from 2014 to 2015 analyses. Additionally, new in-situ beam profile monitors aid in reducing the proton beam profile and off-axis angle contributions.

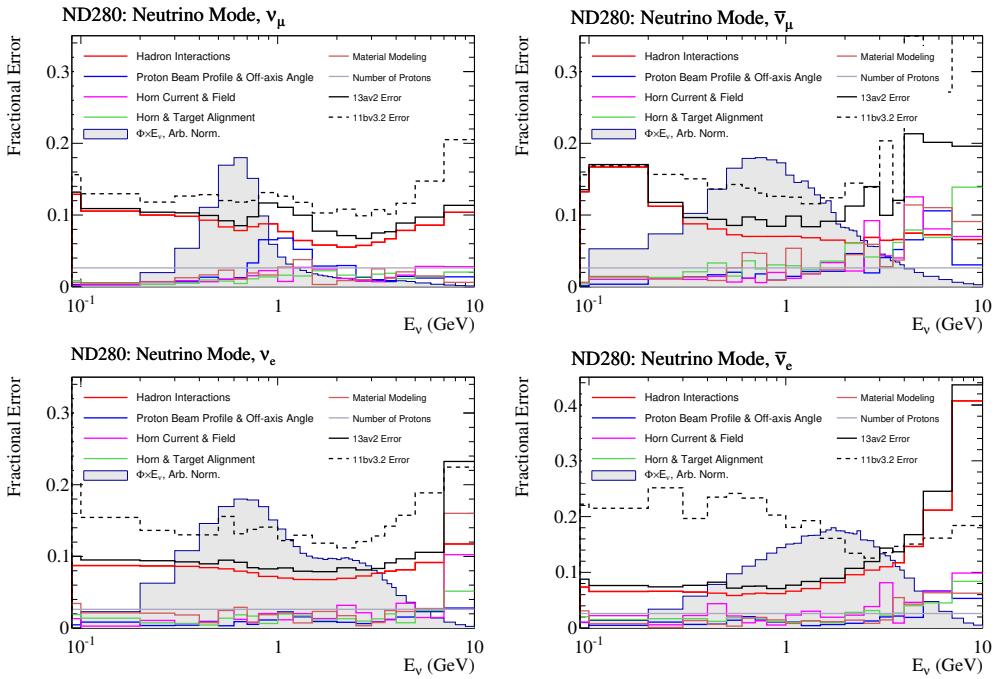


Figure 5.12.: FHC flux uncertainties, "13av2 Error" is used for this analysis

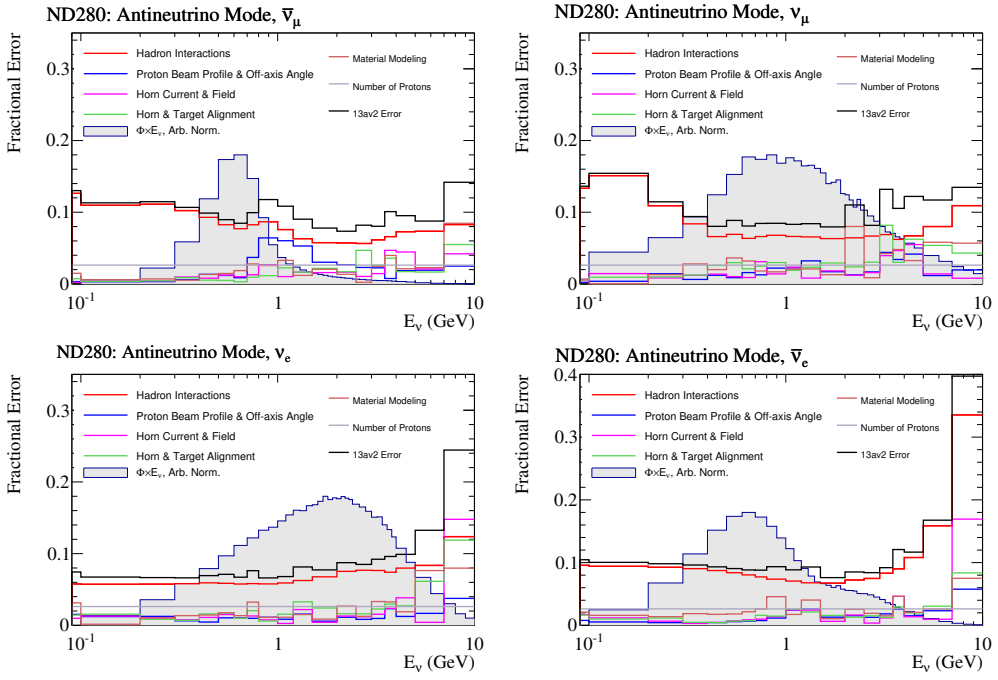


Figure 5.13.: RHC flux uncertainties, "13av2 Error" is used for this analysis

The flux systematics enter the near-detector and oscillation analyses as bin-by-bin normalisations in true neutrino energy, E_ν^{true} (GeV), for four different neutrino species (ν_μ, $\bar{\nu}_\mu$, ν_e, $\bar{\nu}_e$), for each running mode (FHC, RHC), for each detector (ND280, SK). The binning is chosen to reflect the magnitude of the neutrino flux and the changing shape

but simultaneously keeping the number of parameters relatively low. The right-sign and wrong-sign species have the same binning, so ND280 FHC ν_μ is binned the same as ND280 RHC $\bar{\nu}_\mu$:

- ND280, SK FHC ν_μ; ND280, SK RHC $\bar{\nu}_\mu$, :
  E_ν^{true}: 0, 0.4, 0.5, 0.6, 0.7, 1, 1.5, 2.5, 3.5, 5, 7, 30
- ND280, SK FHC $\bar{\nu}_\mu$; ND280, SK RHC ν_μ:
  E_ν^{true}: 0, 0.7, 1, 1.5, 2.5, 30
- ND280, SK FHC ν_e; ND280, SK RHC $\bar{\nu}_e$:
  E_ν^{true}: 0, 0.5, 0.7, 0.8, 1.5, 2.5, 4, 30
- ND280, SK FHC $\bar{\nu}_e$; ND280, SK RHC ν_e:
  E_ν^{true}: 0, 2.5, 30

This procedure brings the total number of flux parameters to 100: 50 for ND280 and 50 for SK. The SK flux parameters are not directly constrained in the ND280-only analysis: the strong correlation between the flux at ND280 and SK indirectly moves SK flux parameters when the ND280 flux parameters move. Hence, all 100 parameters are included in the ND280-only analysis. The flux parameters are highly correlated so the likelihood penalties are evaluated with a covariance matrix, shown in Figure 5.14.

In addition to the variation systematics above, there is also a nominal flux correction applied to each event as a function of its run period (e.g. run 2a), neutrino specie (e.g. $\bar{\nu}_\mu$) and E_ν^{true} (e.g. 0.8 GeV); this is present to correct the nominal flux model with updated measurements. An example of the corrections from run 4a and run 5b is shown in Figure 5.15.

5.3.2. The ND280 Detector

The treatment of ND280 detector systematic uncertainties consists of varying the underlying detector systematics—such as the TPC PID, FGD PID, and TPC Momentum scale—and study the impact on the number of predicted events in each p_μ, $\cos\theta_\mu$ bin.

The parameterisation of near detector systematics are categorised into pure event or selection-changing. The event weighting can be broken down further into efficiencies (with shapes) and normalisation parameters. When FGD-related systematics are concerned—such as pion tagging by Michel electron identification—the two FGDs

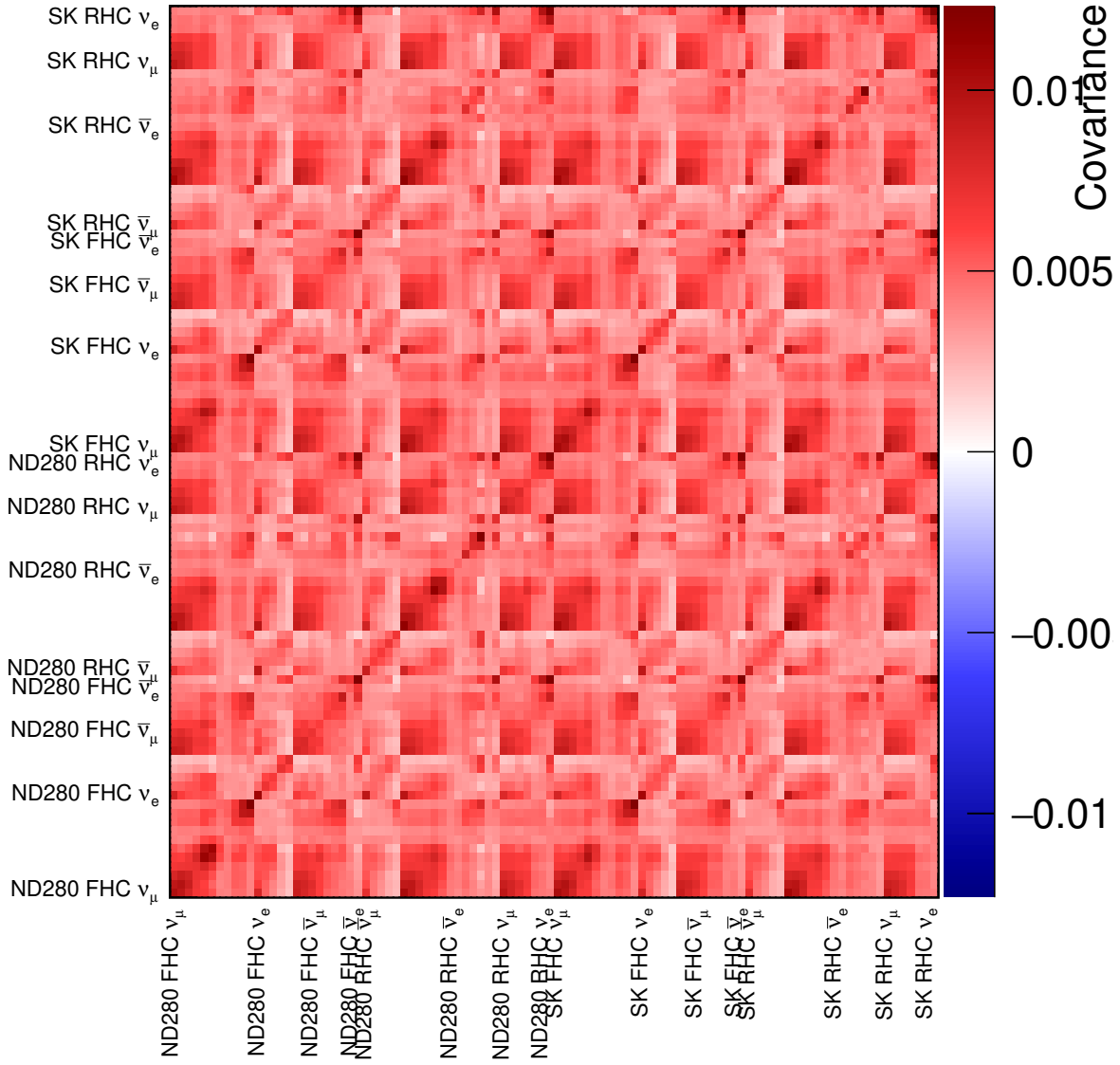
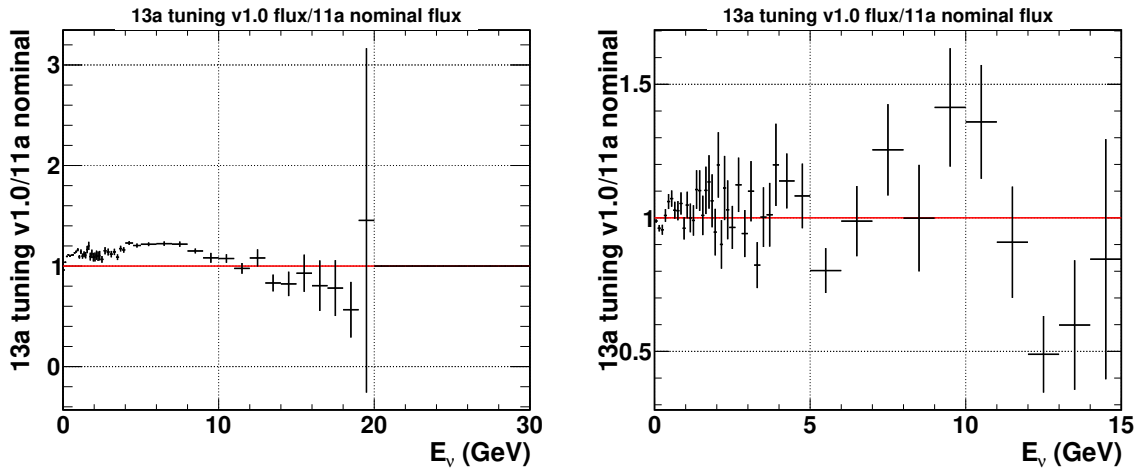


Figure 5.14.: 13av2 neutrino flux covariance matrix, used in this analysis

have separate implementations to account for geometric and compositional differences. The different sources of systematics, their variation type and assumed probability distribution function (PDF) are shown in Table 5.3.

Observable Variation Systematics This group of systematics have the potential to change the reconstructed topology, so allow for migration in and out of selections. They can also switch the reconstructed lepton candidate to a different track in the event. The systematic is applied as a smearing to the reconstructed event variables (e.g. p_μ, $\cos\theta_\mu$) and then reruns the event selection algorithm on the smeared event.



(a) 4a ν_μ **(b)** 5b $\bar{\nu}_\mu$

Figure 5.15.: Neutrino flux corrections that are applied to events interacting in the ND5 (ND280 tracker) plane. 13a tuning denotes the latest tuning, 11a nominal denotes the reference neutrino flux which was used in the simulation

There are two methods with which the smearing is applied:

- If the relevant reconstructed variable has a known true value, the difference between the two is used as a scaling. The updated value of the variable after the variation is then

$$x'_{reco} = x_{true} + \left(x^{MC}_{reco} - x_{true}\right)(s + \alpha \cdot \delta s) \tag{5.9}$$

where α is the random variable from the relevant systematic's PDF in Table 5.3, s is the scaling factor, and δs is its statistical error. The scaling factor is defined as

$$s = \frac{\sigma^{data}}{\sigma^{MC}} \tag{5.10}$$

and

$$\delta s = s \cdot \left| \frac{\delta\sigma^{data}}{\sigma^{Data}} - \frac{\delta\sigma^{MC}}{\sigma^{MC}} \right| \tag{5.11}$$

where σ^{data} is the dispersion observed in data and $\delta\sigma^{data}$ is the error on the dispersion.

- If the MC is corrected to match a data mean value. This correction is needed because the effect of a systematic error ($\delta\Delta\bar{x}$) on the selected event relative the nominal MC is not guaranteed to agree with the corrected MC. The updated

Systematic	Variation	PDF
TPC		
Magnetic Field Distortions	Observable	Flat
TPC Momentum Scale	Observable	Gauss
TPC Momentum Resolution	Observable	Gauss
TPC PID	Observable	Gauss
TPC Cluster Efficiency	Efficiency	Gauss
TPC Tracking Efficiency	Efficiency	Gauss
TPC Charge ID Efficiency	Efficiency	Gauss
FGD-TPC		
TPC-FGD Matching Efficiency	Efficiency	Gauss
FGD		
FGD PID	Observable	Gauss
FGD1-FGD2 Time of Flight	Observable	Gauss
FGD Hybrid Tracking Efficiency	Efficiency	Gauss
Michel Electron Efficiency	Efficiency	Gauss
Backgrounds		
Out-of-Fiducial-Volume	Normalisation	Gauss
Sand Muons	Normalisation	Gauss
Pile-Up	Normalisation	Gauss
MC modelling		
Pion secondary interactions	Normalisation	Gauss
FGD Mass	Normalisation	Gauss

Table 5.3.: ND280 systematics present in the fit

observable is then

$$x'_{reco} = x^{MC}_{reco} + \Delta\bar{x} + \alpha\delta\Delta\bar{x} \tag{5.12}$$

where $\Delta\bar{x} = \bar{x}^{data}_{reco} - \bar{x}^{MC}_{reco}$ and $\bar{x}_{reco}$ is the mean value of the variable x, α is a random variable, and $\delta\Delta\bar{x}$ is the associated uncertainty from the reconstructed data and MC discrepancy,

$$\delta\Delta\bar{x} = \sqrt{\Delta\bar{x}^2 + \left(\delta\bar{x}^{data}_{reco}\right)^2 + \left(\delta\bar{x}^{MC}_{reco}\right)^2} \tag{5.13}$$

Additionally uncertainties from the magnetic field are special cases of the above:

- The TPC laser calibration corrections are used on-top of the B-field mapping corrections, where the latter is applied during reconstruction and the former is treated as an uncertainty. The reconstructed variable is then

$$x'_{reco} = x^{MC}_{reco} + \alpha \left(x^{New}_{reco} - x^{MC}_{reco} \right) \tag{5.14}$$

  where α is the random variable and x^{New}_{reco} is the reconstructed momentum after the updated mapping is applied. This applies to the magnetic field distortion systematic.

- If the observable depends on a scale s that is easy to extract from in-situ measurements

$$x'_{reco} = x^{MC}_{reco} + \alpha \delta x \tag{5.15}$$

  where $\delta x = x^{MC}_{reco} \delta s$ is the uncertainty on the observable and δs is the uncertainty on the scaling variable. The TPC momentum scale systematic uses this parameterisation, in which s is the scale of the magnet current and δs is its uncertainty.

Efficiency Systematics The weight systematics are computed from studies which compare data and MC predictions in well known control samples. The multiple ND280 subdetectors enable cross-checks for tracking and matching efficiencies: e.g. TPC2 tracking efficiency can be computed using tracks with segments in FGD1 and FGD2, which therefore should also have a track in TPC2 which can be cross-checked.

Using the sand muon control sample—defined as having a through-going muon track in most of the detector, where the muon was created in the surrounding sand in the ND280 pit or the magnet—as an example, such muons tend to be very forward-going and high energy. Thus the sample is suitable for alignment studies but not efficiency studies since the p_μ, $\cos\theta_\mu$ phase space is very limited.

The model used to move to some new phase space assumes the ratio between efficiencies in data and MC are the same in analysis and control samples. The efficiency for the data is then

$$\epsilon_{data} = \frac{\epsilon^{Control}_{data}}{\epsilon^{Control}_{MC}} \epsilon_{MC} \tag{5.16}$$

where $\epsilon^{Control}$ is the efficiency in the control sample(s). The statistical uncertainty in $r^{Control} = \epsilon^{Control}_{data}/\epsilon^{Control}_{MC}$ is taken into account as

$$\delta r^{Control} = \sqrt{\left(1 - r^{Control}\right)^2 + \left(\delta r^{Control}_{Stat}\right)^2} \tag{5.17}$$

yielding the predicted efficiency in the data as

$$\epsilon'_{data} = \left(r^{Control} + \alpha \delta r^{Control}\right) \epsilon_{MC} \tag{5.18}$$

where α is the random variable. Finally we define the two weights

$$w_{Eff} = \frac{\epsilon'_{data}}{\epsilon_{MC}} \tag{5.19}$$

for events that identify the track correctly and

$$w_{Ineff} = \frac{1 - \epsilon'_{Data}}{1 - \epsilon_{MC}} \tag{5.20}$$

to propagate the weight systematics on an event-by-event basis.

Normalisation Systematics These systematics are simple one-time weights which change the overall event numbers. The FGD mass error is an example of such a systematic: if the mass of the FGD is larger than the nominal, the overall number of observed events in MC should be increased. The weight w is applied as

$$w = 1 + \alpha \cdot \delta e \tag{5.21}$$

where 1.0 is the nominal weight, α is a random variable, and δe is the systematic error on the source.

Relative Systematic Error Sizes Table 5.4 shows the relative errors from ND280 systematics contribution on the number of predicted events for the different ND280 ν_μ selections for FGD1. The total error for the CC0π selection is 1.66%, CC1π is 3.33% and CCOther 6.47%. For comparison, the flux and cross-section errors are generally $\mathcal{O}\left(10\%\right)$. The error increases with selection because the increased track multiplicity and lower "cleanliness" of the reconstructed events.

The largest contribution to the total error is pion secondary interactions, making up $\sim$90% of the detector systematics. This systematic has the power to migrate events through selections by changing the number of reconstructed pions. For the CCOther selection the TPC tracking efficiency also has a large contribution (1.79%), decreasing to 0.44% for CC1π and 0.27% for CC0π. The FGD mass contributes 0.6% uncertainty for all selections, making up 1/3 of the error on CC0π selection. FGD2 has similarly sized systematic contributions, although slightly modified due to the geometry: e.g. the sand muon error is smaller in FGD2 due to tighter vetos than in FGD1.

Systematic	Percentage error		
	CC0π	CC1π	CCOther
TPC			
Magnetic Field Distortions	0.025	0.063	0.072
TPC Momentum Scale	0.062	0.074	0.230
TPC Momentum Resolution	0.055	0.094	0.286
TPC PID	0.316	0.792	0.616
TPC Cluster Efficiency	0.000	0.000	0.002
TPC Tracking Efficiency	0.259	0.440	1.786
TPC Charge ID Efficiency	0.178	0.270	0.473
FGD-TPC			
TPC-FGD Matching Efficiency	0.148	0.270	0.605
FGD			
FGD PID	0.011	0.034	0.015
FGD1-FGD2 Time of Flight	0.034	0.070	0.017
FGD Hybrid Tracking Efficiency	0.106	0.100	0.532
Michel Electron Efficiency	0.062	0.253	0.008
Backgrounds			
Out-of-Fiducial-Volume	0.391	0.541	0.286
Sand Muons	0.069	0.085	0.031
Pile-Up	0.112	0.112	0.112
MC modelling			
Pion secondary interactions	1.433	3.173	6.118
FGD Mass	0.595	0.595	0.595
Total	1.660	3.329	6.467

Table 5.4.: Integrated systematic errors for FGD1 FHC related systematics

Table 5.5 shows the error table for the RHC selections. The tracking related systematics for the FGD (Michel electron tagging, PID, hybrid tracking) are not present because the number of tracks in the event—defining the 1Trk or NTrk selection—is solely based on the number of TPC tracks present, so has no impact.

The largest systematic by far is the pion secondary modelling (90+%), as is also the case for FHC selections. The impact is larger than for FHC selections since the source of pions are ν_μ interactions (typically higher in E_ν so also higher in multiplicity) and the larger uncertainties on π^- re-interaction probabilities. The FGD mass is the second largest systematic for CC0π, although for the CCNTrk selection the TPC tracking efficiency comes in second, followed by the TPC PID, followed by the FGD mass.

Systematic	Percentage error	
	CC1Trk	CCNTrk
TPC		
Magnetic Field Distortions	0.004	0.165
TPC Momentum Scale	0.049	0.246
TPC Momentum Resolution	0.041	0.123
TPC PID	0.307	0.544
TPC Cluster Efficiency	0.000	0.002
TPC Tracking Efficiency	0.436	1.201
TPC Charge ID Efficiency	0.117	0.115
FGD-TPC		
TPC-FGD Matching Efficiency	0.109	0.394
FGD		
FGD1-FGD2 Time of Flight	0.016	0.009
Backgrounds		
Out-of-Fiducial-Volume	0.336	0.610
Sand Muons	0.153	0.248
Pile-Up	0.240	0.241
MC modelling		
Pion secondary interactions	4.902	9.198
FGD Mass	0.598	0.584
Total	5.371	10.378

Table 5.5.: Integrated systematic errors for FGD1 RHC related systematics

Parameterisation of ND280 Systematics The systematics in Table 5.3 could theoretically be varied on an event-by-event basis. In practice this is unfeasible because: 1) the event selection framework is not sufficiently optimised to guarantee fast reweighting; 2) some values of variation systematics gave rise to discontinuous test-statistics as events migrated from one topology to another. Whereas the former is purely computational, the latter causes problems for finding minima with gradient descent algorithms, employed by the other ND280 fitting group (BANFF).

The systematics are instead parameterised similarly to the flux systematics, which ensures both smoothness and fast reweighting. The systematics listed in Table 5.3 are varied on an event-by-event basis and 500 random variations are chosen according to the prior covariances, and the number of events are binned in the fit-binning from section 5.2. The content of each bin is then treated as a normalisation parameter which is highly correlated with adjacent bins through a covariance matrix. Finally, an MC statistical covariance matrix and a covariance matrix shifting the MC reconstructed lepton momentum of CCQE events by 20 MeV to roughly emulate the differences in Martini and Nieves' 1p1h model to the NEUT CCQE model are added [108]. The central value and uncertainty on the number of events in a bin comes from a the arithmetic mean and the central value is the rms, and a cross-check with a Gaussian fit is done.

Using the fit binning in $p_\mu \cos\theta_\mu$ yields 1624 ND280 parameters, which was reduced to 556 by merging bins with similar features in the underlying bin-by-bin event distributions. The final ND280 detector matrix binning was chosen to be

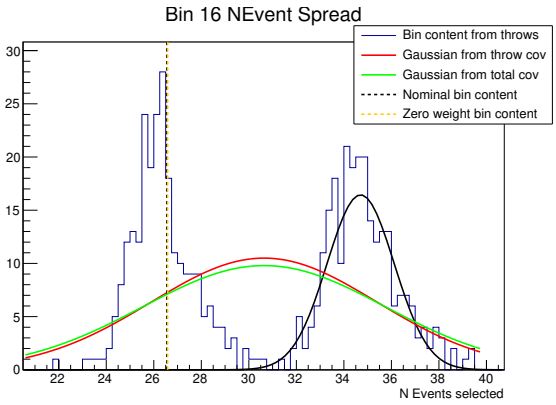
- FHC ν_μ CC0π bin edges:
  p_μ (MeV/c): 0, 1000, 1250, 2000, 3000, 5000, 30000
  $\cos\theta_\mu$: -1, 0.6, 0.7, 0.8, 0.85, 0.94, 0.96, 1
- FHC ν_μ CC1π bin edges:
  p_μ (MeV/c): 0, 300, 1250, 1500, 5000, 30000
  $\cos\theta_\mu$: -1, 0.7, 0.85, 0.9, 0.92, 0.96, 0.98, 0.99, 1
- FHC ν_μ CCOther bin edges:
  p_μ (MeV/c): 0, 1500, 2000, 3000, 5000, 30000
  $\cos\theta_\mu$: -1, 0.8, 0.85, 0.9, 0.92, 0.96, 0.98, 0.99, 1

- RHC $\bar{\nu}_\mu$ CC 1-Track bin edges:
  p_μ (MeV/c): 0, 400, 900, 1100, 2000, 10000
  $\cos\theta_\mu$: -1, 0.6, 0.7, 0.88, 0.95, 0.97, 0.98, 0.99, 1.00
- RHC $\bar{\nu}_\mu$ CC N-Track bin edges:
  p_μ (MeV/c): 0, 700, 1200, 1500, 2000, 3000, 10000
  $\cos\theta_\mu$: -1, 0.85, 0.88, 0.93, 0.98, 0.99, 1.00
- RHC ν_μ CC 1-Track bin edges:
  p_μ (MeV/c): 0, 400, 800, 1100, 2000, 10000
  $\cos\theta_\mu$: -1, 0.7, 0.85, 0.90, 0.93, 0.96, 0.98, 0.99, 1.00
- RHC ν_μ CC N-Track bin edges:
  p_μ (MeV/c): 0, 1000, 1500, 2000, 3000, 10000
  $\cos\theta_\mu$: -1, 0.8, 0.90, 0.93, 0.95, 0.96, 0.97, 0.99, 1.00

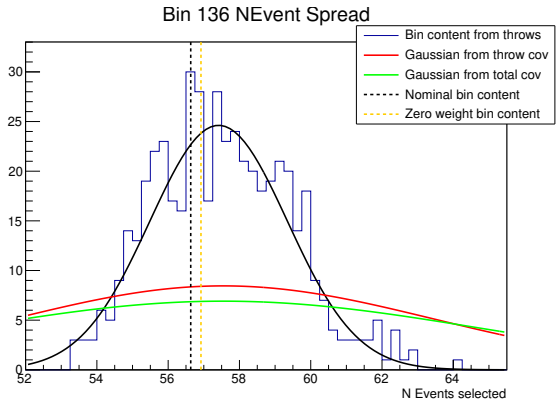
in which we note a large first momentum bin for many selections. This is primarily because the pion secondary interaction systematic is by far dominant, which generally becomes larger at higher p_μ (which in turn correlates with higher E_ν), leading to enough energy to create a pion. Reducing the ND280 detector systematic binning had no discernible effect for flux or interaction parameters in a fit to mock data.

The result of the detector systematics procedure is shown in Figure 5.16 for a selection of bins. In FGD1 CC0π the worst overall example is the p_μ =1250-2000, $\cos\theta_\mu$ =0.6-0.7 bin, which has a double Gaussian bimodal behaviour. This is likely from events migrating in and out of the $\{p_\mu, \cos\theta_\mu, \text{sample}\}$ combination due to pion secondary interactions. Although a simple fit to the bin content chooses one of the peaks, the covariance matrix entry sits in between the two with an error that covers the bimodality.

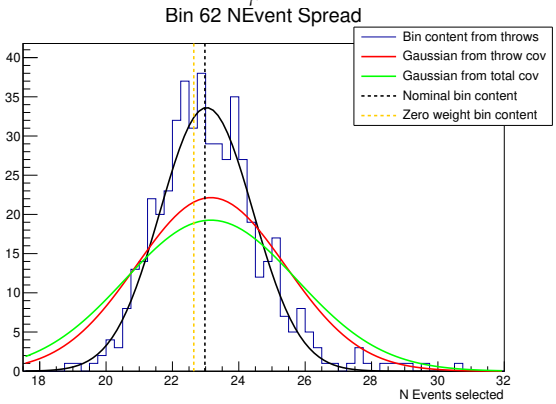
The final covariance matrix is seen in Figure 5.17, where the majority of bins are highly correlated. The correlation matrix is shown in Figure 5.18. We particularly see high correlations for the high-momentum bins (towards the end of each selection). The anti-correlations enter mostly for RHC selections, where bins at low p_μ correlate negatively with bins at high p_μ. There are also anti-correlations in FGD1 vs FGD2 selections.



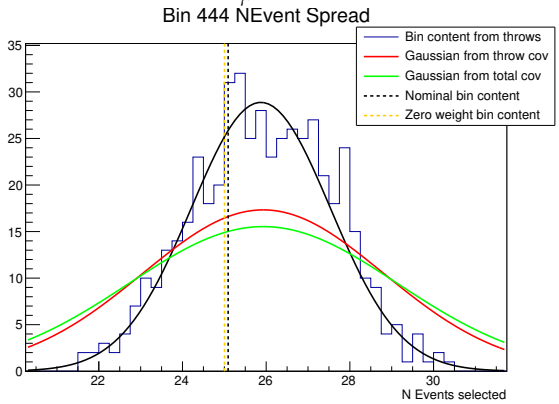
(a) FGD1 CC0π selection, $p_\mu = 1250 - 2000$ MeV/c, $\cos\theta_\mu = 0.6 - 0.7$



(b) FGD1 CC1Trk selection, $p_\mu = 0 - 400$ MeV/c, $\cos\theta_\mu = 0.97 - 0.98$



(c) FGD1 CC1π selection, $p_\mu = 300 - 1250$ MeV/c, $\cos\theta_\mu = 0.9 - 0.92$



(d) FGD2 CCNTrk selection, $p_\mu = 0 - 700$ MeV/c, $\cos\theta_\mu = 0.93 - 0.98$

Figure 5.16.: Number of events in various bins used for parametrising the ND280 detector systematics, using the full suite of uncertainties from the ND280 systematics. Each histogram is representative of the uncertainty in the bin due to detector systematics, and the conservative uncertainties on the distribution are shown (red, green). The populations are also fit with a Gaussian for comparison (black)

5.3.3. The Neutrino-Matter Interaction

The parameterisation of the interaction systematics is frequently updated in T2K analyses to account for new theoretical calculations. This includes nuclear in-medium effects such as RPA and 2p2h corrections [47, 48], initial state models such as Local Fermi Gases [174] and Spectral Functions [50], and parameter tunes of existing models [175].

The T2K and SK experiments both use the custom interaction library NEUT 5.3.3 [53] as their primary neutrino event generator, with cross-checks through GENIE [55]

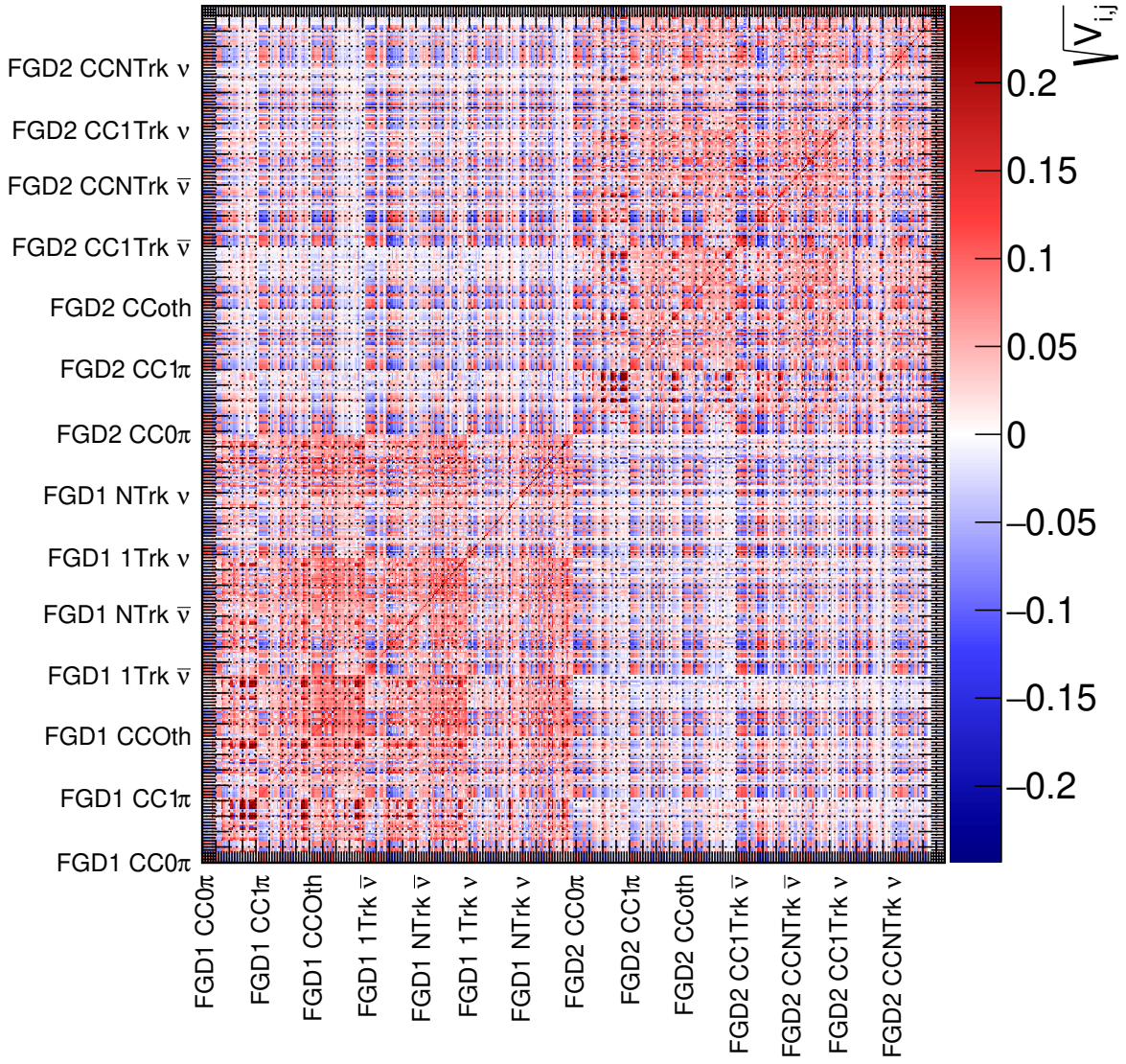


Figure 5.17.: $\mathrm{sgn}(V_{i,j}) \times \sqrt{V_{i,j}}$ for the ND280 systematic parameters

and NuWro [56]. Hence the interaction parameterisation deals with the models implemented in NEUT and are outlined here.

CCQE and CC0π The nominal model is generated with a Spectral Function (SF) from Benhar and others [50] and a 2-particle-2-hole (2p2h) excitation [47,48] for the CCQE/CC0π model. An alternative model uses the Llewellyn-Smith model [176] with a dipole axial form factor and BBBA05 vector form factors [177] coupled to a Smith-Moniz Relativistic Fermi Gas (RFG) [178].

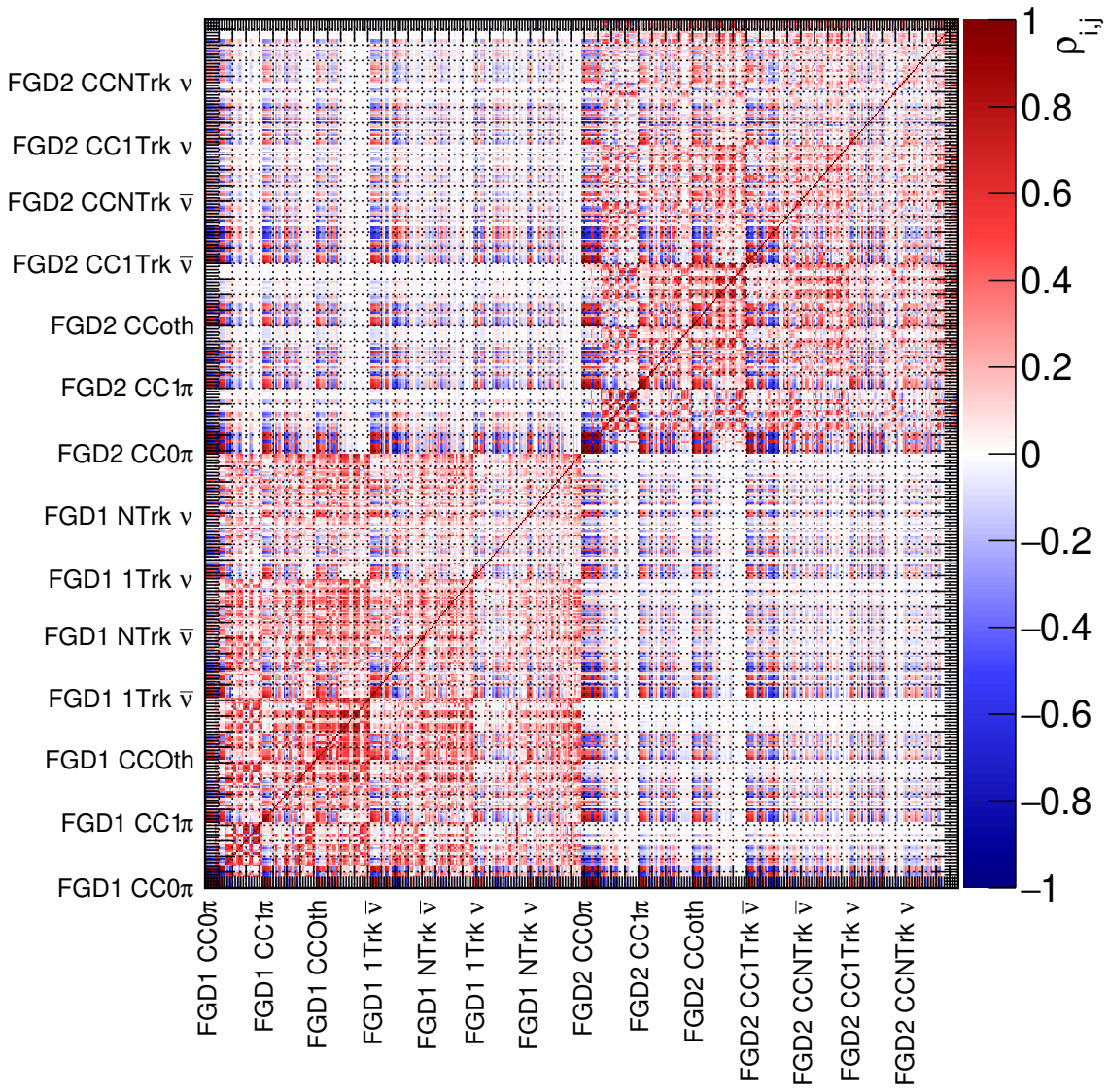


Figure 5.18.: Correlation matrix for the ND280 systematic parameters

When selecting the default CCQE model for T2K analyses, it was found [175] that the SF+2p2h model was inferior to the RFG+2p2h+RPA model when predicting external neutrino and anti-neutrino CCQE scattering data from MiniBooNE [179, 180] and MINERνA [181, 182]. Thus the simpler RFG model was chosen and a one-time weight is applied in p_μ, $\cos\theta_\mu$ to account for the phase space shift.

In the selected CCQE model we have three free parameters to vary: M_A^{QE}, the axial mass in the dipole form factor parameterisation in the Llewelyn-Smith model, and p_F, the Fermi surface momentum, for ^{12}C and ^{16}O coming from the Smith-Moniz model[c].

[c]The binding energy term was found to have a negligible effect in 2015 analyses so was removed

The 2p2h effects first have normalisation parameters separated for neutrino and anti-neutrino, and one for ^{12}C $\rightarrow^{16}$ O scaling. Secondly, there is a 2p2h shape parameter separated for ^{12}C and ^{16}O which is parameterised as a multiplicative weight applied on an event-by-event basis, taking an event's E_ν, q_0, q_3, where E_ν is the true neutrino energy, q_0 is the energy and q_3 is the momentum components of $Q = k_\nu - k_\mu = (q_0, q_3)$. The 2p2h model can be parameterised as having terms with and without pion exchange and the interference between these terms. At a value of -1 the shape uncertainty assigns all the 2p2h to Δ-like and at +1 to nonΔ-like 2p2h, and an interference term soaking up the lost or gained cross-section is included. The systematic has no net effect on the normalisation of 2p2h events.

The net effect of the 2p2h shape parameter on NEUT Monte-Carlo events generated with an ND280 flux on a ^{12}C target in q_0, q_3 is shown in Figure 5.19. The $\sim$300 MeV shift in q_0 from $m_\Delta - m_N$ is evident for the extreme parameter values, whereas the nominal ("Tweak Value = 0") populates both regions.

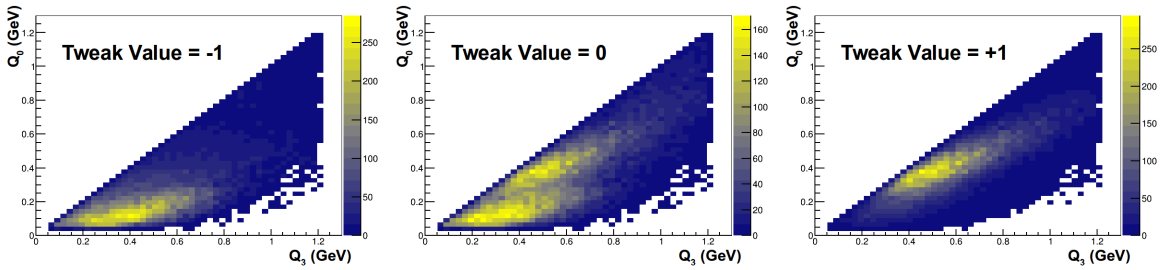
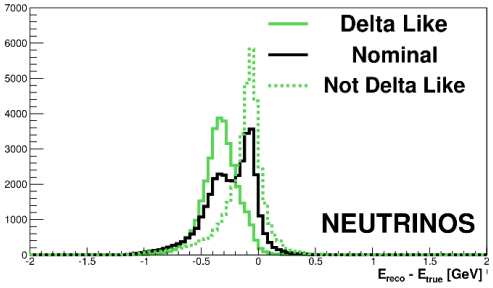


Figure 5.19.: $q_0, q3$ distributions for different values of the 2p2h shape parameter for ν_μ on a ^{12}C target with the ND280 flux

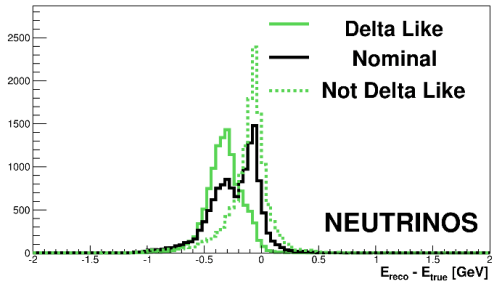
Figure 5.20 shows the $E_\nu^{reco} - E_\nu^{true}$ bias on the same generated events where the reconstructed neutrino energy is

$$E_\nu^{reco} = \frac{m_f^2 - m_i'^2 - m_l^2 + 2m_i' E_l}{2(m_i' - E_l + p_l \cos\theta_{\nu,l})} \tag{5.22}$$

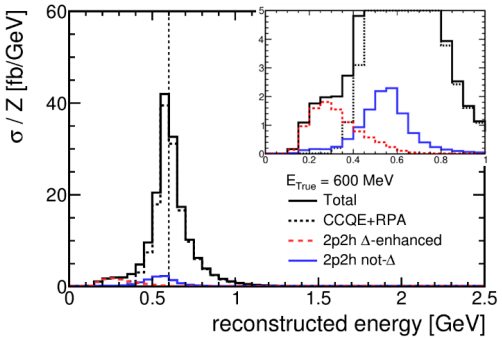
in which m_f is the final state nucleon mass, $m_i' = m_i - E_b$ where m_i is the initial state nucleon mass and E_b is the binding energy of a nucleon in the nucleus (27 MeV for ^{16}O, 25 MeV for ^{12}C), E_l (p_l) is the reconstructed lepton energy (momentum), and $\cos\theta_{\nu,l}$ is the cosine of the reconstructed angle between the incoming neutrino and outgoing lepton.



(a) ^{12}C target



(b) ^{16}O target



(c) E_ν^{reco} for $E_\nu^{true} = 0.6$ GeV events on CH at ND280 for 0π final states

Figure 5.20.: E_ν reconstruction bias for different values of the 2p2h shape parameter for ν_μ with the ND280 flux in NEUT 5.3.3

Figure 5.20c shows the importance of good 2p2h modelling for SK E_ν reconstruction: it biases 2p2h events towards a lower energy and so directly impact oscillation parameter fits.

The second group of nuclear uncertainties for CC0π comes from the Random Phase Approximation [48], which effectively describes correlations between nucleons in a nucleus. The net effect is to modify the 1-particle-1-hole [47] cross-section, which in NEUT is parameterised as a look-up table correction to the CCQE cross-section in E_ν and Q^2. The E_ν dependence was found to be relatively weak, so a Q^2 dependent correction was developed to mimic the uncertainties associated with the model [47]. The correction is parameterised as a third order rising polynomial which switches to a decaying exponential at $Q^2 = 1.2$ GeV^2. A normal polynomial of form $ax^3 + bx^2 + cx + d$ connecting to $\exp(-e(x-f))$ was found to strongly correlate the polynomial and exponential parameters, so a Bernstein polynomial base was chosen instead. The

parameterisation follows

$$w(Q^2) = \begin{cases} A(1-x')^3 + 3B(1-x')^2x' + 3p_1(1-x')x'^2 + Dx'^3, & x \leq U \\ 1 + p_2\exp(-E(x-U)), & x > U \end{cases} \quad (5.23)$$

in which $w(Q^2)$ is the weight from the RPA variation applied to CCQE events, $x = Q^2$, $x' = Q^2/U$, and A, B, D and E are normalisation factors for the four basis functions. Continuity between the functions at $Q^2 = U$ is required and p_1 and p_2 absorb this if

$$p_1 = D + U\frac{E(D-1)}{3}, p_2 = D - 1 \quad (5.24)$$

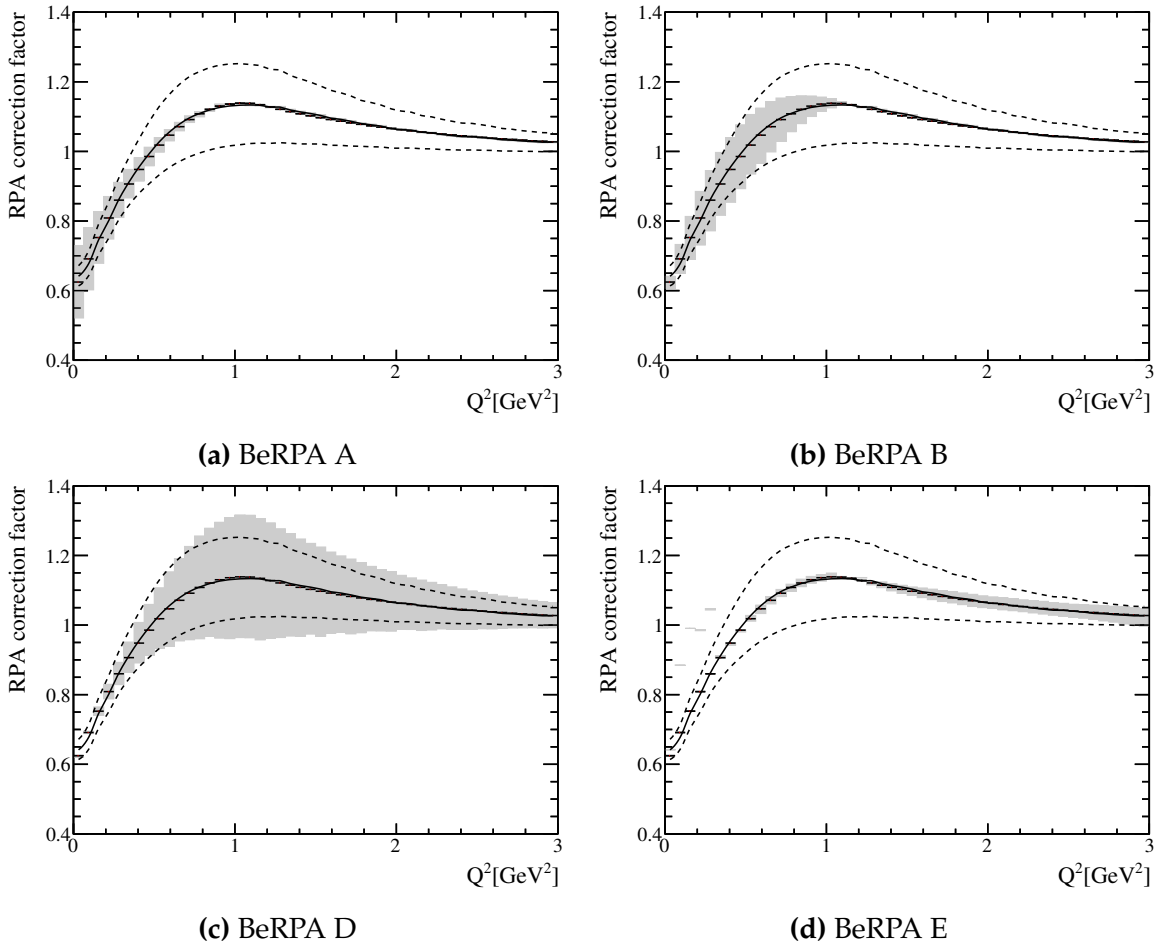
The magnitude and uncertainty of A, B, D and E are then chosen to match the uncertainties provided in [48]. Figure 5.21 shows the uncertainty bands from each parameter for the final values, showing how BeRPA A controls low Q^2, BeRPA B intermediate Q^2, BeRPA D medium Q^2 and BeRPA E high Q^2.

Single Pion Production The single pion production is described with the Rein-Sehgal model [183, 184] with lepton mass corrections [185–187] and modified form factors aimed at the Δ resonance [188–190]. The tunable parameters all relate to the neutrino-nucleon scattering interaction in the Rein-Sehgal model, which are M_A^{RES}, the axial mass for the resonant interaction in the Rein-Sehgal model, $C_5^A(0)$, one of the axial form factors at $Q^2 = 0$ in the Graczyk-Sobczyk form factor parameterisation, and the size of the non-resonant $I_{1/2}$ background in the Rein-Sehgal model.

The single pion production uncertainties were tuned to selected bubble chamber data from ANL [191, 192] and BNL [193–195]. The parameter values were cross-checked with suggested corrections to the ANL and BNL data [196], and compared to results when fitting nuclear data from MiniBooNE [197–199], MINERνA [200–202] and K2K [203] using NUISANCE [1].

Coherent Scattering The coherent scattering model is described by the Rein-Sehgal model [204]. An E_π dependent scaling factor, listed in Table 5.6, is applied to match MINERνA ν and $\bar{\nu}$ CC coherent data [205] and the Berger-Sehgal model [45].

The strength of the coherent interaction is allowed to vary in the fit as normalisation parameters: one for CC ^{12}C, one for CC ^{16}O and one for NC interactions. The prior



(a) BeRPA A **(b)** BeRPA B

(c) BeRPA D **(d)** BeRPA E

Figure 5.21.: BeRPA uncertainties for each separate BeRPA parameter. The dashed line represents the theoretical uncertainties [48], and the correction is applied to CCQE events as a function of Q^2

E_π	Weight
0.00-0.25	0.135
0.25-0.50	0.400
0.50-0.75	0.294
0.75-1.00	1.206

Table 5.6.: Rein-Sehgal coherent scaling in E_π applied as a one-time-weight to coherent events

uncertainty is 30% from inspections of MINERνA data and the two CC parameters are 100% correlated in the prior.

Multi-π and DIS In NEUT the structure functions are taken from the GRV98 parton distribution functions [206] with Bodek-Yang corrections [207]. The details of the

transition can be found in [53] and is inspired by external bubble chamber data on pion multiplicity. The DIS model is directly from PYTHIA 5.7 and JETSET 7.4 [208].

The multi-pi/DIS uncertainties come from measurements by MINOS [209] of DIS interactions on an iron target at higher energies. At $E_\nu = 4.0$ GeV the uncertainties should be $\sim 10\%$, increasing with decreasing E_ν and is parametersied as

$$\delta\left(\sigma_{\text{CC DIS}}\right) = \frac{0.4}{E_\nu} \tag{5.25}$$

NC DIS events do not receive this systematic and are instead controlled by an overall normalisation parameter including other NC interaction modes, explained later.

Subdominant Interactions The interaction modes with smaller cross-sections and/or small effects on 0π selections are controlled by normalisation parameters.

The NC1γ interaction—modelled with the Rein-Sehgal single-pion model with replaced branching fractions—is important because although the cross-section is very small ($\mathcal{O}(10^{-40})$ cm^2/nucleon [210]), the $\gamma \to e^+e^-$ process may mimic ν_e appearance signal at SK. The prior weight is set to 200% of nominal after comparing to a recent theory model [211], and the prior uncertainty is conservative at 100%.

The NC elastic, resonant kaon and eta production, and NC DIS events are all joined together under the "NC other" normalisation parameter. The prior is the nominal value with an uncertainty of 30%.

Electron (Anti-)Neutrinos In principle, there may be unmodelled effects present for ν_e ($\bar{\nu}_e$) and not ν_μ ($\bar{\nu}_\mu$), which could affect parameters extracted from ν_e appearance. Since there is dedicated ν_e selection at ND280[d] to constrain ν_e cross-sections or any high-statistics external data, an uncorrelated 2% uncertainty from radiative corrections and another 2% correlated uncertainty from second class currents are added [212]. These systematics are simply normalisations applied to ν_e and $\bar{\nu}_e$ events separately.

Final State Interactions The hadronic pion final state interactions are handled by a cascade implementation of the Salcedo-Oset model [213] when their momentum is less than 500 MeV. In summary, the interaction probabilities are functions of the

[d] With the exception of the upcoming work presented in [210], included in next year's analysis

hadron's momentum and position in the nucleus, with direction and momentum changes tuned to $\pi - N$ scattering data [214], including in-medium corrections [215]. More details are provided elsewhere [53]. Above 500 MeV, the interaction probabilities are calculated from $\pi^{\pm}$ scattering off free protons and deuteron compiled by the PDG [216]. To avoid double counting and discontinuities, the two models are blended between $400 < p_{\pi} < 500$ MeV.

The systematics are parameterised as the scattering probabilities for the different interaction processes at each microscopic step in the cascade. These are divided into absorption, pion production, quasi-elastic and charge exchange. Furthermore, the two latter are split into low and high energy regions, above and below the model transition at $p_{\pi} = 500$ MeV. For oscillation analyses using the 0π topology, the absorption and elastic probabilities are most important since they can significantly bias energy reconstruction by removing pions from the reconstructed event. The prior values and uncertainties come from tuning to world scattering data: π^{+} $-^{12}$C for the low energy and $\pi^{\pm}$ $-^{12}$C for the high energy parameters.

Importantly, the pion rescattering probabilities are assumed to be independent, which is technically incorrect: an extreme value of the pion absorption probability causes a pion to always be absorbed in FSI, so the effect of the remaining pion FSI parameters should be nothing since there is no longer a pion in the event. Such effects are unaccounted for as it was found to have a $\sim 1\%$ effect on T2K run 2-4 data.

The nucleon final state interactions are modelled using the Bertini model in GEANT4 [149] and are neglected since for water Cherenkov detector such as SK, the nucleon detection threshold is rarely reached. However, when nucleons interact to produce pions these pions are propagated through the simulation and are subject to the above systematics.

Parameterisation The neutrino interaction uncertainties are parameterised as shapes or normalisations and are applied on an event-by-event basis. In the case of normalisations, an event gets attached to normalisation parameters in accordance with their mode: e.g. a CC coherent event on ^{12}C gets the CC coh. C parameter weight: if the parameter is 1.5, the weight is 1.5.

For shape parameters, each event has associated one dimensional "splines"[e], which assumes that the interaction parameters have uncorrelated responses to variations.

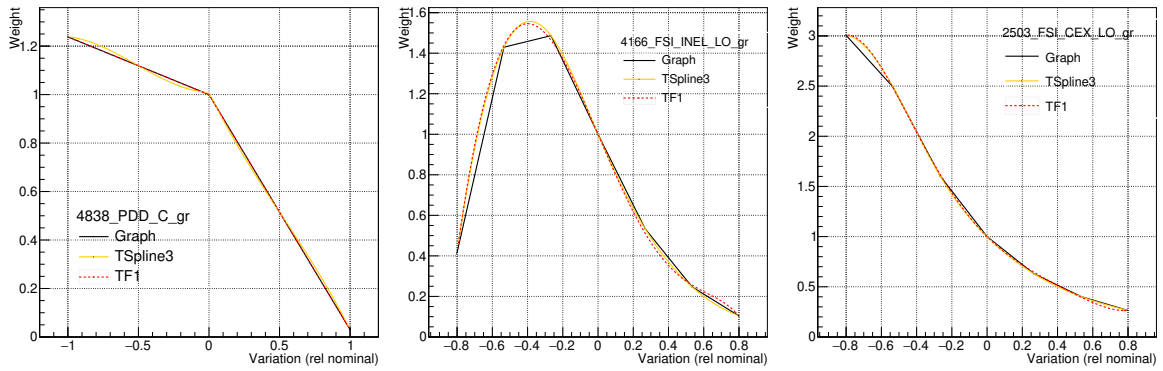
[e]The implementation uses a ROOT `TSpline3` class, reflected in the name

The weights are pre-calculated through a reweighting routine which calculates the weight of parameter variation $x \to x'$, where x is the nominal Monte Carlo parameter value and x' is the new parameter value, as

$$w(x \to x') = \frac{d^n\sigma(x')}{dX^n} / \frac{d^n\sigma(x)}{dX^n} \tag{5.26}$$

for an n dimensional cross-section calculation for X dependent parameters. Similar to the pion FSI parameters, this is not strictly true since for example a simultaneous variation in M_A^{RES} and C_5^A is not equivalent to a variation in M_A^{RES} followed by a variation of C_5^A. This is currently neglected in the analysis since the effect is sub-percent.

An example of three interaction parameters that are parameterised as splines is shown in Figure 5.22. The figure compares parameterising the shape as a third order polynomial `TF1` implementation to a `TSpline3` implementation, both fitted to discrete calculated parameter variations.



(a) 2p2h shape C **(b)** FSI Inel. Lo **(c)** FSI Cex. Lo

Figure 5.22.: `TF1` and `TSpline3` interpolation of three different shape parameters for three random events included in the analysis

Covariance Matrix Figure 5.24 shows the final $\sqrt{\mathbf{V}}$ covariance and correlation matrix. Only a few parameters are correlated in the prior covariance matrix: the 2p2h shape C and O parameters (+0.55), the single pion production parameters (-0.8 to 0.10), the ν_e ($\bar{\nu}_e$)/ν_μ ($\bar{\nu}_\mu$) (-0.71), the CC coherent parameters (+1.00), and the pion FSI parameters (-1.00 to 1.00).

The errors in Figure 5.23 do not include the flat priors on M_A^{QE} and p_F and that BeRPA U is fixed. Furthermore, the errors are expressed as relative their priors: e.g. the

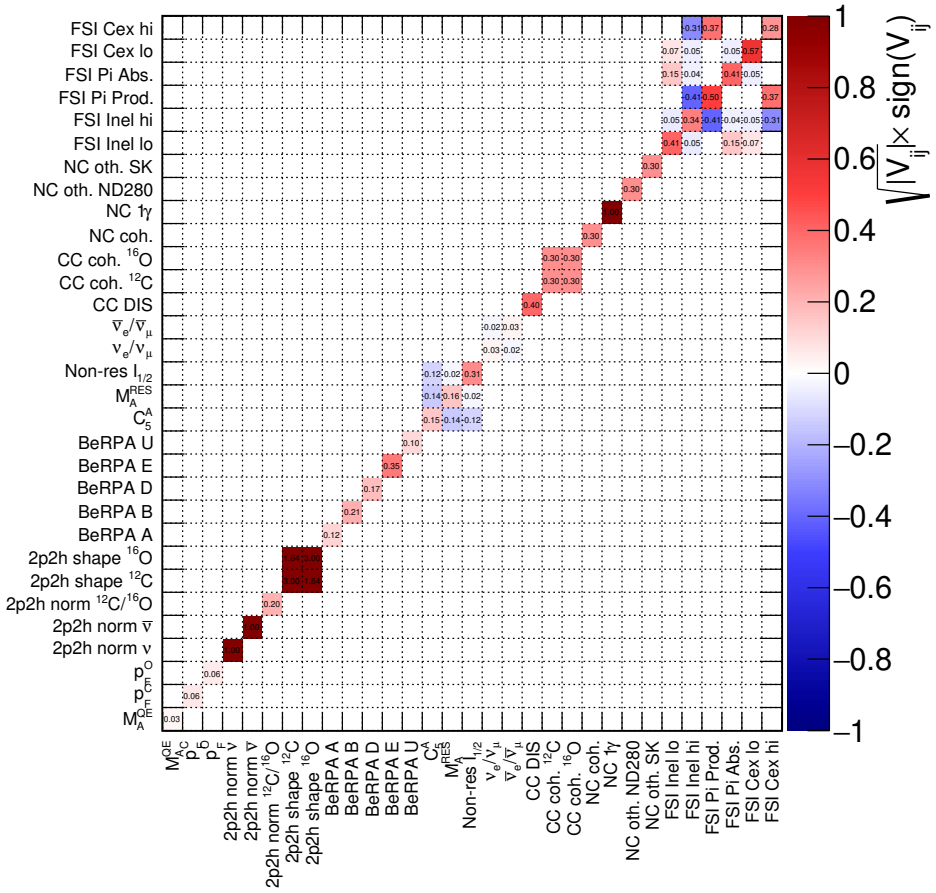


Figure 5.23.: Interaction covariance matrix provided as the prior, $\sqrt{\text{covariance}}$

error on M_A^{RES} is not 0.16 GeV but rather $0.95 \times 0.16 = 0.152$ GeV, where the nominal value is 0.16.

A summary of the interaction parameters applied in this analysis is provided in Table 5.7.

5.4. Building the Monte-Carlo Prediction

T2K has been taking data since 2010 with steadily increasing beam power and protons on target (POT), and is currently on "run 9", shown earlier in Figure 3.6. This analysis uses data from runs 2 to 6: run 1 was omitted because parts of the detector were uninstrumented (and is only $\sim$4% of the run 1-6 data), and run 7 and beyond had not gone through full Monte-Carlo production until summer 2017.

Parameter	Prior (Nominal)	Error	Prior shape	Type
CC0π				
M_A^{QE}	1.21 GeV	0-10	Flat	Shape
M_A^{QE} H	1.03 (1.21) GeV	0	—	Shape
p_F^C	217 MeV	31 MeV	Flat	Shape
p_F^O	225 MeV	31 MeV	Flat	Shape
2p2h norm ν	1.0	1.0	Flat	Norm.
2p2h norm $\bar{\nu}$	1.0	1.0	Flat	Norm.
2p2h norm C to O	1.0	0.2	Gauss	Norm.
2p2h shape C	1.0	3.0	Gauss	Shape
2p2h shape O	1.0	3.0	Gauss	Shape
BeRPA A	0.59	0.118	Gauss	Shape
BeRPA B	1.05	0.210	Gauss	Shape
BeRPA D	1.13	0.170	Gauss	Shape
BeRPA E	0.88	0.352	Gauss	Shape
BeRPA U	1.20	—	—	—
CC1π and DIS				
M_A^{RES}	1.07 (0.95) GeV	0.15 GeV	Gauss	Shape
C_5^A	0.96 (1.01)	0.15	Gauss	Shape
$I_{1/2}$ non-res	0.96 (1.30)	0.40	Gauss	Shape
CC coh ^{12}C	1.0	0.3	Gauss	Norm.
CC coh ^{16}O	1.0	0.3	Gauss	Norm.
CC DIS	1.0	0.4	Gauss	Shape
ν_e and $\bar{\nu}_e$				
CC ν_e/ν_μ	1.0	$\sqrt{2} \times 0.02$	Gauss	Norm.
CC $\bar{\nu}_e/\bar{\nu}_\mu$	1.0	$\sqrt{2} \times 0.02$	Gauss	Norm.
NC				
NC coh	1.0	0.3	Gauss	Norm.
NC 1γ	2.0 (1.0)	1.0	Gauss	Norm.
NC other	1.0	0.3	Gauss	Norm.
Pion FSI				
Pion QE	1.0	0.41	Gauss	Shape
Pion Abs	1.1 (1.0)	0.41	Gauss	Shape
Pion Cex Lo	1.0	0.57	Gauss	Shape
Pion QE Hi	1.8 (1.0)	0.34	Gauss	Shape
Pion Inel	1.0	0.50	Gauss	Shape
Pion Cex Hi	1.8 (1.0)	0.28	Gauss	Shape

Table 5.7.: Interaction parameters for T2K 2017 oscillation analyses

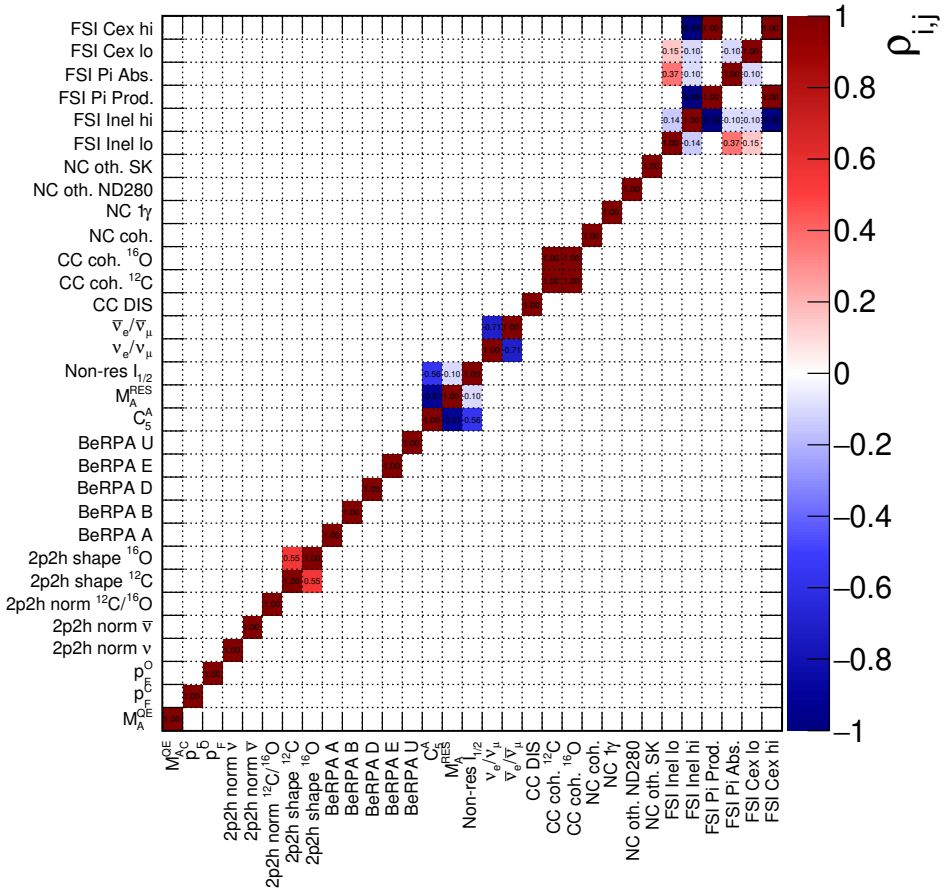


Figure 5.24.: Interaction covariance matrix provided as the prior, correlation

The overall efficiency of ND280 in runs 2 to 6 was approximately 85%, collecting 9.7E20 POT out of 12E20 POT. The good POT[f] per run used in this analysis is listed in Table 5.8.

The Monte-Carlo is generated for the different run-periods to account for beam configurations, ND280 configurations, run-dependent calibrations, and so on. Runs marked "a" and "w" refer to the P0D detector's removable water bags being air filled (a) or water filled (w), which require different ND280 geometries in the simulation.

To build the nominal distribution for direct comparison to data, a number of scalings and weights are applied. For the ND280 analysis all weights are applied on an event-by-event basis[g]:

[f]Defined as POT collected when all ND280 sub-detectors and global DAQ are flagged online
[g]Which is not the case for SK analyses

Run	Data POT (E19)	MC POT (E19)	Sand POT (E19)
2a FHC	3.59337	92.15	10
2w FHC	4.33934	120.15	10
3b FHC	2.17273	44.8	5
3c FHC	13.6447	263	25
4a FHC	17.8271	349.9	30
4w FHC	16.4277	349.65	29.15
5 RHC	4.3468	208.25	20
6b RHC	12.8838	141.03	40
6c RHC	5.07819	53.21	15
6d RHC	7.75302	69.41	20
6e RHC	8.51668	86.72	23
Total FHC	58.00494	1219.65	109.15
Total RHC	38.57849	558.62	118
Total	96.58343	1778.27	227.15

Table 5.8.: Counted and generated proton-on-targets for the T2K ND280 2017 analysis

- **POT weight:**
  A run-by-run scaling factor taking the ratio of total good flagged data to the generated Monte-Carlo POT. An event receives a one-time weight depending on what run it was from and how much MC was generated in the production. These numbers can be read off directly from Table 5.8. This weight w_{POT} is only applied once and is not varied in the fit.

- **Flux weight:**
  A run-by-run correction to the nominal neutrino flux which the full experiment simulation was made in. The weight is applied as a function of E_ν^{true} with $0 < E_\nu^{true} < 30$ GeV , detailed in subsection 5.3.1. An event receives a weight depending on what run it was from and its E_ν^{true}. The weight w_{Flux} is only applied once and is not varied in the fit.

- **Beam variation weight:**
  An event-by-event weight to vary the impact of the flux simulation. An event gets weighted as a function of the neutrino run, the flavour of neutrino (ν_μ, $\bar{\nu}_\mu$, ν_e or $\bar{\nu}_e$) and the neutrino energy, detailed in subsection 5.3.1.

The weight $w_{\vec{b}\to\vec{b}'}$ is applied once to weight the simulation to nominal, and is recalculated for every iteration of the fit.

- **Cross-section variation weight:**
  An event-by-event weight taking the full generated event through the neutrino interaction simulation, calculating a weight to apply to the event. The applied weight is pre-calculated as $w_{\vec{x}\to\vec{x}'} = \sigma(\vec{x}')/\sigma(\vec{x})$ for the generated neutrino interaction systematic parameter values $\vec{x}$ and the modified parameter values $\vec{x}'$. The event can receive a normalisation and a shape parameter depending on its interaction type and the interaction model considered in the analysis, detailed in subsection 5.3.3.

  The weight $w_{\vec{x}\to\vec{x}'}$ is applied once to weight the simulation to nominal, and is recalculated for every iteration of the fit.

- **Detector variation weight:**
  An event-by-event weight from the reconstruction and selection to vary the impact of the detector simulation. It is applied as a function of the event's topology and detector (e.g. FGD1 CC0π), its p_μ and $\cos\theta_\mu$. The weights are normalisation parameters $\vec{d}$ for each bin in the detector covariance matrix, explained in detail in subsection 5.3.2.

  The weight $w_{\vec{d}\to\vec{d}'}$ is applied once to weight the simulation to nominal, and is recalculated for every iteration of the fit.

All weights are parameterised as multiplicative, so for a two parameter variation $x \to x'$ and $y \to y'$ we have the total weight

$$w_{x\to x',y\to y'} = w_{x\to x',y\to y} \times w_{x\to x,y\to y'} \tag{5.27}$$

Defining the beam parameters as $\vec{b}$, cross-section parameters as $\vec{x}$ and detector parameters as $\vec{d}$, we express one rescaled Monte-Carlo event λ_i as

$$\lambda_i\left(\vec{b},\vec{x},\vec{d}\right) = 1 \times w_i^{\mathrm{POT}} \times w_i^{\mathrm{Flux}} \times w_i^{\vec{b}\to\vec{b}'} \times w_i^{\vec{x}\to\vec{x}'} \times w_i^{\vec{d}\to\vec{d}'} \tag{5.28}$$

The effect of each weight on the predictions for each ND280 selection is shown in Table 5.9. The nominal flux and detector weights are 1.0, so are not included in the table.

5.5. Nominal Model Prediction

The neutrino event distributions for the data and nominal model with all mentioned Monte-Carlo scalings and selections are presented in Table 5.10. Generally the Monte-Carlo rates of the CC0π and CC1Track are underestimated (2-3%) , CC1π is overestimated (6-11%) and CCOther is under-estimated (5-15%). $\bar{\nu}_\mu$ selections except FGD2 CC1Track $\bar{\nu}_\mu$ are overestimated by 5%, with FGD1 being 2%. The ν_μ in RHC selections are mildly underestimated. The rates across the FGDs are consistent for all selections.

Figure 5.25 shows the projections of the 2D distributions onto p_μ, where we note generally good modelling. The 2D distributions can be found in Appendix C.

The CC0π selections have a clear transition from under-prediction to overprediction for FGD1 and FGD2 in $0 < p_\mu < 1$ GeV. The FGD1 $\bar{\nu}_\mu$ CC1Track selection appears to show similar behaviour although without the under-prediction at low p_μ, although FGD2 ν_μ CC1Track does not. The CC1π selection shows a nearly consistent over-estimation of $\sim$10% in every bin for both FGDs. The CCOther selection is instead underestimated at the event distribution peak, into the tail ($0.5 < p_\mu < 1.5$ GeV). The CCNtrack distributions are consistent with good modelling due to their large statistical errors.

Figure 5.26 shows the projections of the 2D distributions onto $\cos\theta_\mu$. Again we see consistency across the FGDs, with CC0π showing another oscillatory behaviour, going from underestimation at low $\cos\theta_\mu$ to a good prediction until $\cos\theta_\mu \sim 0.93$, in which the underestimation is back, similar in magnitude. For CC1π we see a similar oscillation but shifted by 10% over-estimation. For ν_μ CCOther we have less consistency, although the bins above $\cos\theta_\mu = 0.93$ are all underestimated, and for FGD2 this continues as $\cos\theta_\mu$ decreases. The most forward bin appears to be well modelled for all ν_μ samples except FGD2 CCOther. For the RHC 1 track selections, we note FGD2 looking similar to the CC0π selection, where FGD1 less so. The NTracks selections look similar for FGD1 and FGD2 with overestimates at high $\cos\theta_\mu$. For RHC ν_μ selections, the NTrack selections appear more consistent than 1 track, with underestimates in the

Sample	Raw MC	POT only	POT+xsec	POT+NDCov	POT+BeamCov
FGD1 CC0π ν_μ	337436	15905.2	15340.2	16246.1	16090.8
FGD1 CC1π ν_μ	84982	4011.58	3819.3	4131.35	4058.36
FGD1 CCOther ν_μ	65286	3071.21	3078.52	3374.21	3107.04
FGD2 CC0π ν_μ	345467	16259.6	15749.8	16415	16449.2
FGD2 CC1π ν_μ	70444	3318.26	3190.09	3321.4	3356.97
FGD2 CCOther ν_μ	63402	2983.43	2995.68	3051.48	3018.22
FGD1 CC1Track $\bar{\nu}_\mu$	54419	3744.02	3430.15	3872.18	3773.79
FGD1 CCNTrack $\bar{\nu}_\mu$	15392	1056.8	986.32	1121.1	1065.2
FGD2 CC1Track $\bar{\nu}_\mu$	55732	3833.73	3506.36	3906.32	3864.24
FGD2 CCNTrack $\bar{\nu}_\mu$	15808	1088.15	1024.77	1122.73	1096.81
FGD1 CC1Track ν_μ in RHC	18146	1246.03	1194.72	1262.00	1255.92
FGD1 CCNTrack ν_μ in RHC	17156	1181.54	1172.13	1257.34	1190.94
FGD2 CC1Track ν_μ in RHC	18052	1231.24	1189.64	1245.3	1241.02
FGD2 CCNTrack ν_μ in RHC	16339	1127.64	1121.37	1150.7	1136.62

Table 5.9.: Neutrino event counts broken down by type of weight applied for the nominal MC samples at ND280

Sample	Data	Nominal MC	Data/MC
FGD1 CC0π ν_μ	17136	16723.80	1.02
FGD1 CC1π ν_μ	3954	4381.47	0.90
FGD1 CCOther ν_μ	4149	3943.95	1.05
FGD2 CC0π ν_μ	17443	16959.30	1.03
FGD2 CC1π ν_μ	3366	3564.23	0.94
FGD2 CCOther ν_μ	4075	3570.94	1.14
FGD1 CC1Track $\bar{\nu}_\mu$	3527	3587.77	0.98
FGD1 CCNTrack $\bar{\nu}_\mu$	1054	1066.91	0.99
FGD2 CC1Track $\bar{\nu}_\mu$	3732	3618.29	1.03
FGD2 CCNTrack $\bar{\nu}_\mu$	1026	1077.24	0.95
FGD1 CC1Track ν_μ in RHC	1363	1272.17	1.07
FGD1 CCNTrack ν_μ in RHC	1370	1357.45	1.01
FGD2 CC1Track ν_μ in RHC	1320	1262.63	1.05
FGD2 CCNTrack ν_μ in RHC	1253	1246.71	1.01
Total	64768	63632.86	0.98

Table 5.10.: Observed and predicted event rates for the different ND280 selections for the 2017 analysis

highest $\cos\theta_\mu$ bin. The 1 track appears consistently underestimated in $0.9 < \cos\theta_\mu < 1$.

The p_μ and $\cos\theta_\mu$ mode distributions in Figure 5.25 and Figure 5.26 after applying the weights scaling show no noticeable differences to the raw MC distributions. Looking at the mode populations in Table 5.11, the selections contain $> 50\%$ of the interaction modes that they were designed to target. We see a 20% contribution of nucleon interaction level CC1π events in CC0π, which come from pion final state interactions, detector thresholds, and misreconstruction.

5.6. Fitting Asimov Data

To internally validate the implementation and evaluate the effectiveness of the fitting framework, we perform studies in which the nominal Monte-Carlo predictions presented in section 5.5 are set to be the data.

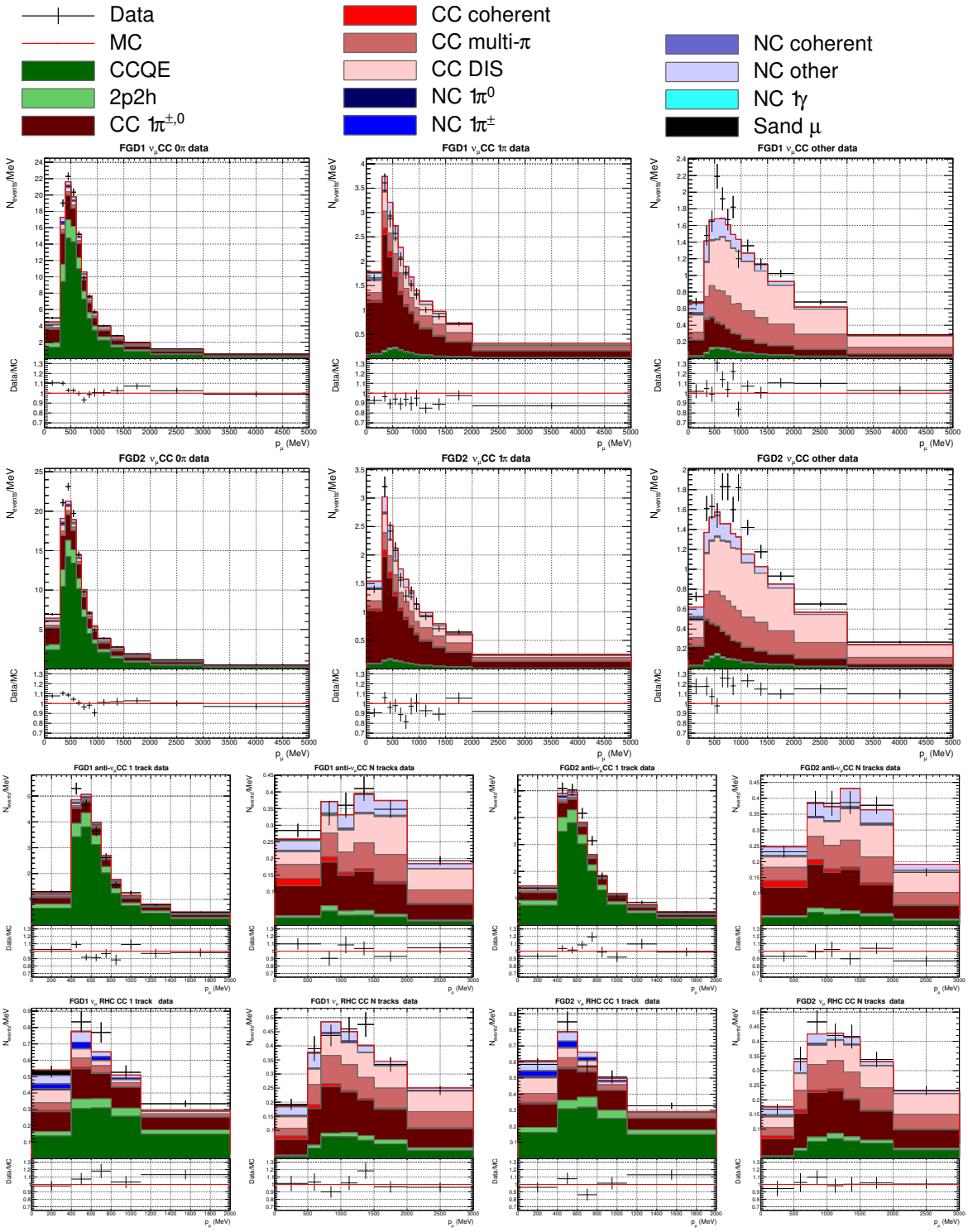


Figure 5.25.: Data and nominal MC distributions selections projected onto p_μ, showing contributions by interaction mode. Bin content is normalised to bin width

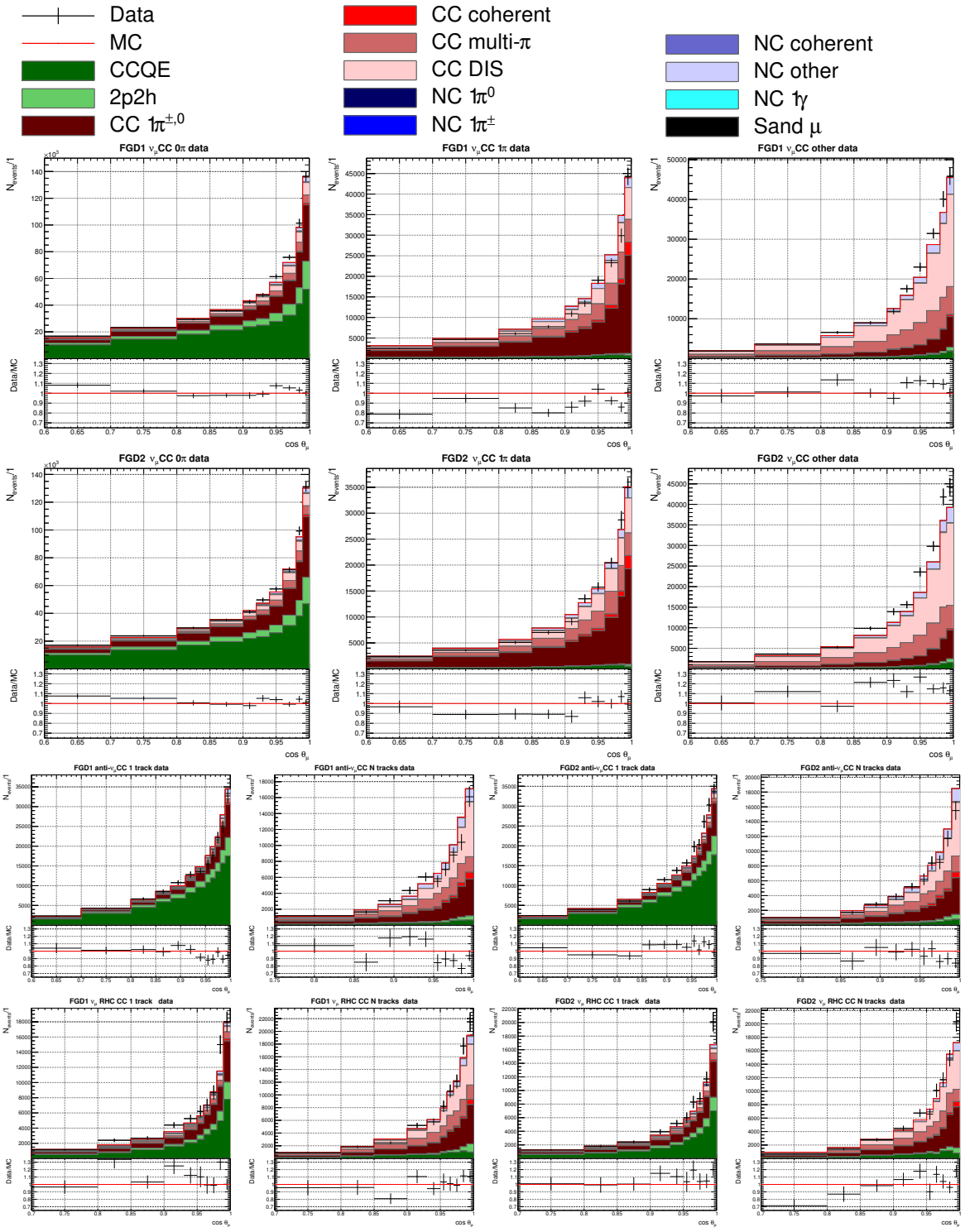


Figure 5.26.: Data and nominal MC distributions selections projected onto $\cos\theta_\mu$, showing contributions by interaction mode. Bin content is normalised to bin width

5.6.1. Log-Likelihood Scans

For the log-likelihood scans each parameter is set to the values recommended by the priors: the same parameter set which produced the nominal model in section 5.5. The

Sample	CCQE	2p2h	CC1$\pi^{\pm,0}$	CC coh	CC multi-π	CC DIS	NC
FGD1 0π	**56.7**	**10.0**	19.8	0.3	4.5	5.1	3.6
FGD2 0π	**55.2**	**9.4**	21.5	0.3	4.9	5.2	3.4
FGD1 1π	5.5	0.8	**49.2**	**2.7**	**18.1**	17.2	6.6
FGD2 1π	5.6	0.7	**48.5**	**2.8**	**18.3**	17.6	6.4
FGD1 Other	4.7	1.0	14.9	0.4	**26.1**	**45.0**	8.1
FGD2 Other	4.9	1.0	15.5	0.3	**25.5**	**44.9**	7.9
FGD1 1Trk	**64.0**	**10.0**	14.7	0.7	2.8	2.5	5.1
FGD2 1Trk	**64.4**	**9.9**	14.7	0.7	2.8	2.6	4.9
FGD1 NTrk	7.6	2.6	**28.4**	**3.3**	**19.8**	**26.9**	11.3
FGD2 NTrk	8.3	2.7	**28.3**	**3.2**	**19.8**	**26.2**	11.6
FGD1 1Trk ν_μ	**43.5**	**8.2**	25.2	0.9	7.2	6.6	8.5
FGD2 1Trk ν_μ	**43.3**	**8.2**	25.2	0.8	7.8	7.1	7.7
FGD1 NTrk ν_μ	12.1	3.1	**28.5**	**1.8**	**20.8**	**26.6**	7.0
FGD2 1Trk ν_μ	11.9	2.7	**29.8**	**1.8**	**21.1**	**26.4**	6.4

Table 5.11.: Percentage mode breakdown for the binned nominal scaled Monte-Carlo samples, boldface indicates interactions targeted by specific selections. The distributions are not bin-width normalised

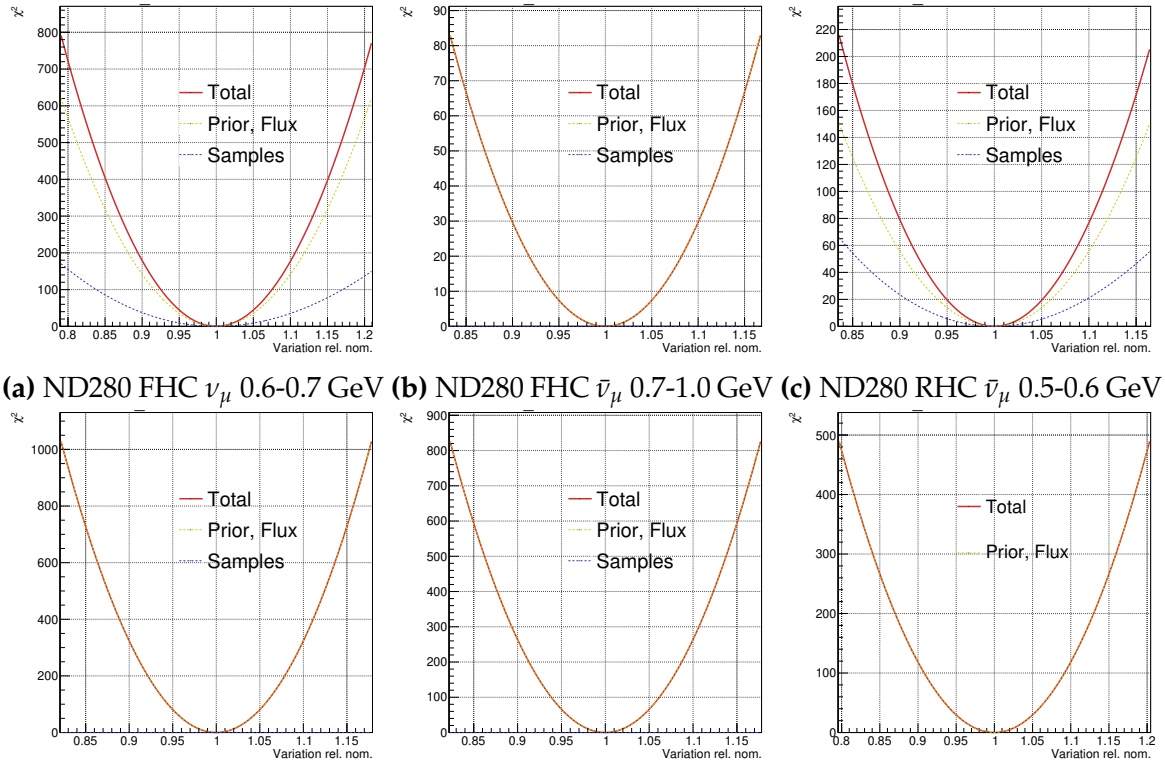
estimated constraints from the one dimensional log-likelihood are not accurate due to the large correlations in the beam and ND280 parameters in the prior covariance matrices; hence $\chi^2 \sim 1$ does not indicate the typical 1σ sensitivity. For sensitivity estimates, it is more appropriate to vary the correlated parameters simultaneously, as is done in the Asimov fit presented later in subsection 5.6.4. The likelihood scans are considered a closure test rather than a sensitivity test.

The parameters are varied one at a time from -2σ to +2σ, where $\sigma = \sqrt{\mathbf{V}_{i,i}}$, where $\mathbf{V}_{i,i}$ is the i^{th} diagonal entry of each group of parameters' covariance matrix[h]. The minimum of the test-statistic occurs when the parameter is equal to the prior. When a parameter has been scanned it is reset to the prior value. Most parameters are expected to have a Gaussian response since the prior probability density function is Gaussian, although a few parameters produce asymmetric responses in the event rate of a given bin.

The scan splits the likelihood into each of the individual contributions presented in Equation 4.3 and shows the total likelihood. For any given likelihood scan, there

[h]So the bounds are unaffected by the correlation when getting the lower and upper bounds of the scan

should only be contributions from the likelihood terms that are being varied: the sample likelihood and the group of systematics to which the parameter belongs.



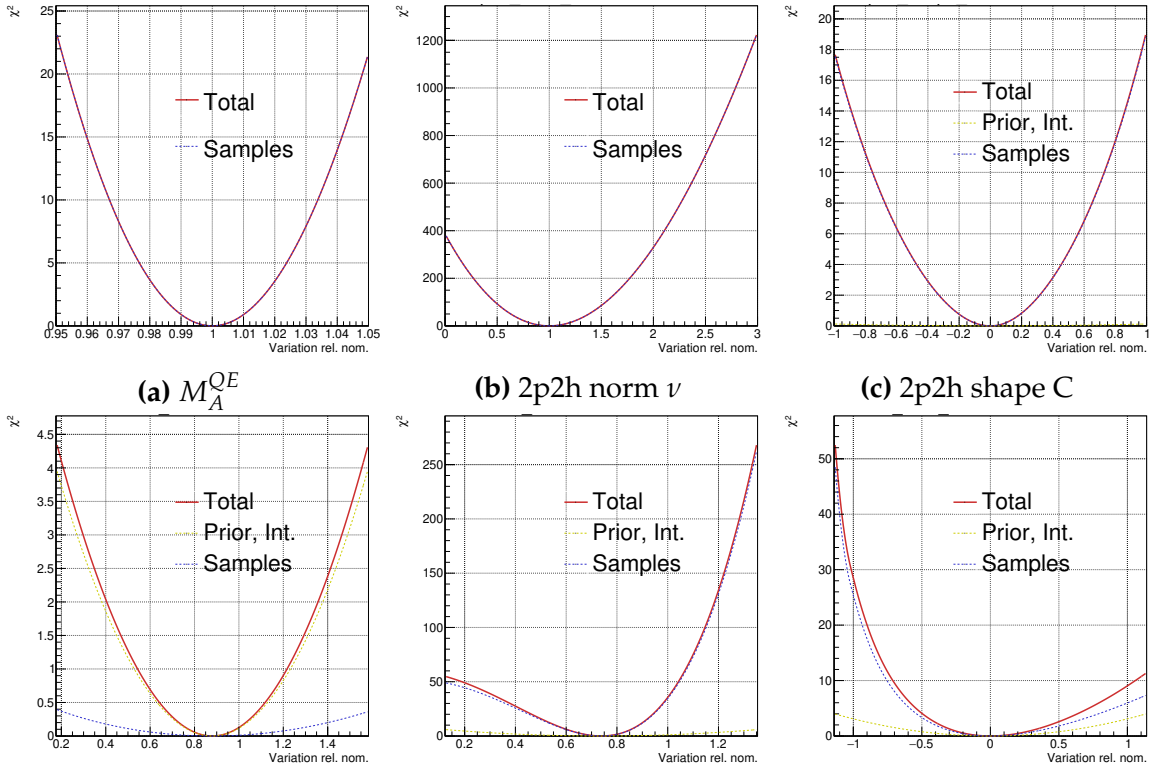
(a) ND280 FHC ν_μ 0.6-0.7 GeV **(b)** ND280 FHC $\bar{\nu}_\mu$ 0.7-1.0 GeV **(c)** ND280 RHC $\bar{\nu}_\mu$ 0.5-0.6 GeV

(d) ND280 FHC ν_e 0.5-0.7 GeV **(e)** ND280 RHC $\bar{\nu}_e$ 0.5-0.7 GeV **(f)** SK FHC ν_μ 0.6-0.7 GeV

Figure 5.27.: Asimov likelihood scans for selected beam parameters

Figure 5.27 shows a selected number of beam parameters. The prior term is dominant, even for the high-statistics ND280 FHC ν_μ $E_\nu = 0.6 - 0.7$ GeV parameter. Many parameters barely see any constraint from the samples—e.g. $\bar{\nu}_\mu$ in FHC running and ν_e in FHC. As expected, the SK flux parameters are only being constrained by the prior and see no contribution from any ND280 selection.

Figure 5.28 shows the likelihood scans for selected cross-section parameters. M_A^{QE} and 2p2h norm ν are both fit without a prior, so only see constraints from the sample likelihood, and 2p2h shape C has an almost flat prior likelihood, mostly constrained by the samples. Some cross-section parameters, like BeRPA E, have a weaker constraint from the MC samples than from the prior, due to the parameter being limited to high Q^2, of which ND280 has few events. We observe some non-Gaussian responses, such as the pion final-state-interaction charge exchange at low pion momentum (FSI CEX LO) and the single pion production non-resonant background parameter, $I_{1/2}^{\text{bkg}}$.



(a) M_A^{QE} **(b)** 2p2h norm ν **(c)** 2p2h shape C

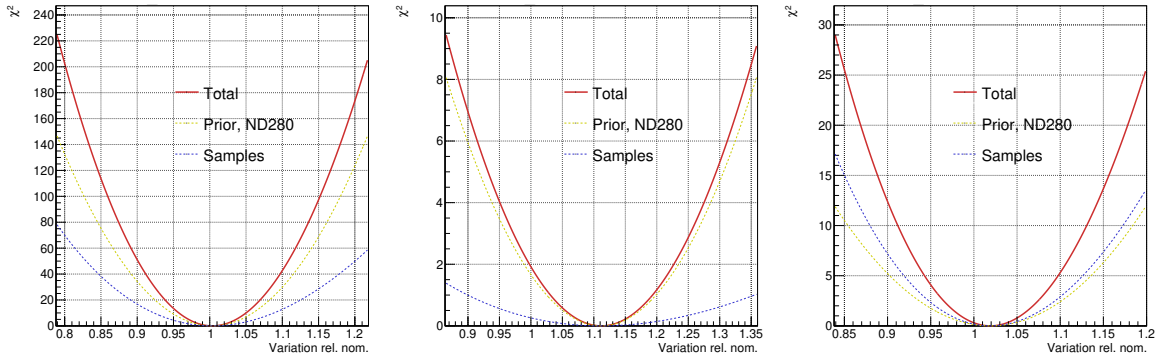
(d) BeRPA E **(e)** $I_{1/2}^{\text{bkg}}$ **(f)** FSI CEX LO

Figure 5.28.: Asimov likelihood scans for selected cross-section parameters

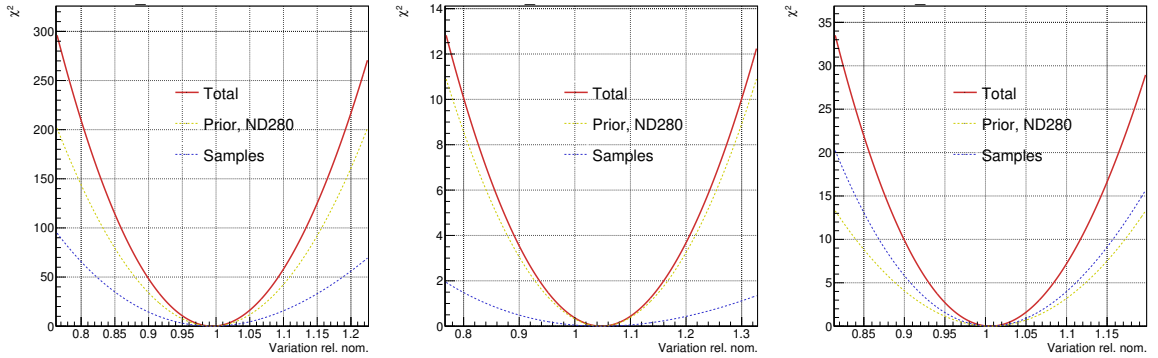
Figure 5.29 shows a selected number of the ND280 parameters for FGD1 and FGD2. The ND280 parameters are more balanced between prior and sample likelihoods. The effect is by design, since the underlying MC events that are being varied when making the detector covariance are the same as those being selected in the fit: the only difference is the analysis binning and the binning used to make the ND280 covariance matrix, covered in subsection 5.3.2.

As expected, the ND280 parameters covering high statistics samples and regions of phase space—such as CC0π, $0 < p_\mu < 1.0$ GeV, $0.6 < \cos\theta_\mu < 0.7$—have higher constraints than low ones—such as CCOther $1.5 < p_\mu < 2.0$ GeV, $0.8 < \cos\theta_\mu < 0.85$.

Comparing top and bottom panels, the responses for equivalent FGD1 and FGD2 parameters are compatible and have similar strength.



(a) FGD1 CC0π {0-1 GeV, 0.6-0.7} **(b)** FGD1 CCOther {1.5-2 GeV, 0.8-0.85} **(c)** FGD1 $\bar{\nu}_\mu$ CC1Trk {0.4-0.9 GeV, 0.6-0.7}



(d) FGD2 CC0π {0-1 GeV, 0.6-0.7} **(e)** FGD2 CCOther {1.5-2 GeV, 0.8-0.85} **(f)** FGD2 $\bar{\nu}_\mu$ CC1Trk {0.4-0.9 GeV, 0.6-0.7}

Figure 5.29.: Asimov likelihood scans for selected ND280 parameters

5.6.2. Parameter Variations

To inspect the effects of the parameterisation we vary the parameters one at a time over one unit of σ, where $\sigma = \sqrt{\mathbf{V}_{i,i}}$ as before, but now looking at the impact on the event distributions for each selection.

The largest effects of the variations on the event rates in each sample is shown in Table 5.12. For the CC0π and 1 track selections the 2p2h normalisation parameters have the largest effect. This is expected because the prior uncertainties are conservative at $\pm 100\%$, so the lower event rate (15139.1 for FGD1 0π) is the event rate when removing any 2p2h ν events. For CC1π selections, the largest effect is the C_5^A parameter, which controls the single pion production model. For the CCOther selections the CC DIS parameter has the largest effect. For the RHC NTrack selections the M_A^{RES} parameter is instead dominant, controlling the single pion production model of which the NTrack selections is dominated by (e.g. FGD1 NTrack 28.4% CC1$\pi^{\pm,0}$ vs 26.9% CCDIS in Table 5.11). The only selection which does not have an interaction parameter as its

largest uncertainity on event rate is FGD1 NTrk ν_μ, where the 29th beam parameter (RHC $\bar{\nu}_\mu$, $E_\nu = 0.7 - 1.0$ GeV)— the second largest effect is from M_A^{RES}, which predicts event rates from +1σ:1291.84 to -1σ: 1432.51. The results are expected and compatible with the likelihood scans in subsection 5.6.1.

Sample	Parameter	+1σ	Nominal	-1σ
FGD1 0π	2p2h norm ν	15139.1	16723.8	18308.6
FGD2 0π	2p2h norm ν	15420.4	16959.3	18498.2
FGD1 1π	C_5^A	4056.67	4381.47	4746.83
FGD2 1π	C_5^A	3307.92	3564.23	3852.44
FGD1 Other	CC DIS	3691.18	3943.95	4196.72
FGD2 Other	CC DIS	3343.41	3570.94	3798.47
FGD1 1Trk	2p2h norm $\bar{\nu}$	3245.72	3587.77	3929.83
FGD2 1Trk	2p2h norm $\bar{\nu}$	3272.86	3618.29	3963.73
FGD1 NTrk	M_A^{RES}	1019.96	1066.91	1126.26
FGD2 NTrk	M_A^{RES}	1028.7	1077.24	1138.62
FGD1 1Trk ν_μ	2p2h norm ν	1178.79	1272.17	1365.55
FGD2 1Trk ν_μ	2p2h norm ν	1170.3	1262.63	1354.97
FGD1 NTrk ν_μ	b29	1282.3	1357.45	1432.61
FGD2 NTrk ν_μ	M_A^{RES}	1184.42	1246.71	1317.25

Table 5.12.: Neutrino event counts for each selection, showing the largest effect of the 1-σ variations

The impact on the $p_\mu \cos\theta_\mu$ distributions for each sample and their respective "highest impact parameters" presented in Table 5.12 are shown in Figure 5.30 and Figure 5.31. We note similar responses in both FGDs and across samples.

5.6.3. Prior Predictive Spectrum

The first statistical test on the Asimov is to investigate the prior predictive spectrum and resulting p-values. This is meant to reflect the compatibility of the prior with the Asimov data. The p-value is expected to be $\sim$50% since the parameter throws are either side of the central value which created the Asimov data set. However, the throws are correlated, so excursions beyond 50% are to be expected. The two p-values are shown in Figure 5.32 and lie around the expected value of 50%.

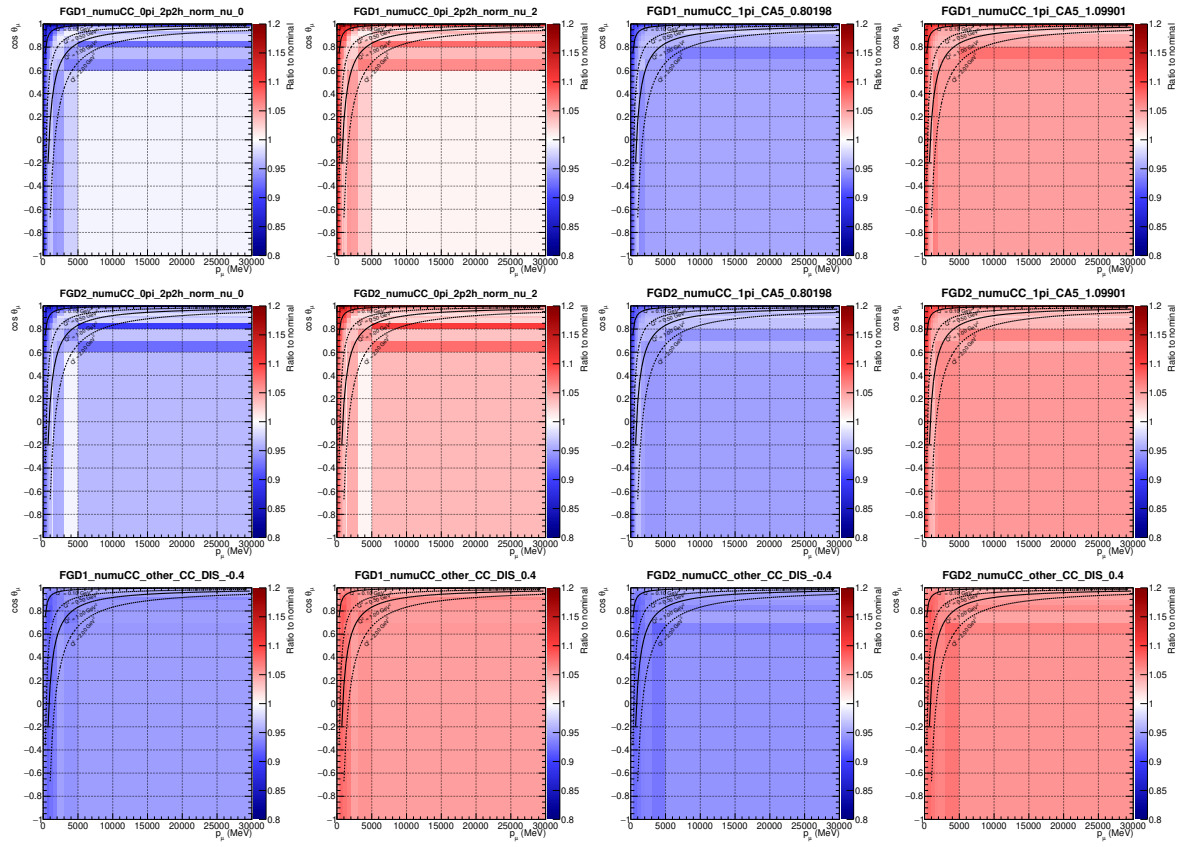


Figure 5.30.: The largest effect of the 1-σ variations on FHC selections' $p_\mu \cos\theta_\mu$

Table 5.13 shows the prior predictive event rates from the prior predictive spectrum to the Asmiov data with uncertainties from all systematics. Using the prior covariances without any fitting to the Asimov data produces 11% uncertainty on the total event rate, with 13% on the CC0π, 11% on CC1π and 13% on CCOther. There is a particularly bad likelihood for the 0π and 1Trk selections for the Asimov prior predictive spectrum. This is likely due to missing correlations between cross-section, flux and near-detector parameters, causing the correlated throws of the systematics to be skewed.

5.6.4. Fitting to Asimov Data

The model is fit to the Asimov data set, defined as the model when set to the central values of the prior constraints. Three different MCMCs are presented, each being a tuned version of the previous in regards to the autocorrelations and acceptance probability of each chain. Table 5.14 shows the chains for the Asimov study. In each case the burn-in is 1/4 of the total chain length.

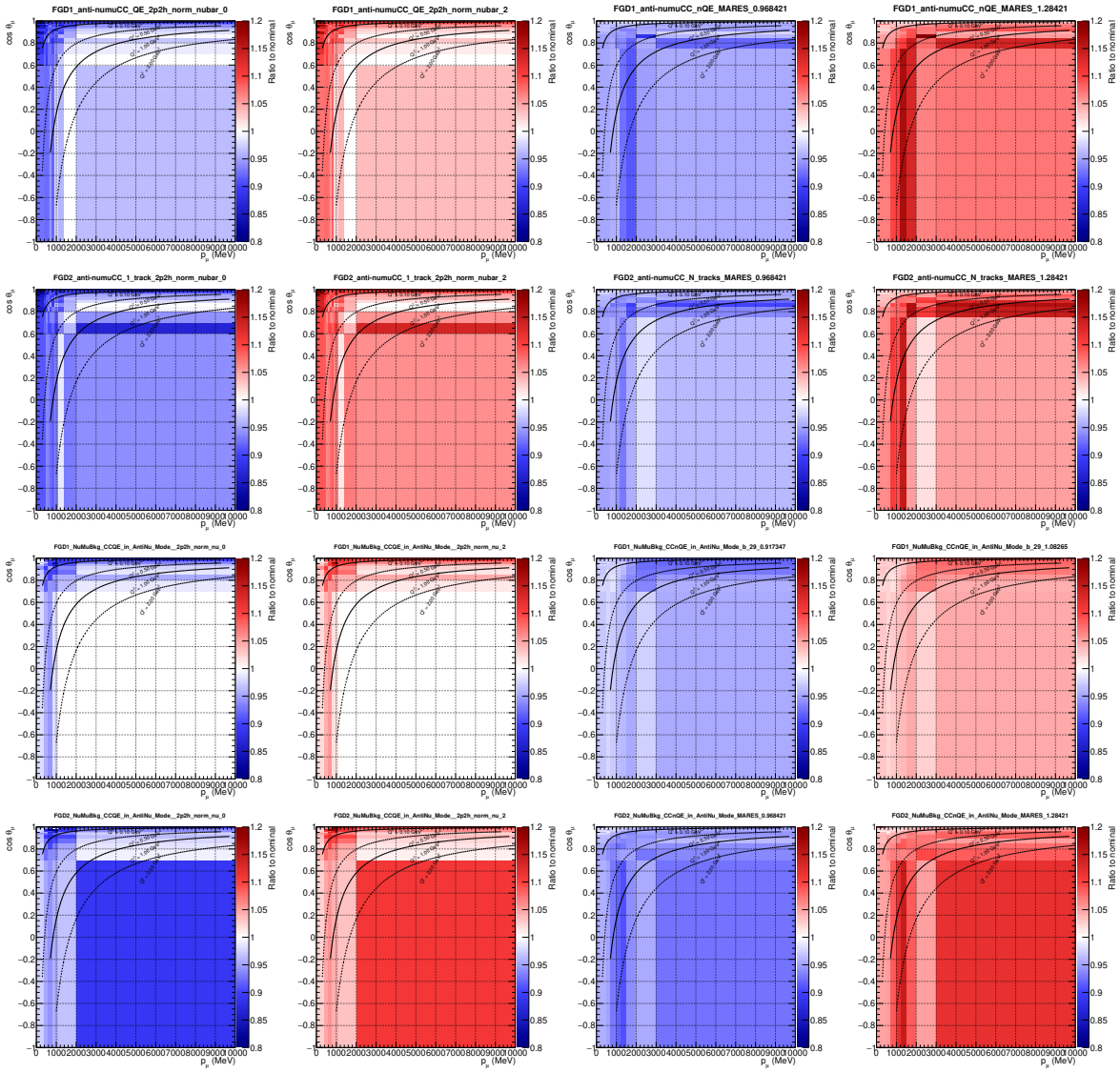


Figure 5.31.: The largest effect of the 1-σ variations on RHC selections' $p_\mu \cos\theta_\mu$

As covered in chapter 4, MCMC methods aren't designed to find a global minimum of the test-statistic, so we revert to defining the "best-fit" parameter set as the marginal posterior over all parameters except the "parameter of interest". Since we're marginalising over ~ 700 parameters, the study is likely to see parameter "biases" from marginalisation effects.

Figure 5.33 demonstrates that the lowest test-statistic is obtained right at the very beginning of the chain, which progresses to move around the minimum, scanning parameter combinations with $\chi^2 \sim 300 - 400$ (here defined as $\chi^2 = -\log \mathcal{L}$), so $-2 \log \mathcal{L}$ roughly coincides with the number of parameters in the fit. The lowest test-statistic after the burn-in is found at step 423471/800000, which has $\chi^2 = 284.49$.

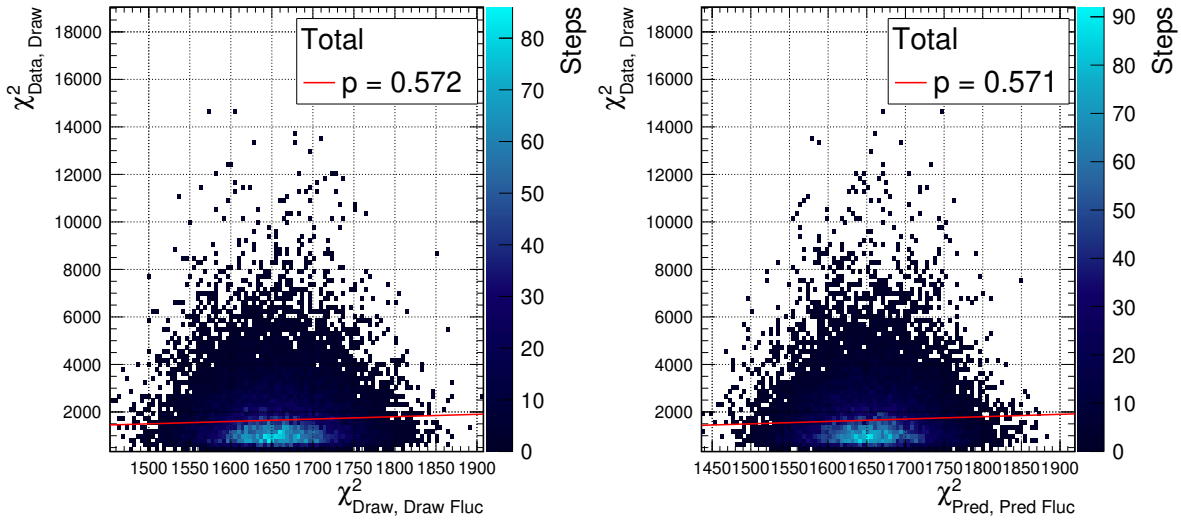


Figure 5.32.: Prior predictive p-values for the Asimov data

Sample	Nominal	Prior Pred	-2log $\mathcal{L}_s$
FGD1 0π	16723.8	17014.5 ± 2244.0	14.58
FGD1 1π	4381.47	4438.3 ± 487.6	0.99
FGD1 Other	3943.95	3984.1 ± 537.5	0.76
FGD2 0π	16959.3	17252.4 ± 2311.7	18.9
FGD2 1π	3564.23	3612.9 ± 398.4	0.92
FGD2 Other	3570.94	3581.5 ± 456.3	0.92
FGD1 1Trk	3587.77	3695.7 ± 481.8	3.99
FGD1 NTrk	1066.91	1093.9 ± 138.4	1.51
FGD2 1Trk	3618.29	3721.1 ± 473.6	3.89
FGD2 NTrk	1077.24	1102.9 ± 131.6	0.90
FGD1 ν_μ 1 Trk	1272.17	1309.4 ± 139.6	1.27
FGD1 ν_μ NTrk	1357.45	1378.0 ± 155.9	0.50
FGD2 ν_μ 1Trk	1262.63	1305.9 ± 140.9	2.09
FGD2 ν_μ NTrk	1246.71	1263.5 ± 139.4	0.49
Total	63632.86	64670.8 ± 7025.2	51.72

Table 5.13.: Event rates broken down by sample after the prior predictive spectrum for the Asimov data

Figure 5.34 (FHC) and Figure 5.35 (RHC) shows the result of the Asimov fit for the ND280 flux parameters. There is a consistent "bias" in estimating the parameter set which generated the Asimov, whose residuals (bottom panel) appear to have a shape, going from over-estimated to under-estimated as E_ν increases. The highest acceptance

	Chain length	Acceptance rate	Accepted steps
Low acceptance	800,000	6.9%	55,200
Mid acceptance	800,000	15.0%	120,000
High acceptance	1,400,000	23.0%	322,000

Table 5.14.: Different MCMC run configurations for the Asimov fit

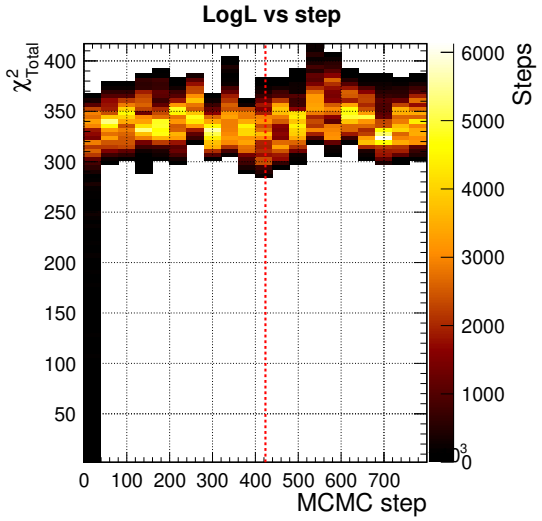


Figure 5.33.: Markov Chain behaviour for the "mid acceptance" MCMC, showing intended behaviour of moving around minimum

chain has less "bias" than the shorter chains, and all the chains are consistent with each other. Many of the parameters with large biases are barely constrained by the ND280 data (e.g. ND280 FHC ν_μ at $E_\nu > 4$ GeV and ND280 RHC $\bar{\nu}_e$) and only get their constraints from the prior covariance matrix. The ND280 to SK correlations are working as expected and the SK flux parameters show almost identical constraint to their ND280 equivalents.

We note improved constraints on all flux parameters from the prior uncertainty. This is strongest around the flux peak at $E_\nu \sim 0.6$ GeV, where the flux parameter uncertainty reduces by 50%.

Figure 5.36 shows interaction parameters after the fit to Asimov data. As for the flux parameters, some parameters are "biased", notably the Fermi momentum parameters (p_F), 2p2h norm $\bar{\nu}$, BeRPA A, CC DIS and NC1γ. Again, the chains are compatible in their estimates of the parameter values, and the reduction of the uncertainties are significant for many parameters. p_F, 2p2h norm C/O, BeRPA E, ν_e/ν_μ, NC coherent, NC1γ and NC oth. SK barely improve due to the fit lacking events

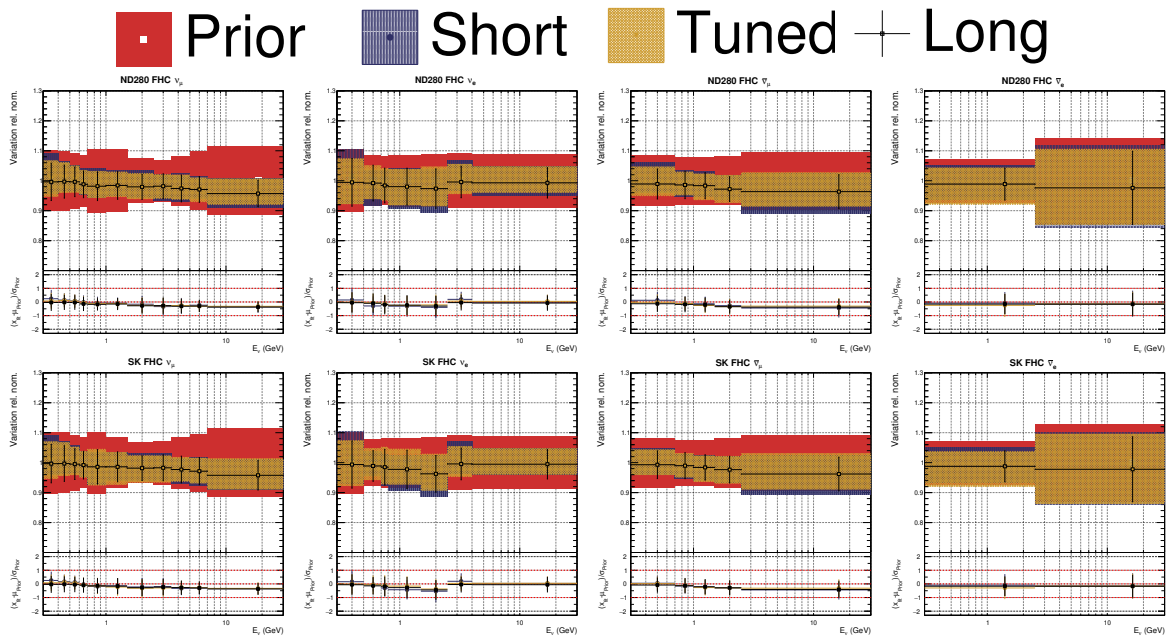


Figure 5.34.: ND280 and SK FHC flux parameters after the Asimov fit for different MCMC chains

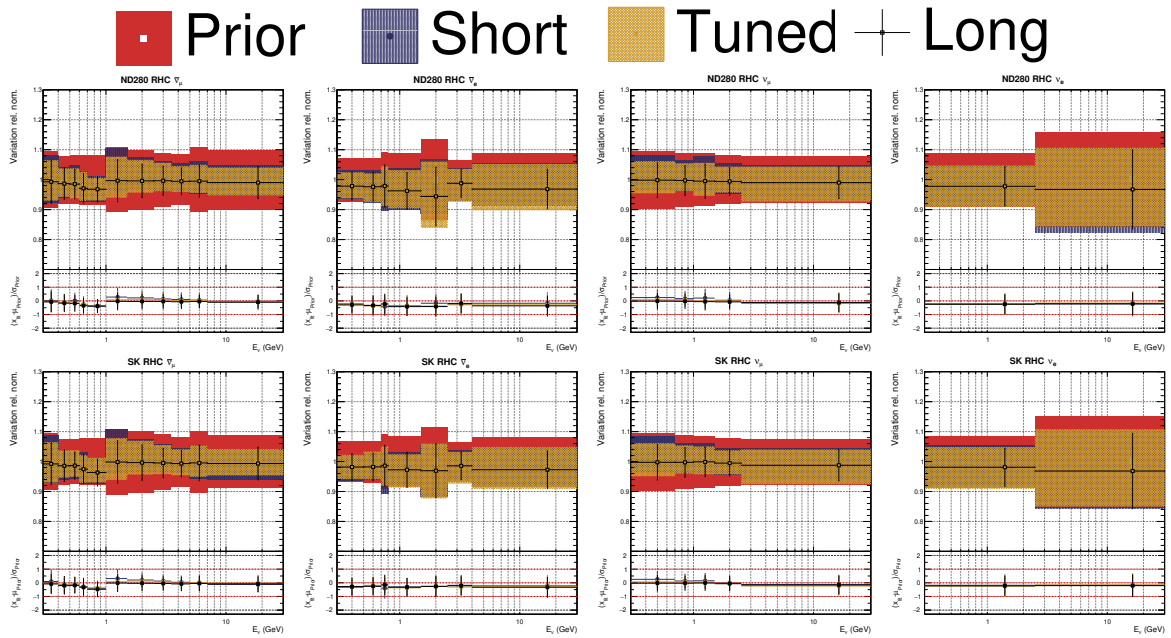


Figure 5.35.: ND280 and SK RHC flux parameters after the Asimov fit for different MCMC chains

which constrain these parameters, and/or the parameter has very little "strength" to change the interaction cross-section.

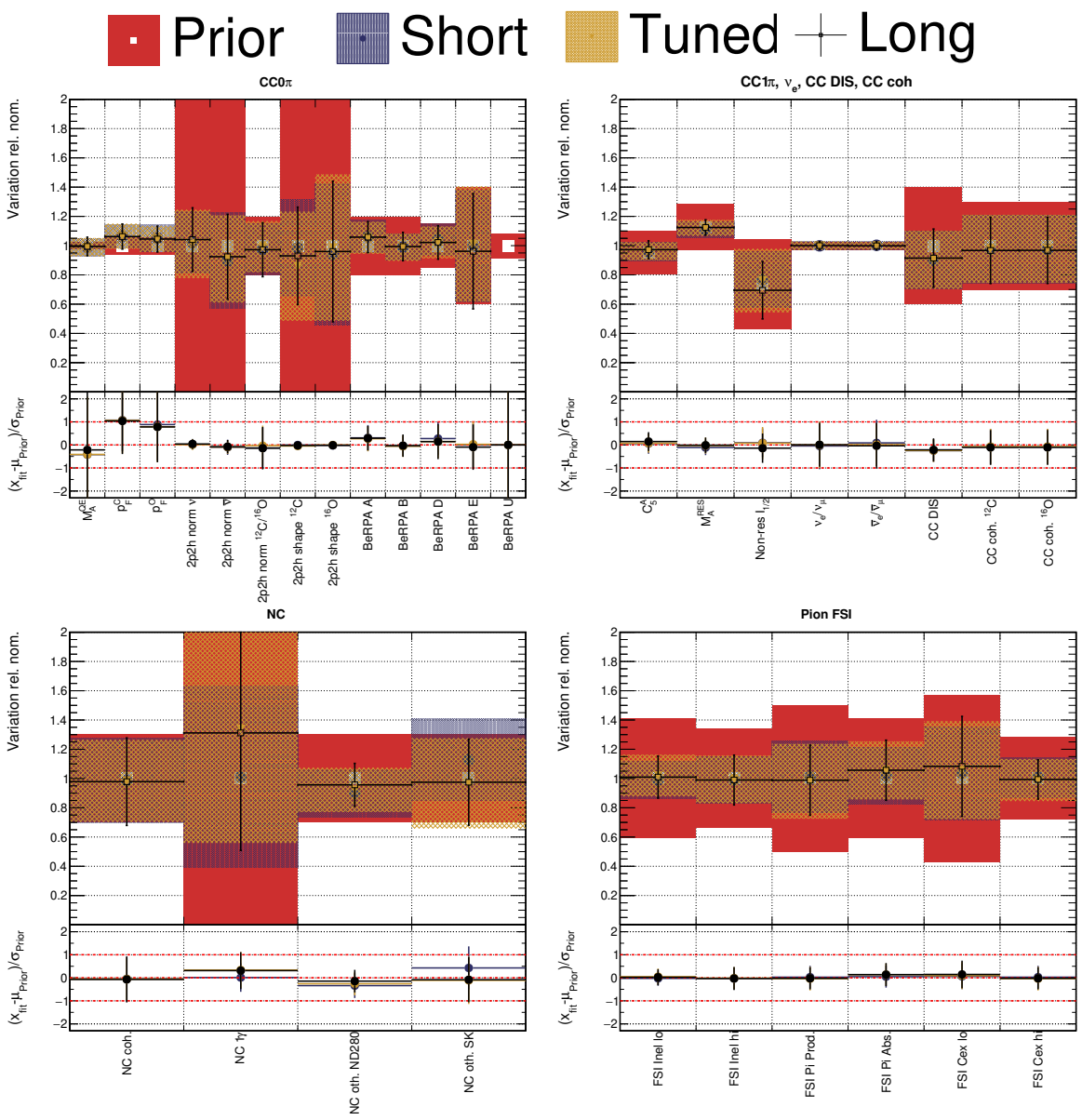


Figure 5.36.: Interaction parameters after the Asimov fit for different MCMC chains

Fitting to Asimov Data, Varying Only Flux Parameters

To investigate if correlations and marginalisation are to blame for the "bias" in subsection 5.6.4, a fit to the Asimov data is done keeping the cross-section and ND280 parameters fixed at their Asimov values, and only the flux parameters are varied. Figure 5.37 shows the parameter values for the ND280 flux parameters using different methods of obtaining the central values. The biases have vanished for all methods, and we conclude that the Asimov parameters are being correctly found when flux parameters alone are being fit.

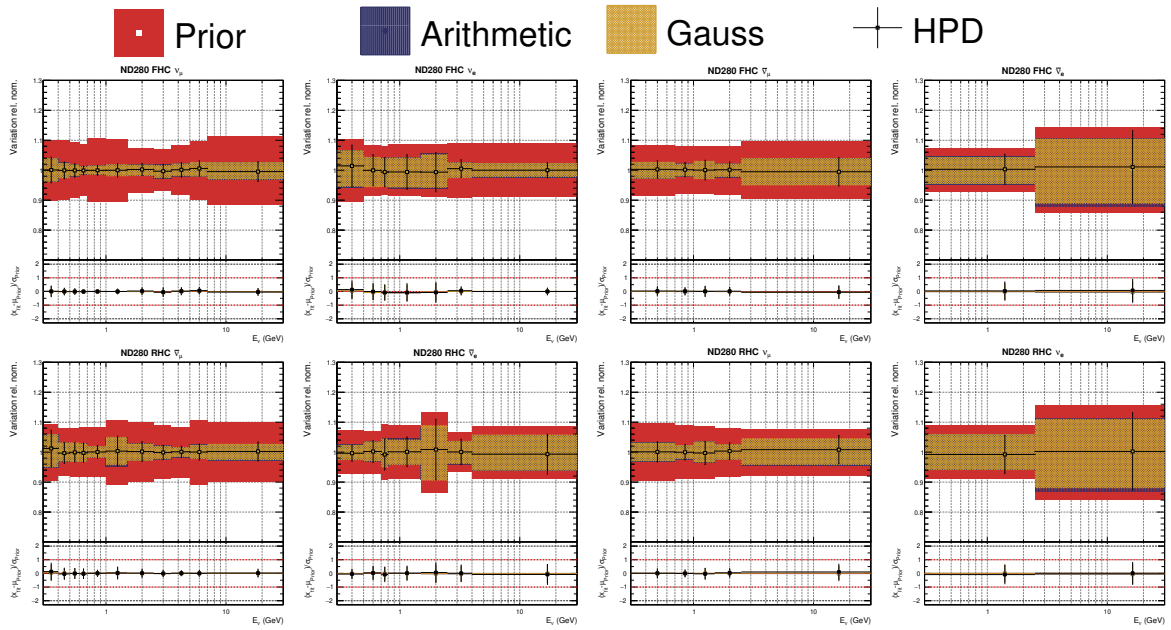


Figure 5.37.: ND280 flux parameters after the Asimov fit, fitting flux only

Marginalisation Effects

Marginalisation effects are expected for parameters which correlate strongly with non-Gaussian parameters, which are present in this analysis. This is investigated further by looking at the two-dimensional marginal posteriors for a few correlated parameters and comparing them to their one-dimensional marginal posterior. For the central value parameter estimates we use the highest posterior density point. We look at three selected parameters that show bias: p_F^C, 2p2h norm $\bar{\nu}$ and the 30th beam parameter ("b_{29}", which is the ND280 RHC $\bar{\nu}_\mu$ parameter for $E_\nu = 0.7 - 1.0$ GeV).

p_F^C bias Figure 5.38 shows the two dimensional marginal posterior of p_F^C with p_F^O and BeRPA D. The p_F parameters are problematic in that the parameterisation allows for non-Gaussian behaviour and additionally have hard cut-offs near the prior value– notably at the lower limit. Marginalising over such parameters causes shifts in the location of the highest posterior density. For the two-dimensional posterior of p_F^C and p_F^O the highest density is very close to the prior input value of 1.0 for both parameters (within one bin-width), whereas marginalising over the other parameter causes the posterior density to shift from $p_F^C = 1.008$ to 1.062. The second inset shows the opposite effect when marginalising p_F^C over BeRPA D, in which the two-dimensional marginal posterior has its highest density at $p_F^C = 1.077$ which shifts to $p_F^C = 1.054$ after the marginalisation.

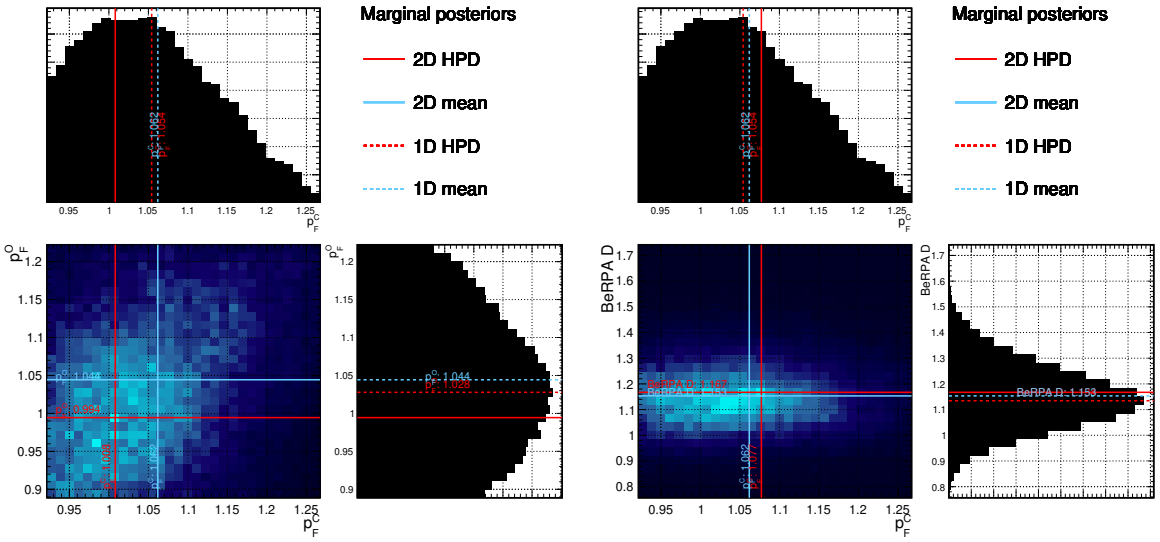


Figure 5.38.: Selected two-dimensional marginal posteriors for p_F^C and 2p2h shape O and BeRPA B, showing the resulting one-dimensional marginal posterior. The left figure demonstrates a large marginalisation effect, while the right shows a small effect

2p2h norm $\bar{\nu}_\mu$ bias Figure 5.39 shows the same marginalisation plots for 2p2h norm $\bar{\nu}$ with 2p2h shape C and BeRPA E. The case of the 2p2h normalisation parameters is slightly different to the p_F parameters: the normalisation is well-behaved across the phase space, looks Gaussian, and does not have any hard cut-offs near the prior[i]. The marginalisation effect instead happens with parameters that 2p2h normalisations correlate with, which may have non-Gaussian posterior shapes. The first example is with 2p2h shape C, in which we notice a tail at low 2p2h shape C and low 2p2h norm $\bar{\nu}$. Marginalising over the parameter brings the posterior density from 2p2h norm $\bar{\nu}$ = 1.061 to 0.842, which is the main cause of the "bias". Again, other parameters have the opposite effect: marginalising over BeRPA E causes the marginal posterior to shift from 2p2h norm $\bar{\nu}$ from 0.769 to 0.842.

***b*_29 bias** The 30th beam parameter ("*b*_29", the ND280 RHC $\bar{\nu}_\mu$ parameter for $E_\nu = 0.7 - 1.0$ GeV) sees similar effects from a number of parameters: "*b*_24" (ND280 FHC $\bar{\nu}_e$ 2.5-30 GeV), "*b*_46" (ND280 RHC ν_μ 1.5-2.5 GeV), "*b*_49" (ND280 RHC ν_e 2.5-30 GeV) and BeRPA D. The marginalisation plots are shown in Figure 5.40 and all the above parameters have identical shifts: *b*_29 moves from 0.988 to 0.974. It is noteworthy that the above parameters are all only weakly constrained by ND280 data:

[i]The cut-off is when the normalisation reaches 0

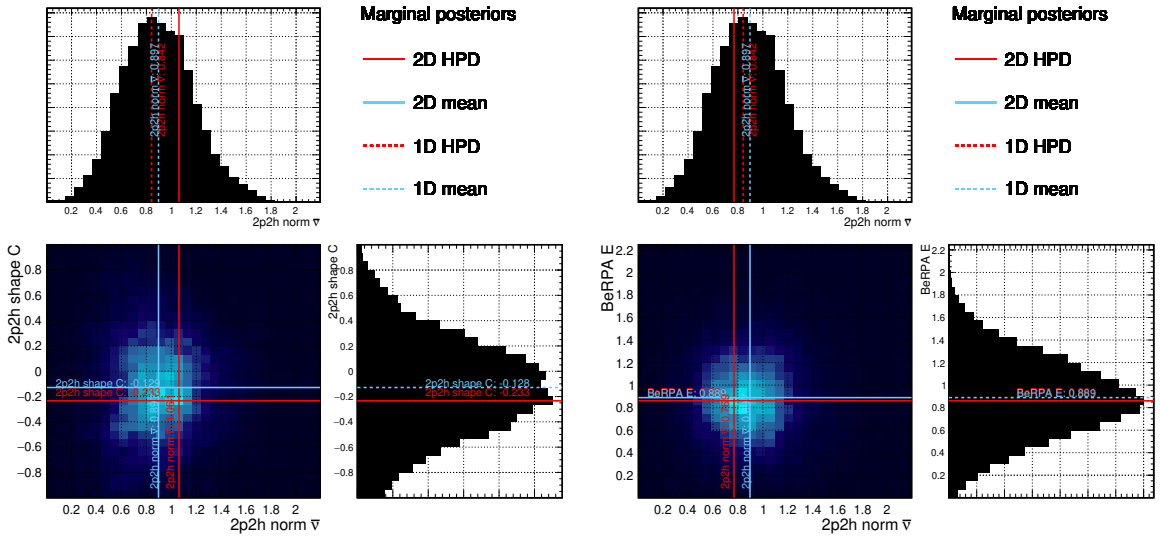


Figure 5.39.: Selected two-dimensional marginal posteriors for 2p2h norm $\bar{\nu}$ with 2p2h shape C and BeRPA E showing the resulting one-dimensional marginal posterior. The left figure demonstrates a negative bias and the right a positive bias

there is no dedicated ν_e selection (b_46 and b_49 control ν_e flux) and high Q^2 CCQE events are sparse (the region for BeRPA D). This can cause non-Gaussianity in that the chain may "wander" the space and explore a largely flat likelihood in parameters.

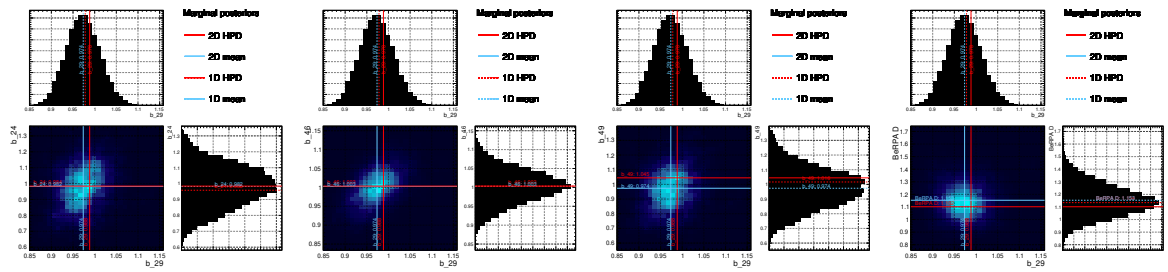


Figure 5.40.: Selected two-dimensional marginal posteriors for "b_29" (ND280 RHC $\bar{\nu}_\mu$ 0.7-1.0 GeV)

In conclusion, the biases present in the Asmiov fit study seems likely to be due to marginalisation effects over parameters that are non-Gaussian, have hard cut-offs or are poorly constrained in the fit. These patterns are expected to arise again when fitting against real data. The frequentist (BANFF) group's validations are provided in Appendix E for comparison, and finds no biases.

5.6.5. Posterior Predictive Spectrum

The next statistical closure test we perform on the Asimov data fit is to calculate the posterior predictive spectra and p-values. The chosen MCMC was presented in subsection 5.6.4 and the longest chain in Table 5.14 was selected for the calculation of these spectra.

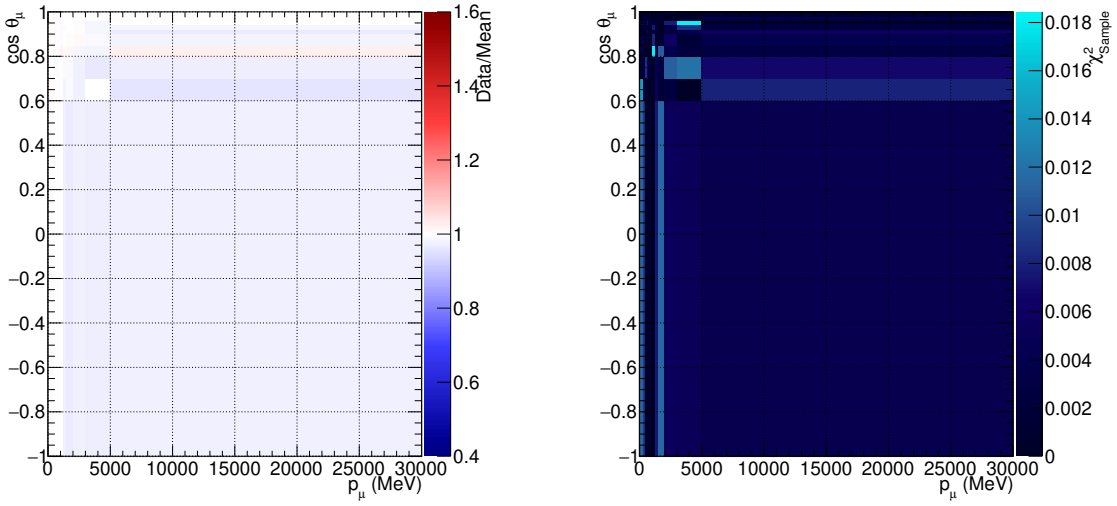
Table 5.15 shows the event rates from the posterior predictive spectrum. There is a consistently low likelihood contribution from all samples ($\sim 0.18 - 0.42$), totalling at 4.43. This is small compared to the number of bins being fit (1624), and the contributions enter primarily in low statistics areas (high p_μ, low $\cos\theta_\mu$), where Poissonian event distributions are expected. The 2D p_μ $\cos\theta_\mu$ distributions and the bin-by-bin likelihood contributions for FGD1 and FGD2 0π are shown in Figure 5.41.

Sample	Nominal	Pos. Pred	-2log $\mathcal{L}_s$
FGD1 0π	16723.8	16730.9 ± 116.8	0.42
FGD1 1π	4381.47	4371.67 ± 57.7	0.29
FGD1 Other	3943.95	3957.1 ± 55.8	0.28
FGD2 0π	16959.3	16952.8 ± 116.7	0.31
FGD2 1π	3564.23	3565.16 ± 49.5	0.36
FGD2 Other	3570.94	3571.43 ± 51.0	0.34
FGD1 1Trk	3587.77	3586.45 ± 52.5	0.28
FGD1 NTrk	1066.91	1075.13 ± 21.4	0.35
FGD2 1Trk	3618.29	3612.19 ± 51.9	0.36
FGD2 NTrk	1077.24	1084.68 ± 21.4	0.20
FGD1 ν_μ 1 Trk	1272.17	1267.04 ± 23.6	0.18
FGD1 ν_μ NTrk	1357.45	1357.19 ± 25.2	0.38
FGD2 ν_μ 1Trk	1262.63	1259.48 ± 22.3	0.35
FGD2 ν_μ NTrk	1246.71	1246.61 ± 25.0	0.33
Total	63632.86	63637.83 ± 253.0	4.43

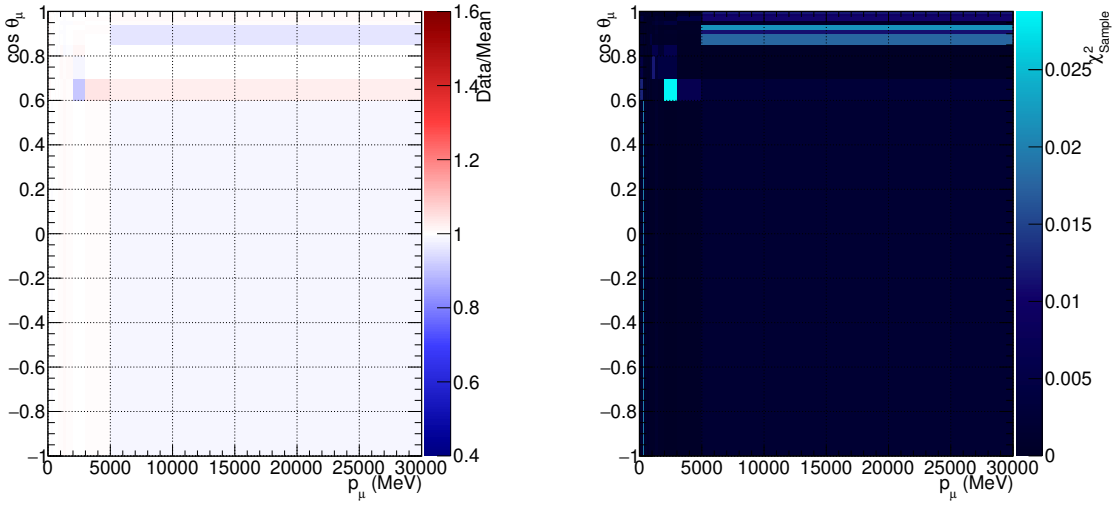
Table 5.15.: Event rates broken down by sample after the posterior predictive spectrum for the fit to Asimov data

Comparing the posterior predictive event rates in Table 5.15 to the prior predictive in Table 5.13 we see a drastic reduction in the event rate uncertainties: 7025.2 to 253.0 for the total rate, and for FGD1 CC0π 2244.0 to 116.8, 487.6 to 57.7 for CC1π and 537.5

to 55.8 for CCOther. This brings the event rate uncertainties below 2% level for all the selections, which sets a benchmark for the fit to data.



(a) FGD1 CC0π



(b) FGD2 CC0π

Figure 5.41.: Posterior predictive $p_\mu \cos\theta_\mu$ spectrum data/post-fit ratios and bin-by-bin likelihood contributions for the fit to Asimov data

The calculated posterior predictive p-value is shown in Figure 5.42, and as expected the p-value is 1.0. Comparing the y-axis ($-2\mathcal{L}_{\text{Sample}}$(Data, Draw)) to the prior predictive spectrum in Figure 5.32, the difference is almost one order of magnitude, essentially reflecting the much tighter model constraint after using ND280 data.

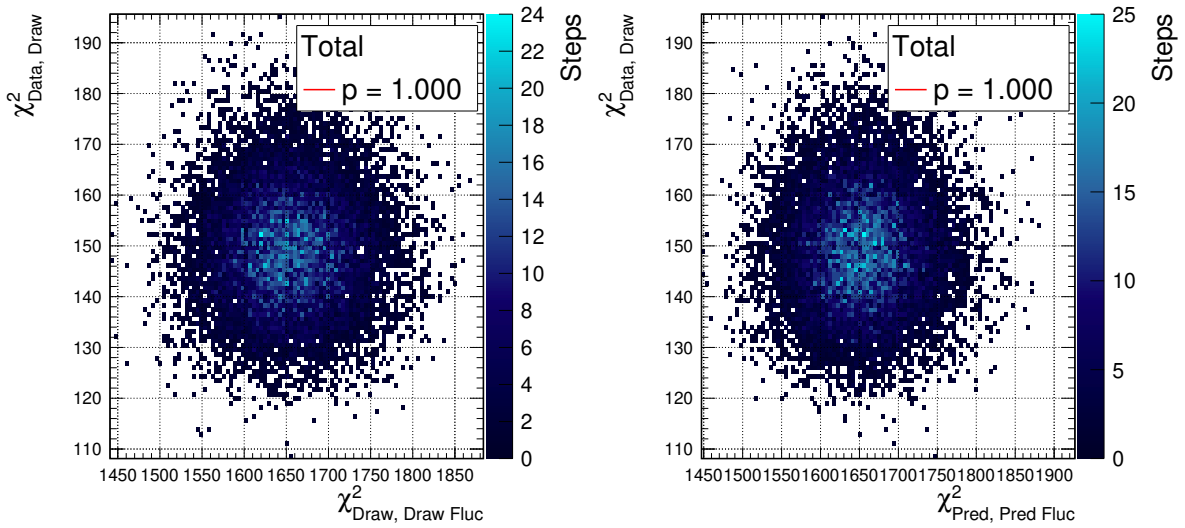


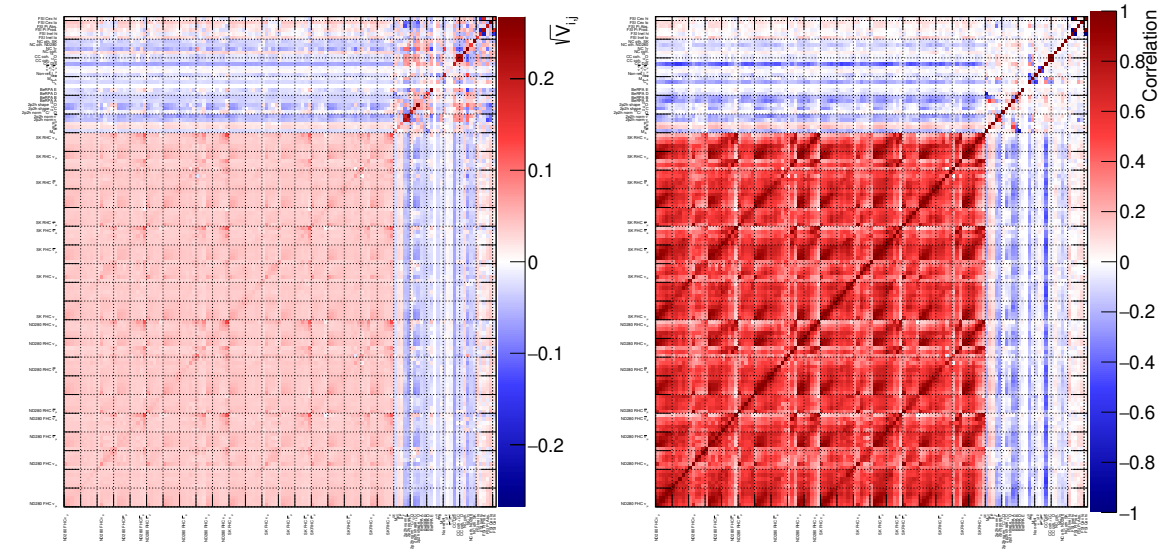
Figure 5.42.: Posterior predictive p-values for the fit to Asimov data

5.6.6. Covariance Matrix from the Asimov Fit

The last consistency check is to inspect the covariance matrix from the fit to Asimov data. The expected covariance matrix contains heavily correlated flux parameters—roughly retaining the covariances from the input covariance matrix—, interaction parameters correlating internally for parameters that affect the same modes and topologies (e.g. M_A^{QE} and BeRPA), and correlations between the flux parameters and cross-section parameters, especially for normalisation parameters.

Figure 5.43 presents the full flux and cross-section parameter square root covariance and correlation matrix. The red square in the bottom left corner is the 100 flux parameters, all strongly internally correlated. The upper right corner is occupied by the cross-section parameters, which has some internal correlations. Many cross-section parameters correlate with the flux parameters, notably CC DIS—which is parameterised as $0.4/E_\nu$, so is essentially a flux normalisation weight for DIS events—, C_5^A—which largely controls the CC1π neutrino-nucleon cross-section normalisation, and the BeRPA parameters—which control the Q^2 correction of CCQE events due to the RPA effects.

A more digestible covariance matrix is shown in Figure 5.44, which excludes the SK flux parameters. The largest post-fit flux uncertainty is from the ν_e flux parameters, and the high energy flux parameters correlate only weakly, as expected from the neutrino production parents at high energies compared to low energies. There are



(a) $\sqrt{\mathbf{V}_{i,j}}$ **(b)** $\rho_{i,j}$

Figure 5.43.: $\sqrt{\mathbf{V}_{i,j}}$ and correlation matrix for the Asimov post-fit, showing the full flux and cross-section parameters

very strong correlations for the low energy FHC ν_μ and RHC $\bar{\nu}_\mu$ parameters, and the low energy flux parameters correlate with all the 2p2h parameters.

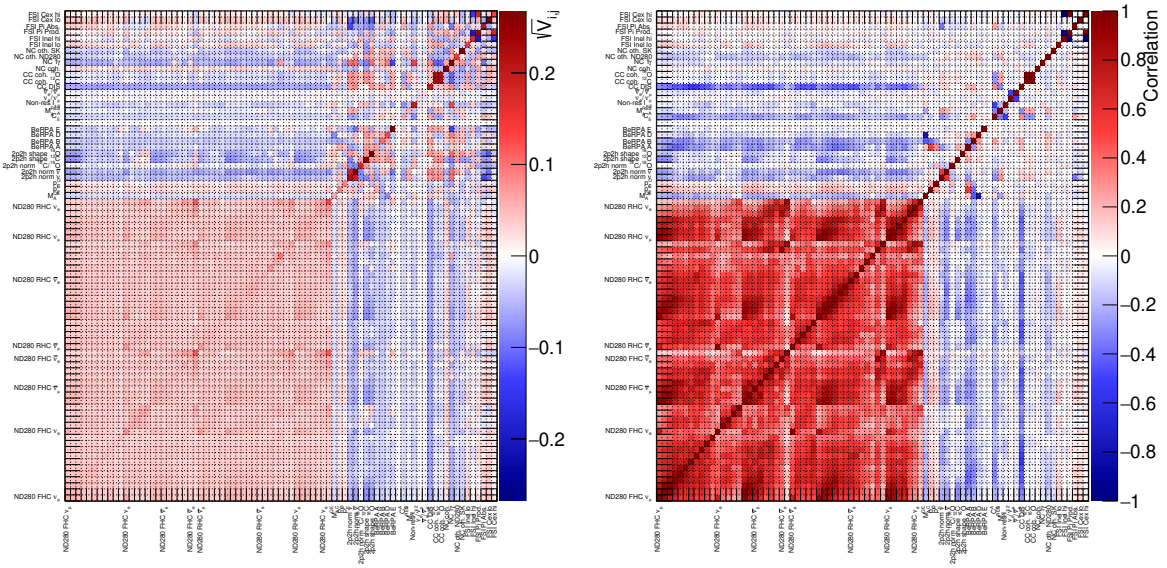
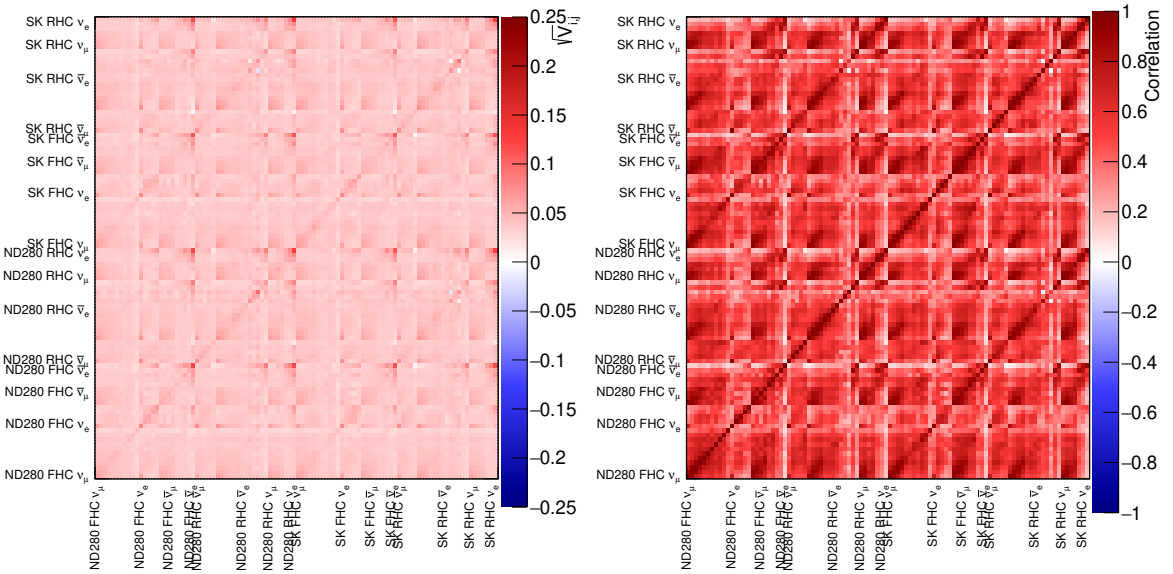


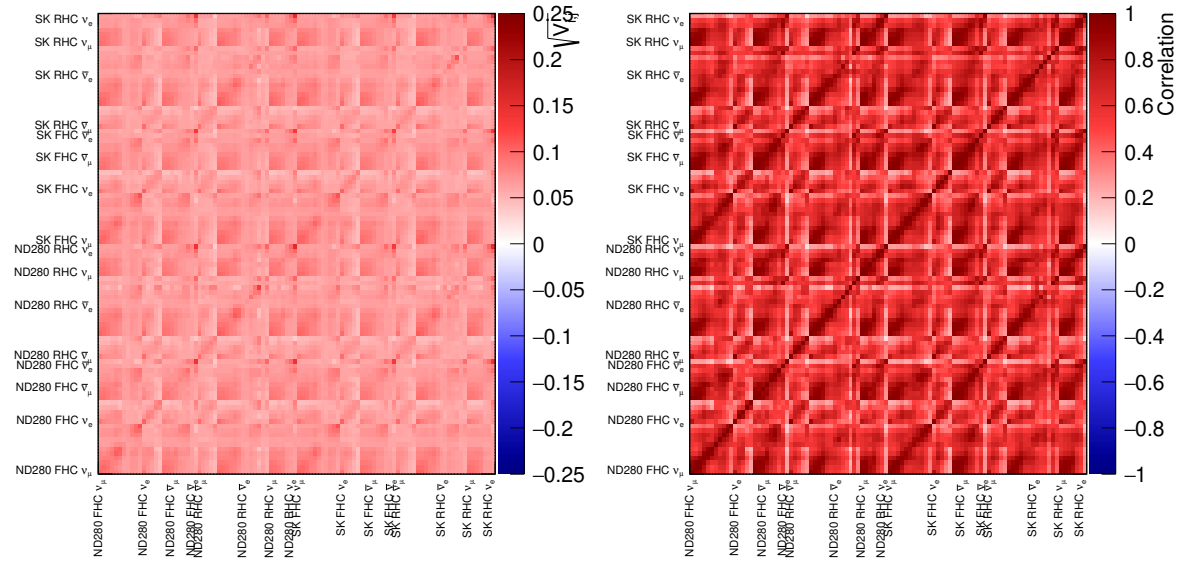
Figure 5.44.: $\sqrt{\mathbf{V}_{i,j}}$ and correlation matrix for the Asimov post-fit, showing ND280 flux and cross-section parameters

Figure 5.45 shows the flux parameters with the input prior covariance matrix and the output post-fit Asimov covariance matrix. The correlations are largely the same,

and we notice the smaller covariance values post-fit compared to the pre-fit, reflecting the tighter constraint on the flux after fitting ND280 data.



(a) Post-fit



(b) Pre-fit

Figure 5.45.: $\sqrt{\mathbf{V}_{i,j}}$ and correlation matrix for the flux parameters pre and post-fit to Asimov data

Figure 5.46 shows the matrices zoomed in on the cross-section parameters (the upper right corner of Figure 5.43 and Figure 5.44). As expected there are correlations between C with O parameters (e.g. CC Coh C and CC Coh O), and ν with $\bar{\nu}$ parameters (e.g. 2p2h norm). The CC0π parameters (lower left block) correlate M_A^{QE}, p_F and

the BeRPA parameters, which affect the CCQE interaction. There are strong internal correlations in the BeRPA parameters, as expected. The single pion parameter block (middle) roughly maintains their prior correlations, as does the FSI block (top right corner). The single pion parameter C_5^A, M_A^{RES} and non-resonant $I_{1/2}$ parameters correlate with the CC coherent parameters—which produce a $1\pi^{\pm}$ final state—and the CC DIS parameter—since they populate the CC1π selection (and to a less extent the CCOther). We note slight correlations between the CCQE and single pion parameters, due to the 20% single pion events in the CC0π selection from final state pion interactions, seen in Table 5.11.

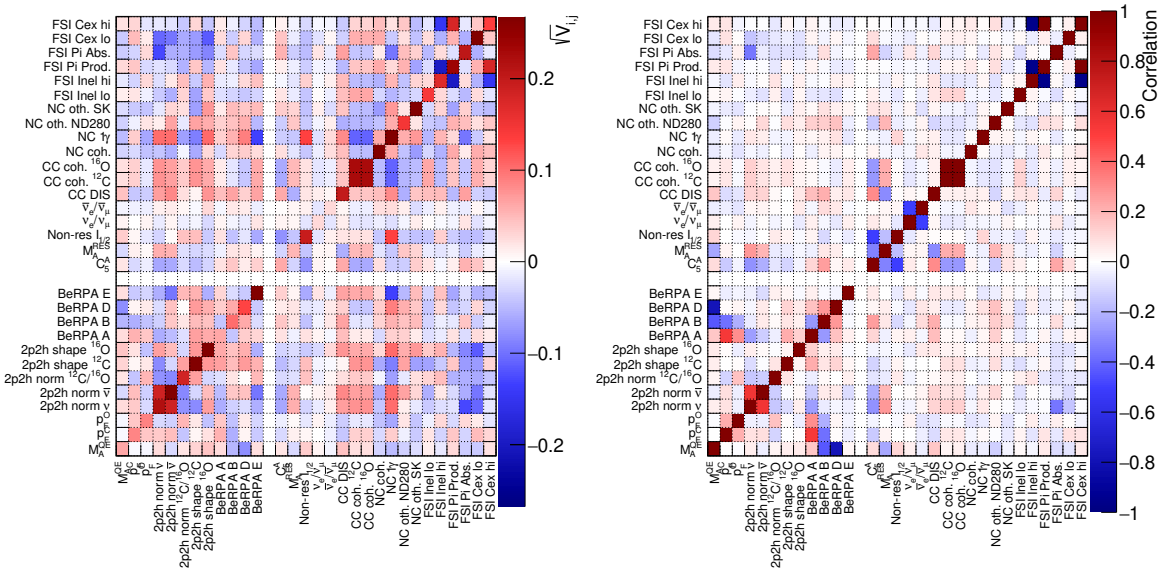


Figure 5.46.: $\sqrt{\mathbf{V}_{i,j}}$ and correlation matrix for the Asimov post-fit, showing cross-section parameters

5.7. Fitting Real Data

The nominal model presented in section 5.5 is fit to data in the reconstructed p_μ $\cos\theta_\mu$ variables, using the 14 ND280 selections with data from run 2-6. Three Markov Chain Monte Carlo samples are presented with different step-size tunings, acceptance rates and sample lengths, and a conservative 1/4 burn-in was used for all chains in Table 5.16.

	Steps	Acceptance (%)	Accepted steps
Tuned	800,000	24.0	192,000
Long	3,000,000	6.2	186,000
Short	1,600,000	6.2	99,200

Table 5.16.: Different MCMC samples run for the data fit

Table 5.17 shows the event rates per sample and total before and after the fit for the tuned MCMC, making correlated throws of the model to build the errors using the prior predictive method and the posterior predictive method, outlined earlier.

There is very good agreement with the overall event rate (64768 data vs 64761.40 post-fit) and the CC0π samples. The total sample contribution to the test-statistic[j] moves from 2539.08 to 1733.08 after the analysis. For the samples, we generally see good effect on targets, with FGD2 CC0π decreasing its test-statistic by factor 2. Interestingly, FGD1 CCOther sees a much smaller improvement than FGD2 (273.39 to 224.02 vs 277.96 to 171.17), indicating tension in the fit. The uncertainty on the predictions decrease remarkably post-fit: for the 0π selections we go from an uncertainty of 2094.2 (12%) to 120.0 (0.7%), and overall from 6666.8 (10%) to 255.2 (0.4%).

Comparing the predictive distributions from the fit to data in Table 5.17 to the Asimov in Table 5.13 and Table 5.15, the data fit uncertainties on the total event rates agree with the expectation. The uncertainties from the data fit are within 10% of the Asimov fit for both the prior and posterior predictive spectra.

Figure 5.47 shows the FHC flux parameters after the fit to data for the different chains. As expected and seen in the Asimov studies, the SK parameters closely follow the ND280 parameters. The parameters mostly sit within 1σ of the prior, with the largest deviation around $E_\nu \sim 0.6 - 1.0$ GeV, where the flux normalisation drops sharply to 0.88 and returns to around nominal at $E_\nu \sim 1.5$ GeV. At higher energies the flux normalisation decreases again to around 0.94. The parameters all sit within one sigma of the central value of the prior.

Figure 5.48 shows the RHC flux parameters post-fit. There is a decrease of $\bar{\nu}_\mu$ for low energies to approximately 0.96, which increases to the nominal at $E_\nu = 0.6$ GeV. At higher energies the normalisations go back to around 0.96. The only increase observed are the $\bar{\nu}_e$ parameters around 0.6 GeV. No parameters leave the 1σ prior.

[j] Which excludes the penalty term from the priors in the likelihood

Sample	Event rate			$-2\ln\mathcal{L}_S$	
	Data	Prior	Posterior	Prior	Posterior
FGD1 0π	17136	16837.5 ± 2094.2	17122.9 ± 120.0	272.22	172.21
FGD1 1π	3954	4438.8 ± 482.9	4053.8 ± 54.3	259.65	164.04
FGD1 Other	4149	3974.0 ± 516.8	4103.7 ± 58.8	273.39	224.03
FGD2 0π	17443	16952.1 ± 2072.6	17501.5 ± 122.5	303.62	166.15
FGD2 1π	3366	3613.8 ± 400.3	3409.4 ± 48.2	228.72	162.71
FGD2 Other	4075	3602.5 ± 463.1	3917.8 ± 50.8	277.96	171.17
FGD1 1Trk	3527	3688.8 ± 477.3	3509.6 ± 50.1	156.19	117.80
FGD1 NTrk	1054	1091.3 ± 134.9	1062.7 ± 21.9	109.59	76.50
FGD2 1Trk	3732	3714.0 ± 471.6	3678.7 ± 51.3	169.17	129.84
FGD2 NTrk	1026	1102.2 ± 129.6	1108.4 ± 23.5	89.74	80.34
FGD1 ν_μ 1Trk	1363	1309.4 ± 142.4	1348.3 ± 23.1	88.42	66.51
FGD1 ν_μ NTrk	1370	1376.8 ± 153.9	1359.1 ± 26.9	99.54	61.75
FGD2 ν_μ 1Trk	1320	1303.4 ± 139.7	1323.1 ± 23.8	92.56	64.34
FGD2 ν_μ NTrk	1253	1263.1 ± 139.7	1265.9 ± 24.2	118.31	75.70
Total	64768	64603.2 ± 6666.8	64761.1 ± 255.2	2539.08	1733.08

Table 5.17.: Neutrino event counts and test-statistic for data, pre-fit MC and post-fit MC broken down by sample

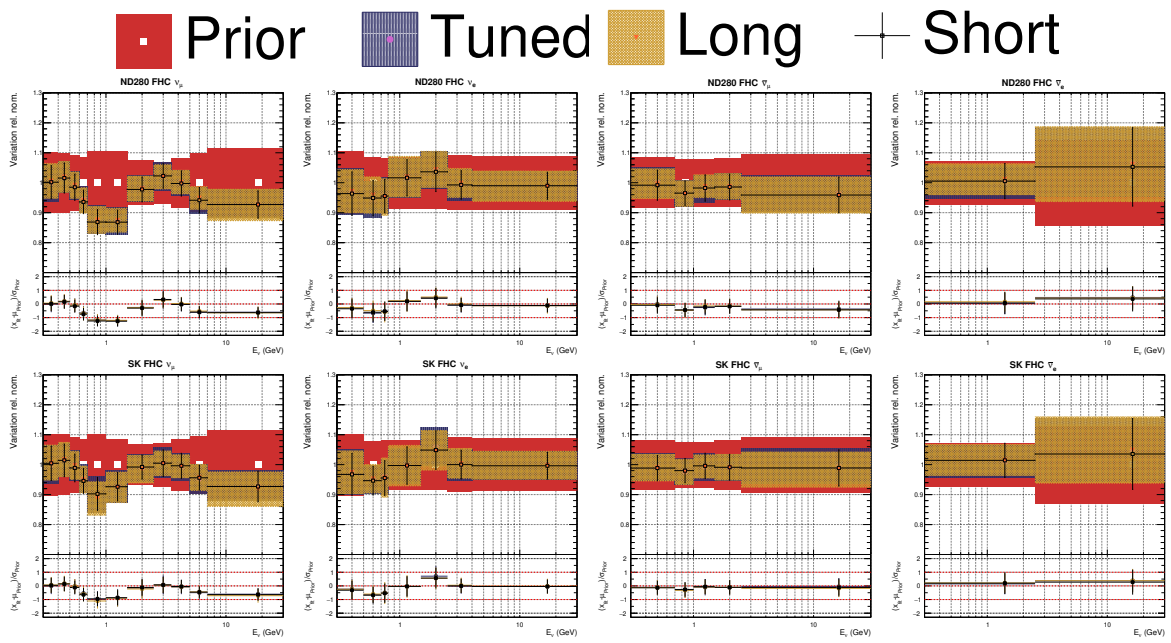


Figure 5.47.: FHC flux parameters after the data fit for different MCMC chains

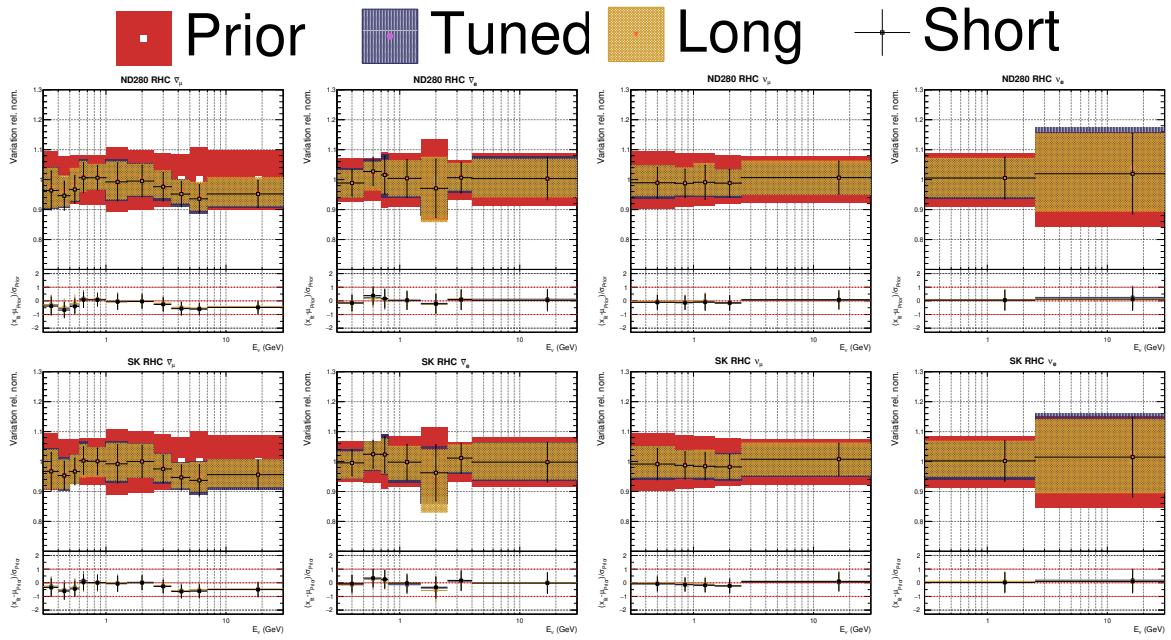


Figure 5.48.: RHC flux parameters after the data fit for different MCMC chains

Figure 5.49 show the interaction parameters after the fit, where we expect most of the parameter movement to happen. The largest deviations from the priors are seen for M_A^{QE} and BeRPA B, followed by M_A^{RES}.

The shift in M_A^{QE} is expected since the "prior" central value is from tunes to nuclear data [175]. M_A^{QE} is historically inflated [179,217–219] when not using adequate nuclear

effects, such as RPA or 2p2h, which leads to it becoming an effective parameter instead. Using a flat prior on M_A^{QE} leads toward values obtained from fits to nucleon data, $M_A^{QE} = 1.069 \pm 0.016$ GeV [220]. The post-fit value here in real units is $M_A^{QE} = 1.12 \pm 0.07$ GeV.

The shifts in BeRPA are more worrisome. The low Q^2 parameters BeRPA A and BeRPA B are being pushed above their pre-fit 1σ uncertainties. This is likely due to ND280 data having more events than the Monte-Carlo at low Q^2, forcing a shape that isn't "RPA-like"[k]. Figure 5.50 shows the BeRPA weight being applied for every step in the data and Asimov chains: the nominal shape of the RPA correction is heavily distorted in the data fit, in which an increased cross-section at low Q^2 is strongly favoured. The weaker constraints above $Q^2 = 0.5\,\text{GeV}^2$ reflect lacking data in the region. This feature may indicate unmodelled nuclear effects being absorbed by the RPA parameters.

The M_A^{RES} parameter is pushed far from the bubble chamber tuned values (from $M_A^{RES} = 1.07 \pm 0.15$ GeV to 0.806 ± 0.04 GeV), whereas the other single pion production parameters are compatible with the prior. The may indicate sweeping up unmodelled nuclear effects (often approximately a function of Q^2) into the M_A^{RES} parameter.

We note interesting 2p2h normalisations, in which the ν normalisation is 1.6, whereas $\bar{\nu}$ normalisation is 0.8, indicating some ν vs $\bar{\nu}$ tension in the nominal 2p2h model. The 2p2h shape parameters are pushed against their upper boundary, preferring a "delta-like" $q_0 - q_3$ phase space, again indicating the 2p2h model does not agree with ND280 data.

To understand why BeRPA is heavily distorted from the nominal form, we look at the reconstructed Q^2 distributions of the CC0π selections in Figure 5.51 and Figure 5.52. For both FGD1 and FGD2 there is a clear deficit at low Q^2 before the fit, about 9% in the first bin. The effect of BeRPA A on the pre-fit distributions is to change the low Q^2 region, and BeRPA B targets slightly higher Q^2. The post-fit distributions are a clear improvement for all of the Q^2 space, and this is the driver behind the BeRPA pulls.

The three different MCMC all produced near identical post-fit parameters.

[k] In the sense that it distorts the Q^2 correction beyond Nieves' [48] work.

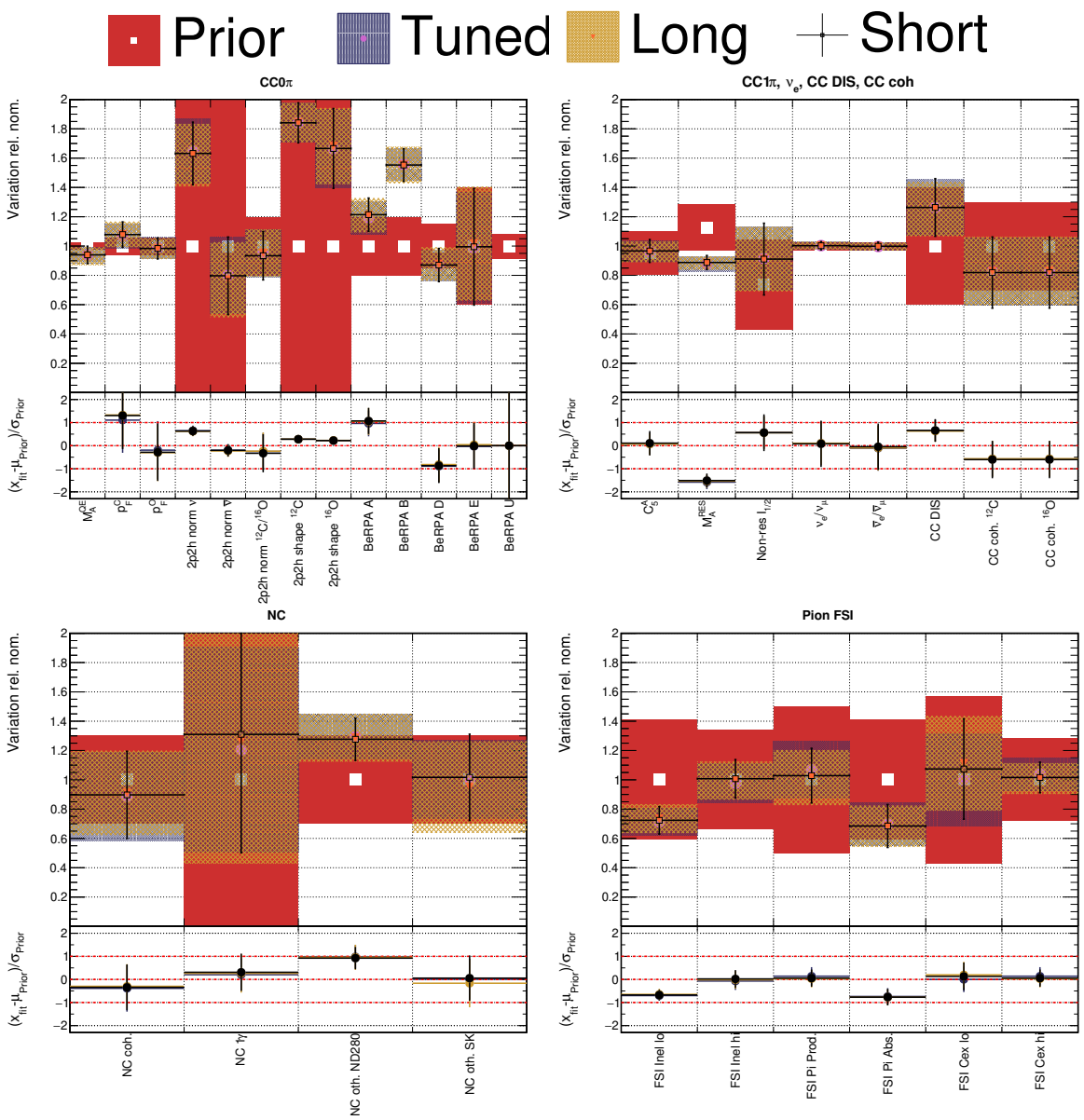
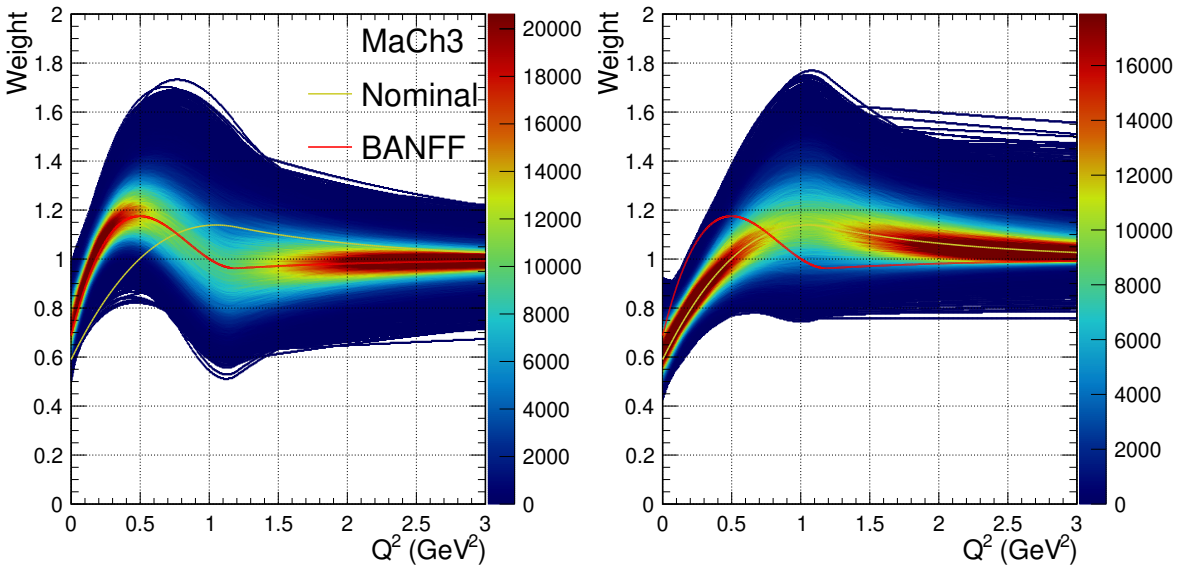


Figure 5.49.: Interaction parameters after the data fit for different MCMC chains

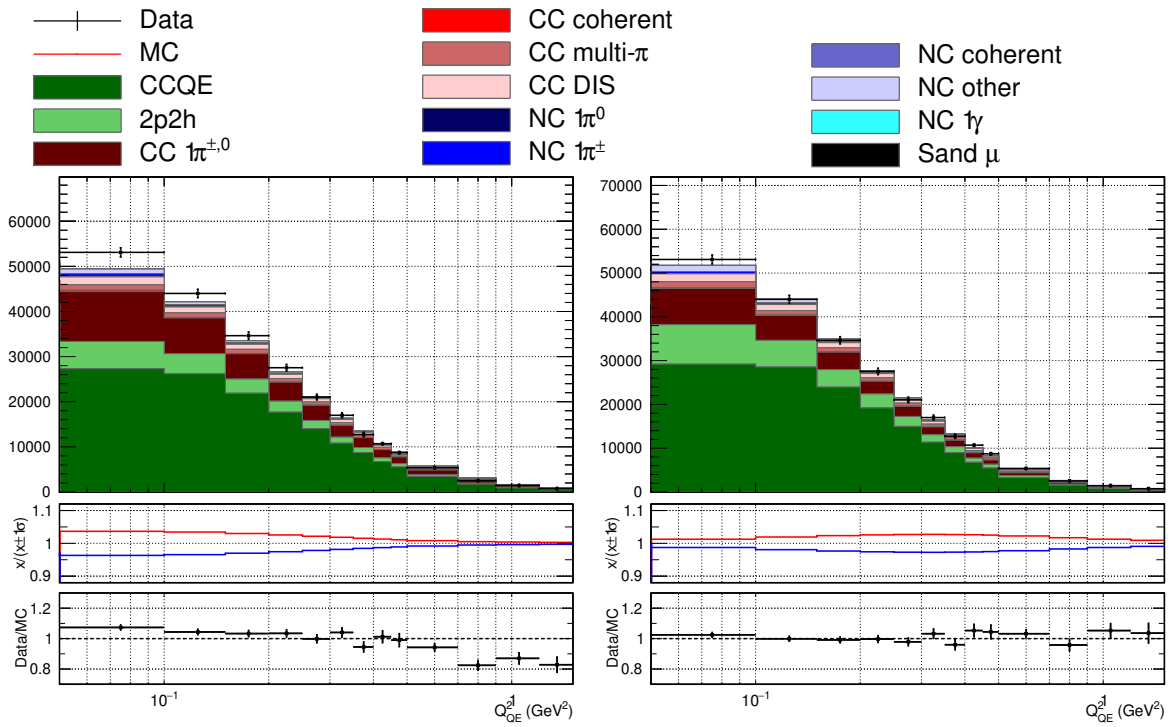
5.7.1. Prior Predictive Spectrum

The prior predictive spectrum and p-values are calculated in the same way as in the fit to Asimov data. Here we expect to see a poor p-value due to the priors' inability to describe ND280 data—the reason for doing the fit. Figure 5.53 shows the prior predictive p-value for all the samples, and as expected not a single parameter variation of the prior model produces a smaller test-statistic against the data than a fluctuation of the prior model does to itself.



(a) BeRPA shape after fitting to data **(b)** BeRPA shape after fitting to Asimov

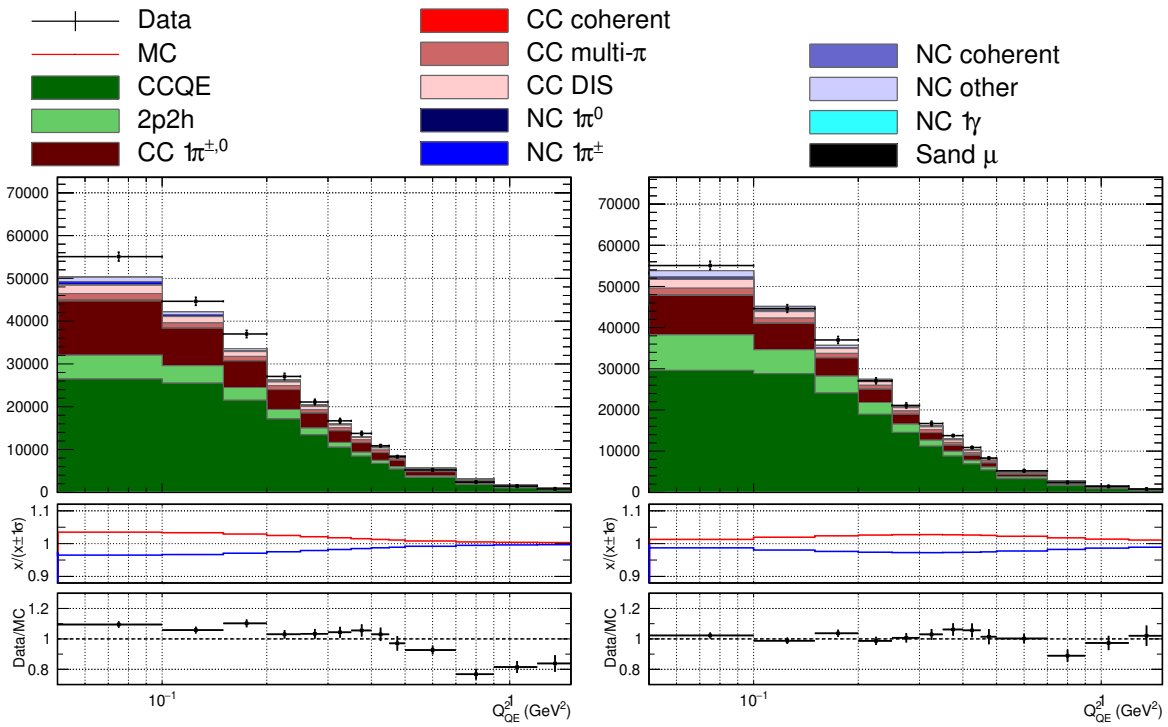
Figure 5.50.: BeRPA shape for each MCMC step for the tuned fits to data and Asimov



(a) Effect of BeRPA A after the fit to Asimov data **(b)** Effect of BeRPA B after the fit to data

Figure 5.51.: FGD1 CC0π in Q^2_{rec}after the fit to data, showing impact of the BeRPA parameters

The table of prior predictive p-values broken down by sample is found in Table 5.18. The high-statistics FHC samples all consistently have $p = 0.000$, and the low-statistics



(a) Pre-fit, BeRPA A effect **(b)** Post-fit, BeRPA B effect

Figure 5.52.: FGD2 CC0π in Q^2_{rec} after the fit to data, showing impact of the BeRPA parameters

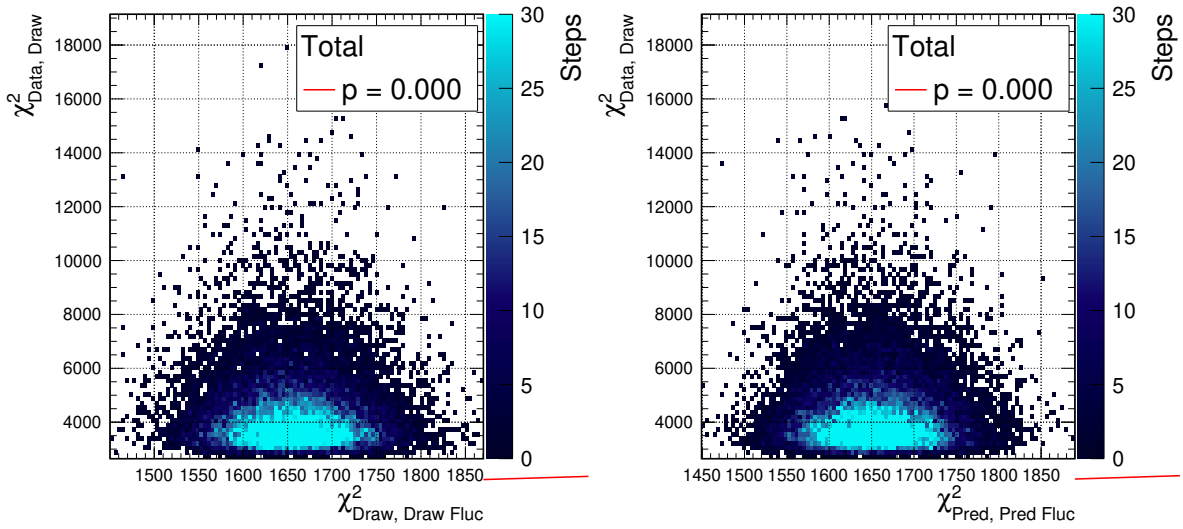


Figure 5.53.: Prior predictive spectrum for the data fit

RHC samples have slightly higher p-values. This is primarily due to statistical fluctuations being a larger effect for the low-statistics samples compared to the size of the systematic fluctuations.

Sample	Draw Fluc.	Pred. Fluc.
FGD1 0π	0.000	0.000
FGD1 1π	0.000	0.000
FGD1 Other	0.000	0.000
FGD2 0π	0.000	0.000
FGD2 1π	0.000	0.000
FGD2 Other	0.000	0.000
FGD1 1Trk	0.002	0.002
FGD1 NTrk	0.003	0.002
FGD2 1Trk	0.001	0.000
FGD2 NTrk	0.017	0.017
FGD1 ν_μ 1Trk	0.004	0.004
FGD1 ν_μ NTrk	0.011	0.009
FGD2 ν_μ 1Trk	0.003	0.003
FGD2 ν_μ NTrk	0.003	0.003

Table 5.18.: Prior predictive p-values for each sample after the data fit

5.7.2. Posterior Predictive Spectrum

The posterior predictive spectrum and p-values are calculated using the model after fitting to data. The resulting test-statistic distribution and p-value is shown in Figure 5.54, where $p = 0.000$[1].

The p-values are broken down into sample contributions, and are presented in Table 5.19. There are good p-values for all samples except FGD1 CCOther, which is zero. Importantly, the FGD2 CCOther selection observes a good p-value.

The test-statistic distributions for FGD1 CCOther are shown in Figure 5.55 with FGD2 inset for comparison. Whereas the x-axes (statistically fluctuated test-statistic in the model) are similar for the two selections, we see a much higher test-statistic on the y-axis, reflecting the post-fit distribution is much less like the data for FGD1 than for FGD2. This was already hinted at in Table 5.17, where FGD1 CCOther showed only a marginal improvement after the fit.

Figure 5.56 shows the likelihood contribution per bin for FGD1 and FGD2 CCOther selections. As expected, there are more bins with high contributions for FGD1 than

[1]A rather discouraging result!

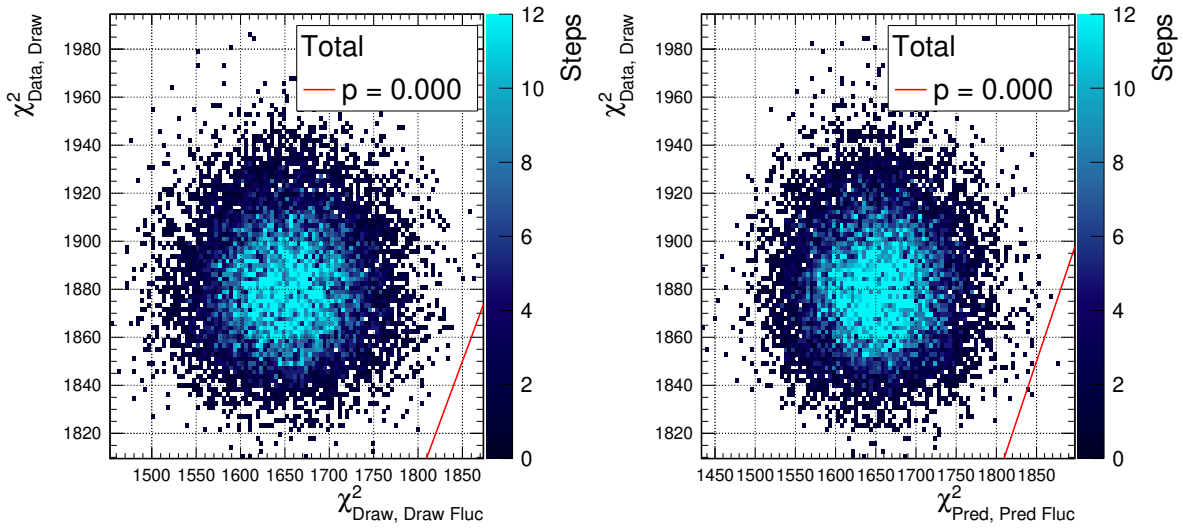


Figure 5.54.: Posterior predictive spectrum for the data fit

Sample	Draw Fluc.	Pred. Fluc.
FGD1 0π	0.062	0.060
FGD1 1π	0.078	0.075
FGD1 Other	0.000	0.000
FGD2 0π	0.115	0.114
FGD2 1π	0.090	0.089
FGD2 Other	0.098	0.103
FGD1 1Trk	0.515	0.515
FGD1 NTrk	0.292	0.289
FGD2 1Trk	0.265	0.260
FGD2 NTrk	0.230	0.219
FGD1 ν_μ 1Trk	0.296	0.293
FGD1 ν_μ NTrk	0.842	0.839
FGD2 ν_μ 1Trk	0.333	0.332
FGD2 ν_μ NTrk	0.587	0.591

Table 5.19.: Posterior predictive p-values for each sample after the data fit

FGD2, especially at low Q^2, which are not present in FGD2. For FGD1 CCOther, the high likelihood contributions sit primarily at $\cos\theta_\mu$ =0.9-0.95, whereas FGD2 has its largest contributions scattered across the phase space.

Inspecting the p_μ $\cos\theta_\mu$ post-fit distributions in Figure 5.57, the p_μ distributions appear different around 500-1000 MeV, where FGD2 sees a consistent underestimation

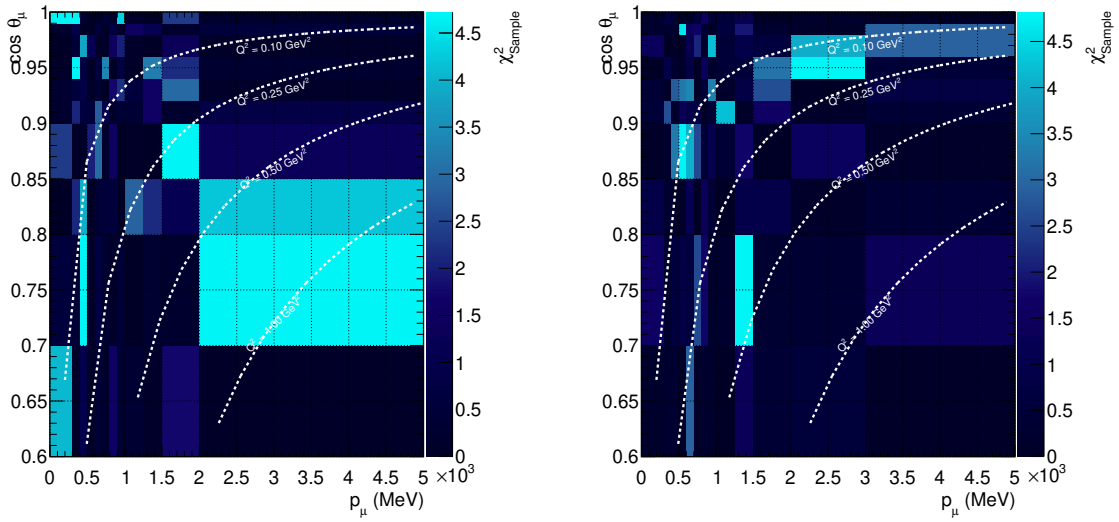
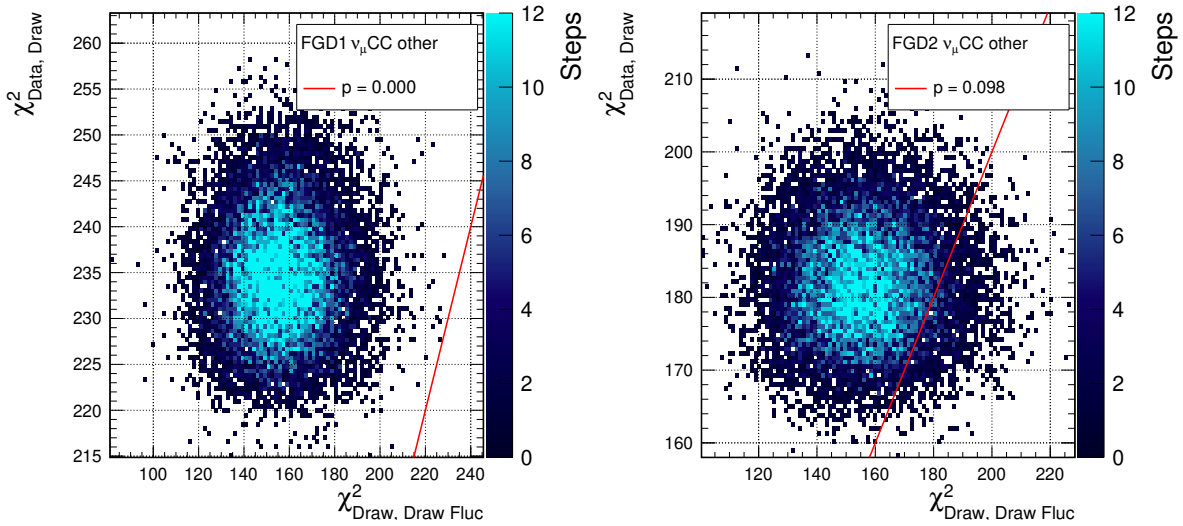


Figure 5.55.: Bin-by-bin likelihood contributions in $p_\mu \cos\theta_\mu$ for the CCOther selections

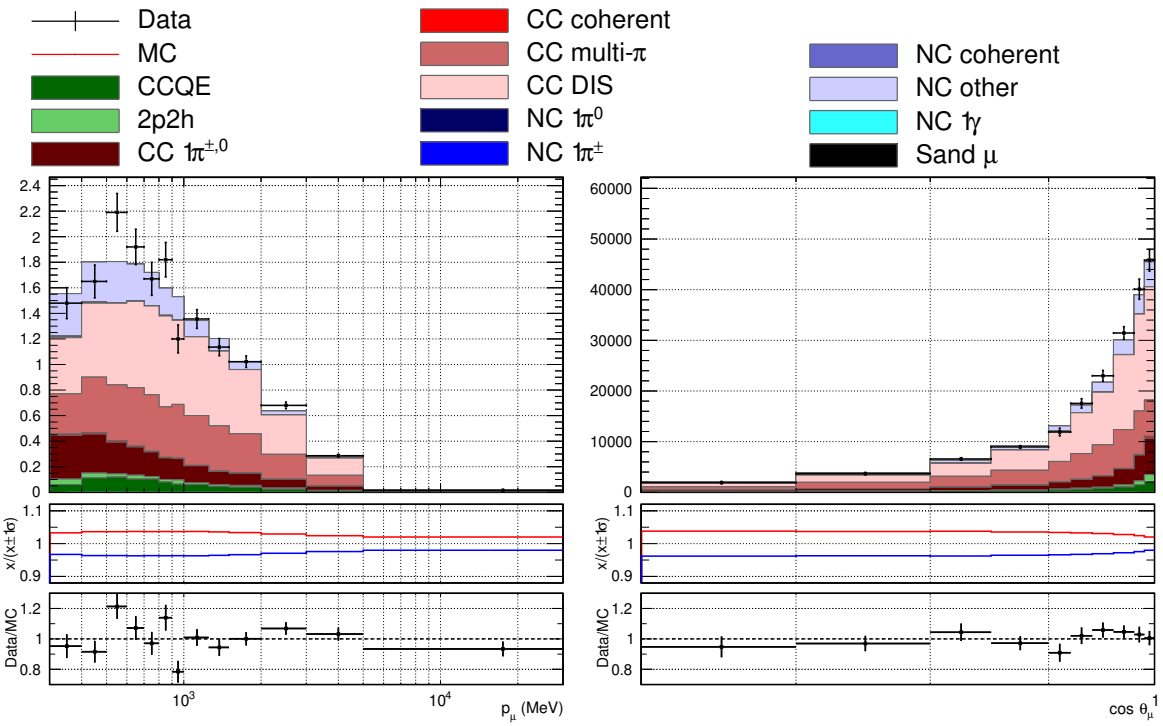


(a) FGD1 CCOther **(b)** FGD2 CCOther

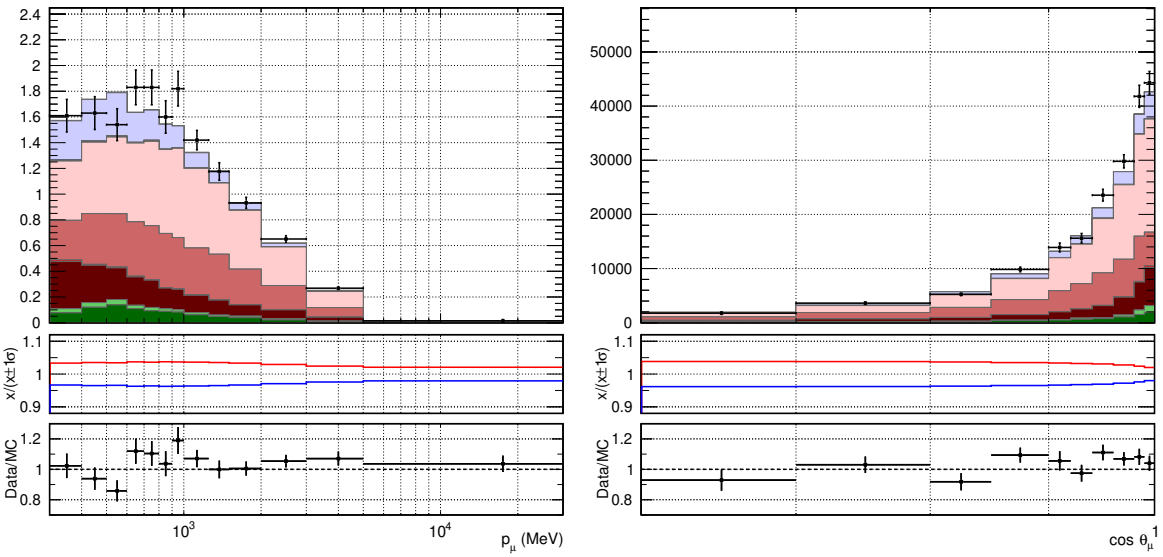
Figure 5.56.: Posterior predictive p-values for the two CCOther selections after the data fit

and FGD1 instead looks statistically fluctuated, slightly above and under 1σ from statistics. The CC DIS parameter is largely a normalisation up to 1.5 GeV. The distributions otherwise look very similar. The $\cos\theta_\mu$ projection of the two FGDs are consistent, with a mild difference at 0.8-0.92, where the bins show opposite behaviour, just outside 1σ. The CC DIS parameter has the smallest effect in the most forward-going bin.

Projecting the post-fit distributions onto Q^2_{rec} and E_ν^{rec} in Figure 5.58 the pattern becomes clearer: FGD1 has a consistent under-estimation in the $Q^2_{rec} = 0.15 - 0.4$ GeV^2 range, whereas FGD2 has a good prediction in all but one bin. Additionally, the CC DIS parameter is almost entirely a normalisation parameter in Q^2, so the fit has little



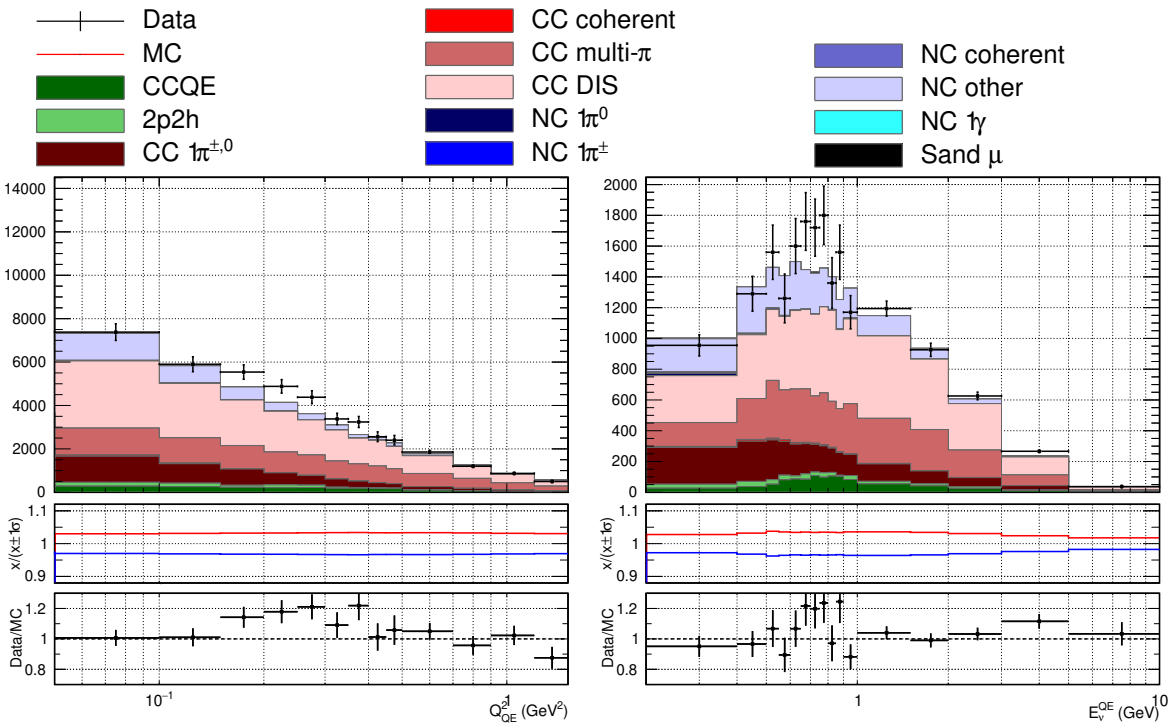
(a) FGD1 CCOther



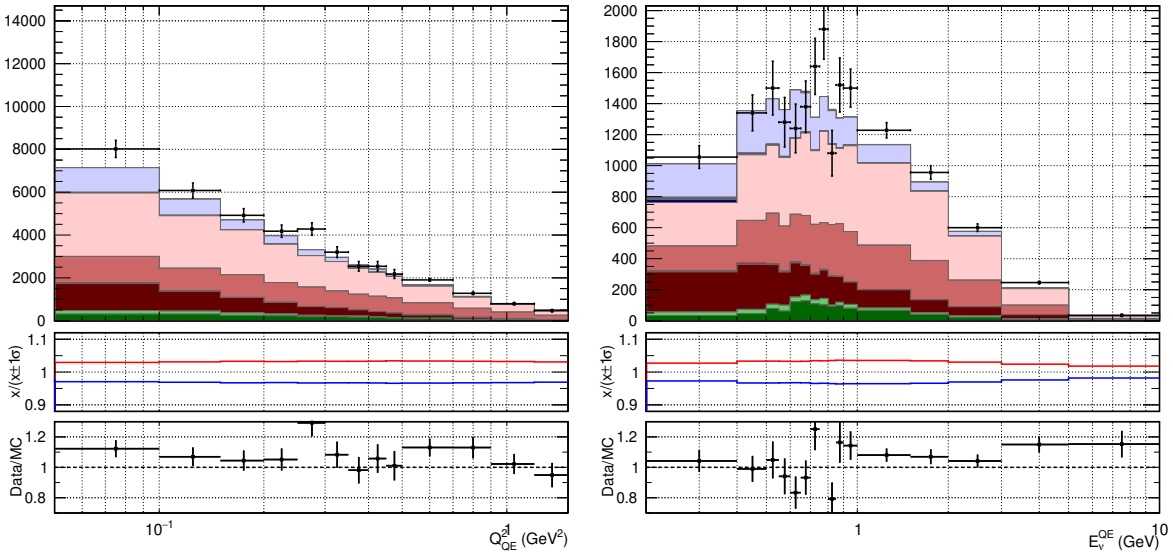
(b) FGD2 CCOther

Figure 5.57.: Post-fit distributions for the CCOther selections in p_μ and $\cos\theta_\mu$, showing the effect of the CC DIS parameter 1σ variation

freedom from this parameter to change the Q^2 shape. Looking at the E_ν^{rec} distribution there is consistency across the two FGDs, although FGD1 is more underestimated in the 0.7-0.9 GeV range, approximately within 1σ statistical uncertainty.



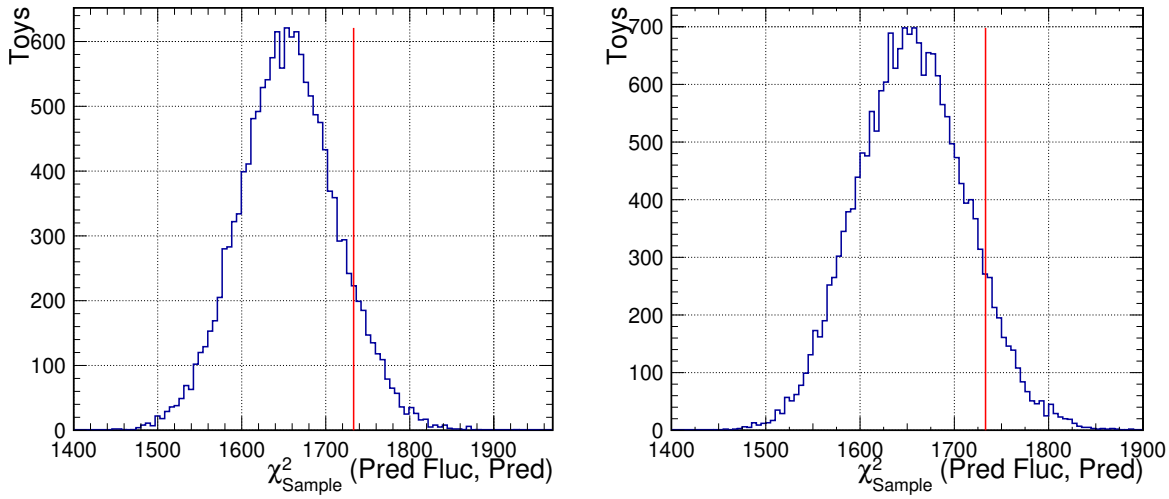
(a) FGD1 CCOther



(b) FGD2 CCOther

Figure 5.58.: Post-fit distributions for the CCOther selections in Q^2_{rec} and E^{rec}_ν, showing the effect of the CC DIS parameter 1σ variation

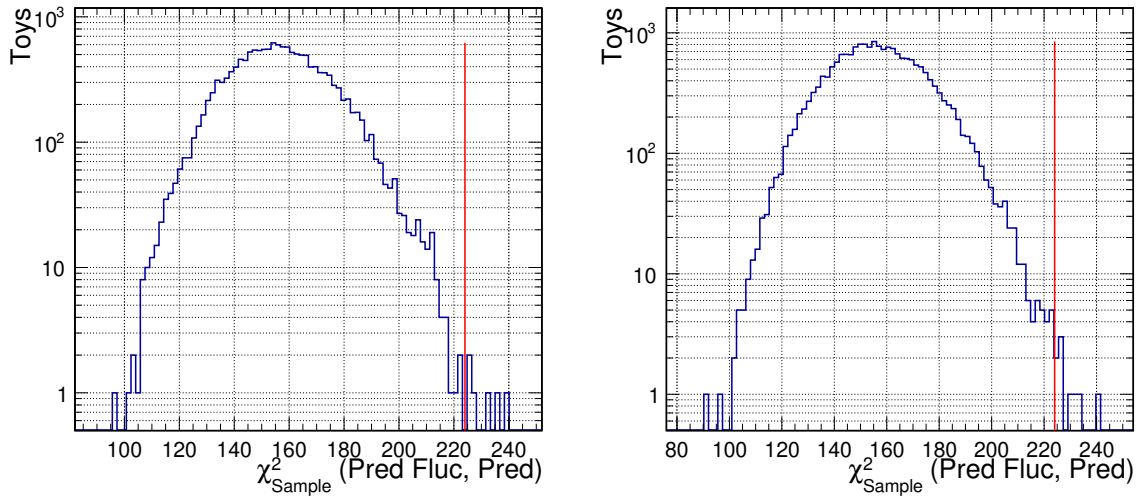
The one-dimensional p-values are shown in Figure 5.59, in which parameter sets from the posterior and prior probability distributions are taken as the reference distributions. The p-values are 10% for both methods.



(a) Posterior predictive, $p = 0.10$ **(b)** Prior predictive hybrid, $p = 0.10$

Figure 5.59.: One-dimensional p-value calculations, applying statistical fluctuations. Toys refer to toy experiments, which here are statistical fluctuations of the simulated distribution. The fraction above the red curve relative the total denotes the p-value. The red line is the realised p-value of the data-fit

The two-dimensional p-values for the FGD1 CCOther selection in Table 5.19 were very poor, and the one-dimensional equivalent in Figure 5.60 concludes similarly. The realised test-statistic falls at the very end of the reference distribution.

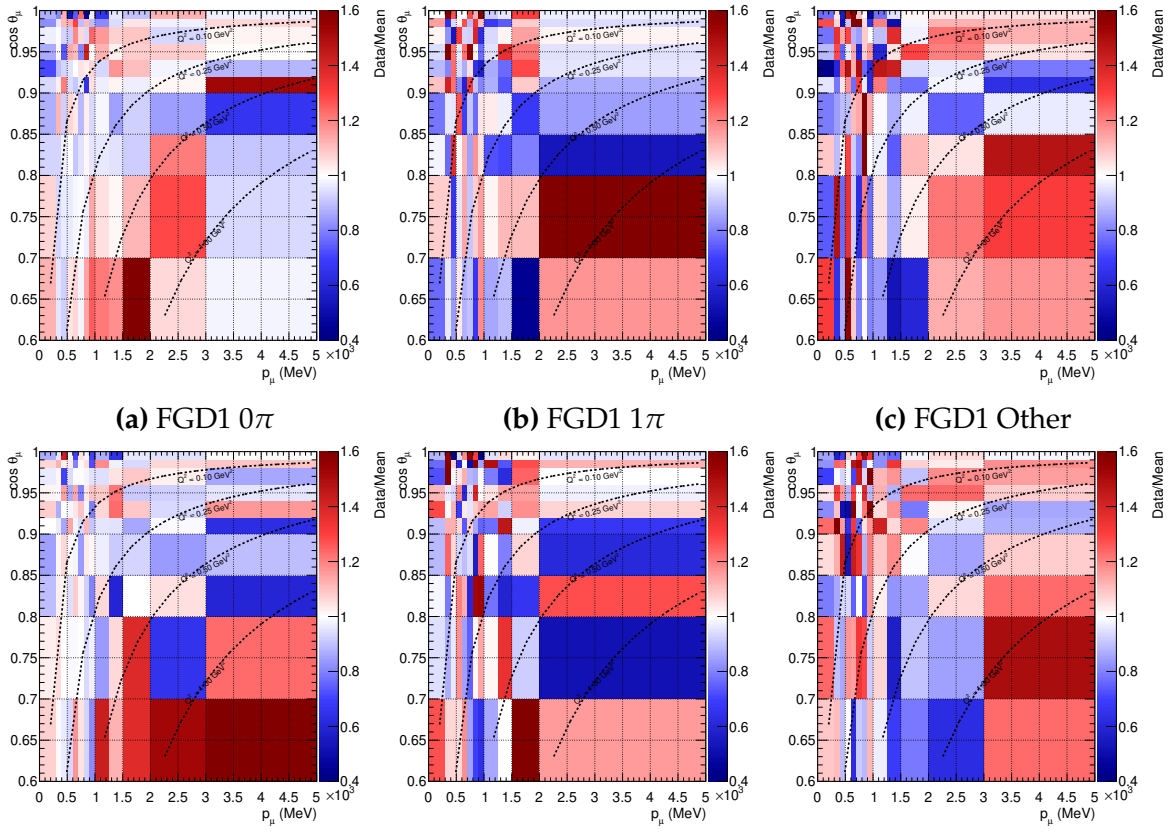


(a) Posterior predictive, $p = 0.00$ **(b)** Prior predictive hybrid, $p = 0.00$

Figure 5.60.: One-dimensional p-value calculations for FGD1 CCOther, applying statistical fluctuations. Toys refer to toy experiments, which here are statistical fluctuations of the simulated distribution. The fraction above the red curve relative the total denotes the p-value. The red line is the realised p-value of the data-fit

5.7.3. Post-fit Distributions

Figure 5.61 and Figure 5.62 shows all the selections' $p_\mu \cos\theta_\mu$ distributions post-fit, using the posterior predictive spectrum as a representation of the post-fit Monte-Carlo. Clearly, the post-fit Monte-Carlo does not describe all distributions, and there are plenty of discrepancies in all selections. The 0π and 1Trk selections generally see good predictions post-fit, especially around the flux peak. The 1π and Other selections are patchier and it's difficult to spot patterns in Q^2, p_μ or $\cos\theta_\mu$. Interestingly, the NTrk predictions are generally better than the 1π and Other selections.

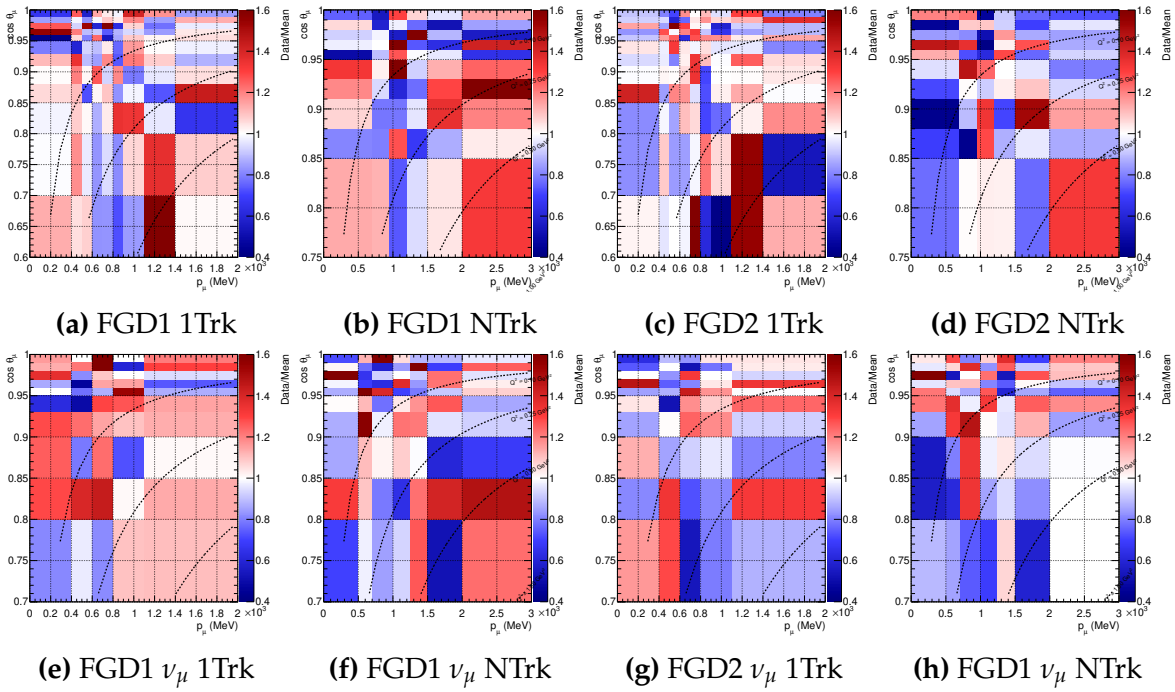


(a) FGD1 0π **(b)** FGD1 1π **(c)** FGD1 Other

(d) FGD2 0π **(e)** FGD2 1π **(f)** FGD2 Other

Figure 5.61.: Data to Posterior predictive $p_\mu \cos\theta_\mu$ spectrum ratios after the fit for FHC selections

To better understand the effect of the fit the distributions are projected onto p_μ and $\cos\theta_\mu$ before and after the fit. Figure 5.63 shows the distributions using the prior and posterior predictive spectrum for FGD1 and 2 CC0π and CC1π selections. The uncertainties are evaluated using the full prior and posterior distributions.



(a) FGD1 1Trk **(b)** FGD1 NTrk **(c)** FGD2 1Trk **(d)** FGD2 NTrk

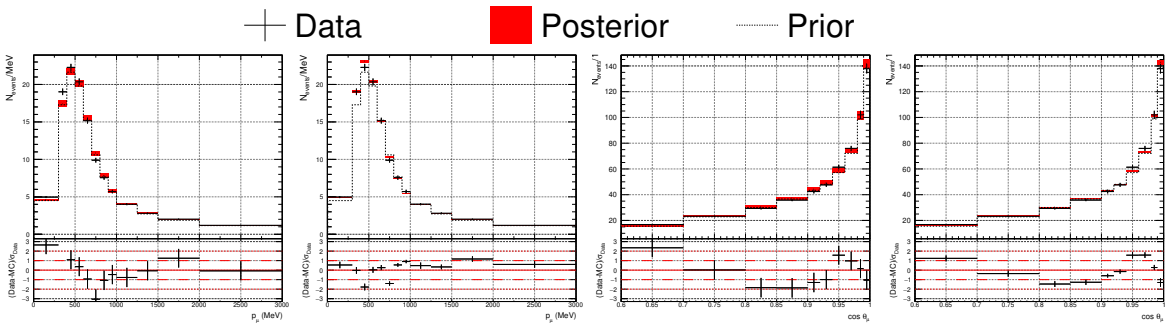
(e) FGD1 ν_μ 1Trk **(f)** FGD1 ν_μ NTrk **(g)** FGD2 ν_μ 1Trk **(h)** FGD1 ν_μ NTrk

Figure 5.62.: Data to Posterior predictive $p_\mu \cos\theta_\mu$ spectrum ratios after the fit for RHC selections

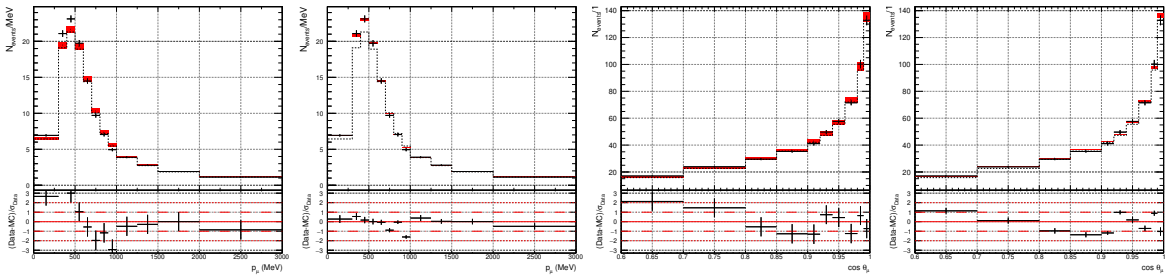
For CC0π, there is some tension between FGD1 and FGD2 in the third p_μ bin for both the prior and posterior predictive distributions where the simulation describes FGD1 well but undershoots FGD2. The post-fit distribution instead fits FGD2 well in this bin and overestimates FGD1. There is good improvement in the first two p_μ bins in the post-fit and a large reduction in the overall error. Moving to the $\cos\theta_\mu$ distributions, there is little change in the central values, where the primary effect appears to be reducing the error band. In $\cos\theta_\mu$ there is good agreement between FGD1 and FGD2.

For CC1π, there is acceptable description in p_μ before the fit except in the first bin. In the 500-1000 MeV region there is consistent over-estimation of the events which gets mostly corrected in the fit. The highest bin is well described for FGD1 and less so for FGD2. The $\cos\theta_\mu$ distributions pre-fit distribution is much worse than p_μ, notably in the 0.8-0.98 region and especially present for FGD1. In the post-fit this is mostly corrected, although the 0.8-0.92 region is approximately 1σ off.

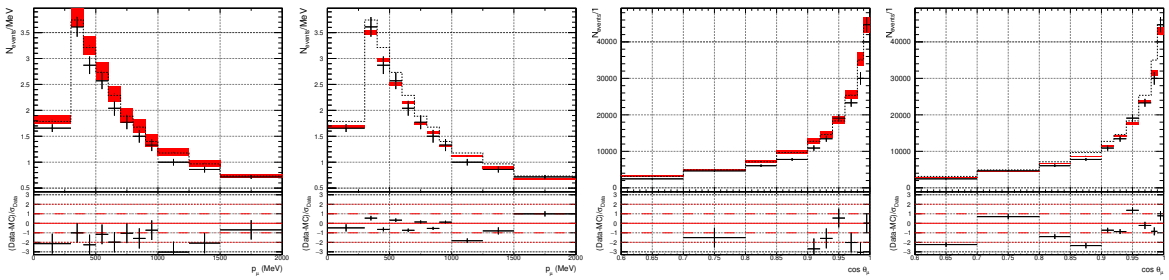
Figure 5.64 shows the projections for FGD1 and FGD2 CCOther before and after the fit. The p_μ distributions are consistent between FGD1 and FGD2, which is underestimated around the maximum. The fit attempts to correct for this but it often settles somewhere in between the two. The $\cos\theta_\mu$ distributions are good up until $\cos\theta_\mu$ =0.85,



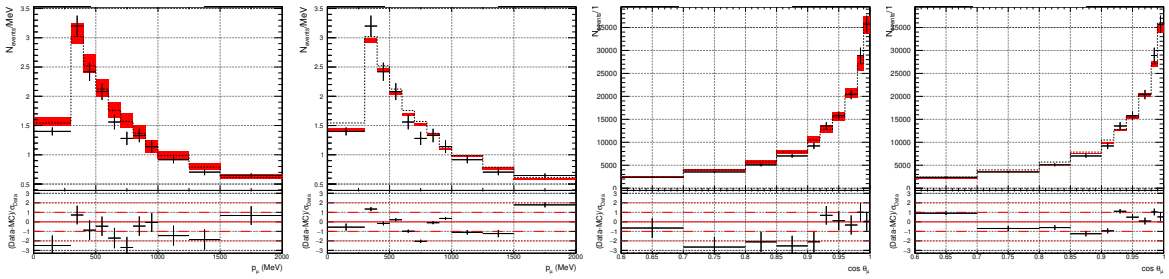
(a) FGD1 0π



(b) FGD2 0π



(c) FGD1 1π

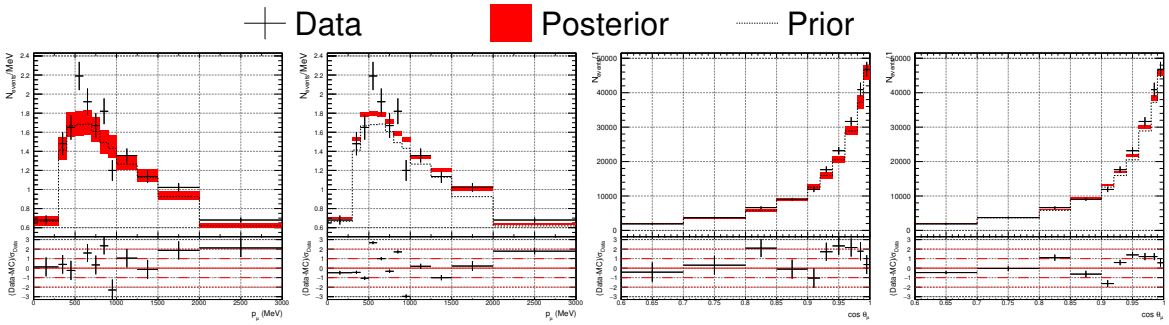


(d) FGD2 1π

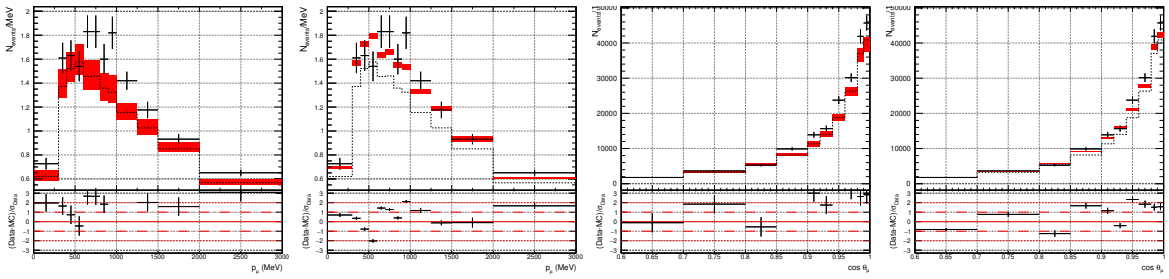
Figure 5.63.: FHC selections p_μ and $\cos\theta_\mu$ projections before and after fit

where the prefit starts to underestimate the data. This underestimation is present after the fit too, and it appears the freedom in $\cos\theta_\mu$ is not sufficient to cover the data.

Figure 5.65 shows the RHC $\bar{\nu}_\mu$ selections for FGD1 and FGD2. For CC1Trk there are differences between FGD1 and FGD2 in the highest bin (around 400MeV), which the simulation describes well for FGD2 but not FGD1. Generally the pre-fit is adequate and the post-fit mostly reduces the error band. The largest difference is around 700



(a) FGD1 Other



(b) FGD2 Other

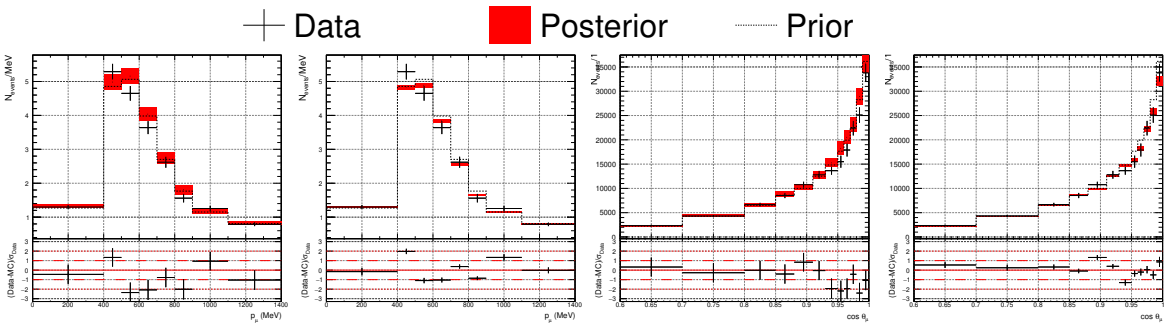
Figure 5.64.: FHC selections p_μ and $\cos\theta_\mu$ projections before and after fit

MeV for FGD2, which is well described for FGD1. For the $\cos\theta_\mu$ distributions, the pre-fit overestimates FGD1 but estimates FGD2 well. The post-fit instead under-estimates FGD2 slightly and estimates FGD1 well.

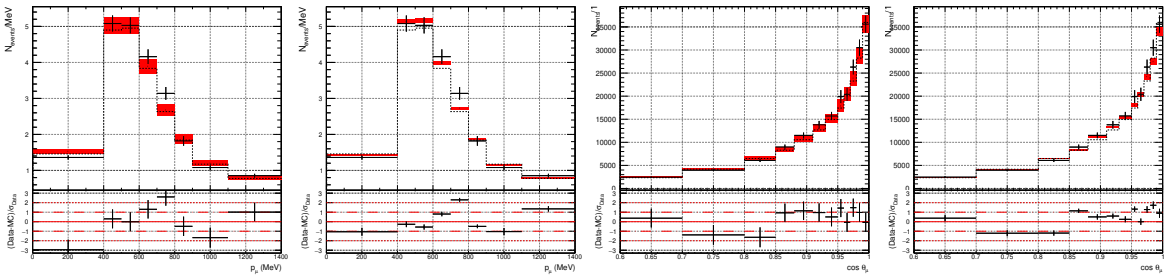
For the CCNTrk distributions the uncertainty from statistics are similar in error to the systematics and the prediction is generally good in p_μ pre-fit and post-fit; again the primary effect of the fit is to reduce the error rather than moving the central value. The $\cos\theta_\mu$ distribution is similarly well described pre-fit, although the second highest $\cos\theta_\mu$ bin is overestimated in FGD1 and the highest $\cos\theta_\mu$ bin is overestimated in FGD2. Post-fit, FGD1 is well described but FGD2 appears consistently over-estimated above $\cos\theta_\mu = 0.96$, and the highest $\cos\theta_\mu$ bin is poorly described.

Figure 5.66 shows FGD1 and FGD2 RHC ν_μ selections, which have comparably low statistics. The 1Trk distributions show consistency for FGD1 and FGD2. The only large discrepancy is found in the second last $\cos\theta_\mu$ bin for FGD1 and the last $\cos\theta_\mu$ bin for FGD2. The post-fit fails to correct for this and in general the fit minimises the error band rather than correcting the central values.

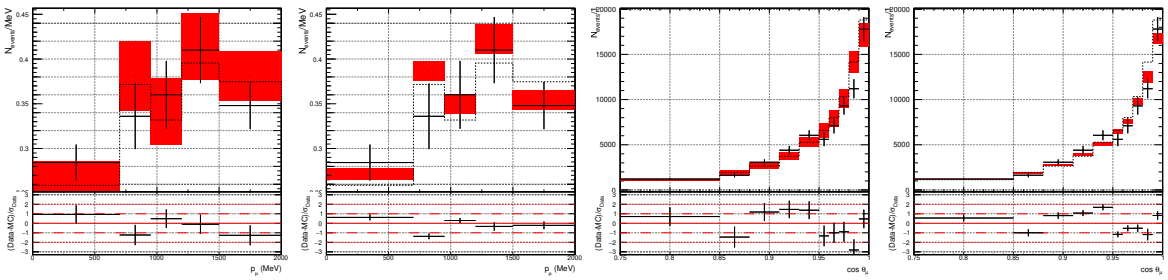
The NTrk distributions tell a similar story: the p_μ distributions are well-described by the prior model and are mildly improved by the post-fit model. For the $\cos\theta_\mu$



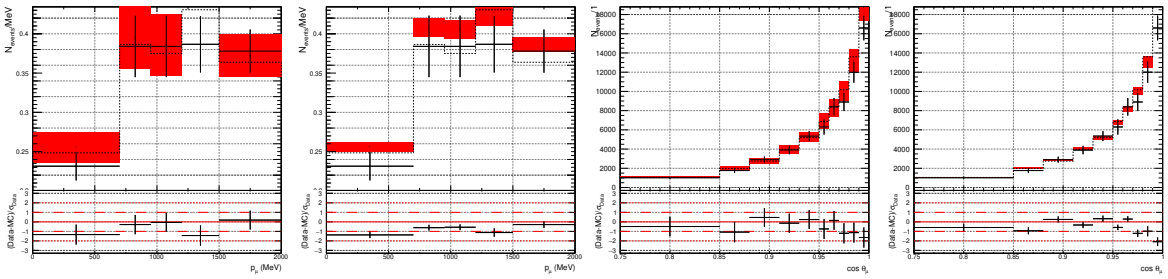
(a) FGD1 1Trk



(b) FGD2 1Trk



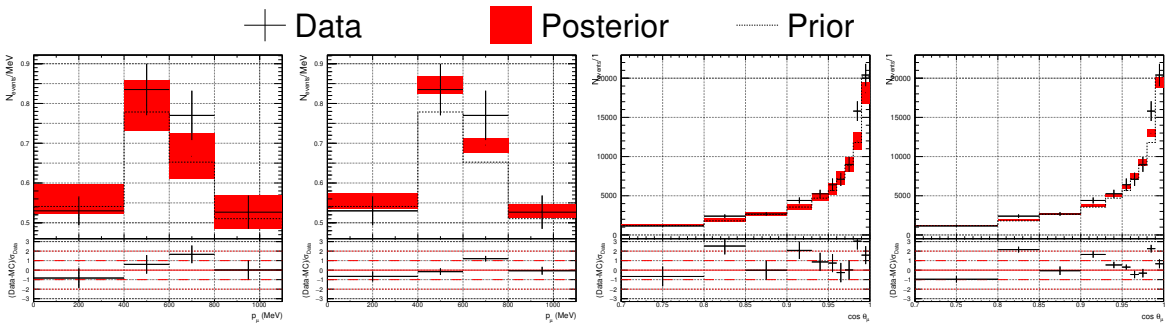
(c) FGD1 NTrk



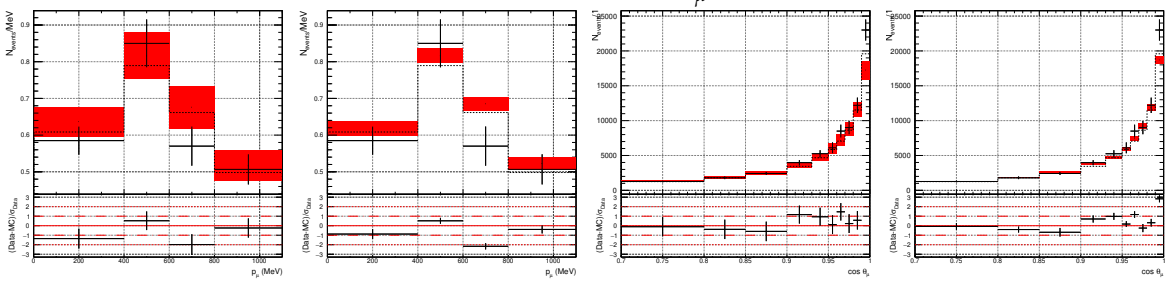
(d) FGD2 NTrk

Figure 5.65.: RHC $\bar{\nu}_\mu$ selections p_μ and $\cos\theta_\mu$ projections before and after fit

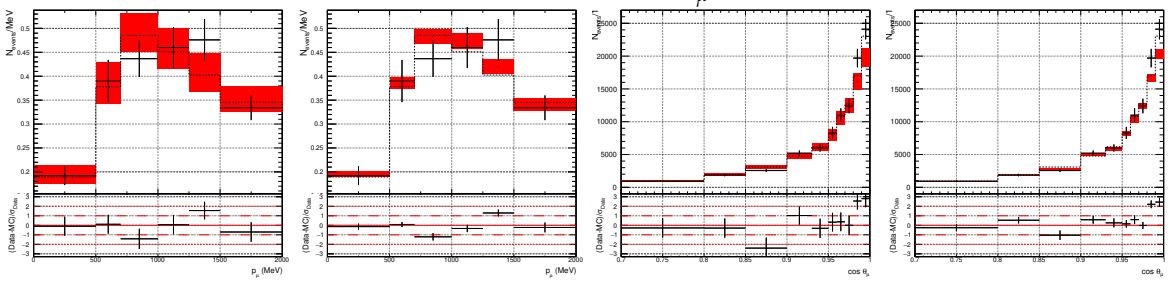
distribution the most forward-going bins are problematic for both FGD1 and FGD2, which the simulation underestimates.



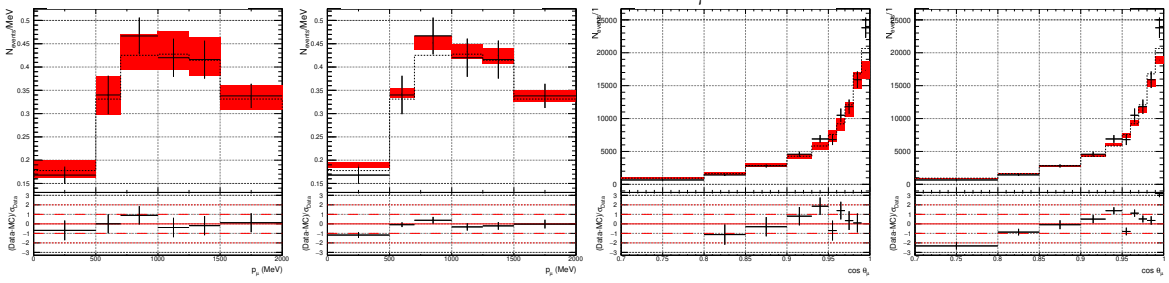
(a) FGD1 1Trk $\bar{\nu}_\mu$



(b) FGD2 1Trk $\bar{\nu}_\mu$



(c) FGD1 NTrk $\bar{\nu}_\mu$



(d) FGD2 NTrk $\bar{\nu}_\mu$

Figure 5.66.: RHC ν_μ selections p_μ and $\cos\theta_\mu$ projections before and after fit

5.7.4. Covariance Matrix from the Data Fit

Figure 5.67 shows the post-fit covariance matrix for the ND280 parameters, directly comparable to the Asimov covariance matrix in Figure 5.44. The bottom row shows the absolute difference in each matrix multiplied by the sign of each matrix. Blue entries have flipped sign, whereas red entries have kept their sign. The flux parameters

never flip signs and generally the covariance barely changes, and the correlation maximally changes by ~ 0.2. The only parameters that swap signs are the cross-section parameters. Although sometimes strong in the correlation plot, looking at the covariance plot the flip happens to parameters with weak constraints.

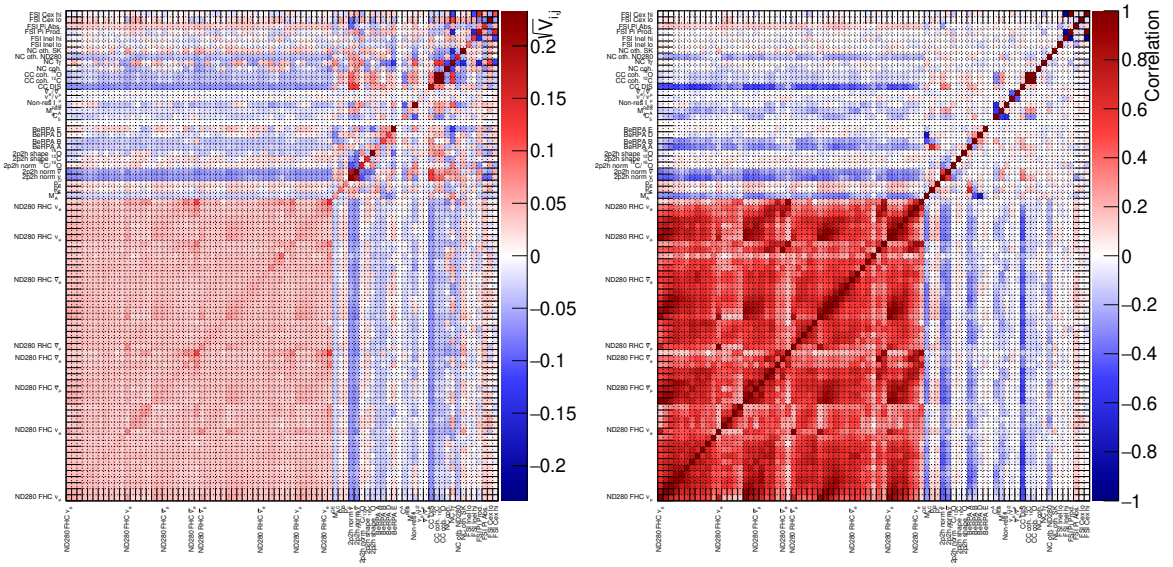
For the cross-section parameters there are stronger correlations between 2p2h shape C and 2p2h norm ν with the flux parameters. Whereas the 2p2h normalisation correlates stronger, the 2p2h shape parameter swaps sign frequently. This may be a result of the parameter being pushed against the boundary in the data fit. Interestingly, 2p2h norm ν correlates very differently with BeRPA A: the sign is swapped and the strength changes. This is largely expected: BeRPA A is pulled far from the value in the Asimov fit, so is entirely possible to correlate differently.

Looking at the correlations, there are strong relations between M_A^{QE}, 2p2h normalisations, BeRPA A, BeRPA B, C_5^A, CC DIS, NC other and the flux parameters. M_A^{QE} correlates strongly with BeRPA A, B and D as expected from their parametrisations[m]. The correlations with the flux parameters is due to ND280 data sitting in a relatively small E_ν, Q^2 region, correlating heavily with E_ν normalisations up to 1 GeV, in turn correlating with the other flux parameters due to the internal flux parameter correlations. The 2p2h normalisations contribute a relatively large portion of the uncertainty on the CC0π sample across E_ν, giving rise to the correlation. For the same reason, the C_5^A correlations enter due to the parameter controlling the normalisation of one of the interaction terms in the single pion model. Finally the CC DIS parameter is parametrised as $0.4/E_\nu$ for DIS events, so directly correlates to any other parameters that are dependent on E_ν, such as the flux parameters.

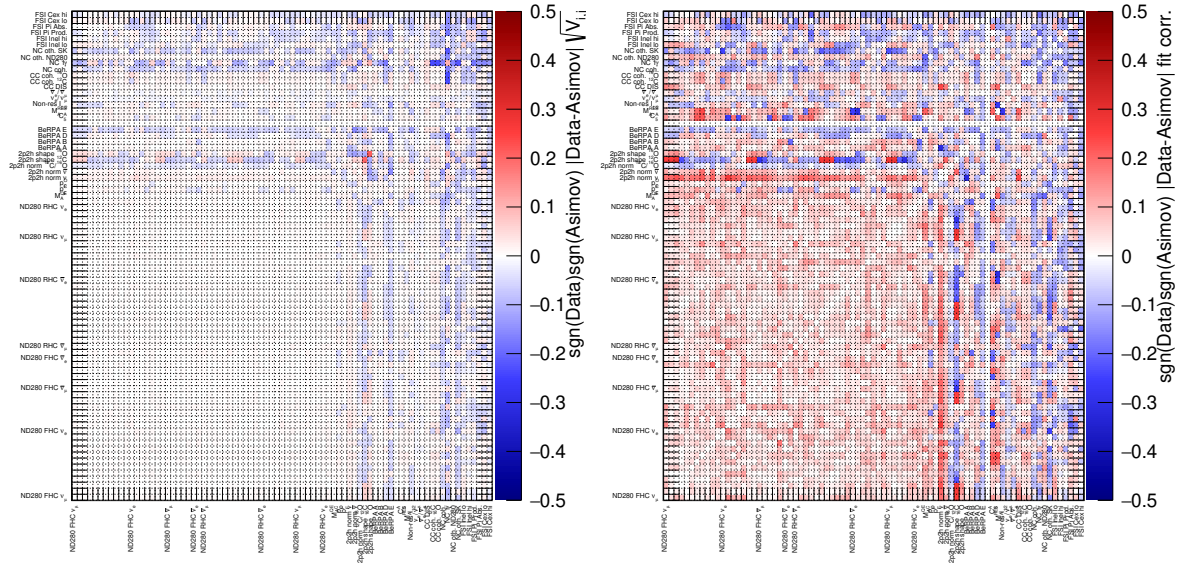
5.7.5. Alternate Model and Compatibility Studies

A number of alternative studies were performed for the 2017 analysis, comparing subruns in data and MC, change of model parameters and their priors, investigating effects of removing runs, and comparing FGD1 and FGD2. The largest impact was found in the FHC versus RHC data fit, and the FGD1 vs FGD2 data fit, with some parameters outside the 1σ range of the full fit to data. The posteriors from these two fits—and a fit using a 2015-like model with BeRPA and 2p2h shape fixed—were propagated to SK and event spectra comparisons were made. Generally, there were

[m] Both being approximately Q^2 shape variations for CCQE interactions



(a) $\sqrt{V_{i,j}}$ and correlation matrices



(b) $\sqrt{V_{i,j}}$ and correlation matrices for Asimov and data

Figure 5.67.: Post-fit covariance matrix for the data fit, showing ND280 related parameters

negligible ($< 1\sigma$) differences with little impact on oscillation analyses. These studies are found in Appendix F.

5.8. Cross-group Validations

Comparisons were made to the frequentist "BANFF" group to further validate the implementation and results of the ND280-only fits. The validation saw differences

to this analysis in all of the flux parameters, and some of the interaction parameters. All were due to marginalisation effects when providing the point estimates from the MCMC. The posterior predictive distributions from the MCMC compared to the "best-fit" distributions from the global minimum agreed well and the validation tests were passed satisfactorily. Appendix E details this procedure and the results

5.9. Impact on T2K Oscillation Analyses

The parameter set from the data fit is used to gauge the impact on the oscillated E^{ν}_{rec} spectrum at SK. The high dimensional posterior distribution sampled by the MCMC running on ND280 data, containing 50 flux and 23 interaction parameters, is propagated without assumptions on the shape of the underlying probability distribution function.

Using the oscillation parameters in Table 5.20, the posterior predictive distribution is formed by taking 10,000 random draws from the ND280-only MCMC running on data, presented in section 5.7. The oscillation parameters and SK detector parameters are fixed to their prior values and are not varied.

Parameter	Value
$\sin^2\theta_{12}$	0.304
$\sin^2\theta_{23}$	0.528
$\sin^2\theta_{13}$	0.0219
Δm^2_{12}	7.53×10^{-5} eV^2
Δm^2_{23}	2.509×10^{-3} eV^2
δ_{cp}	-1.601
POT FHC	1.47341×10^{21} (runs 2 to 8)
POT RHC	7.5573×10^{20} (runs 2 to 8)

Table 5.20.: Oscillation parameters used to produce nominal event rates at SK

The integrated event rates are shown in Table 5.21. As for the ND280 event rates in Table 5.17, the impact of the ND280 posterior compared to the prior is dramatic, reducing uncertainties by 60-80% for all SK selections.

Sample	Event rate		$\delta N/N$ (%)	
	Pre-fit	Post-fit	Pre-fit	Post-fit
1Rμ	249.86 ± 34.96	262.59 ± 8.03	13.99	3.06
1Re	65.62 ± 9.95	72.13 ± 2.88	15.16	3.99
1Re 1de	7.70 ± 0.93	6.73 ± 0.32	12.08	4.75
1Rμ RHC	61.50 ± 7.21	62.57 ± 1.73	11.72	2.76
1Re RHC	7.64 ± 0.95	7.72 ± 0.32	12.43	4.15

Table 5.21.: T2K-SK event rates and uncertainties from flux and interaction systematics with and without near-detector constraints from this analysis (not including SK and oscillation parameter errors)

The E^{ν}_{rec} spectrum using the prior and posterior distributions for each SK selection is shown in Figure 5.68. The 1Rμ distributions are consistent for FHC and RHC, in which there is an enhancement at low E_{rec} up until the oscillation dip, followed by a prediction much in agreement with the prior after $E^{\nu}_{rec} \sim 0.6$ GeV. This impacts Δm^2 and $\sin^2\theta_{23}$, since both the depth and slope of the dip changes.

The fit has little effect on the central value of the 1Re RHC selection but reduces the uncertainty by more than 50% above the oscillation dip. The two FHC 1Re selections show opposite behaviour: 1Re is enhanced throughout the spectrum, sitting approximately on the 1σ of the uncertainty band of the prior, with increasing agreement with the prior with increasing E^{ν}_{rec}. The 1Re1de selection is instead reduced throughout E^{ν}_{rec}, largely due to the ND280 1π selections being overestimated when using the prior values.

The impact of the alternate studies mentioned earlier and their impact on the SK prediction are detailed in Appendix F.

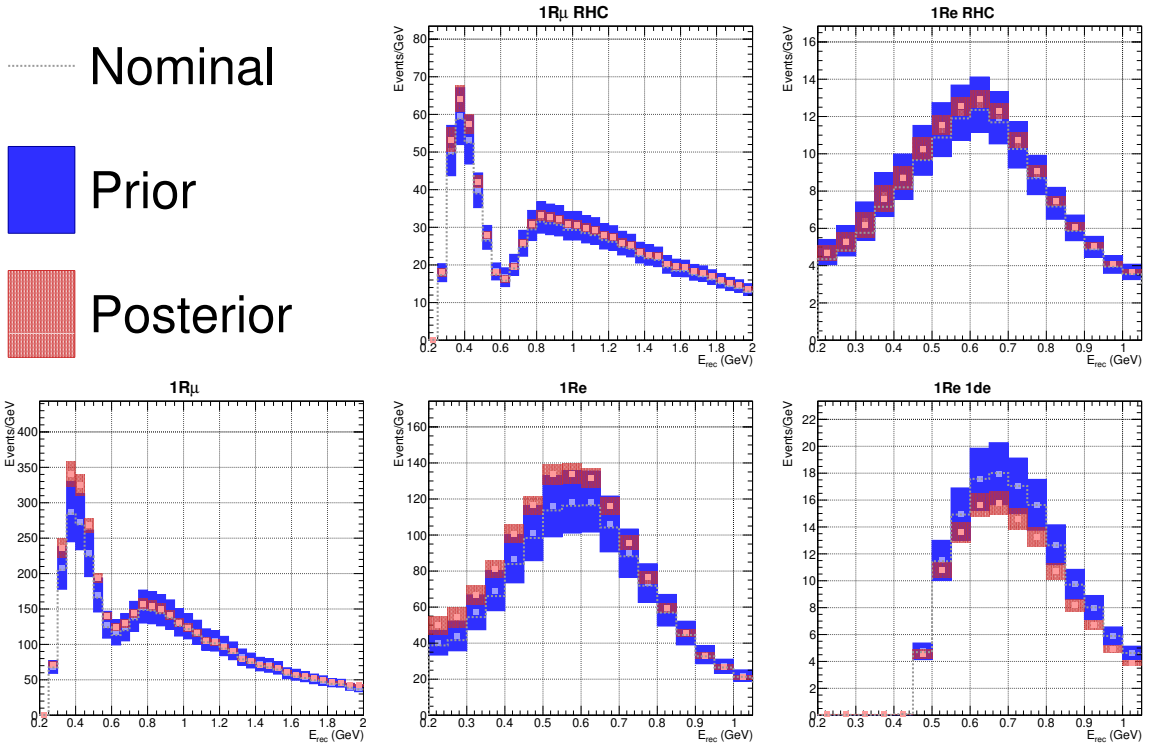


Figure 5.68.: Impact of the full fit on SK spectra compared to the prior

Chapter 6

Updating the Fit to ND280 Data for 2018 and Beyond

The official 2017 near detector data analysis presented in chapter 5 used data from T2K runs 2 through 6, which was collected between 2010 and 2015. It was largely developed for the 2015-2017 oscillation analyses [2, 108], with an inconsistent selection for ν_μ in FHC (of 0π, 1π and Other selection), and in RHC (1Track and NTrack for $\bar{\nu}_\mu$ and ν_μ). The coarse 1Track/NTrack selections and low statistics of the RHC samples led to an uninformative posterior predictive p-values—implying not much could be said about the quality of T2K anti-neutrino modelling. The upcoming 2018 analyses and beyond, presented in this chapter, sees a multitude of updates.

6.1. Adding Run 7 and 8 Data

T2K has been steadily collecting POT since 2015 on top of run 2 to 6 data, making it possible to refine the selections and update the binning for increased parameter sensitivity. As previously shown in Figure 3.6, almost twice the POT were accumulated in run 7 (RHC) and 8 (FHC) thanks to excellent beam performance, culminating at nearly 500 kW and 2E21 POT.

Table 6.1 shows the run-by-run POT breakdown of the data and generated Monte-Carlo, directly comparable to the 2017 analysis' equivalent for run 2 to 6 in Table 5.8. The amount of FHC POT increased by 99% and RHC by 63%.

6.2. Selections

The RHC selections are updated from 1Track/NTrack to 0π, 1π and Other to match the FHC selections. The FHC selections remain identical to what was presented in subsection 5.1.1. The pion counting for the RHC sample is the same as in subsection 5.1.1,

Run	POT (E+19)		
	Data	MC	Sand
2a	3.59337	92.3937	3.7132
2w	4.33765	120.341	4.00035
3b	2.1705	44.7864	2.35053
3c	13.6398	263.227	13.1337
4a	17.8271	349.96	17.4125
4w	16.4277	226.216	15.9801
5	4.3468	229.627	9.07403
6b	12.7301	141.74	25.9187
6c	5.07819	52.7562	10.4626
6d	7.75302	68.83	15.8059
6e	8.51429	85.9439	17.2691
7	24.3683	337.059	50.3961
8a	41.4909	363.054	40.1875
8w	15.8053	264.115	16.1263
Total FHC	115.29232	1724.0931	112.09418
Total RHC	62.791	915.9561	127.92643
Total	178.08332	2640.0492	241.83061
Total FHC x2017	1.9876	—	—
Total RHC x2017	1.6276	—	—
Total x2017	1.8438	—	—

Table 6.1.: Counted and generated protons-on-target for the T2K ND280 2018+ analysis

using either the TPC, FGD Michel electron or FGD isolated track reconstruction for the pion tags.

The only difference in the likelihood cuts from section 5.1 is for the μ^+ selection, which in subsection 5.1.2 required $0.1 < \mathcal{L}_\mu < 0.7$. The upper bound at 0.7 was present to reject low energy μ^- from ν_μ RHC interactions, which was discarded for this analysis after it gave a 4% increase in efficiency with negligible purity change. The μ^- criteria for the ν_μ in RHC selection did not change and still requires a MIP-like track with $0.1 < \mathcal{L}_\mu < 0.8$.

A summary of the 2018 analysis' selection efficiencies and purities is shown in Table 6.2. The FHC selections are near identical to the 2017 analysis, and there is

good agreement for the new 0π samples compared to the old 1 Track samples for RHC selections. The worst performing selection is the CCOther $\bar{\nu}_\mu$ in RHC selections, which have efficiencies of 40% and purities of 25%. For details and discussions, see Appendix B.

Selection	Efficiency (%)	Purity (%)
FGD1 CC0π ν_μ	93.7	75.4
FGD2 CC0π ν_μ	93.0	73.3
FGD1 CC1π ν_μ	83.4	57.3
FGD2 CC1π ν_μ	83.0	56.7
FGD1 CCOther ν_μ	73.0	64.9
FGD2 CCOther ν_μ	73.4	64.9
FGD1 CC0π $\bar{\nu}_\mu$	89.4	75.2
FGD2 CC0π $\bar{\nu}_\mu$	87.8	74.0
FGD1 CC1π $\bar{\nu}_\mu$	65.0	53.5
FGD2 CC1π $\bar{\nu}_\mu$	61.0	49.6
FGD1 CCOther $\bar{\nu}_\mu$	44.1	24.6
FGD2 CCOther $\bar{\nu}_\mu$	41.6	23.6
FGD1 CC0π ν_μ RHC	79.7	55.6
FGD2 CC0π ν_μ RHC	77.8	53.0
FGD1 CC1π ν_μ RHC	65.7	43.4
FGD2 CC1π ν_μ RHC	66.8	43.1
FGD1 CCOther ν_μ RHC	68.8	61.0
FGD2 CCOther ν_μ RHC	69.0	60.8

Table 6.2.: Efficiency and purity summary for all selections with the range $0 < p_{reco} < 3$ GeV/c, directly comparable to Table 5.2

6.3. Binning the Selections

With the large increase in statistics from run 7 and 8 data, a re-binning of the event observables p_μ and $\cos\theta_\mu$ was due. The binning is based on requiring ~ 20 raw MC events per bin, which is roughly equivalent to 1-2 data events. The approximate momentum resolution of ND280 is ~ 50 MeV and the angular resolution $\sim 2^\circ$, seen in Appendix H.

All selections for the 2018 analysis are rebinned using the above criteria, leading to a drastic increase in the number of bins from the 2017 analysis, which had 1624. The total number of bins is now 4238, of which 2942 are FHC (six selections) and 1188 are RHC (12 selections). The CC0π binning alone is more bins than was present in total for 2017 (1682 vs 1624).

- FGD1+2 CC0π: 841 fit bins
  p_μ (MeV/c) = 0, 200, 300, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 2000, 2500, 3000, 5000, 30000.
  $\cos\theta_\mu$ = -1, 0.5, 0.6, 0.7, 0.76, 0.78, 0.8, 0.83, 0.85, 0.88, 0.89, 0.9, 0.91, 0.92, 0.925, 0.93, 0.935, 0.94, 0.945, 0.95, 0.955, 0.96, 0.965, 0.97, 0.975, 0.98, 0.985, 0.99, 0.995, 1.

- FGD1+2 CC1π: 288 fit bins
  p_μ (MeV/c) = 0, 300, 350, 400, 500, 600, 650, 700, 750, 800, 900, 1000, 1100, 1200, 1500, 2000, 3000, 5000, 30000.
  $\cos\theta_\mu$ = -1, 0.6, 0.7, 0.8, 0.85, 0.88, 0.9, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1.

- FGD1+2 CCOther: 342 fit bins
  p_μ (MeV/c) = 0, 300, 400, 500, 600, 650, 700, 750, 800, 900, 1000, 1100, 1250, 1500, 1750, 2000, 3000, 5000, 30000.
  $\cos\theta_\mu$ = -1, 0.6, 0.7, 0.76, 0.8, 0.85, 0.88, 0.89, 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 1.

- FGD1+2 CC0π RHC: 306 fit bins
  p_μ (MeV/c) = 0, 300, 400, 500, 550, 600, 650, 700, 750, 800, 900, 1000, 1100, 1200, 1500, 2000, 4000, 30000.
  $\cos\theta_\mu$ = -1, 0.6, 0.7, 0.8, 0.85, 0.9, 0.92, 0.93, 0.94, 0.95, 0.96, 0.965, 0.97, 0.975, 0.98, 0.985, 0.99, 0.995, 1.

- FGD1+2 CC1π RHC: 48 fit bins
  p_μ (MeV/c) = 0, 500, 700, 900, 1300, 2500, 30000.
  $\cos\theta_\mu$ = -1, 0.7, 0.8, 0.9, 0.94, 0.96, 0.98, 0.99, 1

- FGD1+2 CCOther RHC: 80 fit bins
  p_μ (MeV/c) = 0, 600, 800, 1000, 1250, 1500, 2000, 4000, 30000.
  $\cos\theta_\mu$ = -1, 0.7, 0.8, 0.85, 0.9, 0.93, 0.95, 0.97, 0.98, 0.99, 1.

- FGD1+2 CC0π ν RHC: 120 fit bins
  p_μ (MeV/c) = 0, 300, 500, 700, 800, 900, 1250, 1500, 2000, 4000, 30000.
  $\cos\theta_\mu$ = -1, 0.7, 0.8, 0.85, 0.88, 0.9, 0.92, 0.94, 0.96, 0.97, 0.98, 0.99, 1.
- FGD1+2 CC1π ν RHC: 40 fit bins
  p_μ (MeV/c) = 0, 600, 800, 1500, 30000.
  $\cos\theta_\mu$ = -1, 0.7, 0.8, 0.86, 0.9, 0.94, 0.96, 0.97, 0.98, 0.99, 1
- FGD1+2 CCOther ν RHC: 54 fit bins
  p_μ (MeV/c) = 0, 600, 1000, 1250, 2000, 4000, 30000.
  $\cos\theta_\mu$ = -1, 0.7, 0.8, 0.86, 0.9, 0.93, 0.95, 0.97, 0.99, 1.

6.4. Systematics

The 2018 iteration of the fits to ND280 data was intended primarily as a statistics and reconstruction update, including run 7 and 8 data, and a suitable update to the RHC selections to match the FHC selections. There were therefore comparably few updates to the treatment of systematics.

6.4.1. The Beamline and Neutrino Flux

The flux systematics saw no change and follows that prescribed in the 2017 analysis, detailed in subsection 5.3.1.

6.4.2. The ND280 Detector

The same set of systematics in subsection 5.3.2 are used again in the 2018 analysis. One more systematic was added to the detector response: the proton secondary interaction systematic. It is analogous to the pion secondary interaction but applies to protons, so allows proton tracks to be modified in the reconstruction. Since the proton ID is excellent when $p < 1.4$ GeV, this systematic mostly affects anti-neutrino selections with $p > 1.4$ GeV, where there is a risk of the proton being assigned as the lepton candidate.

The parameterisation of the systematics is also unchanged to subsection 5.3.2—applying correlated normalisation parameters to each $p_\mu \cos\theta_\mu$ bin, merging similar responses when possible. A new ND280 covariance matrix was generated since the number of selections increased from 14 to 18 and the number of bins increased from 1624 to 4238. The number of bins increase by $\sim$2.6: the equivalent increase in detector parameters would bring the number of ND280 parameters to $\sim$1500, so a more aggressive bin merging was applied. A parameterisation of ND280 with identical covariance matrix binning to the analysis binning (4238 ND280 parameters) was also tested and is presented throughout this analysis as a reference.

The bin merging was approached in two ways: bins were merged if 1) there was a <5% difference in effect from ND280 systematics, or 2) the effect of the systematics was >5% but the bin content was less than one and the effect of the systematics on the bin was to change the number of events by less than one. This bin strategy merges bins with weak response to the ND280 systematics and bins of low statistics.

The number of final ND280 detector parameters decreased from 4238 to 1076 and favours finer binning for the high statistics FHC selections, as intended. The binning is:

- FGD1 and FGD2 CC0π: 272 detector bins (841)
  p_μ (GeV/c): 0, 200, 300, 400, 450, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1400, 5000, 30000
  $\cos\theta_\mu$: -1, 0.5, 0.6, 0.7, 0.76, 0.8, 0.83, 0.85, 0.88, 0.965, 0.97, 0.975, 0.98, 0.985, 0.99, 0.995, 1

- FGD1 and FGD2 CC1π: 110 detector bins (288)
  p_μ (GeV/c): 0, 300, 350, 400, 500, 600, 650, 700, 1100, 3000, 5000, 30000
  $\cos\theta_\mu$: -1, 0.6, 0.7, 0.8, 0.85, 0.88, 0.9, 0.92, 0.93, 0.94, 1

- FGD1 and FGD2 CCOther: 72 detector bins (342)
  p_μ (GeV/c): 0, 300, 400, 600, 650, 1750, 2000, 5000, 30000
  $\cos\theta_\mu$: -1, 0.6, 0.93, 0.94, 0.95, 0.96, 0.98, 0.99, 0.995, 1

- FGD1 and FGD2 CC0π RHC: 49 detector bins (306)
  p_μ (GeV/c): 0, 300, 400, 500, 550, 2000, 4000, 30000
  $\cos\theta_\mu$: -1, 0.6, 0.7, 0.8, 0.85, 0.9, 0.96, 1

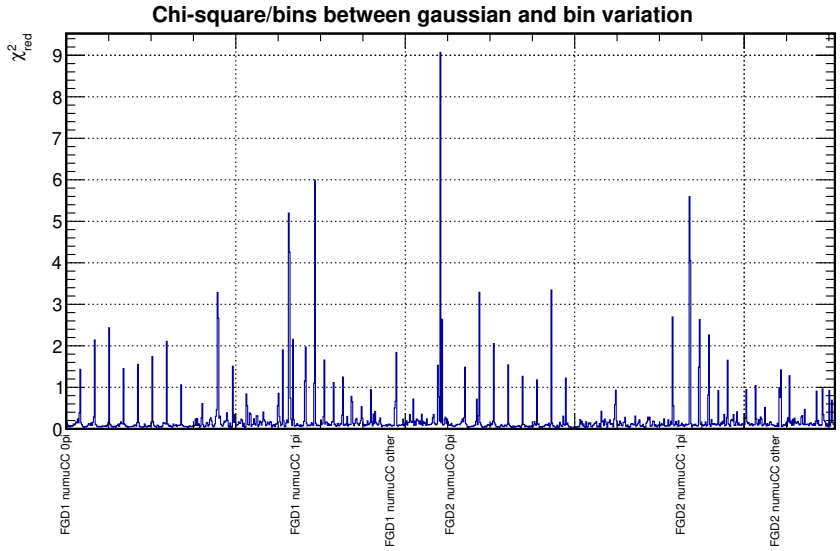
- FGD1 and FGD2 CC1π RHC: 4 detector bins (48)
  p_μ (GeV/c): 0, 500, 30000.
  $\cos\theta_\mu$: -1, 0.7, 1
- FGD1 and FGD2 CCOther RHC: 6 detector bins (80)
  p_μ (GeV/c): 0, 600, 800, 30000.
  $\cos\theta_\mu$: -1, 0.7, 1.
- FGD1 and FGD2 CC0π RHC ν: 15 detector bins (120)
  p_μ (GeV/c): 0, 300, 500, 700, 800, 30000.
  $\cos\theta_\mu$: -1, 0.7, 0.8, 1.
- FGD1 and FGD2 CC1π RHC ν: 6 detector bins (40)
  p_μ (GeV/c): 0, 600, 800, 30000.
  $\cos\theta_\mu$: -1, 0.7, 1
- FGD1 and FGD2 CCOther RHC ν: 4 detector bins (54)
  p_μ (GeV/c): 0, 600, 30000.
  $\cos\theta_\mu$: -1, 0.7, 1.

The underlying systematic parameters are varied and produce event distributions in each $p_\mu \cos\theta_\mu$ bin. The content is fit with a Gaussian and the χ^2/nbins for the fit is shown in Figure 6.1. There are clear outliers which follow a repeating pattern: the lowest momentum bin often has bimodal distributions due to pions being reconstructed or not, causing the entire event to enter and exit the selection. In most cases the difference in the 1σ between the distribution and the fitted Gaussian is < 0.1 events, although modelling the distribution as a Gaussian is not ideal.

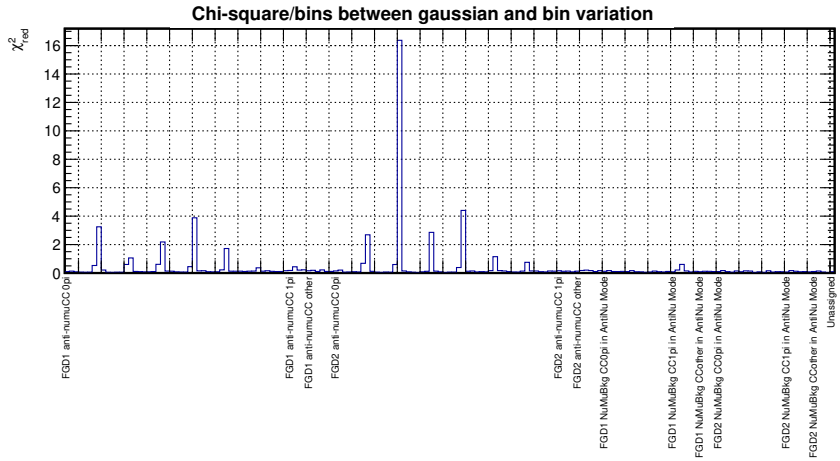
Figure 6.2 shows the event distributions for the three worst bins in Figure 6.1 (upper panel), which are all bimodal. All cases are low statistics so in reality have a low impact on the robustness of the analysis. The bottom panel shows a random selection of other bins which are better behaved.

The new ND280 covariance matrix is shown in Figure 6.3 and Figure 6.5[a]. There are many similarities with the 2017 matrices: the largest systematics are $\sim$50%, notably

[a]The matrices are not directly comparable to the 2017 matrices in Figure 5.17 and Figure 5.18 because the dominant parameter switched from p_μ to $\cos\theta_\mu$. Hence parameters now run $\{\cos\theta^i_\mu, p^j_\mu\} = \{\cos\theta^1_\mu, p^1_\mu\}, \{\cos\theta^1_\mu, p^2_\mu\}, ...\{\cos\theta^1_\mu, p^N_\mu\}$ instead of $\{\cos\theta^i_\mu, p^j_\mu\} = \{\cos\theta^1_\mu, p^1_\mu\}, \{\cos\theta^2_\mu, p^1_\mu\}, ...\{\cos\theta^N_\mu, p^1_\mu\}$



(a) FHC selections



(b) RHC selections

Figure 6.1.: χ^2/nbins for the reduced detector covariance matrix bins when fitting the bin's content distribution to a Gaussian

in the high-momentum bins. FGD1 and FGD2 are correlated as expected, and the selections in each FGD is correlated more.

6.4.3. The Neutrino-Matter Interaction

The interaction systematics are similar to subsection 5.3.3 with two changes: a new FSI tune was developed, and an uncertainty from radiative corrections was applied.

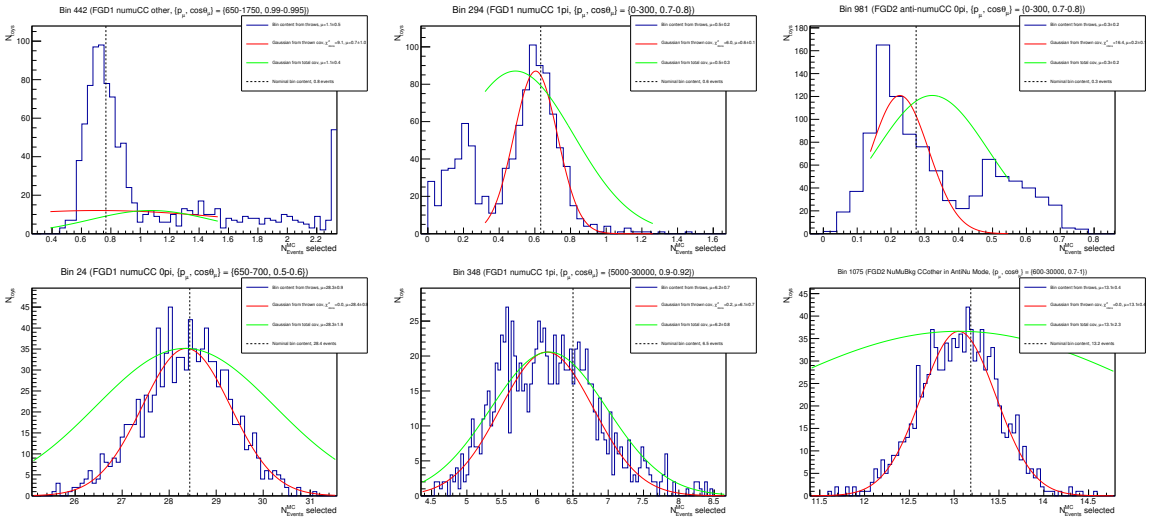


Figure 6.2.: Number of events in various bins used for parametrising the ND280 detector systematics, using the full suite of uncertainties from the ND280 systematics. Each histogram is representative of the uncertainty in the bin due to detector systematics, and the conservative uncertainties on the distribution are shown (red, green)

Pion Final State Interactions The new FSI tune [57] did not change the underlying model, but applied a more robust fitting method, extending the scattering data to heavier targets such as oxygen, aluminium, iron and lead, including new data from the dedicated DUET experiment [221]. It also removed the high energy charge exchange pion final state parameter. The new constraints are shown in Table 6.3, and is a reduction for the quasi-elastic and charge exchange parameters, but is otherwise an inflation. The inelastic and quasi-elastic high energy probabilities double in uncertainty. The absorption and charge exchange central values move by $\sim 1\sigma$ from the old recommendation. However, the correlations are now fully evaluated and the expectation is to reduce the systematics by about 50%.

Parameter	2017 value	Best-fit $\pm 1\sigma$
FEFQE	1.0 ± 0.41	1.069 ± 0.313
FEFABS	1.1 ± 0.41	1.404 ± 0.432
FEFCX	1.0 ± 0.57	0.697 ± 0.305
FEFINEL	1.0 ± 0.50	1.002 ± 1.101
FEFQEH	1.8 ± 0.34	1.824 ± 0.859

Table 6.3.: New pion final state interaction central values and uncertainties introduced for 2018 analyses

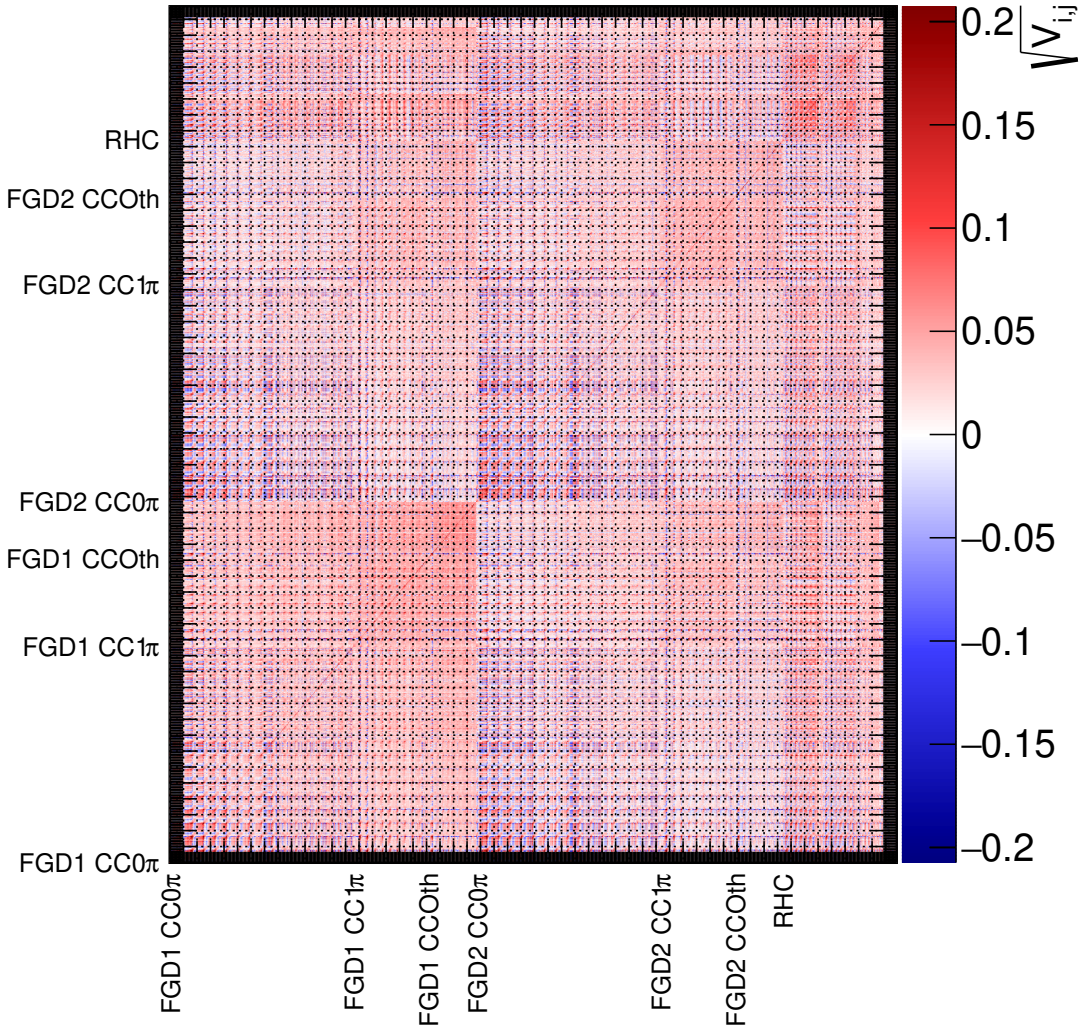
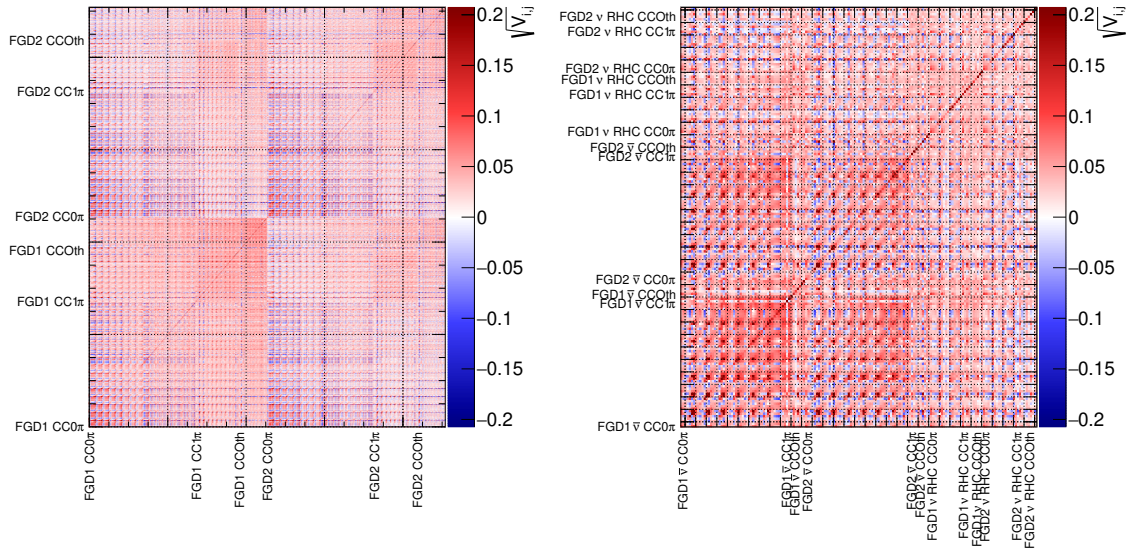


Figure 6.3.: $\sqrt{\text{Covariance}}$ for the reduced ND280 parameterisation, showing all selections. The covariance split by FHC and RHC are shown in Figure 6.4, also giving the order of RHC selections

Coulomb Corrections As the anti-neutrino statistics increased, a systematic uncertainty was included to account for the Coulomb effect of a lepton leaving the interaction nucleon and nucleus. μ^+/e^+ receive a repulsive force and μ^-/e^- an attractive force, and is simply modelled by shifting the true lepton momentum by a fixed amount, dependent on the target nucleus [222, 223]. The shifts are shown in Table 6.4. Since the relative effect of the Coulomb shift is smaller at higher momentum the systematic is only applied to CC inclusive ν and $\bar{\nu}$ with $0.4 < E_\nu < 0.6$ GeV. A correlated 3% total uncertainty (2% on ν, 1% on $\bar{\nu}$ negatively correlated with each other) is imposed.



(a) FHC selections **(b)** RHC selections

Figure 6.4.: $\sqrt{\text{Covariance}}$ for the reduced ND280 parameterisation, split into FHC and RHC selections

V_C (MeV)	μ^-	μ^+
^{12}C	-3.6	$+2.6$
^{16}O	-4.3	$+3.3$

Table 6.4.: Lepton momentum shifts as a result of Coulomb corrections

Covariance Matrix The new covariance and correlation matrices for the interaction systematics are presented in Figure 6.7. As previously noted, the pion FSI parameters (upper right corner) are now fully correlated which gives rise to smaller overall pion FSI systematics. The new CC normalisation from uncertainties in the Coulomb correction are visible in the middle of the matrix.

6.5. Nominal Model Prediction

Using the nominal model and applying the multiplicative nominal weights, the event rates in Table 6.7 are obtained.

As expected from the large increase in POT, run 7 and 8 almost doubles the data for the FHC and RHC selections. There are now 82,000 CC0πevents—67,000 FHC ν_μ, 13,000 RHC $\bar{\nu}_\mu$, and 5,000 RHC ν_μ. The FGD1 and FGD2 selections are consistent and

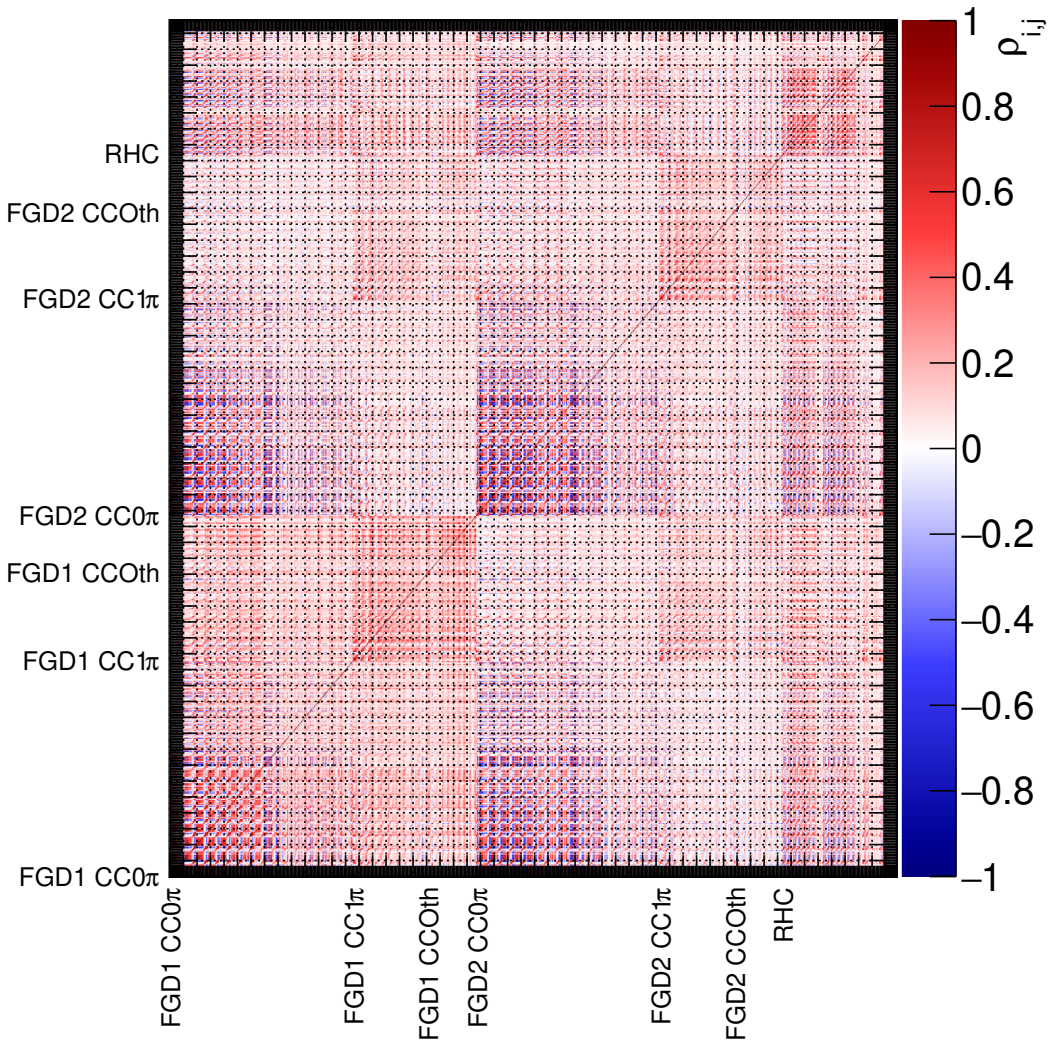
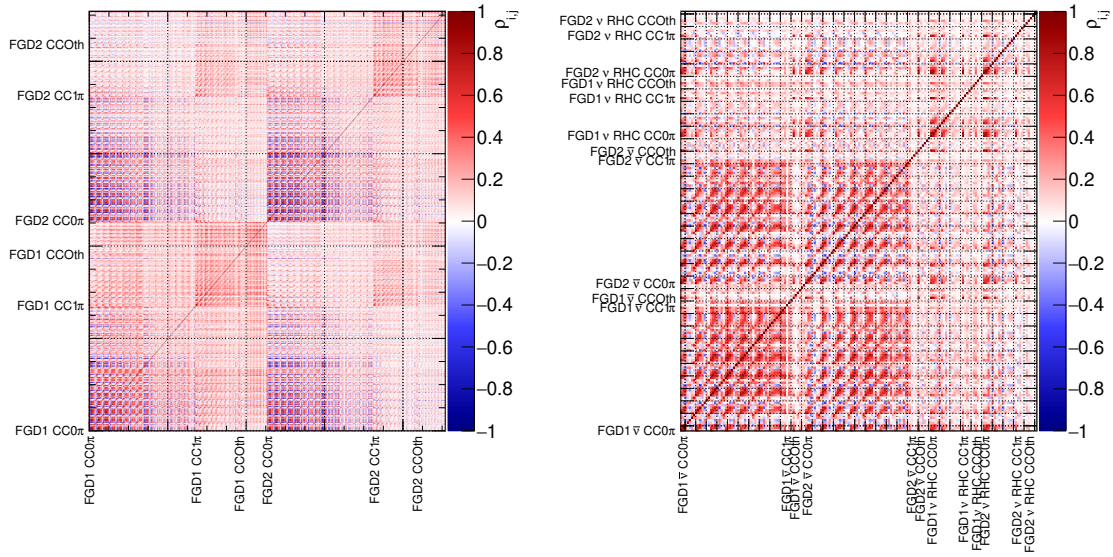


Figure 6.5.: Correlation for the reduced ND280 parameterisation, showing all selections. The correlation split by FHC and RHC are shown in Figure 6.6, also giving the order of RHC selections

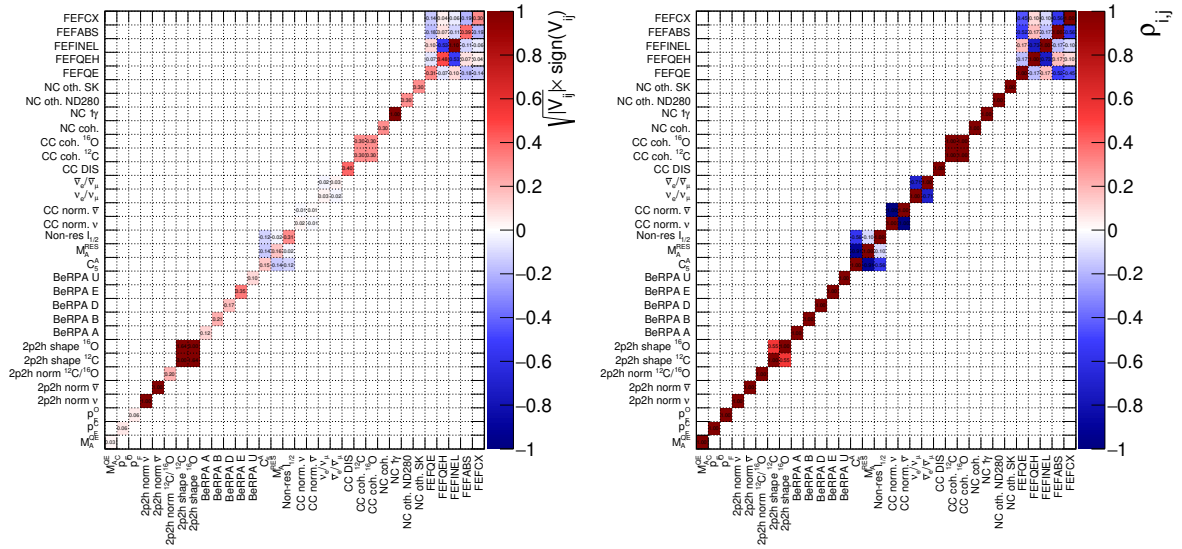
CC0π selections are generally underestimated by 0-6%, CC1π overestimated by the same amount, and CCOther underestimated by 9-26% for all FGDs and beam running modes. Averaging over all the selections, the nominal model underestimates the data by 6%.

The 2D $p_\mu \cos\theta_\mu$ distributions are found in Appendix D. The nominal 2D distributions are projected onto p_μ and $\cos\theta_\mu$ separately, showing their mode contributions in the nominal model. The distributions for FGD1 FHC ν_μ selections are shown in Figure 6.8. CC0π sees the same behaviour as in 2017 (Figure 5.25), with the low



(a) FHC selections **(b)** RHC selections

Figure 6.6.: Correlation for the reduced ND280 parameterisation, split into FHC and RHC selections



(a) $\sqrt{\text{Covariance}}$ **(b)** Correlation

Figure 6.7.: Covariance and correlation matrix for the interaction parameter priors in the 2018 analysis

momentum underestimated until the event peak at 500 MeV, which after 1 GeV is mostly well modelled. The CC1π distributions look marginally more consistent with data compared to 2017, but there is a consistent overestimation in $0.5 < p_\mu < 1$ GeV, with a slight underestimation at the event peak. The CCOther distributions are again

Sample	Data	Nominal MC	Data/MC
FGD1 0π	33553	31529.3	1.06
FGD1 1π	7757	7998.1	0.97
FGD1 other	8068	6793.68	1.18
FGD2 0π	33462	31734	1.05
FGD2 1π	6133	6419.04	0.96
FGD2 other	7664	6562.75	1.17
FGD1 $\bar{\nu}_\mu$ 0π	6368	6371.34	1.00
FGD1 $\bar{\nu}_\mu$ 1π	535	533.253	1.00
FGD1 $\bar{\nu}_\mu$ other	1102	1023.36	1.08
FGD2 $\bar{\nu}_\mu$ 0π	6451	6283.35	1.03
FGD2 $\bar{\nu}_\mu$ 1π	465	483.508	0.96
FGD2 $\bar{\nu}_\mu$ other	1032	943.956	1.09
FGD1 ν_μ RHC 0π	2707	2485.51	1.09
FGD1 ν_μ RHC 1π	847	855.911	0.99
FGD1 ν_μ RHC other	1015	804.647	1.26
FGD2 ν_μ RHC 0π	2648	2553.51	1.04
FGD2 ν_μ RHC 1π	693	679.99	1.02
FGD2 ν_μ RHC other	932	792.166	1.18
Total	121432	114847	1.06
Total x2017	1.87	1.80	

Table 6.5.: Observed and predicted event counts for the different ND280 selections for the 2018 analysis

grossly underestimated throughout p_μ between 10 and 20%. FGD2 looks particularly like a normalisation effect.

The new RHC $\bar{\nu}_\mu$ distributions' projections are shown in Figure 6.9, where the 0π selection shows a similar pattern to the FHC ν_μ projections and the 1Trk distributions from 2017. Namely, there is an underestimation at low p_μ, which goes to an overestimation at ~ 0.5 GeV, which returns to a satisfactory prediction above 1 GeV. FGD1 appears to see this more than FGD2, as was the case in 2017. For the 1π selections, the prediction is mostly within 1σ of the data—excluding systematic errors on the Monte Carlo. For the CCOther distributions there are hints of a consistent underestimation for both FGDs—particularly at lower momentum—but the majority of the prediction is still within statistical error of the data.

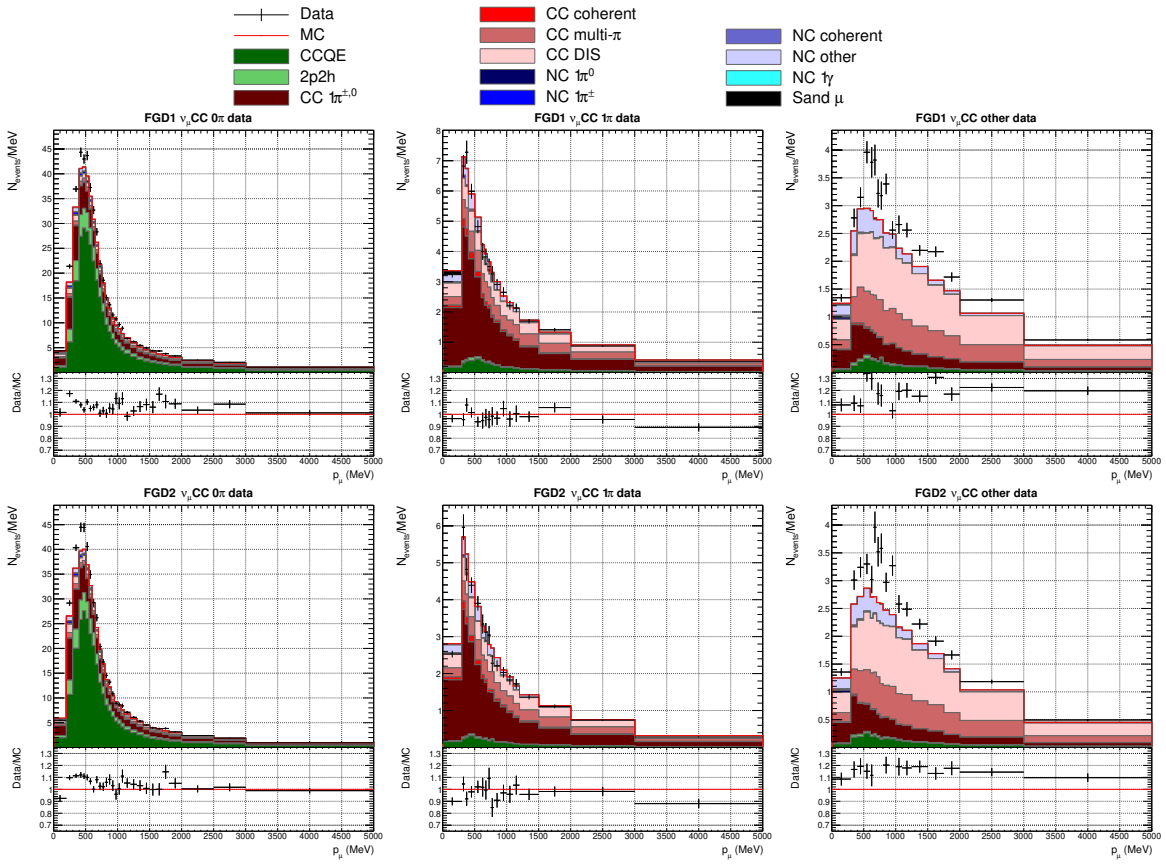


Figure 6.8.: Data and nominal MC distributions for FHC ν_μ selections projected onto p_μ, showing contributions by interaction mode. Bin content is normalised to bin width

Turning to the p_μ projections of the RHC ν_μ distributions in Figure 6.10, there is consistent underestimation of the FGD1 CC0π selection, with FGD2 overestimating the lowest and two adjacent bins at 700 MeV. The CC1π distributions are compatible across the two FGDs, where the low momentum bin is slightly overestimated and the two next bins underestimated. The CCQE and 2p2h contributions are highest at low momentum for the ν_μ selection, and barely present for the $\bar{\nu}_\mu$ selection. The CCOther selection—which contains more $\bar{\nu}_\mu$ CCOther than ν_μ CCOther—verifies the consistent picture of grossly underestimating the data, constant with p_μ at 10-25%.

Turning attention to the $\cos\theta_\mu$ projections of the FHC ν_μ selections in Figure 6.11, the additional data and bins compared to 2017 highlights the underestimation of much of the high-angle data until $\cos\theta_\mu \sim 0.85$ for both FGDs. The overestimation is present in almost every bin above that and outside the statistical error of the data. The small overestimation from 2017 when $0.8 < \cos\theta_\mu < 0.93$ appears gone. The 1π selections repeat the wavy Data/MC pattern of 2017 in both FGDs, with overestimates at high-angles and underestimates the most forward going angles. As was the case for

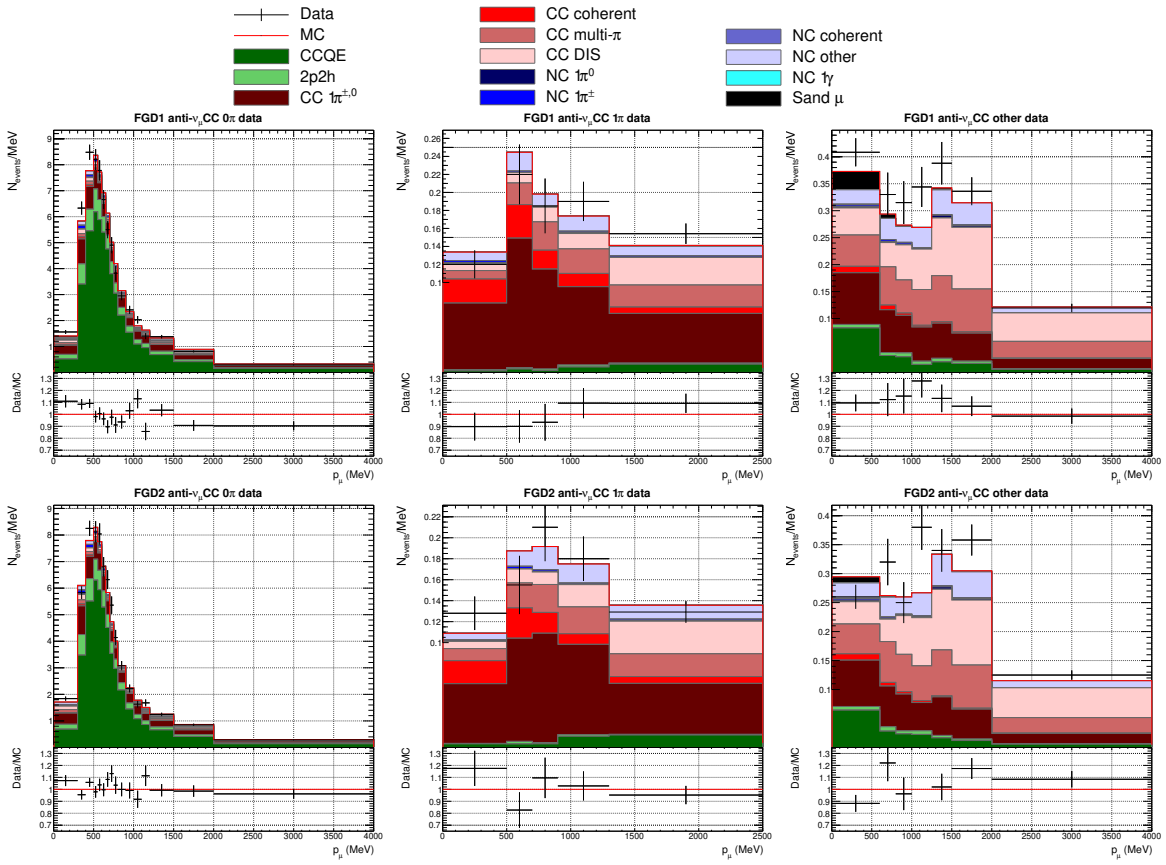


Figure 6.9.: Data and nominal MC distributions for RHC $\bar{\nu}_\mu$ selections projected onto p_μ, showing contributions by interaction mode. Bin content is normalised to bin width

the p_μ projections, the CCOther $\cos\theta_\mu$ distributions are underestimated throughout $\cos\theta_\mu$, between 10-30%.

The RHC $\bar{\nu}_\mu$ selections' $\cos\theta_\mu$ projections are shown in Figure 6.12, where the CC0π echoes the ν_μ equivalent: underestimation at high angles up to $\cos\theta_\mu = 0.85$, above which FGD1 oscillates between over and under estimate, and FGD2 consistently underestimates at 10%, just within the statistical error of the data. As for the p_μ projection, the 1π selections are mostly within statistical error except for the most forward-going bin, which sees an underestimate for FGD1 and an overestimate for FGD2. The CCOther distributions are both almost consistently underestimated with the exception of three bins in total, although each of the three are within statistical error of the data.

Finally the RHC ν_μ $\cos\theta_\mu$ projections are shown in Figure 6.13, where we note a large 1π and multi-π contribution to the 0π selection, especially in the forward-going region. The data is underestimated above $\cos\theta_\mu = 0.9$, similar to what was seen in the RHC $\bar{\nu}_\mu$ and FHC ν_μ distributions, compatible with 2017. The 1π selection is harder to

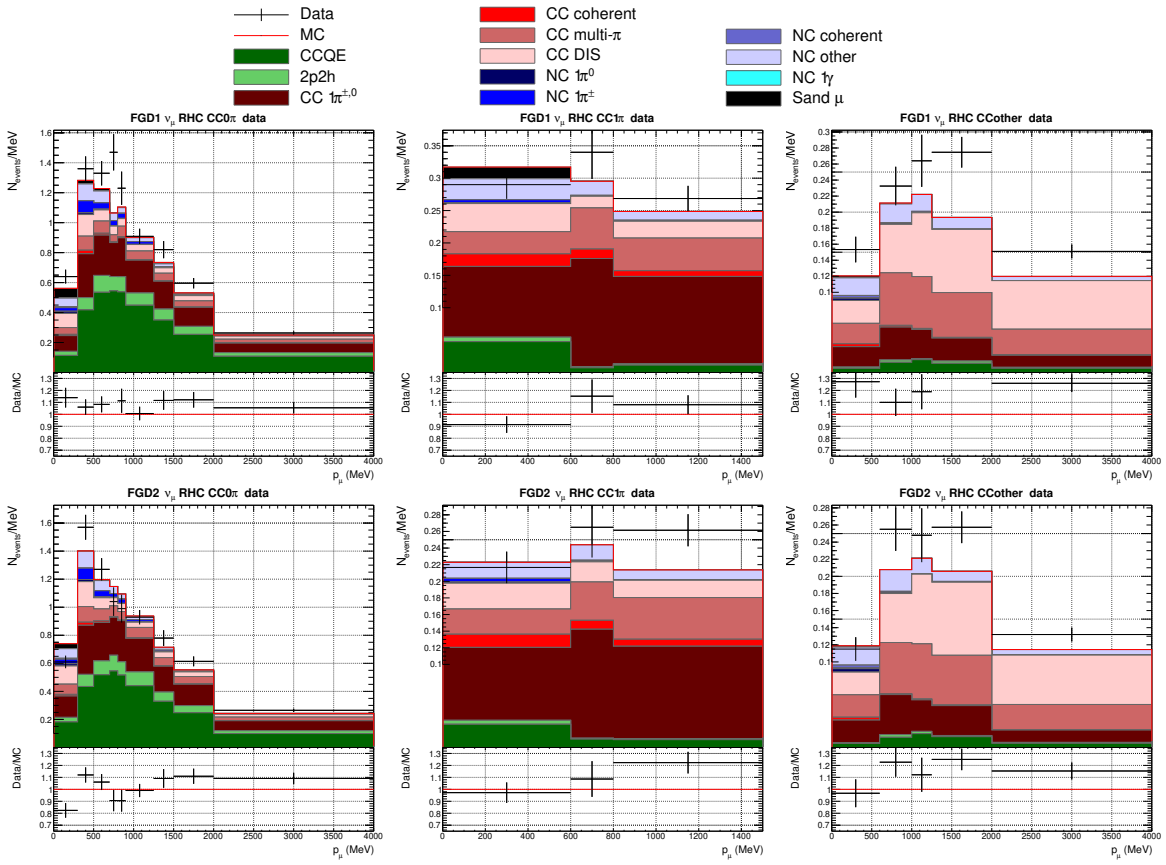


Figure 6.10.: Data and nominal MC distributions for RHC ν_μ selections projected onto p_μ, showing contributions by interaction mode. Bin content is normalised to bin width

draw conclusions from, although both FGDs are mostly consistent in underestimating the most forward-going bins and overestimating the high-angle and backwards. The CCOther distributions are again consistent with the other CCOther selections: constant underestimation of the data between 10-30%.

The summary of each mode contribution to the selections is shown in Table 6.6. They are in general very similar to the 2017 results for FHC ν_μ selections. The RHC $\bar{\nu}_\mu$ 0π selection has 3% less CCQE events, likely due to the change in the muon likelihood cut, and has a larger contamination of 1π, multi-π and DIS events. The NC contribution is largest for the new RHC $\bar{\nu}_\mu$ CCOther selections at 13.5%. Generally, the selections perform satisfactorily at targeting the appropriate interaction modes, reaching above 55% throughout.

Table 6.7 shows the effect on the overall event rate in applying the different classes of weights.

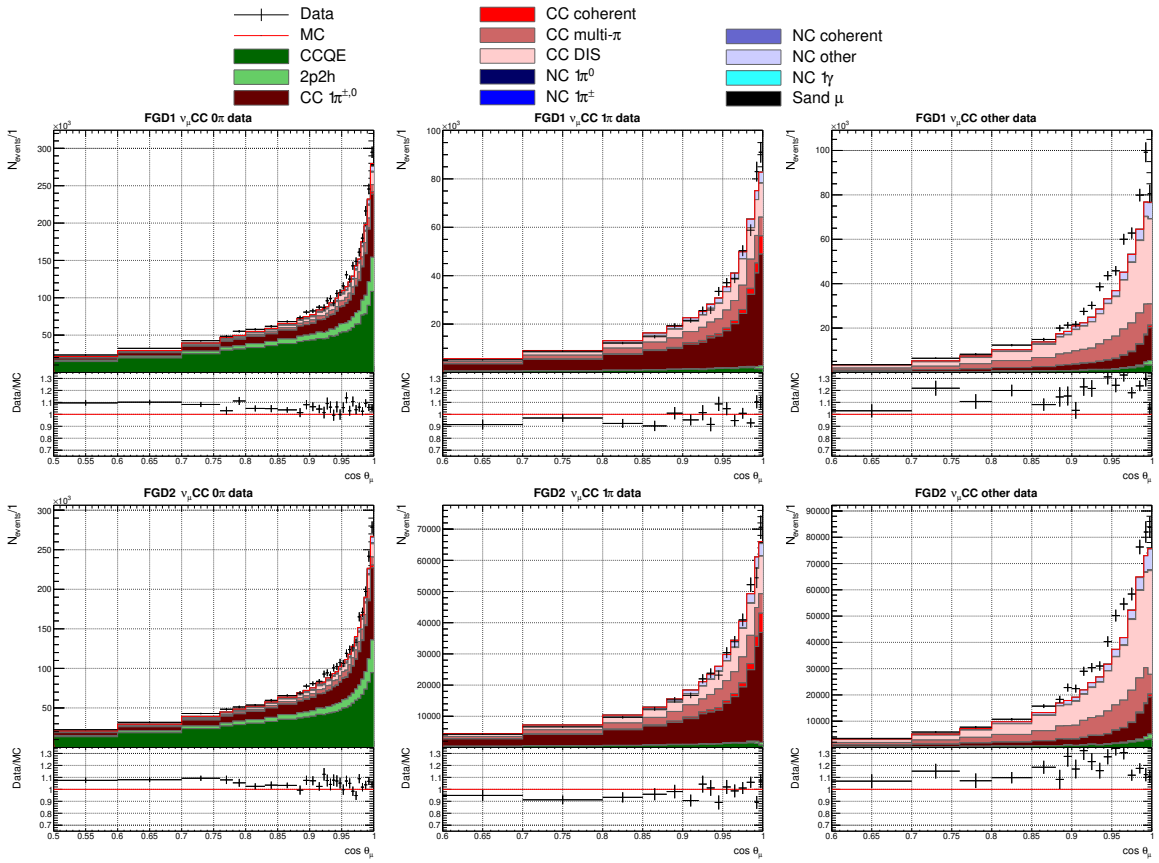


Figure 6.11.: Data and nominal MC distributions for FHC ν_μ selections projected onto $\cos\theta_\mu$, showing contributions by interaction mode. Bin content is normalised to bin width

6.6. Fitting Asimov Data

In this section the nominal model predictions outlined above are set to be the data. Closure tests and expected sensitivity studies are made, as was done in section 5.6 for the 2017 analysis.

6.6.1. Log-Likelihood Scans

Since new selections, double the data and finer binning have been added in the 2018 analysis, relatively large increases in sensitivity to many parameters are expected in the likelihood scans.

Figure 6.14 shows the same beam parameters as Figure 5.27. There is a significantly stronger constraint on the sample likelihoods, with no change to the prior as expected.

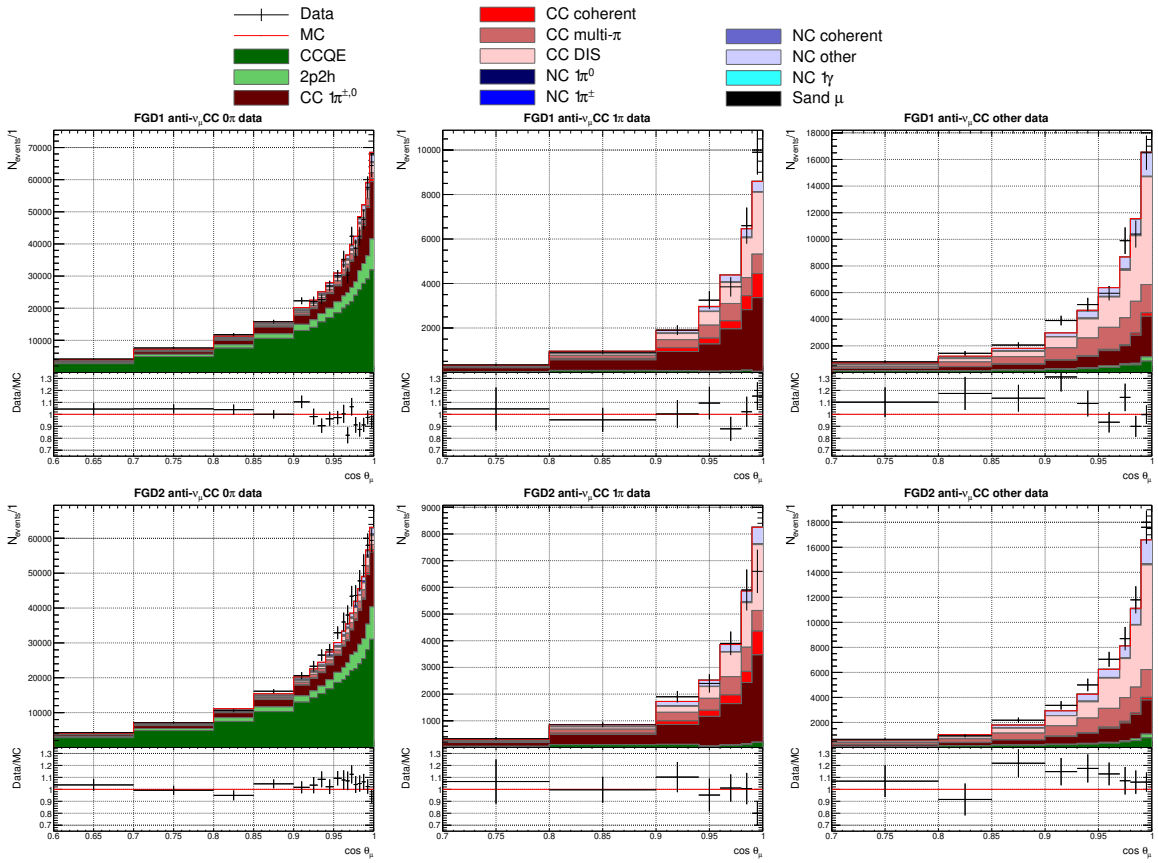


Figure 6.12.: Data and nominal MC distributions for RHC $\bar{\nu}_\mu$ selections projected onto $\cos\theta_\mu$, showing contributions by interaction mode. Bin content is normalised to bin width

The ND280 FHC ν_μ 0.6-0.7 GeV parameter moves from a sample $-2\ln\mathcal{L}$ contribution of $\sim$175 to $\sim$350, and all the flux parameters have Gaussian responses.

Figure 6.15 shows the likelihood scans for the same interaction parameters as Figure 5.28. Again, there are large increases in the likelihood responses for many parameters. Notably M_A^{QE} almost doubles in sensitivity, as does 2p2h shape and the non-resonant $I_{1/2}$ single pion parameter.

To facilitate direct comparisons between 2017 and 2018 analyses, the likelihood response for some of the parameters with the largest improvement are compared in Figure 6.16. All the listed interaction parameters increase by a factor $1.5 \sim 2.0$, whereas the flux parameters don't improve more than 25%. Finally the effect coming from the new data versus the new data with rebinning and new selections is shown in Figure 6.17. The flux parameters all varied less than 5%, whereas some interaction parameters see larger changes. The amount of data is clearly the larger contributor to the increased sensitivity compared to the binning and selection.

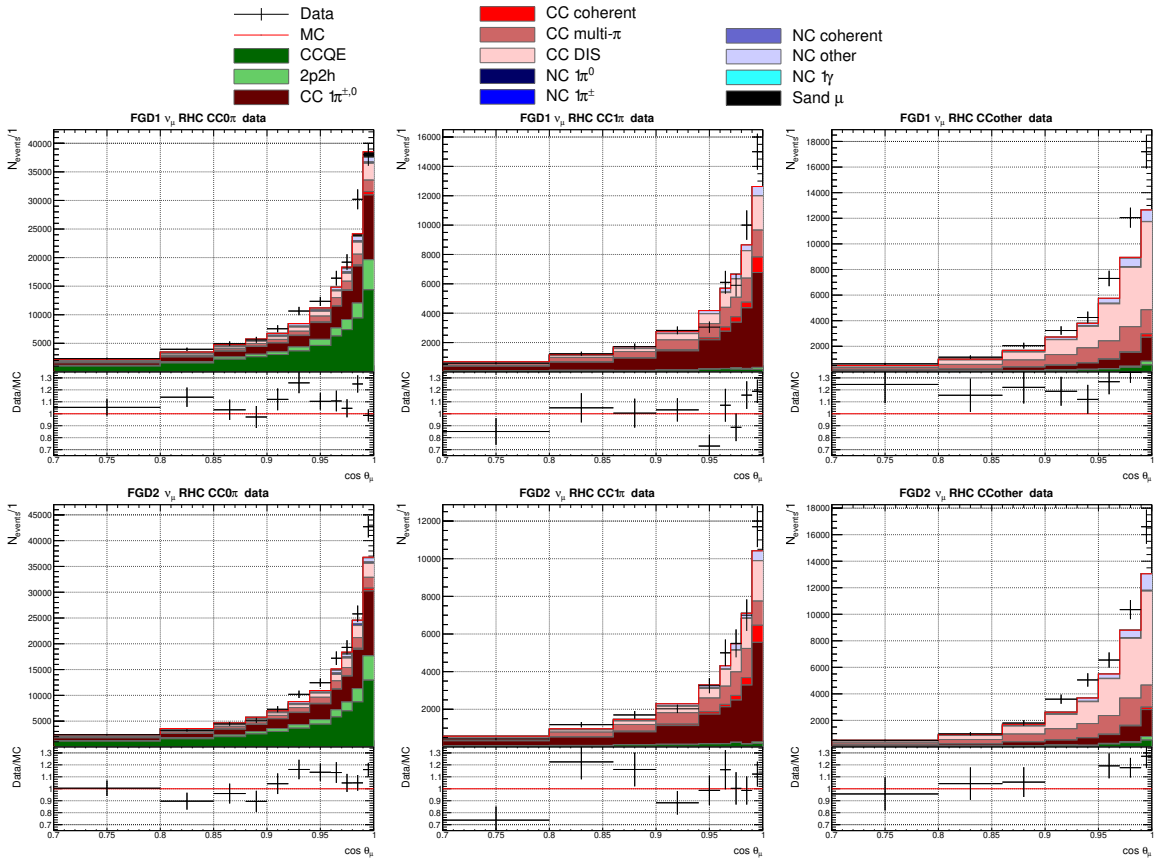


Figure 6.13.: Data and nominal MC distributions for RHC ν_μ selections projected onto $\cos\theta_\mu$, showing contributions by interaction mode. Bin content is normalised to bin width

Sample	CCQE	2p2h	CC1$\pi^{\pm,0}$	CC coh	CC multi-π	CC DIS	NC
FGD1 0π	**56.7**	**9.9**	19.5	0.3	4.6	5.2	3.7
FGD2 0π	**54.8**	**9.3**	21.6	0.3	5.0	5.3	3.7
FGD1 1π	6.2	1.0	**48.1**	**2.8**	**17.7**	17.4	6.8
FGD2 1π	5.7	0.8	**47.7**	**2.8**	**18.1**	18.1	6.8
FGD1 Other	5.0	1.0	14.9	0.4	**25.6**	**44.8**	8.3
FGD2 Other	5.1	1.0	15.4	0.4	**25.1**	**45.0**	8.0
FGD1 $\bar{\nu}_\mu$ 0π	**61.5**	**10.1**	15.6	0.8	3.3	3.2	5.5
FGD2 $\bar{\nu}_\mu$ 0π	**61.7**	**9.8**	15.8	0.7	3.6	3.3	5.1
FGD1 $\bar{\nu}_\mu$ 1π	3.8	0.8	**44.4**	**8.7**	**15.3**	18.7	8.3
FGD2 $\bar{\nu}_\mu$ 1π	5.5	0.6	**42.2**	**7.8**	**15.7**	18.9	9.3
FGD1 $\bar{\nu}_\mu$ Other	9.4	1.5	19.8	1.4	**22.4**	**32.0**	13.5
FGD2 $\bar{\nu}_\mu$ Other	8.7	1.5	19.4	1.4	**22.3**	**33.4**	13.3
FGD1 ν_μ RHC 0π	**42.3**	**8.7**	24.1	0.7	7.8	8.3	8.1
FGD2 ν_μ RHC 0π	**40.2**	**8.2**	26.2	0.7	8.8	8.7	7.2
FGD1 ν_μ RHC 1π	6.7	1.1	**45.1**	**3.8**	**20.3**	15.7	7.3
FGD2 ν_μ RHC 1π	6.4	0.8	**44.7**	**3.9**	**21.5**	16.8	5.9
FGD1 ν_μ RHC Other	4.9	1.0	14.2	0.8	**26.3**	**44.7**	8.1
FGD2 ν_μ RHC Other	4.6	0.9	15.2	0.6	**25.9**	**44.5**	8.3

Table 6.6.: Percentage mode breakdown for the binned nominal scaled Monte-Carlo samples, boldface indicates interactions targeted by specific selections. Directly comparable to 2017 results in Table 5.11

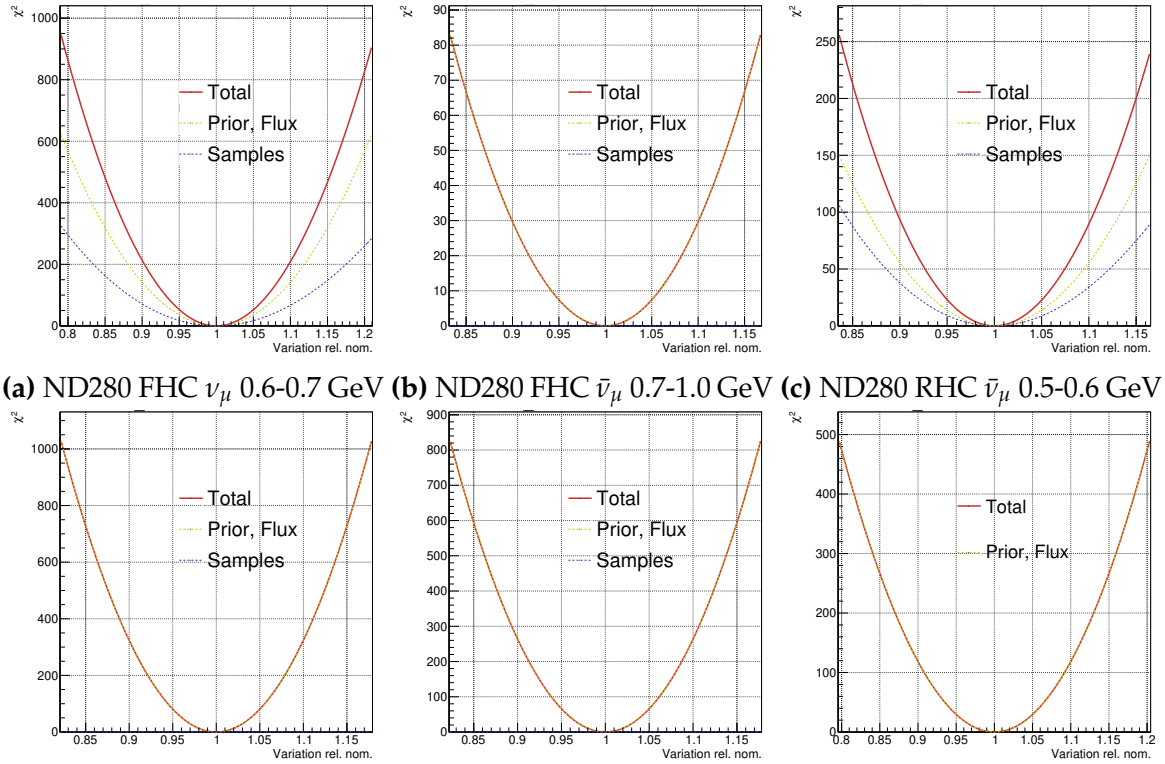
However, the pion final state parameters and 2p2h shape parameters change shape considerably, and see increased sensitivity by $\times 1.5 - 2.0$. A clear example of this is the pion final-state-interaction parameters in Figure 6.18. They all see the largest increase in response, often with a complex shape. This is somewhat expected since the RHC selections in 2017 only had one selection per FGD with reconstructed pions (NTrack) but had two per FGD for the FHC selections (1π and Other). Many of the FSI parameters get stronger constraints from using both π^+ and π^- data, which is better satisfied with the new RHC selections.

6.6.2. Prior Predictive Spectrum

As was done in subsection 5.6.3 the prior's predictive power on the Asimov data set is checked. The predictive spectrum is expected to largely agree with the Asimov data

Sample	MC	POT	Flux	Xsec	Det	ND280 Cov	All
FGD1 0π	470176	31464.6	34153.7	30106.3	30298	31487.9	31529.3
FGD1 1π	119835	8059.92	9106.99	7496.3	7690.07	7959.59	7998.1
FGD1 Other	92630	6224.1	7389.73	6096.03	5917.31	6144.95	6793.68
FGD2 0π	471140	31215.4	33881.4	30017.8	30361.4	31198.7	31734
FGD2 1π	95498	6303.05	7148.91	5919.55	6121.38	6184.34	6419.04
FGD2 Other	87931	5839.17	6933.53	5723.31	5697.16	5773.49	6562.75
FGD1 $\bar{\nu}_\mu$ 0π	96602	6784.14	6983.08	6247.83	6723.77	6794.13	6371.34
FGD1 $\bar{\nu}_\mu$ 1π	9133	639.595	658.426	536.293	623.687	635.372	533.253
FGD1 $\bar{\nu}_\mu$ Other	15046	1066.91	1113.12	1012.08	1044.37	1055.12	1023.36
FGD2 $\bar{\nu}_\mu$ 0π	95597	6692.82	6897.34	6185.55	6578.7	6715.93	6283.35
FGD2 $\bar{\nu}_\mu$ 1π	8165	568.917	587.274	491.61	553.376	557.22	483.508
FGD2 $\bar{\nu}_\mu$ Other	13849	970.796	1015.05	927.283	954.657	961.134	943.956
FGD1 ν_μ RHC 0π	34950	2457.88	2646.97	2392.37	2379.95	2449.2	2485.51
FGD1 ν_μ RHC 1π	12352	871.792	952.669	814.909	838.532	871.141	855.911
FGD1 ν_μ RHC Other	10894	764.374	854.448	750.775	733.571	764.374	804.647
FGD2 ν_μ RHC 0π	35180	2458.01	2645.96	2408.41	2419.07	2458.01	2553.51
FGD2 ν_μ RHC 1π	9714	676.737	740.521	632.743	662.563	676.737	679.99
FGD2 ν_μ RHC Other	10421	732.441	819.628	718.565	720.726	732.441	792.166
Total	1689110	113791	124529	108478	110318	113420	114847

Table 6.7.: Event rates broken by type of weight applied



(a) ND280 FHC ν_μ 0.6-0.7 GeV **(b)** ND280 FHC $\bar{\nu}_\mu$ 0.7-1.0 GeV **(c)** ND280 RHC $\bar{\nu}_\mu$ 0.5-0.6 GeV

(d) ND280 FHC ν_e 0.5-0.7 GeV **(e)** ND280 RHC $\bar{\nu}_e$ 0.5-0.7 GeV **(f)** SK FHC ν_μ 0.6-0.7 GeV

Figure 6.14.: Asimov likelihood scans for selected beam parameters

and produce large uncertainties and thereby p-values around 0.5. Figure 6.19 shows the resulting two-dimensional p-value, agreeing with the expectation.

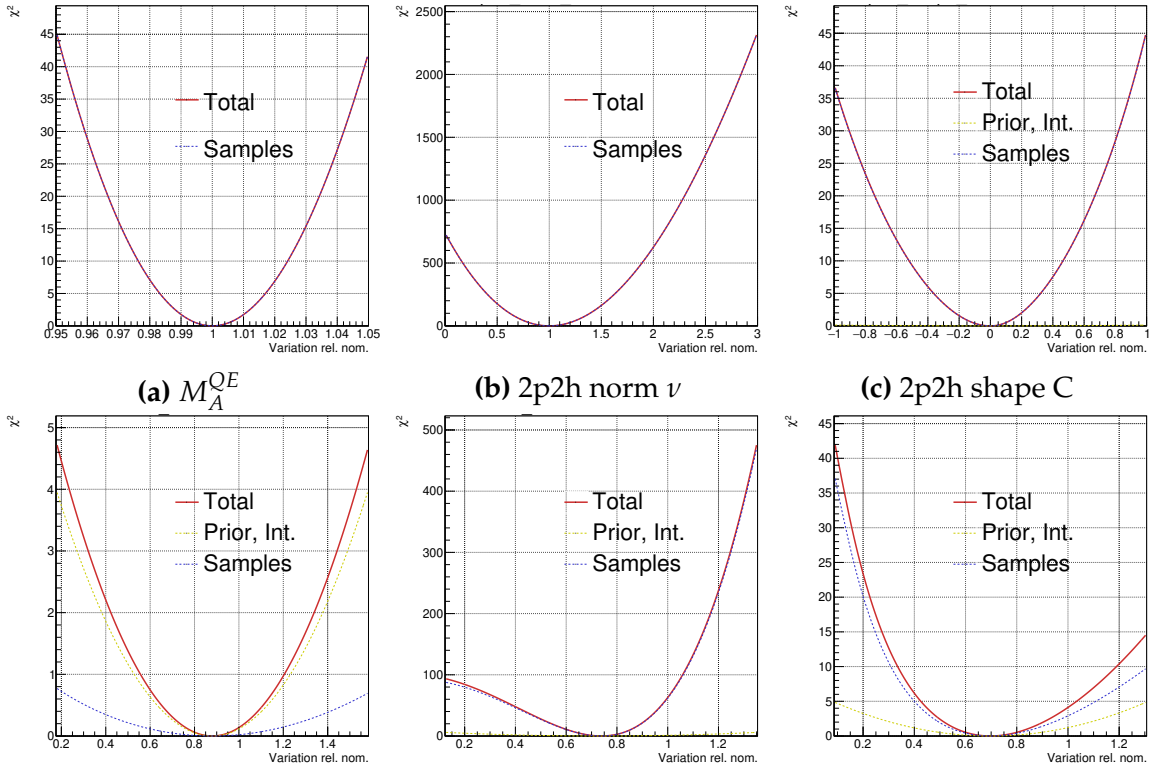
Table 6.8 shows the event rates after making correlated throws of the systematics with the test-statistic against Asimov data. Comparing to the 2017 equivalent in Table 5.13 there is a similar level of uncertainty as expected. As in 2017, the prior predictive on Asimov data produces an 11% uncertainty on the total event rate, with 13% on the CC0π, 11% on CC1π and 13% on CCOther. There is a consistent skew in the prior predictive "best-fit" where it overestimates the Asimov data set, likely due to missing covariance matrices between cross-section, flux and ND280 parameters. The contribution to the test-statistic is much larger compared to 2017, primarily due to the many more bins in each selection.

6.6.3. Fitting to Asimov Data

The model is fit to the Asimov data to estimate the sensitivity and perform closure tests. Fits using different ND280 covariance matrices are also done, comparing results

Sample	Nominal	Prior Pred.	$-2\log\mathcal{L}_S$
FGD1 0π	31529.3	32347.8 ± 4135.7	26.1
FGD1 1π	7998.1	8106.6 ± 8899.4	2.1
FGD1 Other	6793.7	6894.1 ± 858.4	3.1
FGD2 0π	31734.1	32572.8 ± 4012.8	26.5
FGD2 1π	6419.0	6511.2 ± 702.4	2.0
FGD2 Other	6562.8	6652.8 ± 786.4	3.0
FGD1 $\bar{\nu}_\mu$ 0π	6371.3	6541.9 ± 847.5	7.7
FGD1 $\bar{\nu}_\mu$ 1π	533.3	530.5 ± 115.5	4.6
FGD1 $\bar{\nu}_\mu$ Other	1023.4	1037.4 ± 191.1	0.8
FGD2 $\bar{\nu}_\mu$ 0π	6283.4	6444.7 ± 830.5	8.4
FGD2 $\bar{\nu}_\mu$ 1π	483.5	486.8 ± 102.3	0.6
FGD2 $\bar{\nu}_\mu$ Other	944.0	953.2 ± 204.1	0.6
FGD1 ν_μ RHC 0π	2485.5	2543.8 ± 429.5	3.3
FGD1 ν_μ RHC 1π	855.9	861.9 ± 107.7	0.5
FGD1 ν_μ RHC Other	804.7	790.9 ± 159.8	1.9
FGD2 ν_μ RHC 0π	2553.5	2503.5 ± 395.5	6.3
FGD2 ν_μ RHC 1π	680.0	687.3 ± 88.1	0.8
FGD2 ν_μ RHC Other	792.2	780.4 ± 136.5	1.9
Total	114834	117541.9 ± 12383.0	100.2

Table 6.8.: Prior predictive event rates for the Asimov data



(a) M_A^{QE} **(b)** 2p2h norm ν **(c)** 2p2h shape C

(d) BeRPA E **(e)** $I_{1/2}^{\text{bkg}}$ **(f)** FSI CEX LO

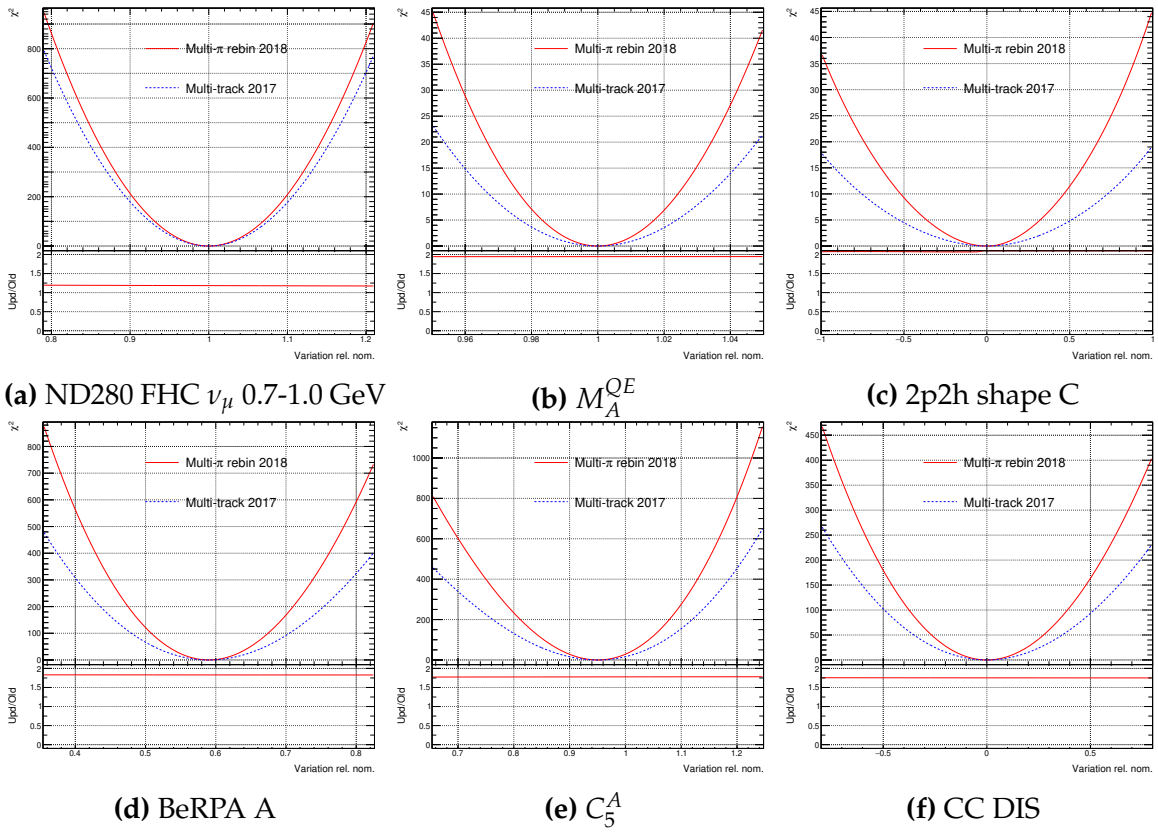
Figure 6.15.: Asimov likelihood scans for selected cross-section parameters

using the nominal binned multi-track to the binned multi-π RHC selections, and fitting without any detector parameters.

The MCMC parameters for the different studies are shown in Table 6.9. The primary reason behind the low acceptance for the "Full cov" (using the full ND280 covariance matrix) is the 4238 parameters being fit simultaneously, making MCMC parameter tuning cumbersome.

Name	Step length	Acceptance	Accepted steps
Nominal cov	3,900,000	12.1%	471,900
Full cov	1,367,502	5.8%	79,315
Multi-track	3,000,000	10.8%	324,272
No det	3,900,000	24.3%	947,700

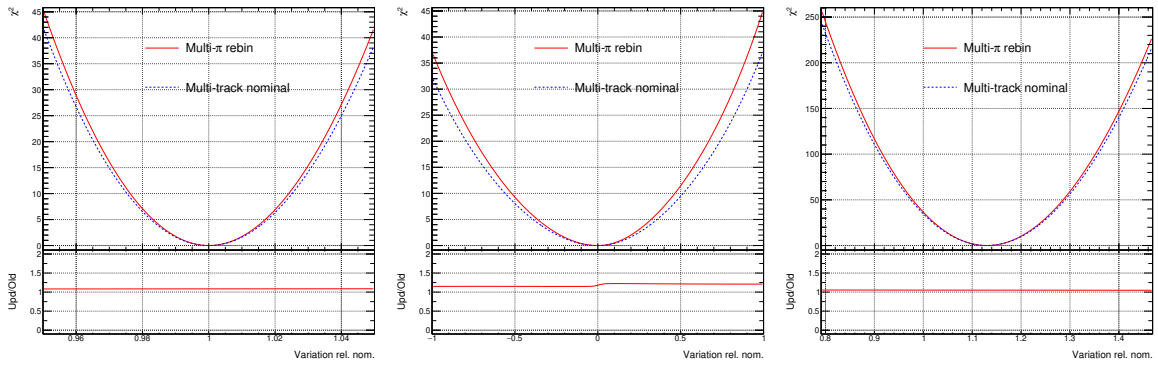
Table 6.9.: Markov Chain parameters for the various Asimov fits in subsection 6.6.3



(a) ND280 FHC ν_μ 0.7-1.0 GeV **(b)** M_A^{QE} **(c)** 2p2h shape C

(d) BeRPA A **(e)** C_5^A **(f)** CC DIS

Figure 6.16.: Asimov likelihood scans for 2017 and 2018 analyses

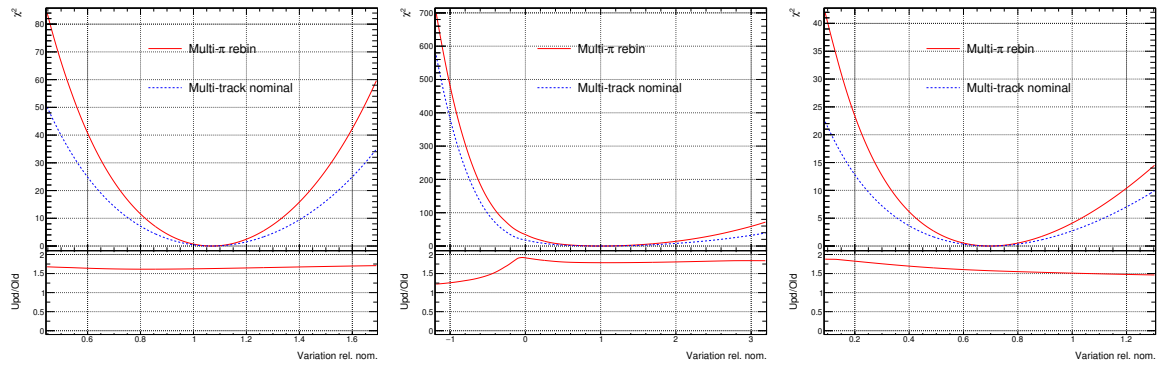


(a) M_A^{QE} **(b)** 2p2h shape C **(c)** BeRPA D

Figure 6.17.: Asimov likelihood scans for multi-π and rebinned samples versus the unchanged multi-track sample from 2017 with run 2 to 8 statistics

Full and Reduced ND280 Detector Systematics Parameterisation

A fit to the Asimov data using the reduced ND280 covariance matrix (1076 ND280 detector parameters) is compared to one using the full ND280 covariance matrix (4238 ND280 detector parameters). Figure 6.20 and Figure 6.21 show the flux parameters after the fit to Asimov data using the two matrices. There is a consistent bias in all



(a) Quasi-elastic **(b)** Inelastic **(c)** Charge exchange

Figure 6.18.: Asimov likelihood scans for multi-π and rebinned samples versus the unchanged multi-track sample from 2017 with run 2 to 8 statistics, for some pion FSI rescatering parameters

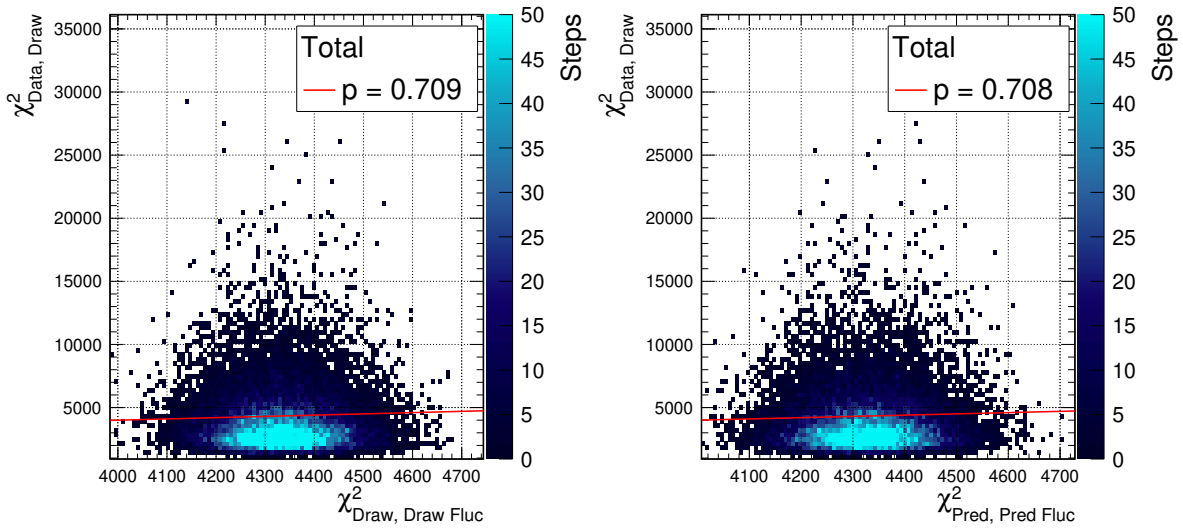
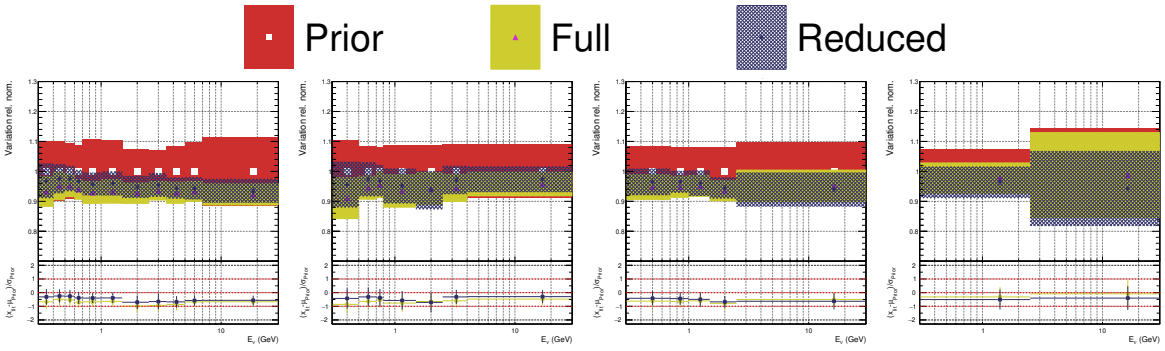


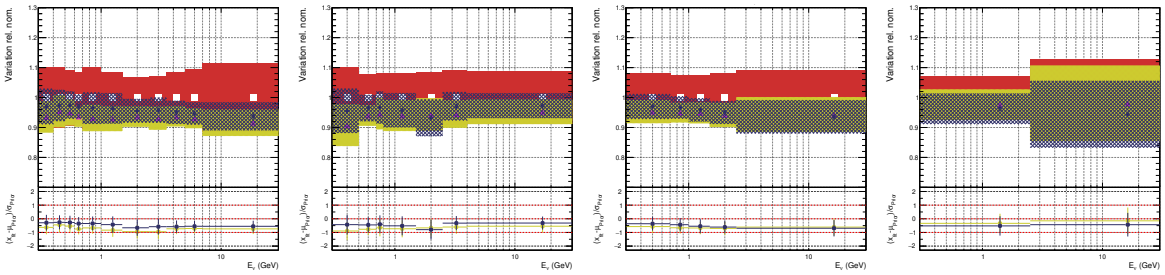
Figure 6.19.: Prior predictive p-values for the Asimov data in 2018

flux parameters for both ND280 and SK, which appear to be a normalisation offset by 1-2% for the reduced parameterisation, and 5% for the full parameterisation. The ND280 and SK parameters follow the same pattern throughout, and the uncertainties are approximately halved compared to the prior, similar to the 2017 fit to Asimov data.

Looking at the cross-section parameters after the fit to Asimov data in Figure 6.22, the pattern is much the same as in 2017. p_F, 2p2h normalisation $\bar{\nu}$, 2p2h shape, BeRPA A and BeRPA E appear biased from the Asimov parameter value, whereas the rest seem to find it satisfactorily.



(a) ND280



(b) SK

Figure 6.20.: FHC flux parameters, comparing Asimov fits with full and reduced ND280 covariance matrices

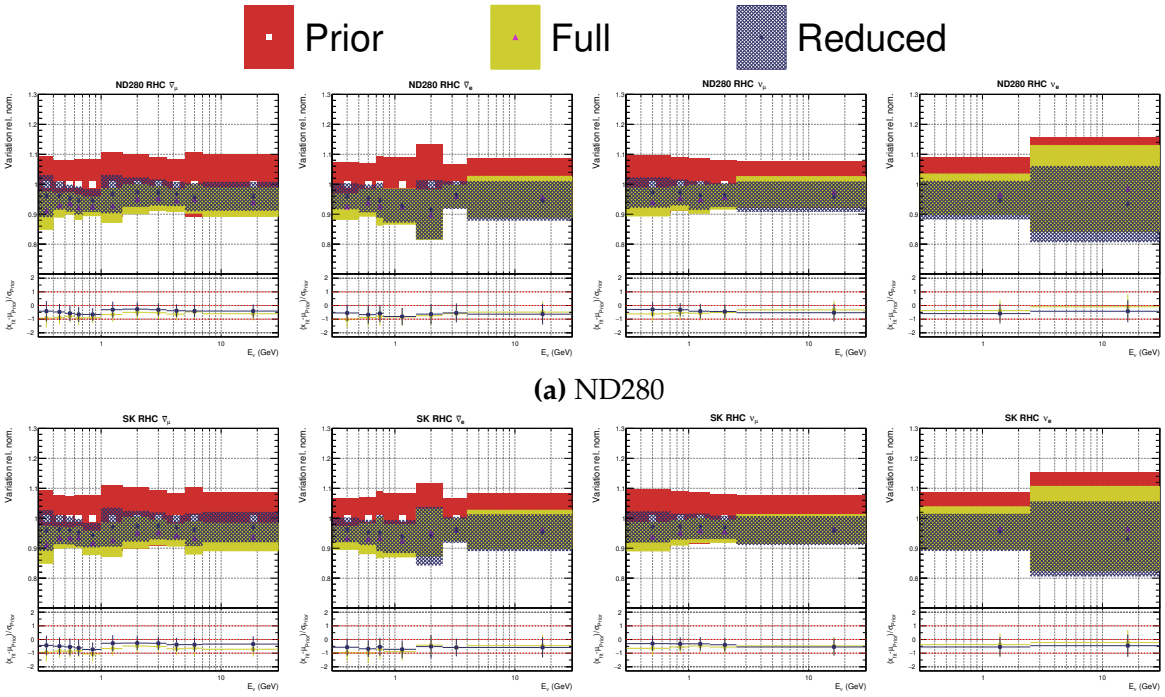
The reduction in parameter uncertainties from doubling of data is particularly noticeable for 2p2h normalisation which is reduced to 10% from 20%, M_A^{QE} which now has a similar uncertainty to fits from bubble chamber data, the 2p2h shape parameters which now have 20% uncertainty, and the BeRPA B parameter, whose uncertainty more than halves compared to the prior. Single pion production and pion FSI parameters also see reduced uncertainties by between 1/2 to 1/4 of the prior.

The two ND280 covariance matrices are compatible for the interaction parameters.

Comparing to the Multi-Track Selection

As with the likelihood scan, the impact of the selection is studied. Results from using the updated selections for the 2018 analysis are compared to the selection and binning of 2017 (RHC multi-track, no rebinning), including the new run 7 and 8 data for both.

In the light of the biases in the flux parameters from using the multi-π samples, and the biases observed in the 2017 analysis' Asimov study (subsection 5.6.4), it is particularly interesting to see the flux parameters largely unbiased for the multi-track



(a) ND280

(b) SK

Figure 6.21.: RHC flux parameters, comparing Asimov fits with full and reduced ND280 covariance matrices

selection in Figure 6.23 and Figure 6.24. The two selections follow a similar pattern, although the fit using the multi-π is offset by 1-2%. The ND280 and SK flux parameters echo each other and are compatible. Comparing the size of the errors from the two selections, there is a marginally smaller error for the multi-π selection for the flux parameters, although barely discernible.

The interaction parameters in Figure 6.25 are entirely compatible and neither of the two fits show unexpected biases. Comparing the size of the errors on the parameters the multi-π selection and rebinning has a larger impact than on the flux parameters. Many parameters reduce by as much as 20-30%, as expected from the earlier likelihood scans. The 2p2h and pion FSI parameters see the largest reductions.

Asimov Without Varying ND280 Detector Parameters

To investigate the biases in the flux parameters seen in the fit to Asimov data for both the new ND280 covariance matrices for the multi-π selections, a fit to the Asimov data is made without varying the ND280 parameters. This should help answer if there are issues related to dimensionality in the new fit, since the new parameterisation fits

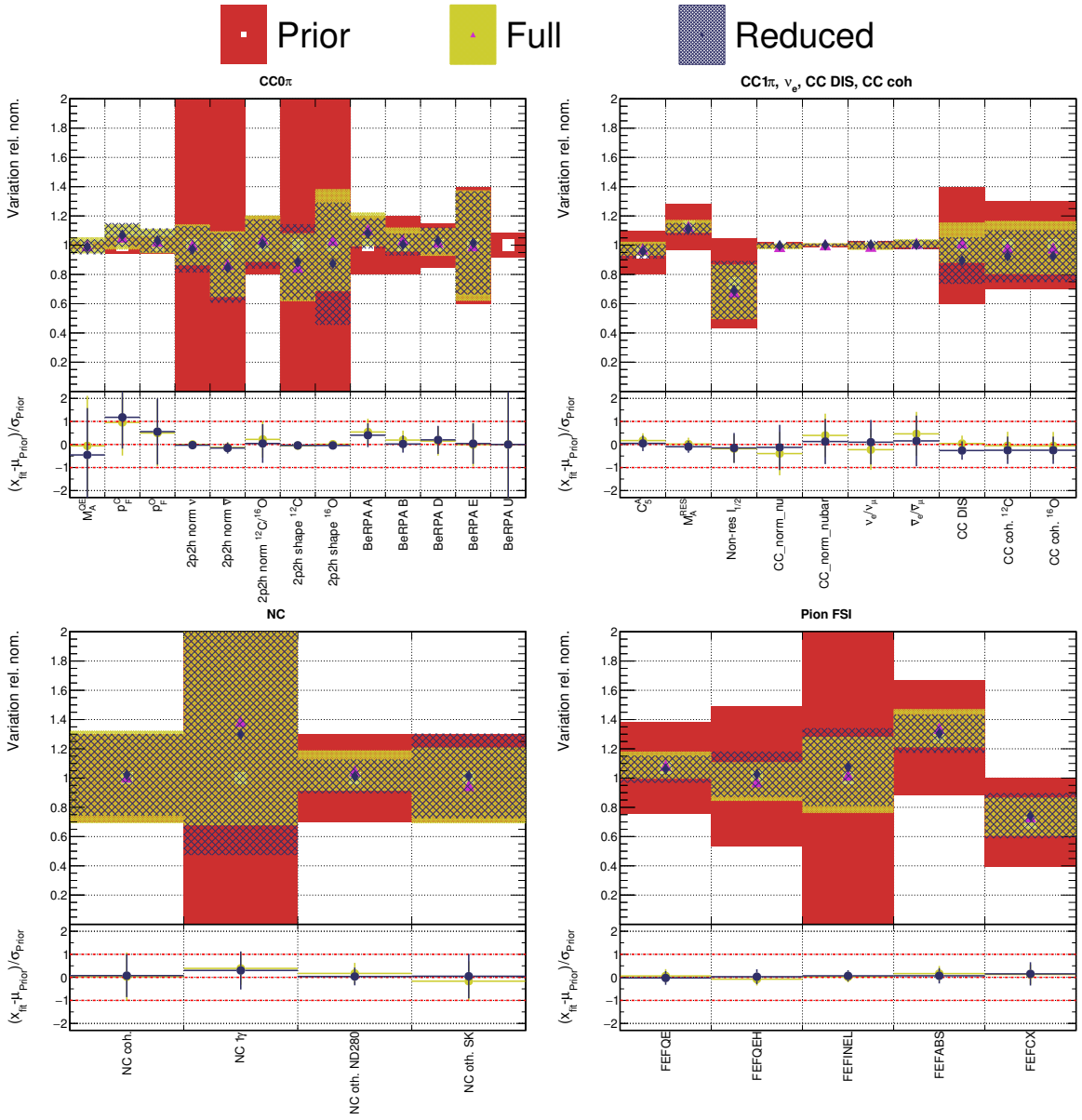
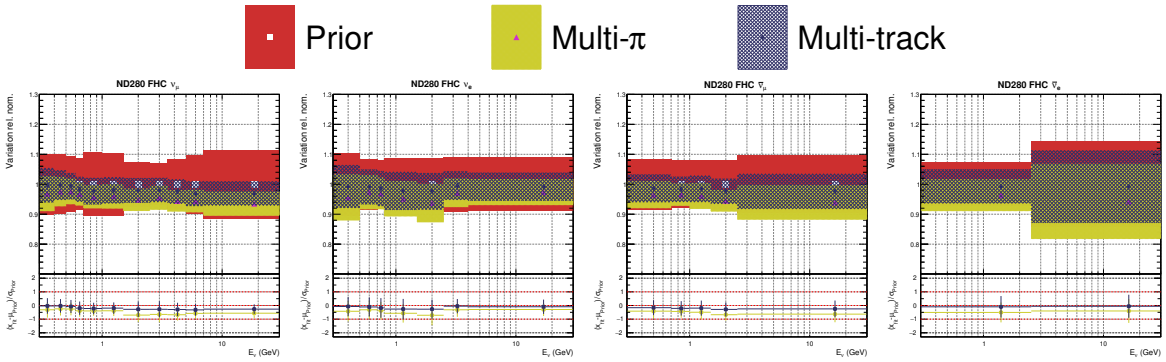


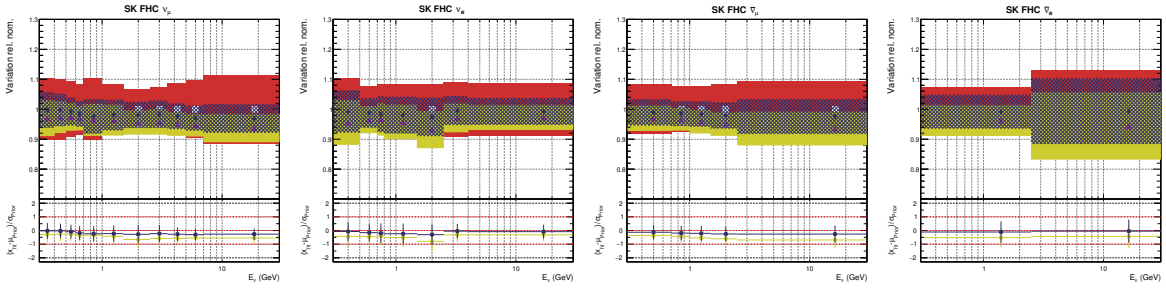
Figure 6.22.: Interaction parameters, comparing Asimov fits with Full and reduced ND280 covariance matrices

approximately double the number of parameters (1307 vs 781) with a large increase in number of bins (4238 vs 1624).

Figure 6.26 presents the FHC flux parameters without varying the ND280 systematics. All of the biases seen in the previous Asimov fit are gone. The uncertainties are reduced due since the ND280 systematics are not varied. The same is true for the RHC parameters in Figure 6.27.



(a) ND280



(b) SK

Figure 6.23.: FHC flux parameters, comparing Asimov fits with rebinned multi-π to 2017 binned multi-track

The interaction parameters in Figure 6.28 are similarly less biased than in the full fit, although the effect is less extreme than for the flux parameters.

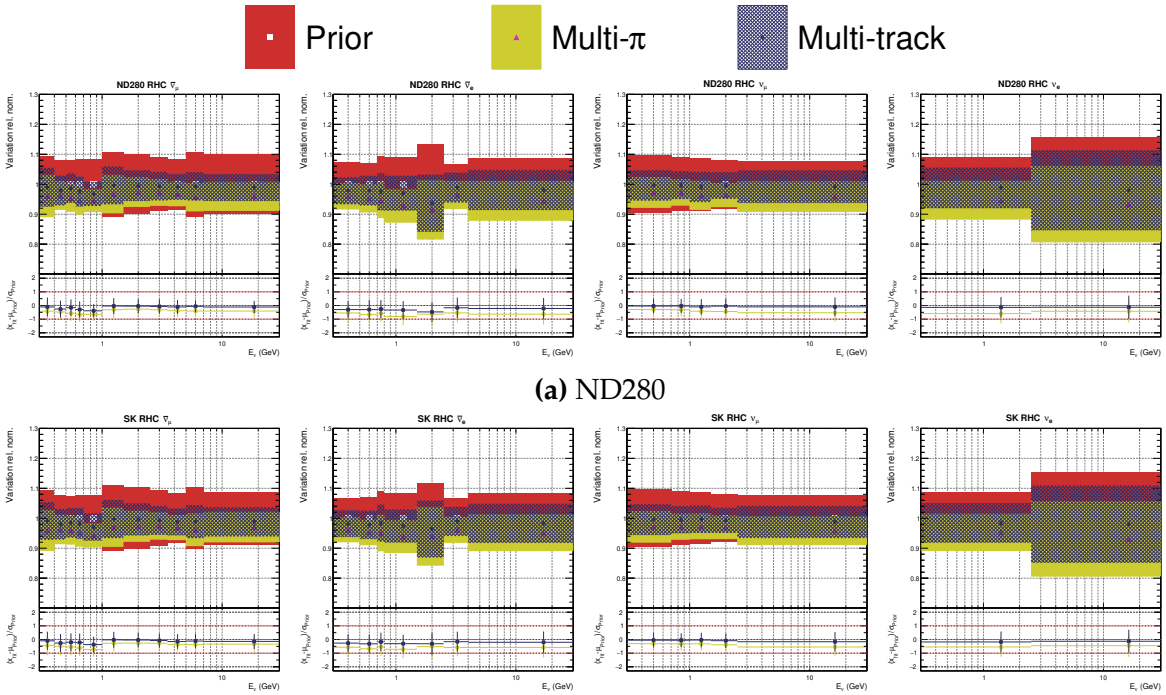
In conclusion, it appears that the bias in the flux parameters come from the ND280 parameters.

6.6.4. Covariance Matrix from the Asimov Fit

The fit to Asimov data is here used to deduce the parameter correlations.

Figure 6.29 shows the full $\surd$ covariance and correlation matrix for the flux (bottom left corner) and cross-section (upper right corner) parameters. Many patterns are repeated from the 2017 case (Figure 5.43): the flux parameters are internally correlated and the cross-section parameters are mostly separated into categories of correlations (e.g. CC0π-CC0π correlations are strong, CC0π-CC1π are not).

Looking at the more digestible version in Figure 6.30—which excludes the SK flux parameters—the strongest correlations between the flux and interaction parameters



(a) ND280

(b) SK

Figure 6.24.: RHC flux parameters, comparing Asimov fits with rebinned multi-π to 2017 binned multi-track

are for the CC DIS parameter and the M_A^{QE}, BeRPA A, BeRPA B and C_5^A parameter in which the latter group are especially strong around the flux peak.

For the flux parameters, the largest uncertainties are seen for the high energy wrong-sign ν_e parameters, since there is barely any data to constrain it. Furthermore, the production processes leading to such neutrinos are only weakly correlated with the lower energy right-sign processes, which is why it is weakly constrained by the prior covariance.

The largest interaction uncertainties are the 2p2h normalisations, BeRPA E, CC coherent normalisations and the NC parameters, which again is expected since the lack of data of such processes and Q^2, compatible with the 2017 fit.

Comparing the flux covariances before and after the fit in Figure 6.31, it's clear that ND280 reduces the uncertainty but maintains the parameters correlations. However, four of the ND280 and SK $\bar{\nu}_e$ parameters appear weakly negatively correlated (-0.15), which is not present in the prefit covariance.

Finally comparing the cross-section correlation matrix from the fit to Asimov data from 2017 to the 2018 results in Figure 6.32 there are generally very few changes. The

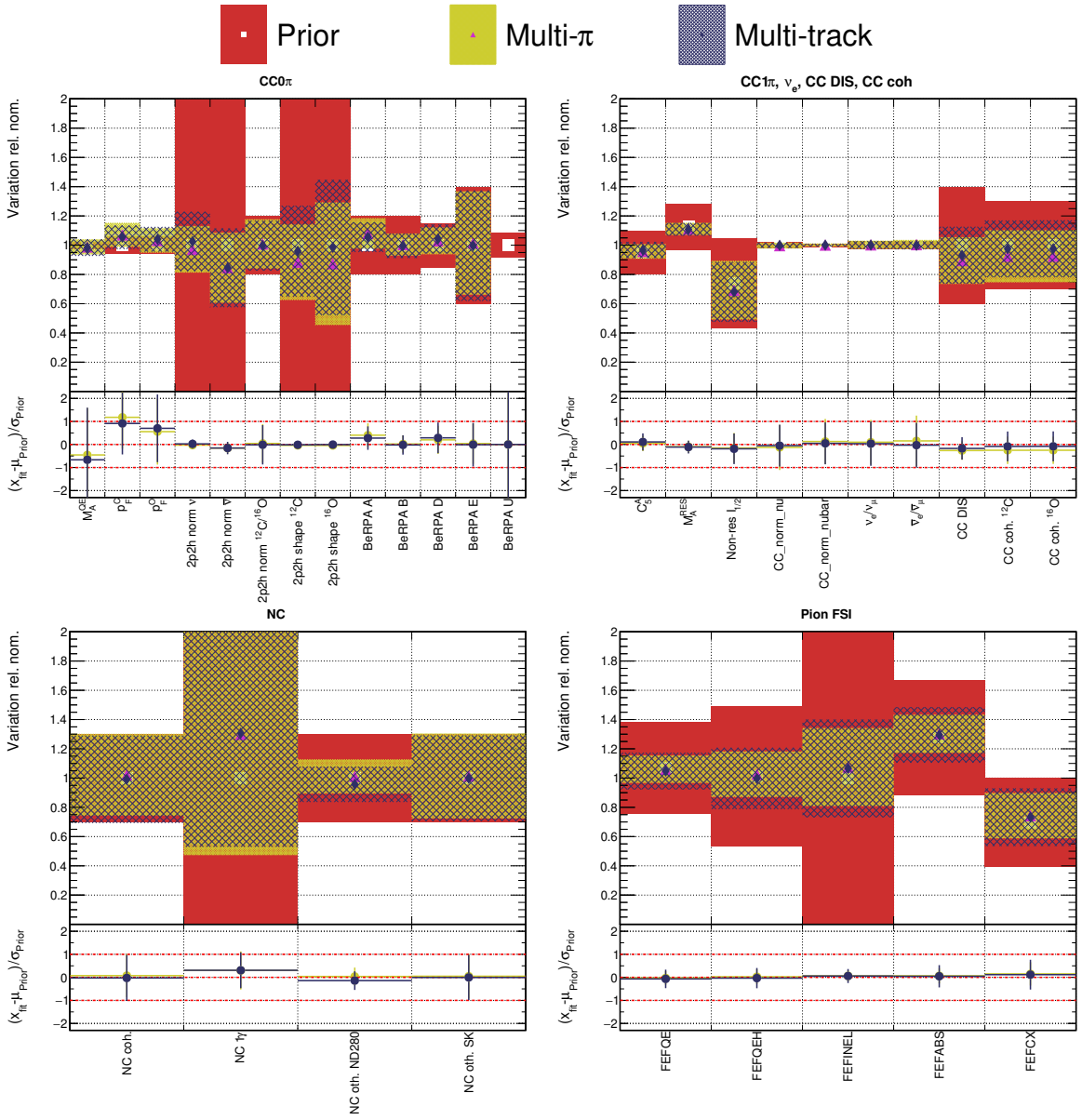
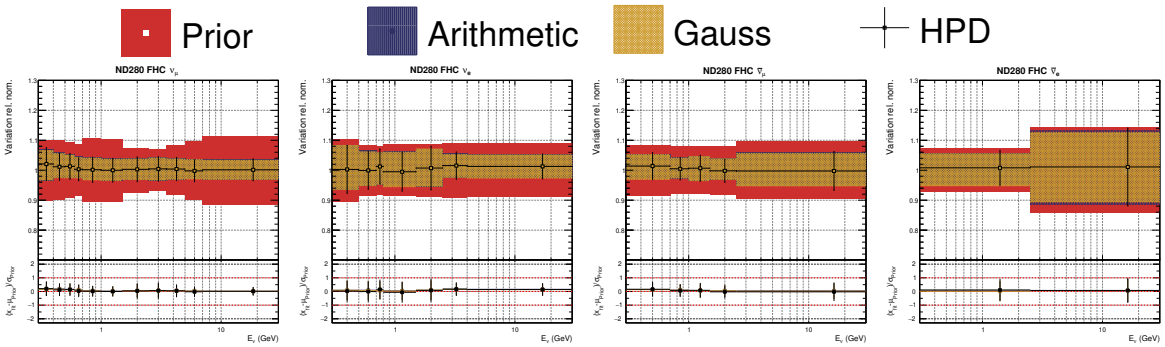
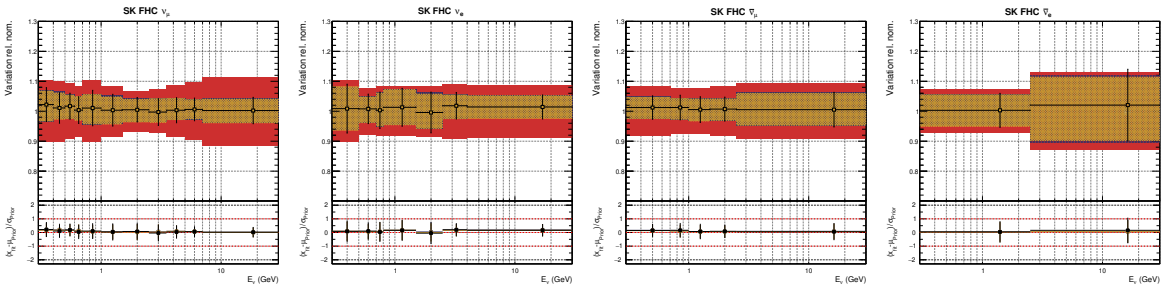


Figure 6.25.: Interaction parameters, comparing Asimov fits with rebinned multi-π to 2017 binned multi-track

CC0π parameters (bottom left corner) are unchanged, and the single pion parameters are correlating marginally more with the coherent parameters. The single pion and 2p2h normalisation correlations are slightly reduced, possibly due to the new RHC $\bar{\nu}_\mu$ samples. The pion FSI parameters (upper right corner) are clearly less correlated in 2018, owing to the updated covariance matrix.

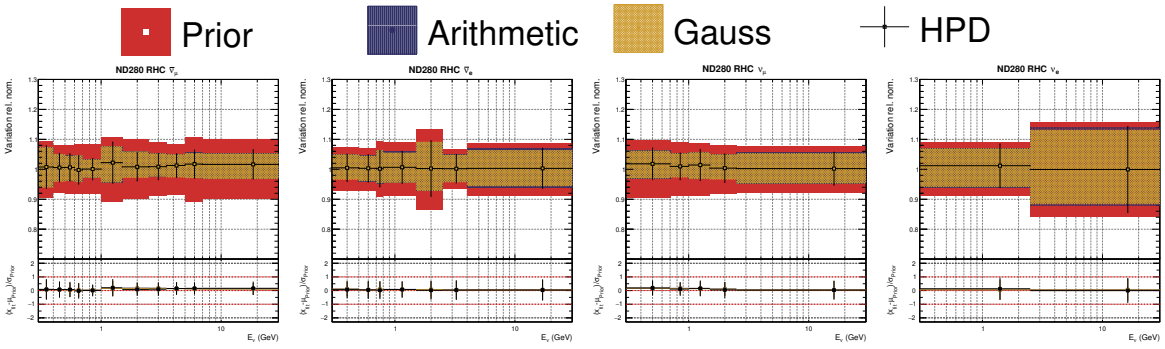


(a) ND280

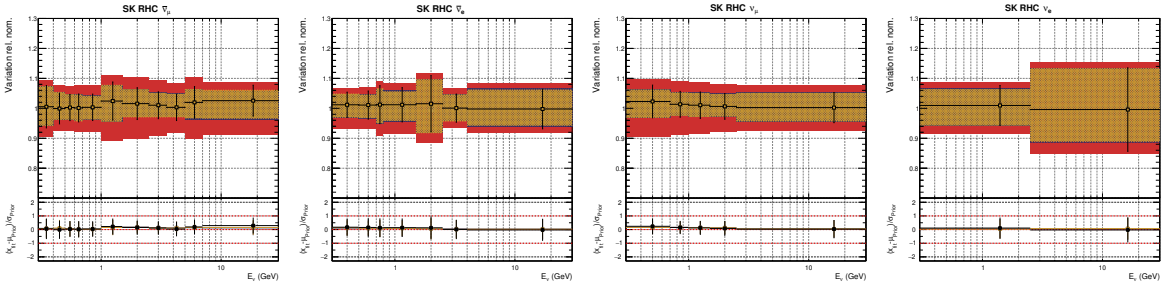


(b) SK

Figure 6.26.: FHC flux parameters, fitting to Asimov without varying detector parameters



(a) ND280



(b) SK

Figure 6.27.: RHC flux parameters, fitting to Asimov without varying detector parameters

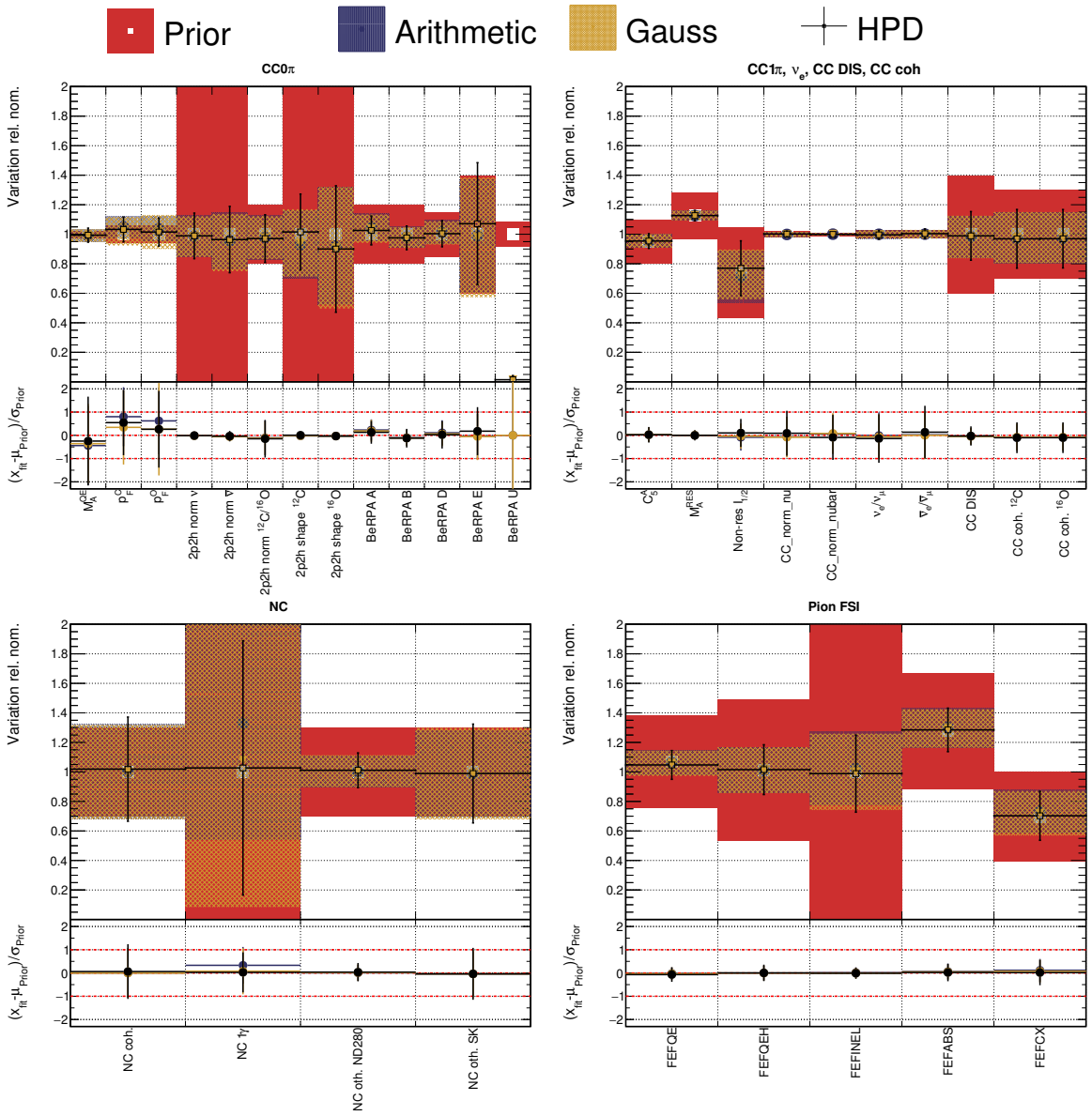
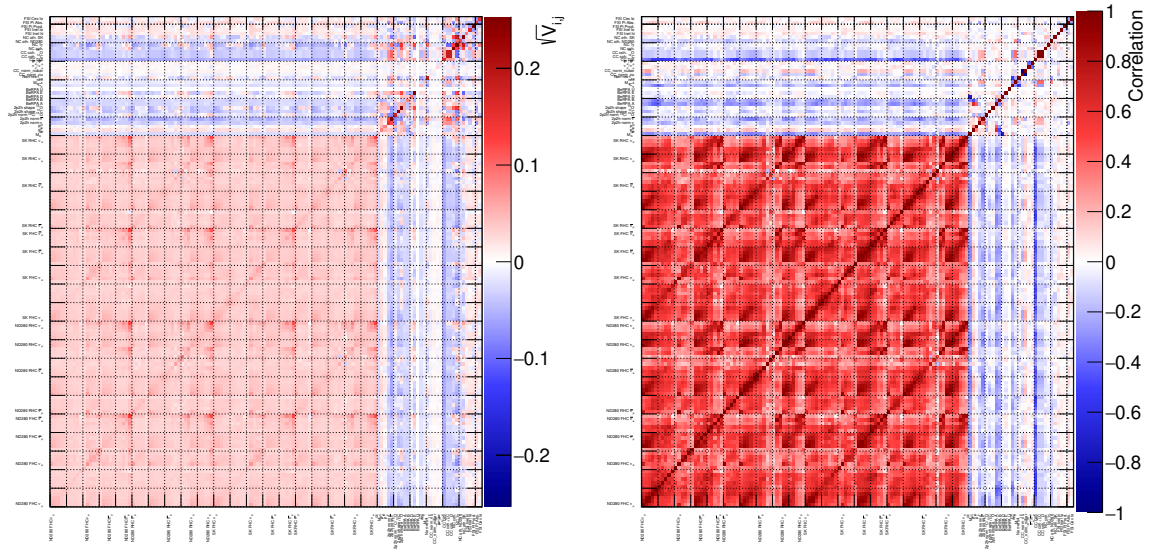


Figure 6.28.: Interaction parameters, fitting to Asimov without varying detector parameters

6.6.5. Posterior Predictive Spectrum

The posterior predictive $p_\mu \cos\theta_\mu$ spectrum and p-values are calculated using the MCMC with the reduced covariance matrix outlined above. The calculation proceeds as described in subsection 5.6.5 using 20,000 randomly chosen steps after a conservative 1/4 burn-in (corresponding to 975,000 steps). The expectation for the p-values is close to 1.0 since the post-fit constraints are small relative statistical fluctuations, as was the case in the 2017 analysis. Figure 6.33 shows the two p-values, which both are exactly 1.0.



(a) $\sqrt{\mathbf{V}_{i,j}}$ **(b)** $\rho_{i,j}$

Figure 6.29.: $\sqrt{\mathbf{V}_{i,j}}$ and correlation matrix for the Asimov post-fit, showing the full flux and cross-section parameters

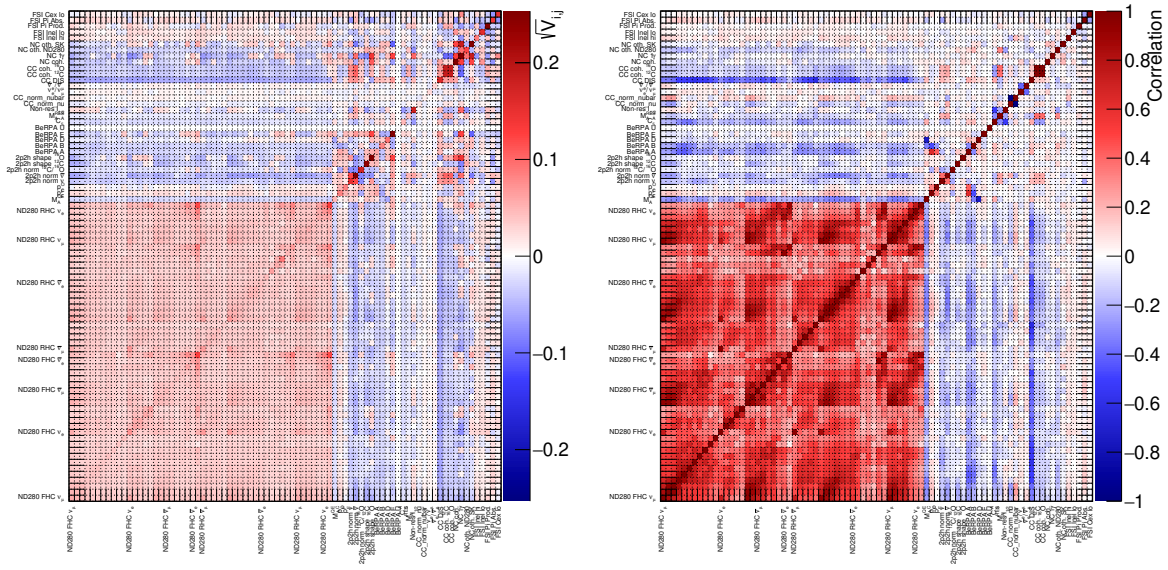
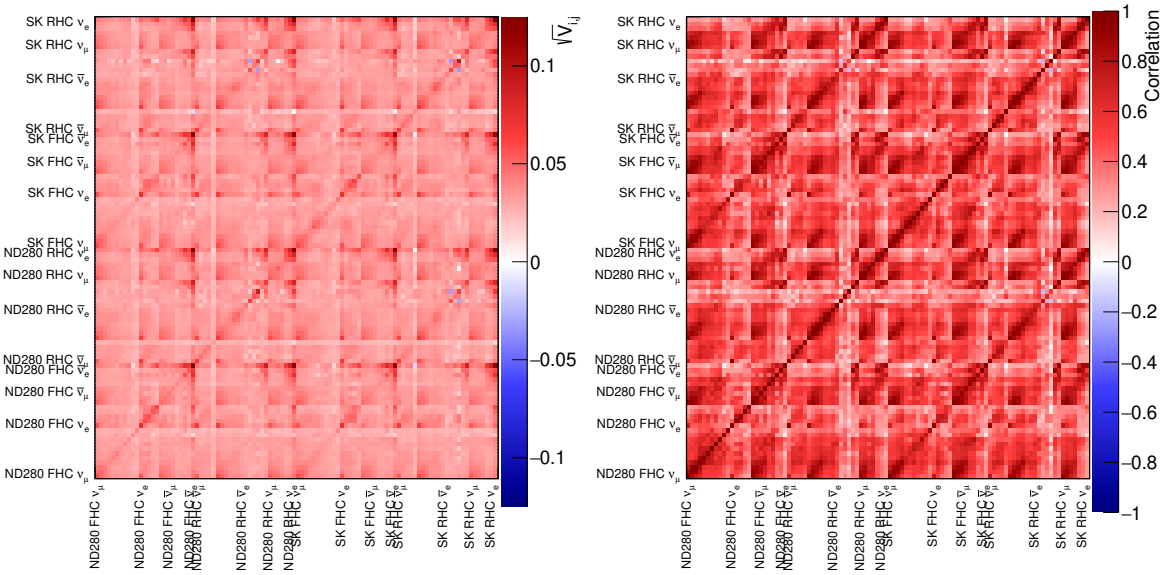
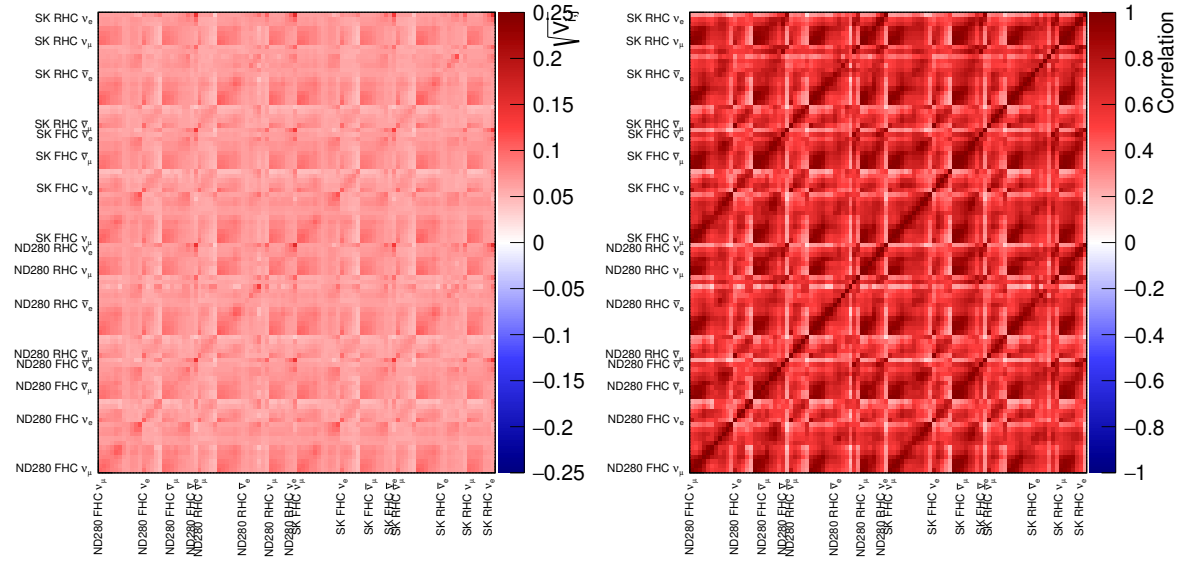


Figure 6.30.: $\sqrt{\mathbf{V}_{i,j}}$ and correlation matrix for the Asimov post-fit, showing ND280 flux and cross-section parameters

Moving attention to the posterior predictive spectrum's event rates in Table 6.10, the reduction in the post-fit event rate uncertainty compared to the prior predictive is large: from 12411.1 to 340.9 overall, 4245.4 to 168.3 for CC0π, 891.8 to 76.6 for CC1π and 827.6 to 76.2 for CCOther. The reduction in uncertainty is comparable to 2017, and overall doubling the statistics has the effect of moving the percentage uncertainty from 0.40% to 0.30%. The CC0π uncertainty moves from 0.70% to 0.53%, CC1π from



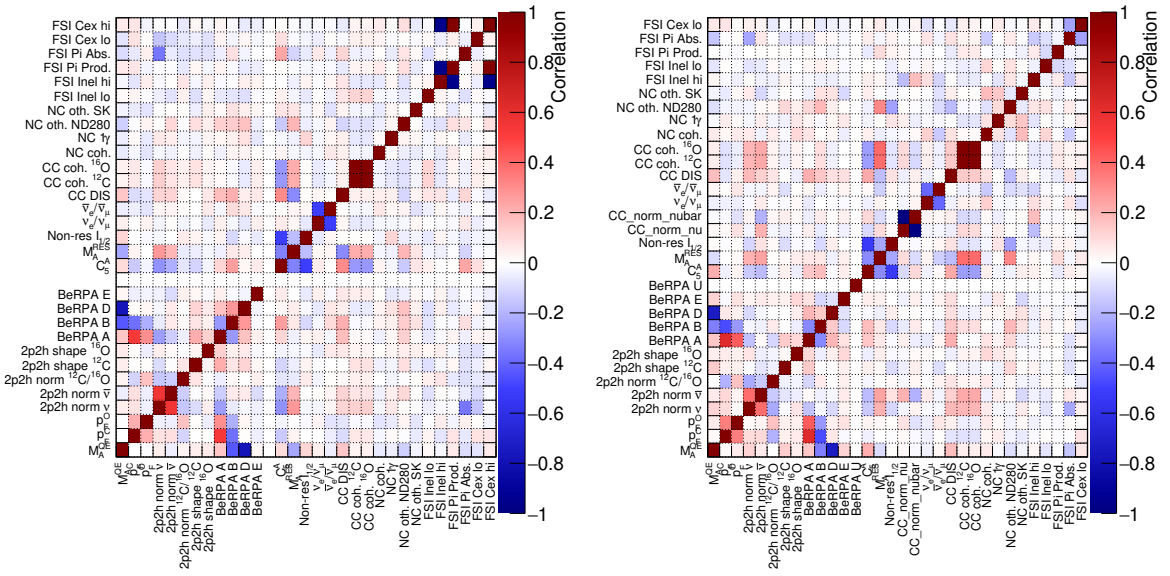
(a) Post-fit



(b) Pre-fit

Figure 6.31.: $\sqrt{\mathbf{V}_{i,j}}$ and correlation matrix for the flux parameters pre and post-fit to Asimov data

1.3% to 0.96%, CCOther from 1.4% to 1.1%. The anti-neutrino CC0π selections moves from 1.5% to 1.1%.



(a) 2017 **(b)** 2018

Figure 6.32.: Correlation matrix for the Asimov post-fit, showing cross-section parameters for 2017 and 2018 fits

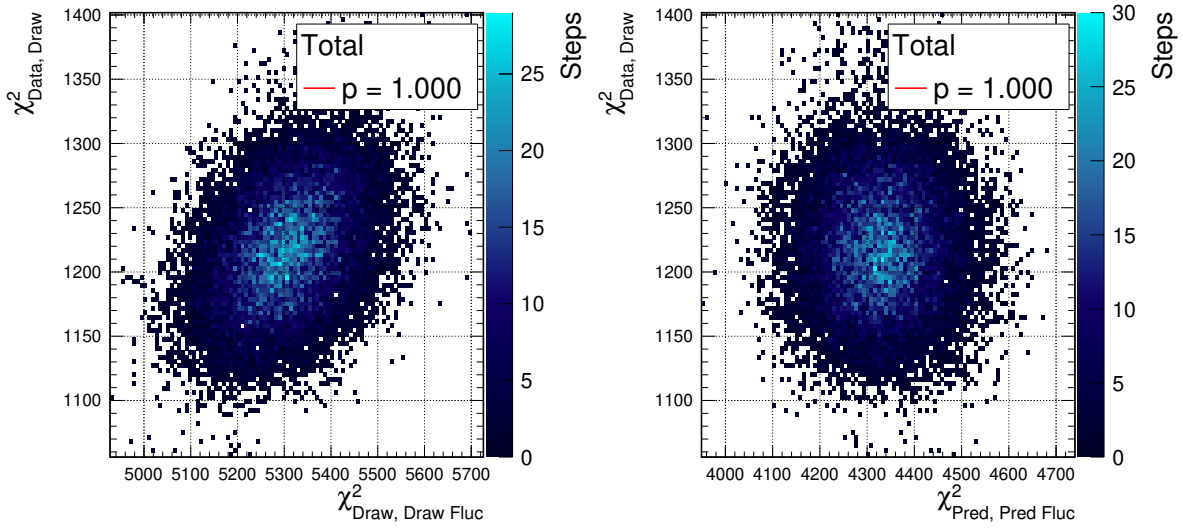


Figure 6.33.: Posterior predictive p-values for the Asimov data in 2018

6.7. Fitting Real Data

The Asimov fit stumbled upon one main issue: the 1-2% flux normalisation bias when using the multi-π selection. With this in mind the multi-track selection (which showed no bias) is fit to data, as is the multi-π selection with the reduced ND280 covariance matrix, and the multi-π selection with the full ND280 covariance matrix.

Sample	Nominal	Pos. Pred.	$-2\log\mathcal{L}_S$
FGD1 0π	31529.3	31545.3 ± 168.3	1.22
FGD1 1π	7998.1	8015.68 ± 76.6	0.70
FGD1 Other	6793.68	6804.29 ± 76.2	0.48
FGD2 0π	31734	31713.9 ± 166.9	0.99
FGD2 1π	6419.04	6428.7 ± 68.2	0.44
FGD2 Other	6562.75	6554.53 ± 71.9	0.38
FGD1 $\bar{\nu}_\mu$ 0π	6371.34	6369.71 ± 71.4	0.56
FGD1 $\bar{\nu}_\mu$ 1π	533.25	537.87 ± 20.8	0.12
FGD1 $\bar{\nu}_\mu$ Other	1023.36	1027.79 ± 29.3	0.05
FGD2 $\bar{\nu}_\mu$ 0π	6283.35	6287.51 ± 71.2	0.47
FGD2 $\bar{\nu}_\mu$ 1π	483.51	487.22 ± 20.7	0.04
FGD2 $\bar{\nu}_\mu$ Other	943.96	946.53 ± 28.1	0.04
FGD1 ν_μ RHC 0π	2485.51	2513.19 ± 42.1	0.47
FGD1 ν_μ RHC 1π	855.91	844.85 ± 13.7	0.22
FGD1 ν_μ RHC Other	804.65	795.18 ± 13.3	0.14
FGD2 ν_μ RHC 0π	2553.51	2529.72 ± 32.2	0.35
FGD2 ν_μ RHC 1π	679.99	669.64 ± 9.8	0.18
FGD2 ν_μ RHC Other	792.17	783.53 ± 13.1	0.12
Total	114847	114855.4 ± 340.9	6.98

Table 6.10.: Posterior predictive event rates after fitting to the Asimov data

The details for the MCMC that were obtained for the different fits are shown in Table 6.11. The lower acceptance probability is a direct result of the number of ND280 parameters being fit: the multi-track fit has 556, the nominal covariance has 1076 and the full covariance has 4238 ND280 parameters. All chains were monitored for stability and often converged within 1/8th of the total steps requested. For the parameter plots a conservative burn-in of 1/4 is used.

Name	Step length	Acceptance	Accepted steps
Nominal cov	3,900,000	12.2%	470,956
Full cov	5,869,504	6.1%	358,039
Multi-track	1,500,000	17.1%	257,439

Table 6.11.: Markov Chain parameters for the various data fits in section 6.7

The summary for the predictive distributions against data are shown in Table 6.12 using the reduced ND280 covariance matrix. In contrast to the 2017 analysis, the prior predictive distribution predicts the RHC 1π and Other distributions relatively well. This is reflected in the marginal decrease in the test-statistic after the fit, e.g. 42.7 to 40.0 for FGD2 $\bar{\nu}_\mu$ 1π. The same distribution barely sees a change in the prediction from the prefit to the postfit, although the uncertainties are effectively reduced (542.7 ± 127.2 to 542.9 ± 21.6). However, the prior predictive distributions do a poor job predicting the central values of the data of the high statistics FHC distributions, although often inside the 1σ uncertainties from the priors.

Inspecting the sample test-statistic, the $-2\log\mathcal{L}_S/\text{nBins}$ goes from 1.50 to 1.12, a deterioration from the 2017 value of 1.07. The fit is driven by the FHC 0π distributions (35.9% of the statistic) and the 1π, Other and RHC 0π making up between 10-15% for both FGDs. The new RHC 1π and Other distributions make small contributions to the statistic in total. Interestingly, the test statistics of FGD1 $\bar{\nu}_\mu$ 1π and Other both increase by one unit after the fit, implying the prior model is similar to the posterior model in these studies.

Table 6.13 shows the fit uncertainties broken down by sample and systematic source. As in 2017, the fit reduces the uncertainties on the total event rate by at least one order of magnitude; from 9.9% to 0.3% for the total, 11.7% to 0.5% for 0π, 1.7% to 1.0% for 1π, 12.2% to 1.0% for Other. For the new RHC selections there are reductions of 12.9% to 1.1% for 0π, 23.1% to 4.0% for 1π and 16.3% to 2.8% for Other. Comparing to the 2017 result in Table 5.17, the overall uncertainty on the event rate moves from 0.39% to 0.3% for the total event rate, 0.7% to 0.5% for FGD1 0π, 1.3% to 1.0% for 1π and 1.4% to 1.0% for Other.

Before fitting to data, the largest systematic for the FHC selections is flux and interactions at 8-10%, which generally agrees with cross-section measurements (e.g. [179,202]). After the fit to ND280 data the impact on the uncertainty of the event rate more than halves. For the RHC selections with one or more pions, the detector systematics are the dominant systematic in the prior model: e.g. for FGD1 $\bar{\nu}_\mu$ 1π the prior detector uncertainty is 21% whereas the FHC equivalent is 5.3%. This is primarily due to the larger wrong-sign background in RHC and the proton background in selecting the lepton candidate. After the fit these selections are still dominated by detector systematics, albeit it at a 5% level. For the FHC selections the ND280 uncertainties lay around 1.5-3%, and are instead controlled equally by flux and interaction systematics.

Sample	Event rate			$-2\ln\mathcal{L}_S$	
	Data	Prior	Posterior	Prior	Posterior
FGD1 0π	33553	32149.5 ± 3905.0	33622.9 ± 172.8	1094.4	833.9
FGD1 1π	7757	8086.0 ± 862.5	7971.4 ± 76.3	408.9	310.5
FGD1 Other	8068	6882.7 ± 838.0	7855.6 ± 79.8	714.6	458.6
FGD2 0π	33462	32348.0 ± 3813.5	33402.4 ± 172.8	1158.3	866.2
FGD2 1π	6133	6466.7 ± 665.5	6282.8 ± 67.4	412.2	303.9
FGD2 Other	7664	6636.2 ± 768.7	7473.9 ± 76.3	653.9	414.1
FGD1 $\bar{\nu}_\mu$ 0π	6368	6541.7 ± 839.8	6337.3 ± 70.2	409.8	358.9
FGD1 $\bar{\nu}_\mu$ 1π	535	540.1 ± 124.8	544.3 ± 22.0	53.7	54.8
FGD1 $\bar{\nu}_\mu$ Other	1102	1017.1 ± 166.2	1089.6 ± 30.0	87.0	87.0
FGD2 $\bar{\nu}_\mu$ 0π	6451	6389.3 ± 781.9	6452.2 ± 70.0	441.7	406.3
FGD2 $\bar{\nu}_\mu$ 1π	465	474.3 ± 93.3	471.7 ± 20.8	42.7	40.0
FGD2 $\bar{\nu}_\mu$ Other	1032	932.5 ± 183.0	1036.9 ± 30.6	118.3	104.0
FGD1 ν_μ RHC 0π	2707	2481.0 ± 391.1	2693.8 ± 46.5	185.5	126.5
FGD1 ν_μ RHC 1π	847	857.0 ± 103.3	863.9 ± 25.3	60.7	52.3
FGD1 ν_μ RHC Other	1015	800.4 ± 166.3	1006.9 ± 30.6	153.8	63.4
FGD2 ν_μ RHC 0π	2648	2540.6 ± 413.3	2697.1 ± 46.9	187.6	136.5
FGD2 ν_μ RHC 1π	693	684.7 ± 85.6	690.6 ± 22.7	80.0	66.5
FGD2 ν_μ RHC Other	932	787.3 ± 138.7	937.6 ± 28.6	90.7	53.1
Total	121432	116575.3 ± 11548.9	121431.3 ± 349.1	6353.6	4736.61

Table 6.12.: Event rate and test-statistic for data, pre-fit MC and post-fit MC broken by sample, using the reduced ND280 covariance matrix

Sample	$\delta N/N(\%)$							
	Flux		ND280		Interaction		All	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
FGD1 0π	8.0	3.6	3.3	1.7	9.9	3.4	11.7	0.5
FGD1 1π	7.6	3.1	5.3	2.2	6.5	2.8	10.7	1.0
FGD1 Other	8.1	3.4	6.6	2.4	7.7	2.9	12.2	1.0
FGD2 0π	8.0	3.6	2.8	1.4	9.6	3.3	11.8	0.5
FGD2 1π	7.6	3.1	4.8	2.4	6.6	2.8	10.3	1.1
FGD2 Other	8.1	3.4	5.5	2.4	7.8	2.9	11.6	1.0
FGD1 $\bar{\nu}_\mu$ 0π	7.7	4.4	6.3	2.5	9.5	3.5	12.9	1.1
FGD1 $\bar{\nu}_\mu$ 1π	6.8	4.6	21.0	5.2	7.1	3.2	23.1	4.0
FGD1 $\bar{\nu}_\mu$ Other	6.8	4.5	15.2	3.8	7.4	2.9	16.3	2.8
FGD2 $\bar{\nu}_\mu$ 0π	7.6	4.3	6.7	2.4	9.3	3.6	12.2	1.1
FGD2 $\bar{\nu}_\mu$ 1π	6.8	3.8	19.3	5.1	7.3	3.3	19.7	4.4
FGD2 $\bar{\nu}_\mu$ Other	6.8	4.4	19.0	4.0	7.6	2.9	19.6	3.0
FGD1 ν_μ RHC 0π	7.1	4.3	14.3	3.1	8.4	3.2	15.8	1.7
FGD1 ν_μ RHC 1π	7.2	3.9	8.7	4.1	6.4	2.8	12.1	2.9
FGD1 ν_μ RHC Other	7.7	4.4	19.3	4.4	7.3	2.8	20.8	3.0
FGD2 ν_μ RHC 0π	7.1	4.2	14.9	3.0	8.3	3.0	16.3	1.7
FGD2 ν_μ RHC 1π	7.2	4.2	9.2	4.2	6.8	2.9	12.5	3.3
FGD2 ν_μ RHC Other	7.7	4.3	16.1	4.6	7.3	2.8	17.6	3.1
Total	7.5	4.6	3.4	2.1	7.2	3.0	9.9	0.3

Table 6.13.: Event rate uncertainties from the prior and posterior model broken down by selection and systematic types

Varying the combined full parameter set has a much greater impact on reducing the event rate uncertainties due to the very strong correlations between the flux, interaction and ND280 parameters: e.g. the FGD1 FHC 0π event rate uncertainty from the flux is 3.6%, ND280 is 1.7% and interaction is 3.4% but put together they amount to 0.5%.

6.7.1. Full and Reduced ND280 Systematic Parameterisation

As for the Asimov case in subsection 6.6.3, the data fit is here compared to results using the full and reduced ND280 covariance. Similar results are expected for the two fits, as only small differences were found in the fit to Asimov data.

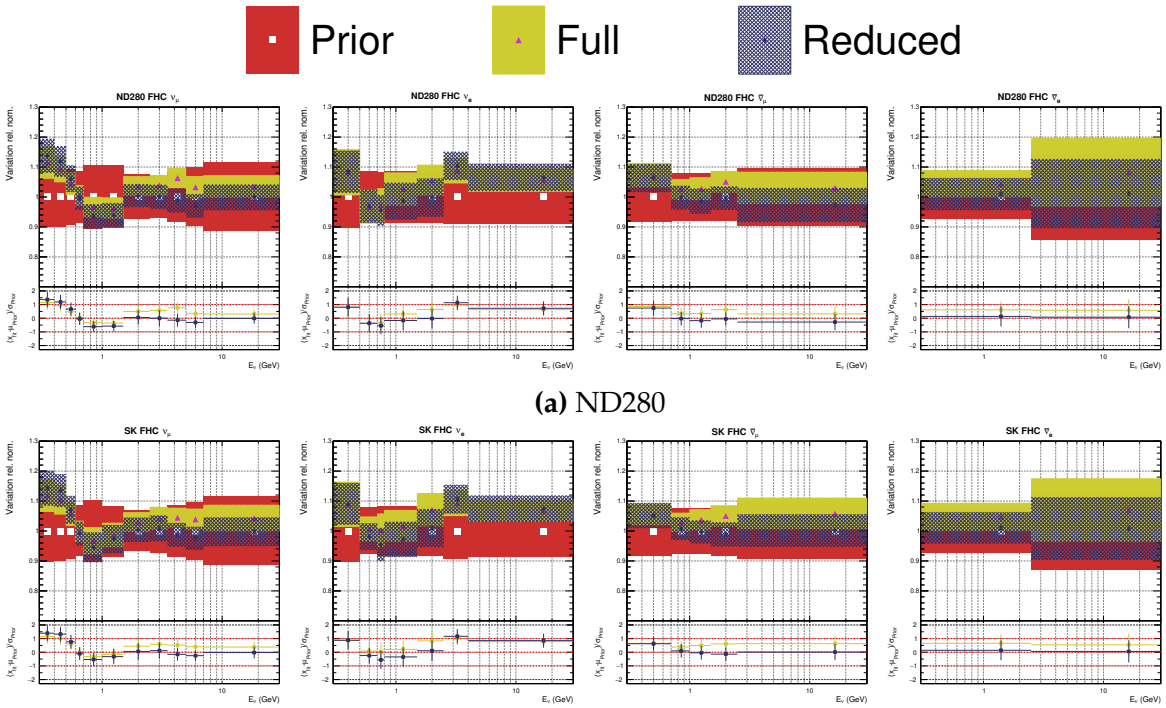
Figure 6.35 shows the FHC flux parameters for the two fits, which are compatible. The divergence happens at higher energies, where it's likely that the reduced covariance matrix (which was binned to mostly ignore shapes at higher energies) is missing freedom to vary events which the full ND280 parameterisation can.

Looking at the shape of the FHC flux parameters, there is a tendency towards the nominal at higher energies and the majority of the shape change happens below 1 GeV. The shape below 1 GeV increases the flux by 10% below 0.5 GeV, sitting at nominal at the flux peak of 0.6 GeV, and then decreases the flux by 7% at 0.7 GeV up until 1.5 GeV. A similar shape was observed in the 2017 analysis in Figure 5.47, although the results are slightly shifted upwards relative the 2017 analysis.

For the RHC flux parameters in Figure 6.35 there is good agreement using the two covariance matrices. However, an entirely different shape of the flux is present in the RHC parameters compared to the FHC parameters. The low energy normalisation is nominal, falling to 95% below 0.5 GeV and then sharply increasing to 10% above 1 GeV, which then sinks to nominal at 3 GeV. The wrong-sign flux is however increased at low energies, similar to the FHC flux parameters.

Additionally, ND280 and SK see a slightly different behaviour around 2 GeV, where ND280 prefers a 10% increased normalisation and the SK parameters sits at nominal. Otherwise the parameters are well mirrored. The two detector parametrisations are only visually different for the $\bar{\nu}_e$ portion of the flux and the 0.7-1.0 GeV $\bar{\nu}_\mu$ bin.

Figure 6.36 shows the interaction parameters which are mostly compatible. Interestingly, M_A^{QE} is pulled even further than in 2017, now fully compatible with bubble



(a) ND280

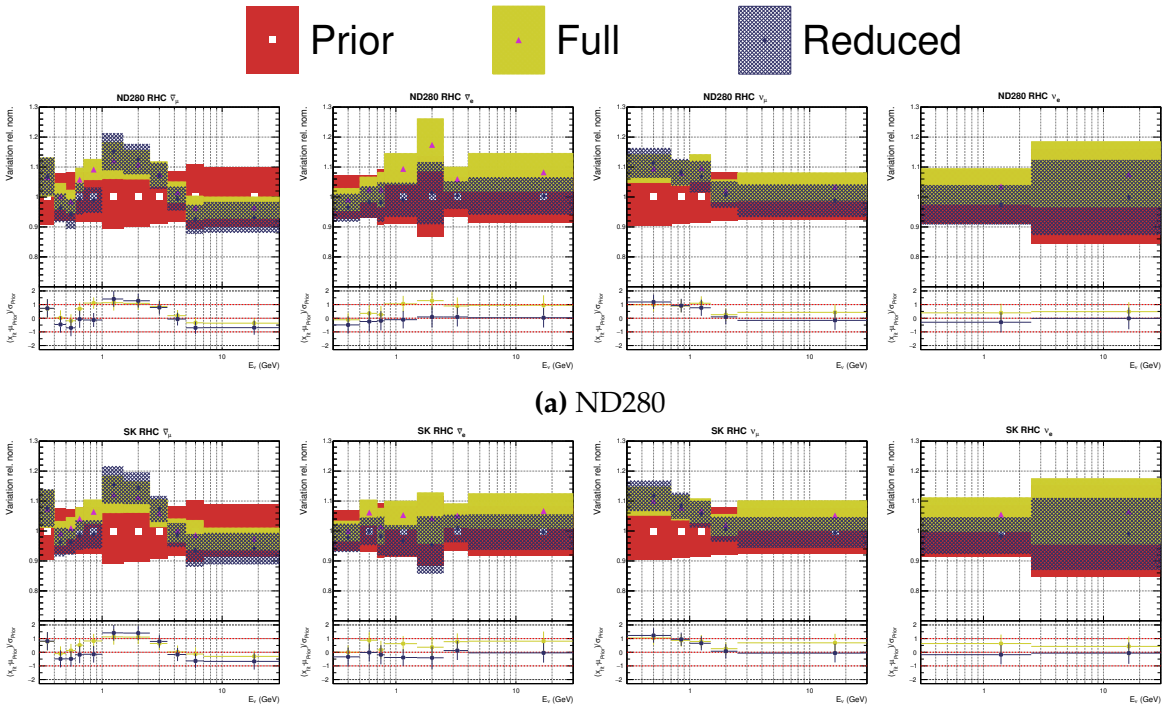
(b) SK

Figure 6.34.: FHC flux parameters, fitting to data with different ND280 matrices

chamber results ($M_A^{QE} = 1.04 \pm 0.06$ GeV). The 2p2h ν normalisation parameter is decreased relative 2017 (1.64 ± 0.21 to 1.31 ± 0.17), whereas the $\bar{\nu}$ parameter increases (0.80 ± 0.26 to 0.91 ± 0.23). Importantly, the 2p2h normalisation parameter for $\bar{\nu}$ is the largest difference for using the two ND280 covariance matrices, just outside of each other's 1σ uncertainty. The 2p2h shape parameters now fit very similar values for ^{12}C and ^{16}O and are not pushed towards any boundaries. The BeRPA A parameter is much closer to the nominal compared to the 2017 result, although the opposite is true for BeRPA B, which now sits 2-3σ from the prior central value.

The single pion parameters largely agree with 2017, albeit with smaller parameter errors. C_5^A decreases slightly but still agrees with the prior uncertainty. The non-resonant $I_{1/2}$ background increases outside the 1σ range but has a large associated error. The CC DIS parameter moves away from the prior—which would be expected if the poor CC Other disagreement still exists. The NC other parameter remains at the upper boundary of the prior 1σ uncertainty.

The pion FSI—which received new priors for this analysis—interestingly depart from their priors and are close to their fitted values in the 2017 analysis. The reduction



(a) ND280

(b) SK

Figure 6.35.: RHC flux parameters, fitting to data with different ND280 matrices

in uncertainty is the most dramatic for the pion FSI parameters, also observed in the Asimov fit.

The two detector parameterisations are mostly compatible, consistently within 1σ of each other.

Figure 6.37 shows how the BeRPA weight has changed relative the 2017 weight and the nominal Nieves RPA. The behaviour below $Q^2 \sim 0.4$ GeV^2 is similar to 2017, although at $Q^2 = 0.5$ GeV^2 the enhancement is much stronger. The BeRPA parameterisation dies off more in accordance with the nominal than in 2017. In general though, the nominal RPA weight is heavily modified in the fit.

6.7.2. Comparing to the Multi-Track Selection

A fit to data using the old multi-track selection with the 2017 binning using the run 2 to 8 data was made. The fit to Asimov data in subsection 5.6.4 showed less bias in the flux parameters for the multi-track selection, although the likelihood scans showed the new selection with updated binning to be more sensitive to parameters.

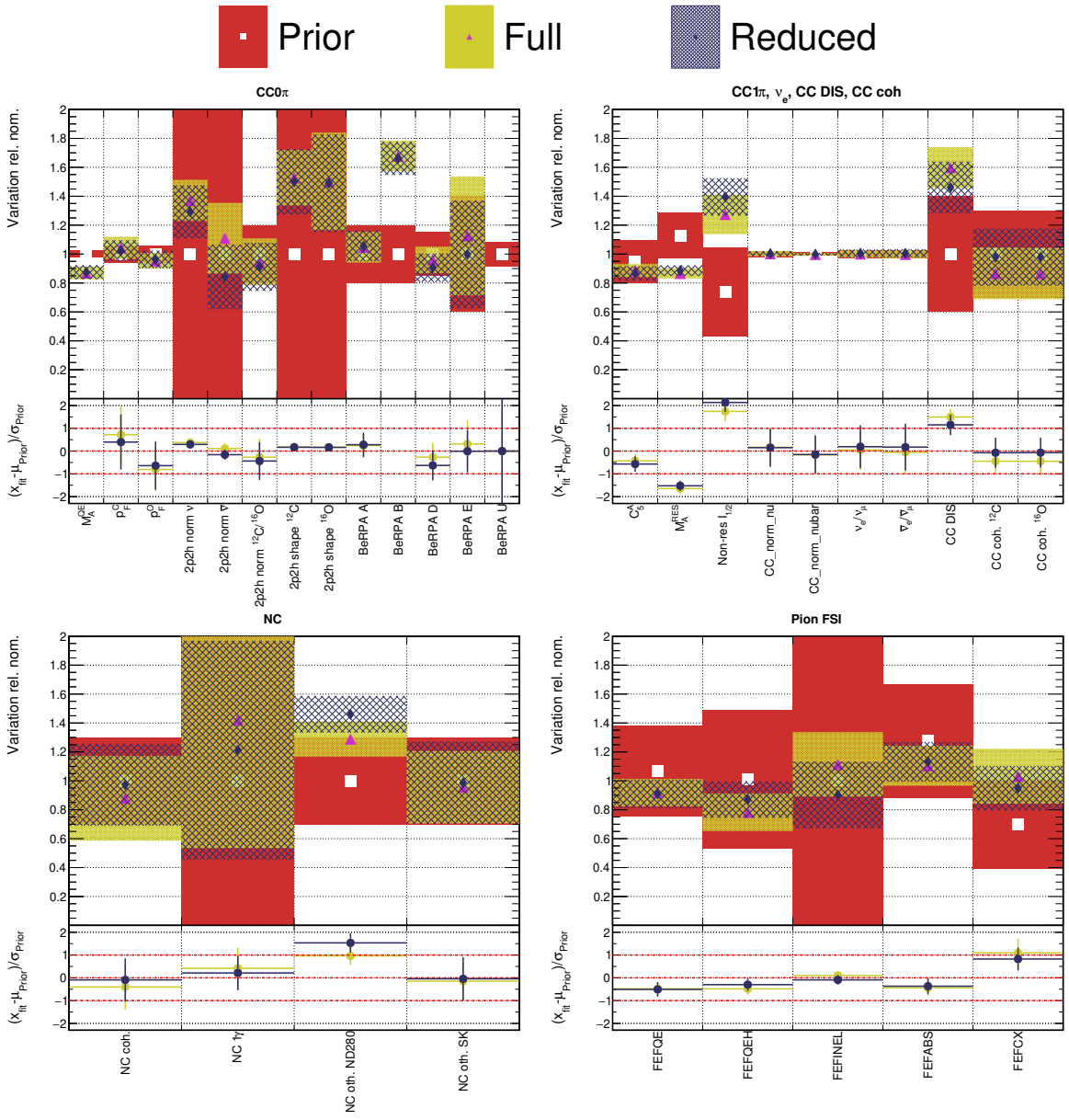


Figure 6.36.: Interaction parameters, fitting to data with different ND280 matrices

Figure 6.38 shows the FHC flux parameters after the data fit. For FHC ν_μ, the multi-track selection sits closer to the priors and does not display the "wavy" oscillation between 0 and 1 GeV, seen when using the multi-π selection. The multi-track FHC ν_μ flux parameters do however increase to $\sim$108% above 1.5 GeV, whereas the multi-π is closer to nominal. The SK FHC ν_μ parameters echo this behaviour. Looking at the other flux parameters (ν_e, $\bar{\nu}_\mu$, $\bar{\nu}_e$), there is a trend for the first bin in the multi-π fit to have a 10% normalisation—likely driven by the low E_ν ν_μ parameters—, which then

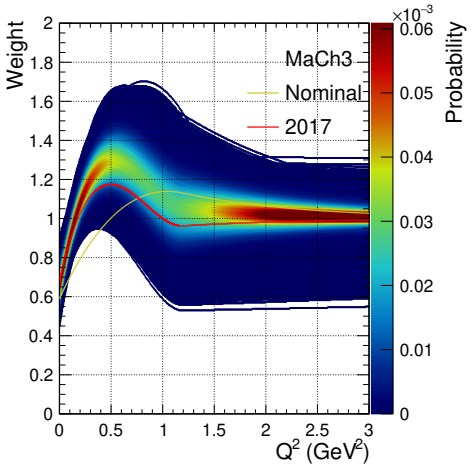
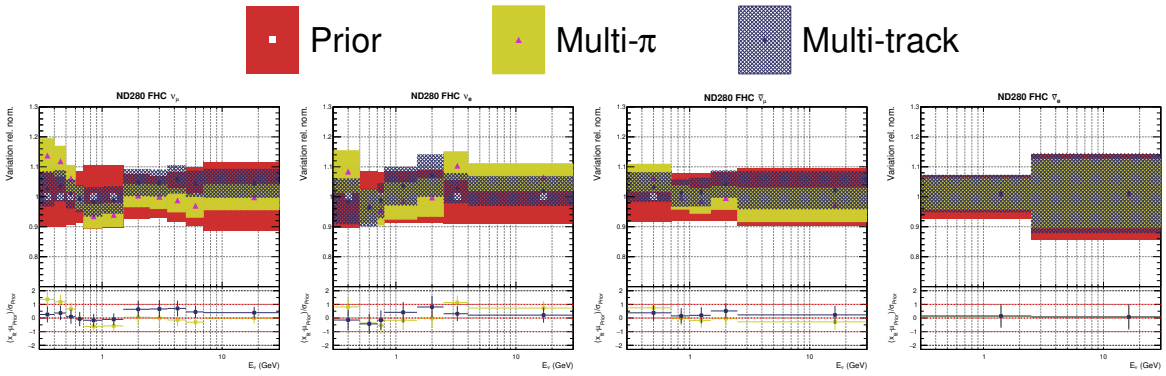
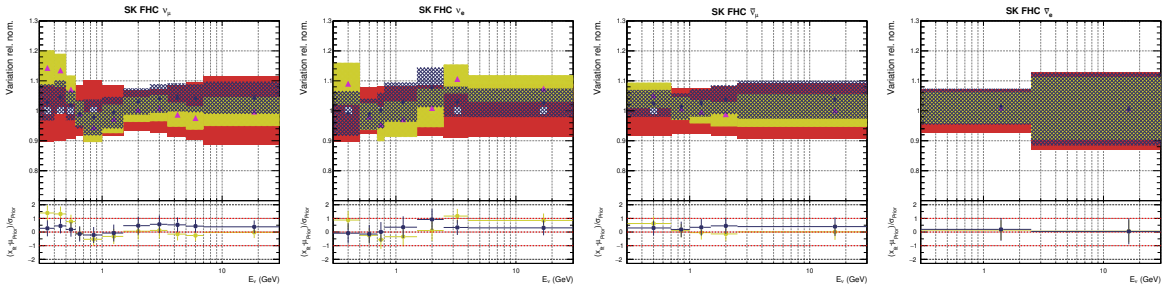


Figure 6.37.: BeRPA weight applied to CCQE events after the fit to data

sits $\mathcal{O}(1\%)$ below the multi-track parameter value. However, for ν_e there is a flip at 2.5 GeV, where multi-π fits a higher flux.



(a) ND280



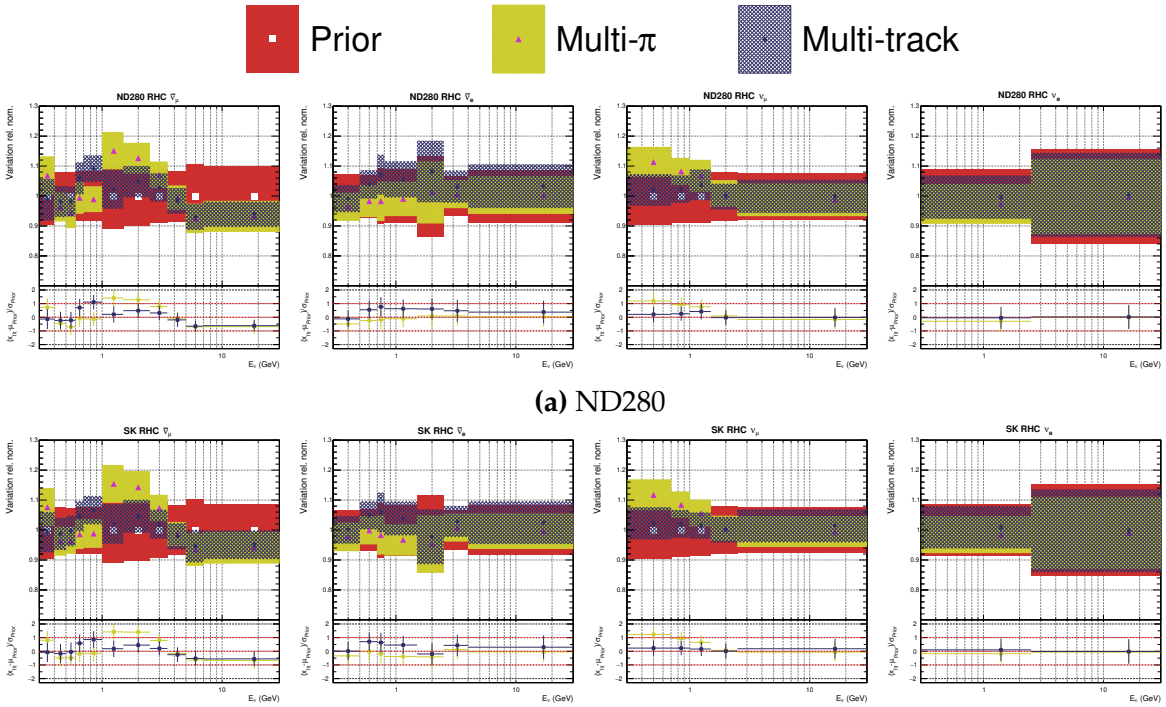
(b) SK

Figure 6.38.: FHC flux parameters, fitting to data with different selections

Inspecting the $\bar{\nu}_\mu$ RHC parameters in Figure 6.39, a similar pattern emerges: the "wavy" nature at low energy is slightly stronger for multi-π. Above 1 GeV, the multi-track sits close to the prior throughout, and the multi-π follows a similar shape but

with more extreme parameter values. For $\bar{\nu}_e$ flux parameters the multi-π is closer to the nominal.

For the wrong-sign component, the multi-π is again enhanced at low E_ν, moving towards the nominal and the multi-track parameter values with increasing E_ν. Again, the two fits consistently sit (just) within 1σ of the prior uncertainty and each other's uncertainties.



(a) ND280

(b) SK

Figure 6.39.: RHC flux parameters, fitting to data with different selections

Finally the cross-section parameters in Figure 6.40 show consistent behaviour up until 2p2h normalisation for $\bar{\nu}$. The multi-π selection prefers a normalisation in line with the prior and has a reduced uncertainty, where the multi-track agrees more with the value of the 2p2h normalisation for ν. Compared to the 2017 value of ~ 0.8, the multi-π selection seems more compatible. The 2p2h ^{12}C $\rightarrow^{16}$ O normalisation parameter shows a preference for the nominal in the multi-track, whereas the multi-π pushes it half-way down the prior.

The 2p2h shape C parameter sits close to the upper boundary for multi-track—similar to results in the 2017 fit—within 1σ of the multi-π selection's central value. The 2p2h shape O parameter shows a similar level of disagreement, with the multi-track

sitting on the prior value, whereas the multi-π selection shows a similar value to the carbon parameter.

The BeRPA parameters are particularly interesting: the multi-track selection is very similar to the 2017 values, whereas the multi-π sees a "softening" of the effect at low Q^2 (BeRPA A), but a larger move from the prior for BeRPA B, even higher than in 2017 in Figure 6.37. This is likely tied to the difference in flux parameters seen previously in Figure 6.39 and how in 2017 fixing BeRPA to nominal caused a flux increase $\mathcal{O}(10\%)$ in Appendix F.

The single pion parameters are entirely compatible and similar to 2017, with the small exception of the non-resonant $I_{1/2}$ background which increases outside the prior. The CC coherent normalisations are similar to 2017 from the multi-track, sitting just on the 1σ boundary, whereas the multi-π agrees with the nominal. Of the NC parameters, only NC other (encompassing NC1π and NC DIS) sees a difference: the multi-track fit prefers a value in agreement with the prior's central value, whereas the multi-π selection pulls outside the prior uncertainty. The pion FSI parameters are expected to show large differences, as was the case in the Asimov studies of the fit and likelihood scans. This comes both from a rebinning of the FHC 1π samples and the splitting of the NTrack into 1π and Other for RHC selections. The parameters are largely compatible, and the multi-π consistently sits close to the prior value. Both fits still sit within the 1σ uncertainty of the prior.

6.7.3. ND280 Detector Systematic Parameters

In this section the fitted ND280 detector parameters are inspected in detail to shine light on how they are being varied in the fit to data. Only the reduced ND280 covariance matrix is considered, which is here fit to data. The parameters are sliced into bins of $\cos\theta_\mu$ (primary) and p_μ (secondary).

Figure 6.41 shows the FGD1 0π parameters, which all sit within the 1σ of the prior uncertainty. There's a general trend to increase the normalisation with p_μ up until $\cos\theta_\mu = 0.965$. Additionally, most of the parameter uncertainties are approximately 2/3 of their prior uncertainties. In the most forward-going region the parameters all favour their prior value. Low $\cos\theta_\mu$, high p_μ have the largest prior uncertainties of 25-60% up until $\cos\theta_\mu = 0.88$, where the uncertainties become approximately similar throughout p_μ at 5-10%. Looking at FGD2 in Figure 6.42 the patterns of FGD1 are essentially

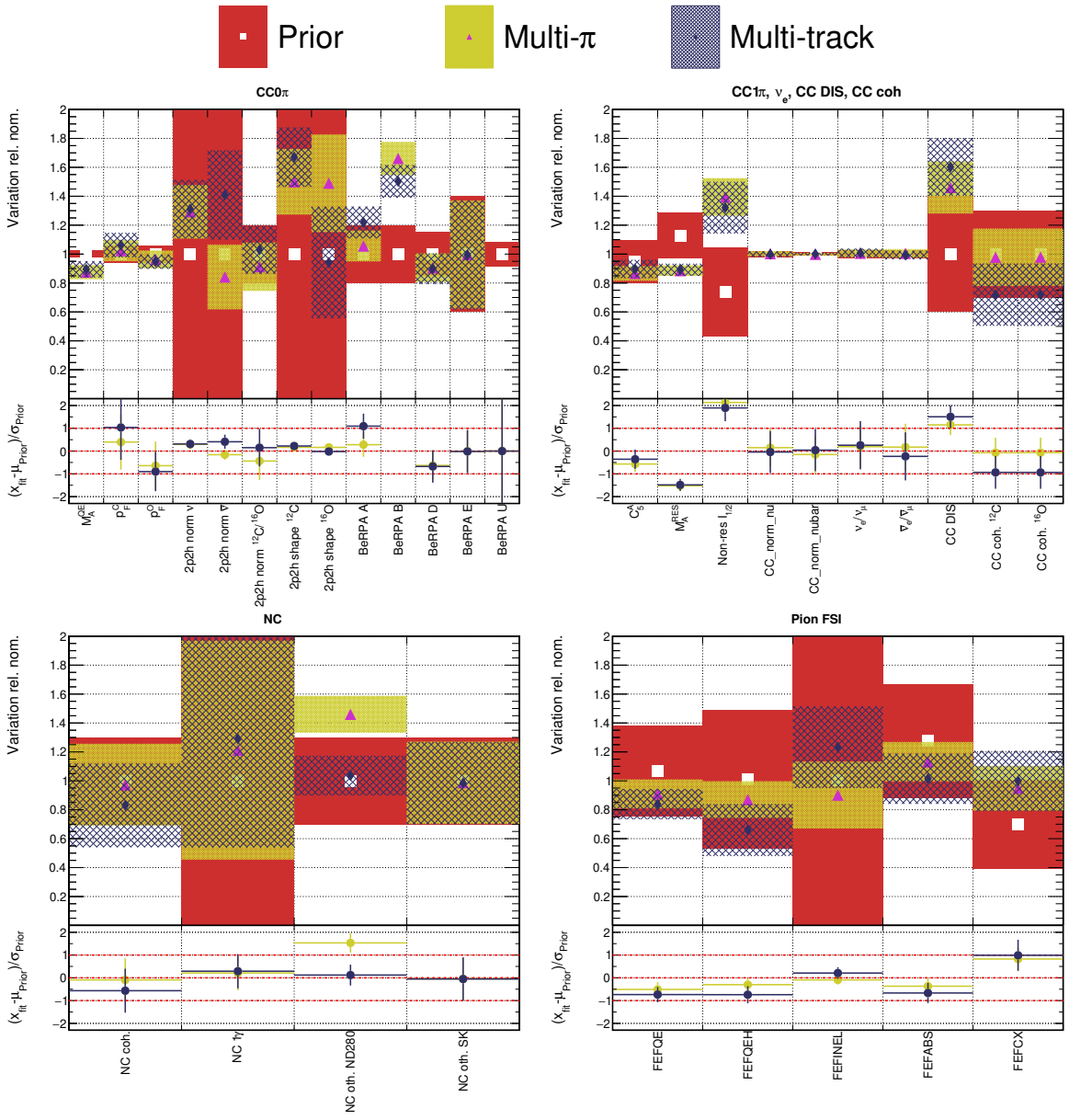


Figure 6.40.: Interaction parameters, fitting to data with different selections

repeated. The FGD1 1π parameters in Figure 6.43 don't show the same increase with p_μ and no clear pattern is visible. The parameters sit within 1σ of the prior and the postfit uncertainty is again roughly 2/3 of the prior. The low $\cos\theta_\mu$, high p_μ events again dominate the prior uncertainty up until $\cos\theta_\mu = 0.90$, similar to the 0π selection. The FGD2 1π selection in Figure 6.44 shares most features with FGD1, although a pattern of increasing normalisation with p_μ in the low $\cos\theta_\mu$ region is evident, similar to the 0π selections. Otherwise the parameters appear well behaved. The FGD1 Other parameters are shown in Figure 6.45, which show the largest difference to the prior of

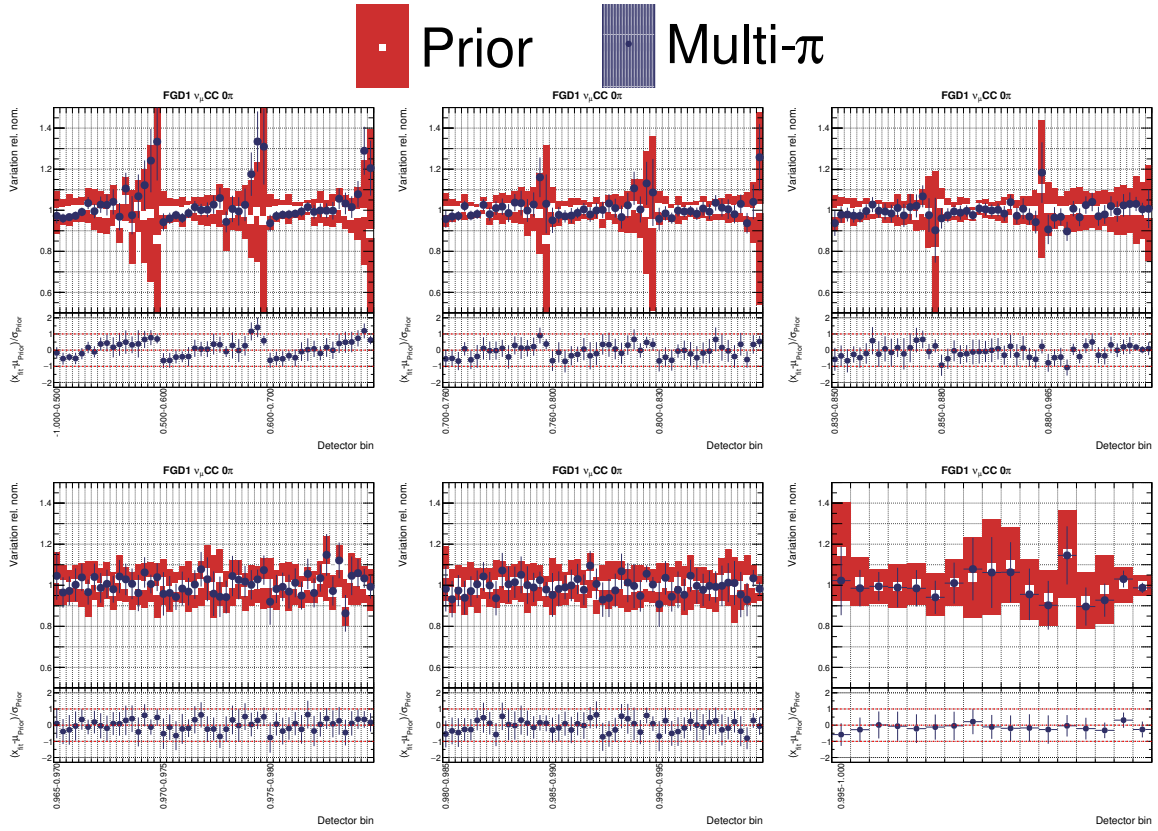


Figure 6.41.: FGD1 0π detector parameters

the detector parameters. This is anticipated due to the poor prior model description of the data, also seen in 2017. The model generally underestimated the data, and so the detector parameters are often pushed to $+1\sigma$ and beyond. The FGD2 Other selection in Figure 6.46 sees a similar pattern to FGD1, although less pushed to $+1\sigma$. The FGD1 RHC 0π parameters presented in Figure 6.47, show similar behaviour to their FHC counterparts: large uncertainties at low $\cos\theta_\mu$, high p_μ; high normalisation at high p_μ; and a forward-going region which largely agrees with the prior. FGD2 $\bar{\nu}_\mu$ 0π in Figure 6.48 follows the FGD1 equivalent, although some priors have different central value, notably at high p_μ. The $\bar{\nu}_\mu$ 1π and other selections for FGD1 in Figure 6.49 have parameters mostly consistent with the prior and well within 1σ. Notably, the high momentum bins are well constrained because the aggressive bin-merging strategy. The FGD2 equivalent selection in Figure 6.50 shows quite different behaviour to FGD1, with 1σ excursions common for both 1π and Other selections. The FGD1 ν_μ RHC 0π selections in Figure 6.51 are again consistent with the previous 0π selections, in which the low $\cos\theta_\mu$, low p_μ regions have a decreased normalisation, moving upward with p_μ and resetting for the next $\cos\theta_\mu$ bin. Overall, the normalisation effect prefers a value less than the prior. The highest $\cos\theta_\mu$, p_μ bin is very well constrained after the fit due

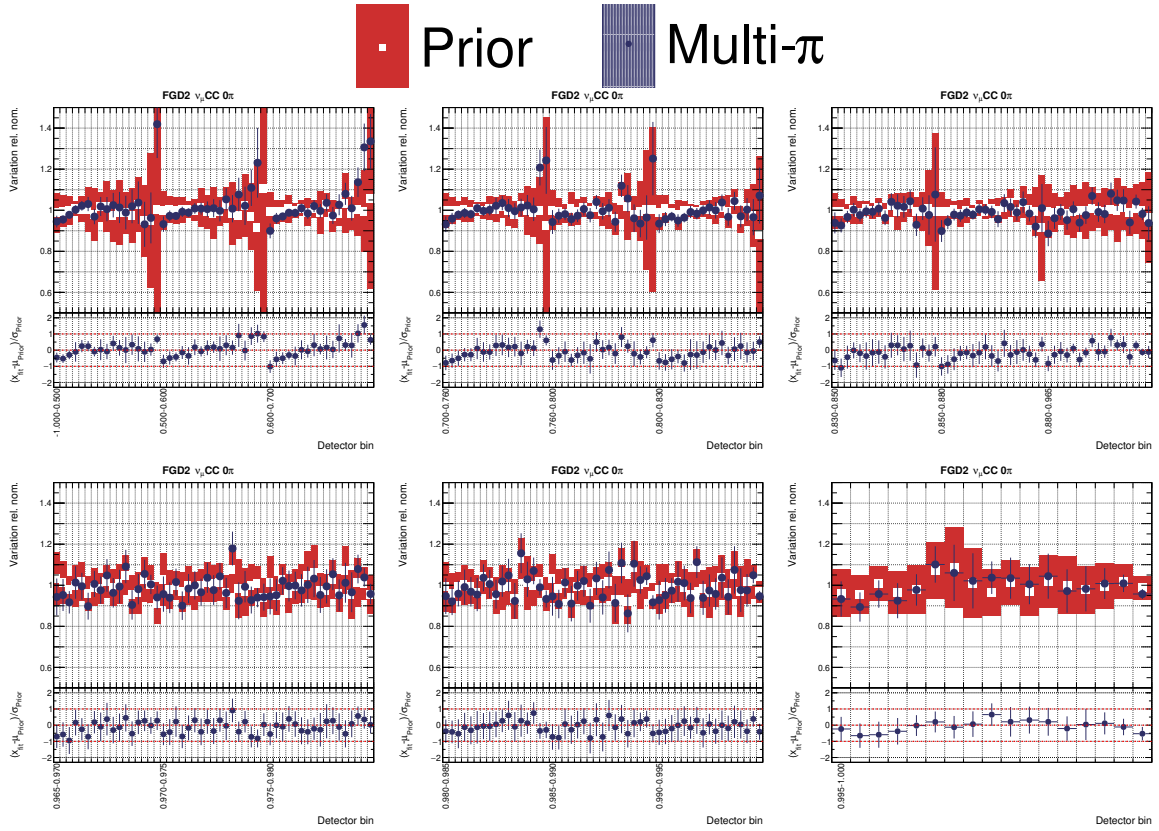


Figure 6.42.: FGD2 0π detector parameters

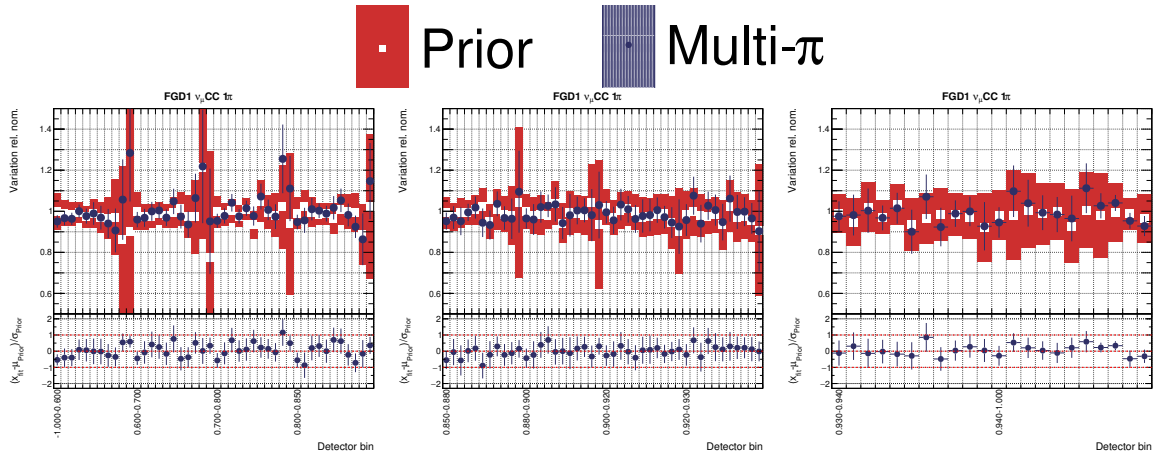


Figure 6.43.: FGD1 1π detector parameters

to the relatively high statistics in this region. The parameters are all consistent with 1σ, although excursions exceeding that happen twice. The 1π and Other distributions are consistent with the prior and the data barely constrains the parameters. Finally the FGD2 ν_μ RHC parameters in Figure 6.52 show consistency with FGD1 both in the prefit and the postfit. In conclusion, the ND280 parameter plots in this section showed

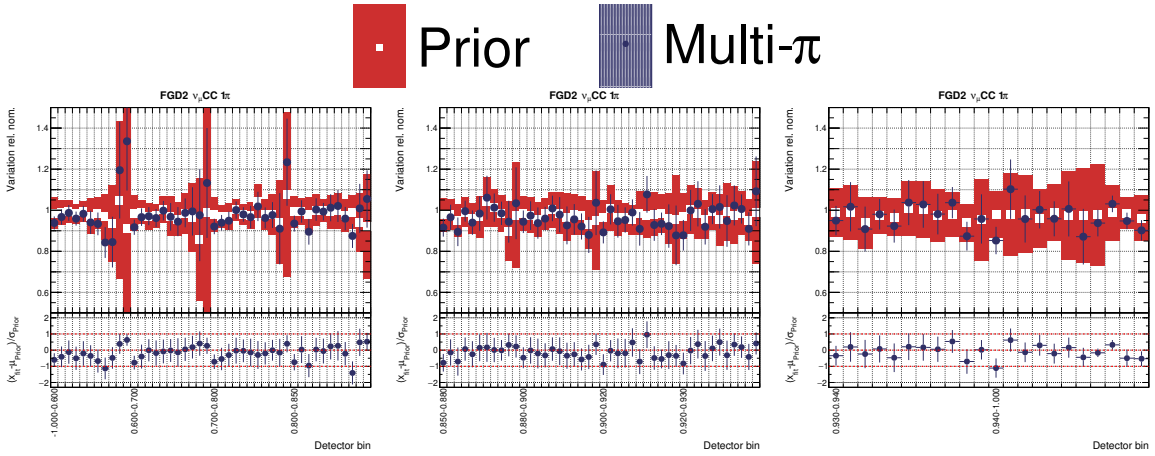


Figure 6.44.: FGD2 1π detector parameters

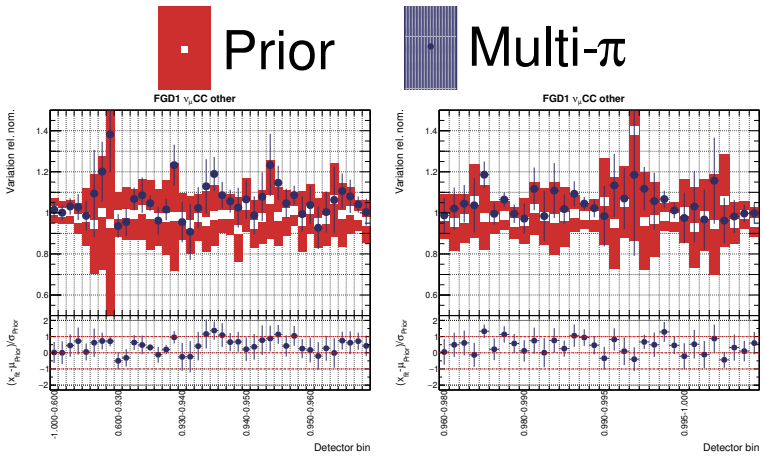


Figure 6.45.: FGD1 Other detector parameters

that:

- In the low $\cos\theta_\mu$ region, the 0π selections uniformly prefer a normalisation less than the prior for low p_μ, moving to a larger normalisation at higher p_μ. This behaviour ceases around $\cos\theta_\mu$ =0.9 for all 0π selections
- The largest prior uncertainties are in low $\cos\theta_\mu$ and high p_μ, where the error is between 20-60%. This is seen in 0π, 1π and Other selections
- The FHC ν_μ CCOther parameters all sit above the prior in the postfit to data, likely to remedy the large underestimation of the data in the prior model
- FGD1 and FGD2 are consistent in both the prior and posterior parameter values
- Most detector parameters are within 1σ of the uncertainty given by the prior, and only small excursions happen beyond that

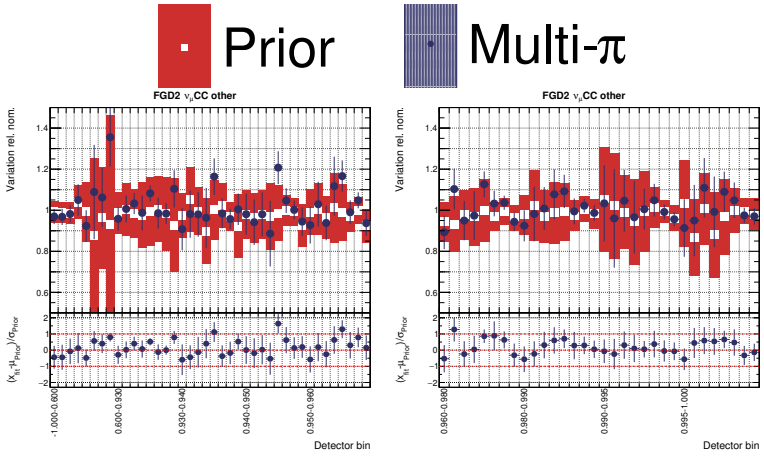


Figure 6.46.: FGD2 Other detector parameters

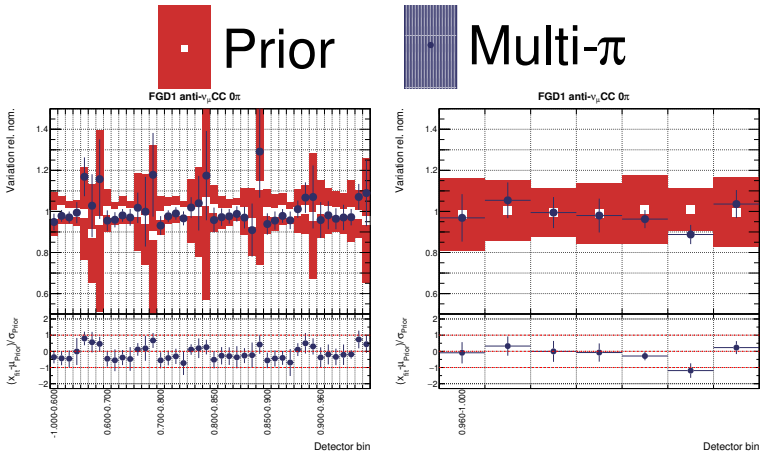


Figure 6.47.: FGD1 $\bar{\nu}_\mu$ 0π detector parameters

The unbinned full detector matrix largely agrees with the above conclusions but has increased granularity to make minute adjustments in each fit bin.

6.7.4. Covariance Matrix from the Data Fit

The postfit covariance and correlation matrix are found in Figure 6.53. Many familiar features from the 2017 analysis are present: M_A^{QE}, BeRPA and 2p2h normalisation correlating with the 0.5-1.0 GeV flux, CC DIS correlating strongly with the flux. M_A^{QE} is strongly negatively correlated with BeRPA B and D and BeRPA A is correlated with p_F. The 2p2h shape parameters are correlating with the 2p2h normalisation parameters and the p_F parameters with each other. The single pion parameters are

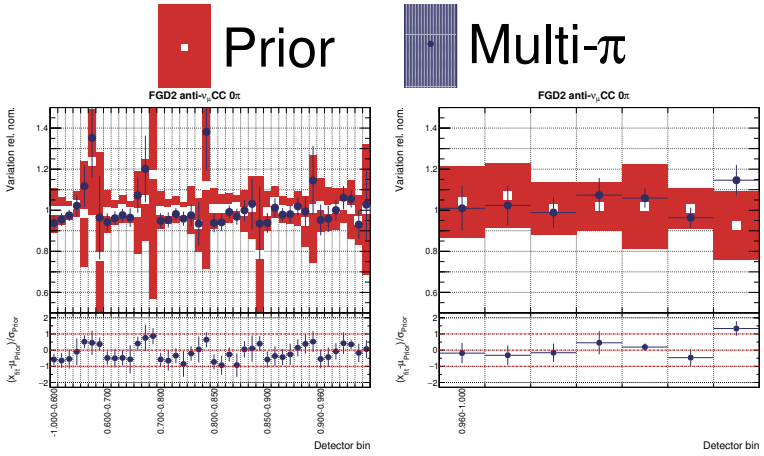
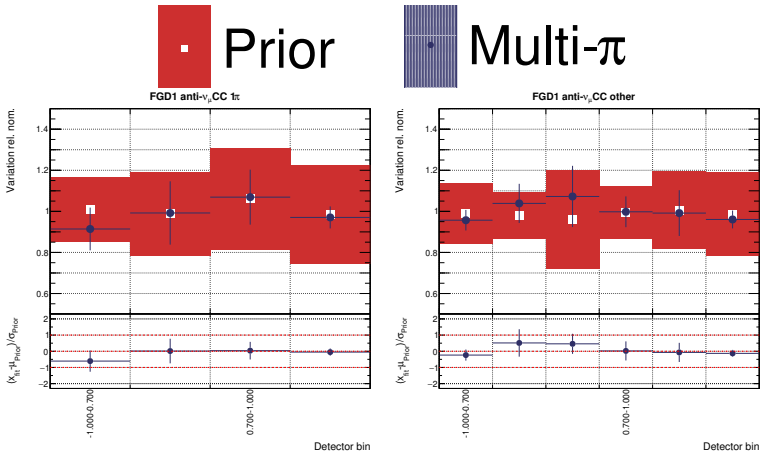


Figure 6.48.: FGD2 $\bar{\nu}_\mu$ 0π detector parameters

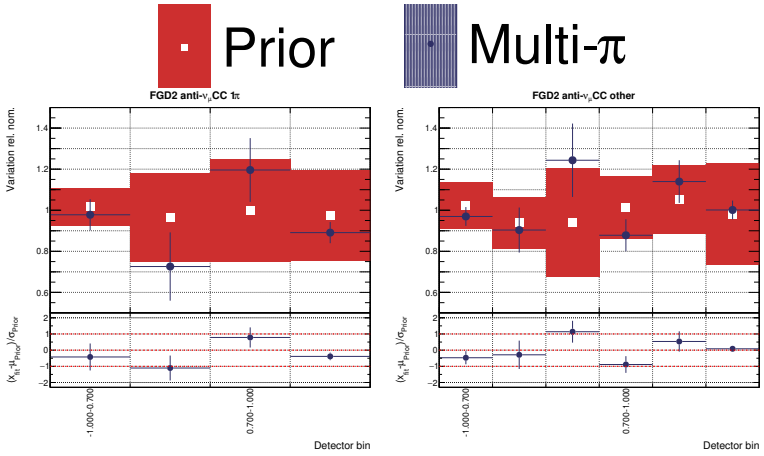


(a) 1π **(b)** Other

Figure 6.49.: FGD1 $\bar{\nu}_\mu$ 1π and Other selections detector parameters

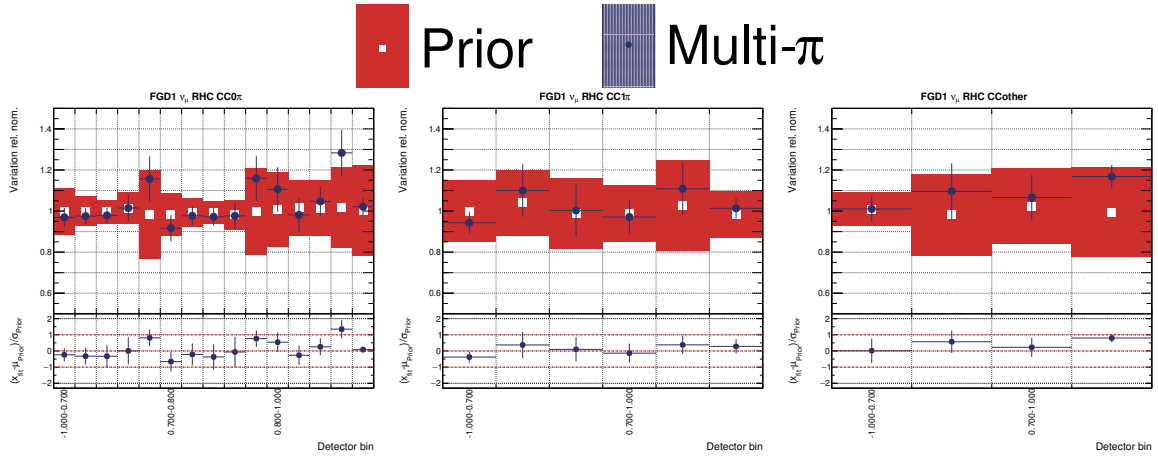
in turn correlating with each other and the CC coherent, and the FSI parameter correlations are washed out.

Figure 6.54 shows the ratio of 2018 to 2017 covariance and correlation matrices. A few parameters flip sign (negative entries), although they are a minority. The majority of bins decrease in uncertainty (cyan in $V_{i,j}$), although 2p2h shape C and O increase; however this is due to the parameters not being pushed against boundaries in the 2018 analysis. Most correlations decrease in strength—notably the flux parameters and 2p2h normalisations—although BeRPA and 2p2h shape drastically increase as BeRPA loses correlations with most flux parameters. M_A^{RES} interestingly loses correlations with most of the flux parameters.



(a) 1π **(b)** Other

Figure 6.50.: FGD2 $\bar{\nu}_\mu$ 1π and Other selections detector parameters



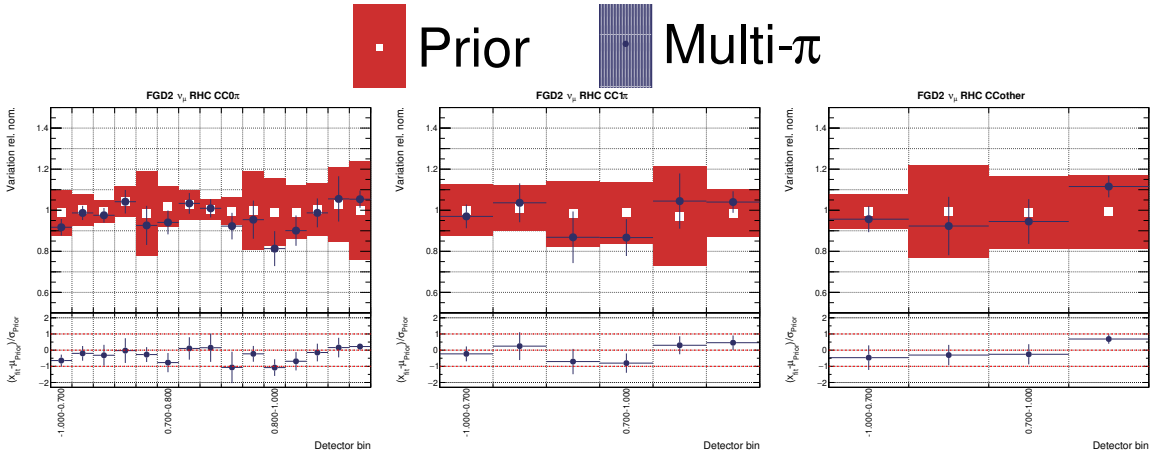
(a) 0π **(b)** 1π **(c)** Other

Figure 6.51.: FGD1 ν_μ RHC selections detector parameters

6.7.5. Prior Predictive Spectrum

In 2017 the prior model was deemed entirely incompatible with the data (subsection 5.7.1), resulting in p-values of 0.0 for every selection. The study is repeated here for the 2018 analysis.

Table 6.14 shows the prior predictive p-values, which are zero for all the high-statistics samples. The new lower statistics RHC 1π and Other selections have non-zero p-values, largely due to the sample size being less than 1000 in Table 6.12. The new RHC $\bar{\nu}_\mu$ 1π selections are all above 5%, as is FGD1 $\bar{\nu}_\mu$ Other and FGD2 ν_μ RHC Other. Generally, small improvements on the 2017 results are present. The p-values



(a) 0π **(b)** 1π **(c)** Other

Figure 6.52.: FGD2 ν_μ RHC selections detector parameters

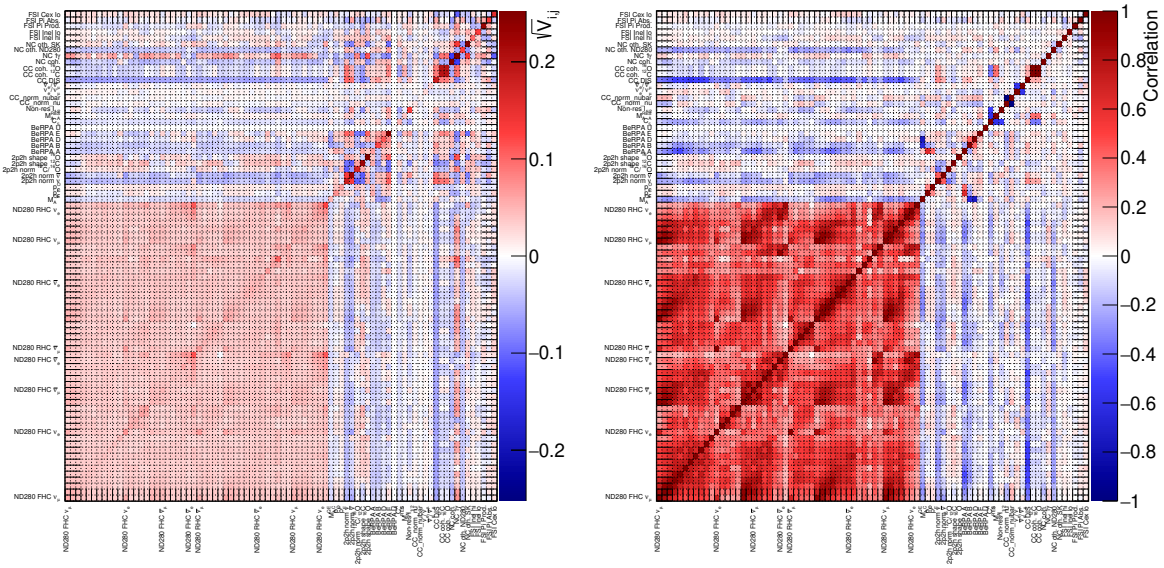


Figure 6.53.: $\sqrt{V_{i,j}}$ and correlation matrices for 2018

are in agreement with the 2017 values in Table 5.18, and the statistics increase of anti-neutrino data has brought the 1Trk (aka 0π) p-value to zero.

6.7.6. Posterior Predictive Spectrum

The posterior predictive p-values are shown by sample in Table 6.15. Four selections have bad p-values: FGD1 CCOther, FGD2 CCOther, FGD2 $\bar{\nu}_\mu$ CC0π, and FGD2 ν_μ RHC 1π. Whereas the first was expected from the 2017 results in Table 5.19, the two latter are new. We also see a low p-value ($\sim$1%) for the FGD1 $\bar{\nu}_\mu$ 0π selection.

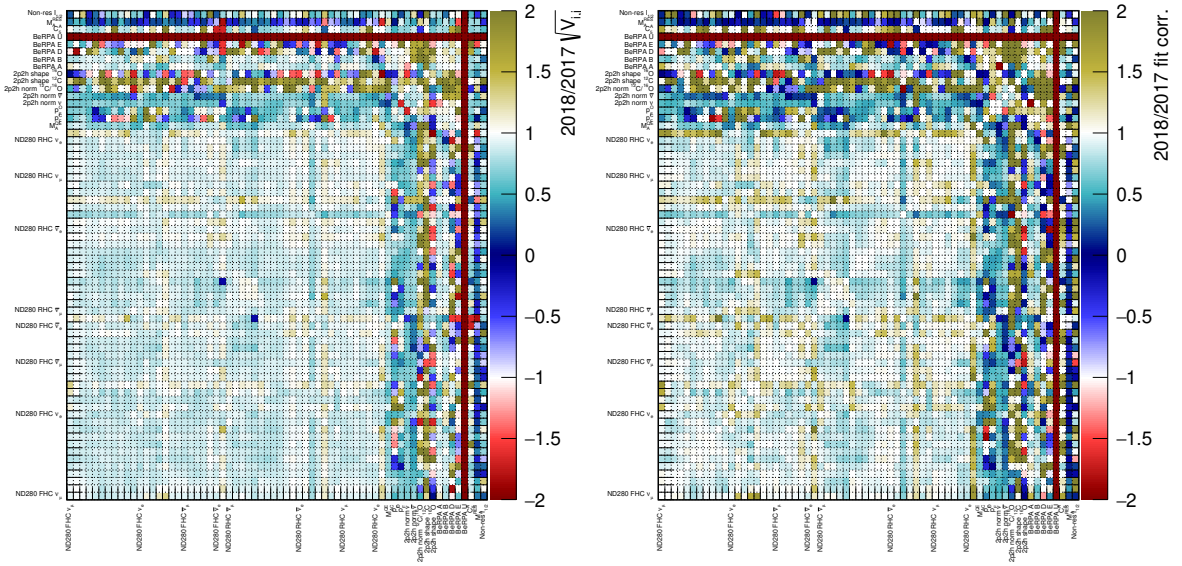


Figure 6.54.: $\sqrt{V_{i,j}}$ and correlation matrix ratios for 2018/2017

The $\bar{\nu}_\mu$ 0π selections are particularly worrying—in 2017 these were 0.515 and 0.265 respectively. There are large differences between FGD1 and FGD2 p-values, e.g. FHC CC0π for FGD1 is 0.215 whereas for FGD2 is 0.072, or FGD1 $\bar{\nu}_\mu$ Other is 0.271 whereas FGD2 is 0.041.

The multi-track p-values are shown in Table 6.16, in the poor CCOther p-value for FGD1 and FGD2 strikes again. Curiously, FGD2 CC0π is now the good p-value whereas FGD1 CC0π has a poor p-value. Furthermore, FGD1 1π has a bad p-value, not seen in the multi-π case.

The RHC 1Trk selections are all well described with p-values of 0.156 (FGD1) and 0.162 (FGD2), in contrast to the CC0π equivalents in the multi-π samples above. This is likely due to the low statistics.

Looking at the p-values from using the full detector covariance matrix in Table 6.17, FGD1 CCOther, FGD2 CCOther, FGD2 $\bar{\nu}_\mu$ 0π and FGD2 ν_μ RHC 1π have the lowest p-values although FGD2 CCOther and FGD2 ν_μ RHC 1π are acceptable. The p-values are roughly $2 \sim 4$ of those from the reduced detector matrix in Table 6.15, with similar relative strengths.

This indicates that the ND280 parametrisations for both the reduced and multi-track versions fail to capture the nuances required in the simulation to adequately describe the data. However, increasing the dimensionality is problematic, as convergence times

Sample	Draw Fluc.	Pred. Fluc.
FGD1 0π	0.000	0.000
FGD1 1π	0.000	0.000
FGD1 Other	0.000	0.000
FGD2 0π	0.000	0.000
FGD2 1π	0.000	0.000
FGD2 Other	0.000	0.000
FGD1 $\bar{\nu}_\mu$ 0π	0.000	0.000
FGD1 $\bar{\nu}_\mu$ 1π	0.073	0.071
FGD1 $\bar{\nu}_\mu$ Other	0.058	0.062
FGD2 $\bar{\nu}_\mu$ 0π	0.000	0.000
FGD2 $\bar{\nu}_\mu$ 1π	0.209	0.211
FGD2 $\bar{\nu}_\mu$ Other	0.002	0.002
FGD1 ν_μ RHC 0π	0.003	0.003
FGD1 ν_μ RHC 1π	0.010	0.009
FGD1 ν_μ RHC Other	0.018	0.019
FGD2 ν_μ RHC 0π	0.000	0.000
FGD2 ν_μ RHC 1π	0.000	0.000
FGD2 ν_μ RHC Other	0.072	0.072

Table 6.14.: Prior predictive p-values for each sample after the data fit

for the full detector matrix parametrisation are an order of magnitude longer, making such fits unfeasible for joint ND280+SK fits.

Finally, the event rates with the reduced and full detector matrix fits are shown in Table 6.18. As expected from the p-values, the sample test-statistic for the posterior predictive distribution is ~ 400 units improved, bringing the $-2\ln\mathcal{L}_S/\text{nBins}$ to 1.02. Although in many cases the overall event rate is better for the reduced covariance matrix, the test-statistic is worse (except for FGD2 $\bar{\nu}_\mu$ 1π).

The contributions to the sample test-statistic by $p_\mu \cos\theta_\mu$ is shown in Figure 6.55. It confirms that the dominant likelihood contributions are shared in both fits, but the full matrix has a slightly smaller contribution, and the smaller contributions are washed out more.

The one-dimensional p-values are calculated and the results are shown in Figure 6.56. The calculation confirms the two dimensional p-values: using the full

Sample	Draw Fluc.	Pred. Fluc.
FGD1 0π	0.215	0.211
FGD1 1π	0.071	0.075
FGD1 Other	0.000	0.000
FGD2 0π	0.072	0.072
FGD2 1π	0.128	0.132
FGD2 Other	0.002	0.002
FGD1 $\bar{\nu}_\mu$ 0π	0.012	0.012
FGD1 $\bar{\nu}_\mu$ 1π	0.241	0.239
FGD1 $\bar{\nu}_\mu$ Other	0.271	0.271
FGD2 $\bar{\nu}_\mu$ 0π	0.000	0.000
FGD2 $\bar{\nu}_\mu$ 1π	0.773	0.764
FGD2 $\bar{\nu}_\mu$ Other	0.041	0.042
FGD1 ν_μ RHC 0π	0.267	0.265
FGD1 ν_μ RHC 1π	0.079	0.079
FGD1 ν_μ RHC Other	0.170	0.170
FGD2 ν_μ RHC 0π	0.121	0.120
FGD2 ν_μ RHC 1π	0.007	0.007
FGD2 ν_μ RHC Other	0.479	0.476

Table 6.15.: Posterior predictive p-values for each sample after the data fit

parameterisation appears a good representation of the data with $p = 0.400 - 0.428$ depending on method. The reduced parameterisation has $p = 0.000$, and lies in the very tail of the two distributions.

6.7.7. Post-fit Distributions

The prefit and postfit distributions are here compared for the full fit to data using the reduced covariance matrix. The likelihood contributions in $p_\mu \cos\theta_\mu$ are inspected rather than event distributions.

Two-dimensional Likelihood Distributions

Figure 6.57 shows the post-fit distributions in $p_\mu \cos\theta_\mu$. For the 0π selections the largest contributions are above 1 GeV for FGD1, whereas FGD2 has plenty contributions in the

Sample	Draw Fluc.	Pred. Fluc.
FGD1 0π	0.053	0.056
FGD1 1π	0.007	0.006
FGD1 Other	0.000	0.000
FGD2 0π	0.413	0.417
FGD2 1π	0.251	0.252
FGD2 Other	0.000	0.000
FGD1 $\bar{\nu}_\mu$ 1Trk	0.156	0.153
FGD1 $\bar{\nu}_\mu$ NTrk	0.391	0.387
FGD2 $\bar{\nu}_\mu$ 1Trk	0.162	0.164
FGD2 $\bar{\nu}_\mu$ NTrk	0.353	0.350
FGD1 ν_μ RHC 1Trk	0.208	0.205
FGD1 ν_μ RHC NTrk	0.693	0.691
FGD2 ν_μ RHC 1Trk	0.029	0.031
FGD2 ν_μ RHC NTrk	0.629	0.626

Table 6.16.: Posterior predictive p-values for each sample after the data fit, using the multi-track selection

$p_\mu = 0 - 1$ GeV, $\cos\theta_\mu = 0.9 - 1.0$ area. FGD2 also has three large contributions in the $\cos\theta_\mu = 0.6 - 0.75$ area: one in the first momentum bin, and two in the 2-5 GeV area, which aren't present in FGD1. For 1π selections, the picture is more consistent. At high angle, low p_μ, FGD1 has dotted contributions whereas FGD2 has barely any. There is possibly some constant Q^2 behaviour, especially at low Q^2. The largest contributions to the likelihood happen above 1.5 GeV, with the exception of the high-angle events with $p_\mu \sim 0.6$ GeV. The Other distributions are less clear, with large likelihood contributions dotted around the distributions: there is potentially some shape in Q^2 around 0.1 GeV^2 and 1-3 GeV appears to contain the largest contributions for FGD2, whereas for FGD1 it continues in lower p_μ.

The new RHC multi-π selections are shown in Figure 6.58. The 0π selections have their largest contributions in 0.4-1 GeV region, and FGD1 additionally has large contributions in the forward-going high p_μ bins. There are some contributions scattered around 1.5 GeV, indicating weakness in the proton modelling[b]. The 1π selections barely have any contributions to the test-statistic due to the sample size and relatively good modelling. There are two isolated bins in FGD1 with high contributions, which

[b]Since the lepton candidate at this energy has a high probability of being a proton

Sample	Draw Fluc.	Pred. Fluc.
FGD1 0π	0.521	0.514
FGD1 1π	0.196	0.197
FGD1 Other	0.000	0.001
FGD2 0π	0.216	0.220
FGD2 1π	0.162	0.165
FGD2 Other	0.053	0.055
FGD1 $\bar{\nu}_\mu$ 0π	0.080	0.077
FGD1 $\bar{\nu}_\mu$ 1π	0.504	0.505
FGD1 $\bar{\nu}_\mu$ Other	0.444	0.439
FGD2 $\bar{\nu}_\mu$ 0π	0.004	0.003
FGD2 $\bar{\nu}_\mu$ 1π	0.551	0.550
FGD2 $\bar{\nu}_\mu$ Other	0.092	0.088
FGD1 ν_μ RHC 0π	0.412	0.407
FGD1 ν_μ RHC 1π	0.187	0.187
FGD1 ν_μ RHC Other	0.287	0.288
FGD2 ν_μ RHC 0π	0.233	0.229
FGD2 ν_μ RHC 1π	0.026	0.026
FGD2 ν_μ RHC Other	0.506	0.510

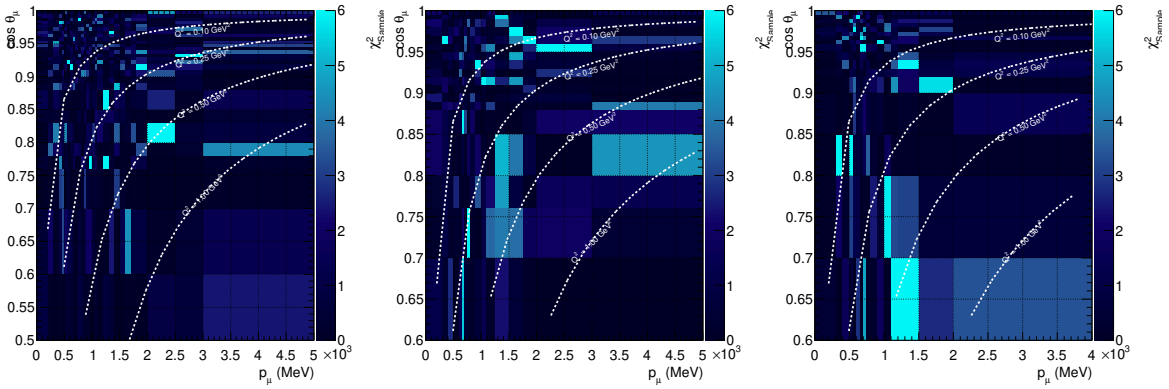
Table 6.17.: Posterior predictive p-values for each sample after the data fit, using the full ND280 covariance matrix

aren't present in FGD2. The Other selections have larger contributions, focussed in the 0.5-1.5 GeV region. FGD1 additionally has a large contribution in a high p_μ bin which FGD2 does not.

The ν_μ RHC likelihood contributions are shown in Figure 6.59. For 0π, FGD2 has a large contribution at low p_μ where FGD1 barely has any. FGD1 sees some contributions between 1-2 GeV which are also present in FGD2. The low statistics 1π selections are again barely contributing to the test-statistic other than FGD2 in one bin. For the Other distribution there is consistent behaviour for FGD1 and FGD2, moving diagonally downwards in $p_\mu \cos\theta_\mu$.

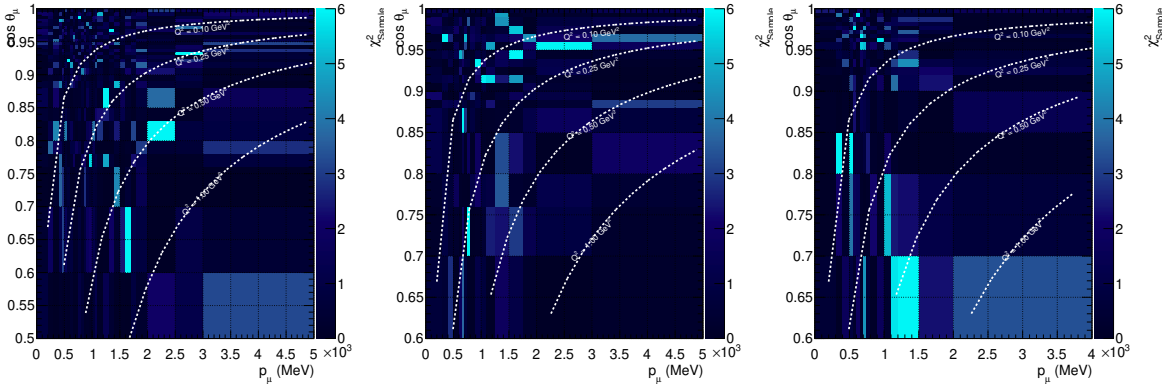
Sample	Post Pred. Event rate			$-2\ln\mathcal{L}_S$	
	Data	Reduced	Full	Reduced	Full
FGD1 0π	33553	33622.9 ± 172.8	33570.7 ± 168.4	833.9	770.9
FGD1 1π	7757	7971.4 ± 76.3	7973.8 ± 70.0	310.5	291.5
FGD1 Other	8068	7855.6 ± 79.8	7854.9 ± 69.6	458.6	410.9
FGD2 0π	33462	33402.4 ± 172.8	33477.7 ± 168.1	866.2	810.5
FGD2 1π	6133	6282.8 ± 67.4	6327.5 ± 55.6	303.9	299.1
FGD2 Other	7664	7473.9 ± 76.3	7421.0 ± 69.4	414.1	361.3
FGD1 $\bar{\nu}_\mu$ 0π	6368	6337.3 ± 70.2	6393.1 ± 61.7	358.9	324.5
FGD1 $\bar{\nu}_\mu$ 1π	535	544.3 ± 22.0	540.9 ± 11.2	54.8	45.0
FGD1 $\bar{\nu}_\mu$ Other	1102	1089.6 ± 30.0	1142.3 ± 17.9	87.0	75.9
FGD2 $\bar{\nu}_\mu$ 0π	6451	6452.2 ± 70.0	6337.9 ± 60.1	406.3	365.6
FGD2 $\bar{\nu}_\mu$ 1π	465	471.7 ± 20.8	506.8 ± 10.6	40.0	44.4
FGD2 $\bar{\nu}_\mu$ Other	1032	1036.9 ± 30.6	1049.6 ± 15.7	104.0	93.6
FGD1 ν_μ RHC 0π	2707	2693.8 ± 46.5	2693.5 ± 33.8	126.5	114.3
FGD1 ν_μ RHC 1π	847	863.9 ± 25.3	870.4 ± 13.2	52.3	46.2
FGD1 ν_μ RHC Other	1015	1006.9 ± 30.6	948.0 ± 16.2	63.4	55.7
FGD2 ν_μ RHC 0π	2648	2697.1 ± 46.9	2711.6 ± 32.7	136.5	124.3
FGD2 ν_μ RHC 1π	693	690.6 ± 22.7	687.6 ± 10.8	66.5	59.3
FGD2 ν_μ RHC Other	932	937.6 ± 28.6	912.5 ± 14.8	53.1	49.1
Total	121432	121431.3 ± 349.1	121419.4 ± 345.0	4736.6	4342.2

Table 6.18.: Event rates and test-statistic for the different ND280 covariance matrix after fitting to data



FGD1 0π FGD2 Other FGD2 $\bar{\nu}_\mu$ 0π

(a) Reduced ND280 covariance matrix



FGD1 0π FGD2 Other FGD2 $\bar{\nu}_\mu$ 0π

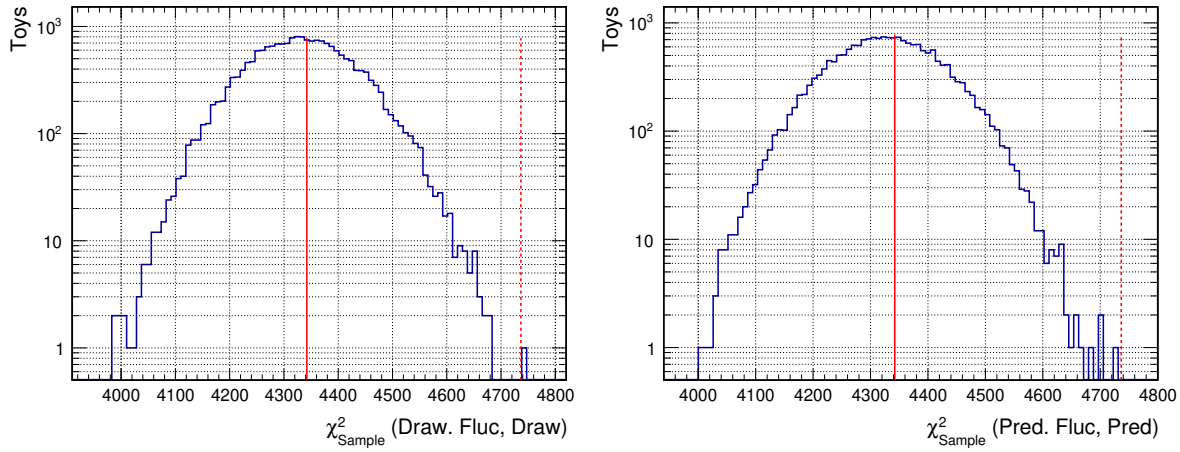
(b) Full ND280 covariance matrix

Figure 6.55.: Likelihood contributions by p_μ $\cos\theta_\mu$ bin for a few selections using the full and reduced ND280 covariance matrices

One-dimensional p_μ, $\cos\theta_\mu$ Distributions

For a more digestible breakdown of the post-fit results the one dimensional projected prior and posterior predictive p_μ and $\cos\theta_\mu$ distributions are compared.

Figure 6.60 shows the FHC 0π and 1π selections. Much like the 2017 results, the post-fit is a clear improvement both in central value. The first, second and fifth p_μ bins are the largest deviations for FGD1 and FGD2. FGD2 appears better described than FGD1, although the p-value in previous section disagrees. The general shape of p_μ is encapsulated well and mostly inside the statistical 1σ of the data, and there doesn't seem to be any particular shape to the disagreements. The $\cos\theta_\mu$ distribution is under-estimated by 1σ until $\cos\theta_\mu = 0.76$ after which oscillation between over and under-prediction occurs. The pattern is somewhat clearer for FGD2, where the low



(a) Fluctuations of the drawn distributions from the MCMC, calculating the χ^2 of a draw and its statistical fluctuation: $p_{full} = 0.400$ (in solid red), $p_{red} = 0.000$ (in dashed red)

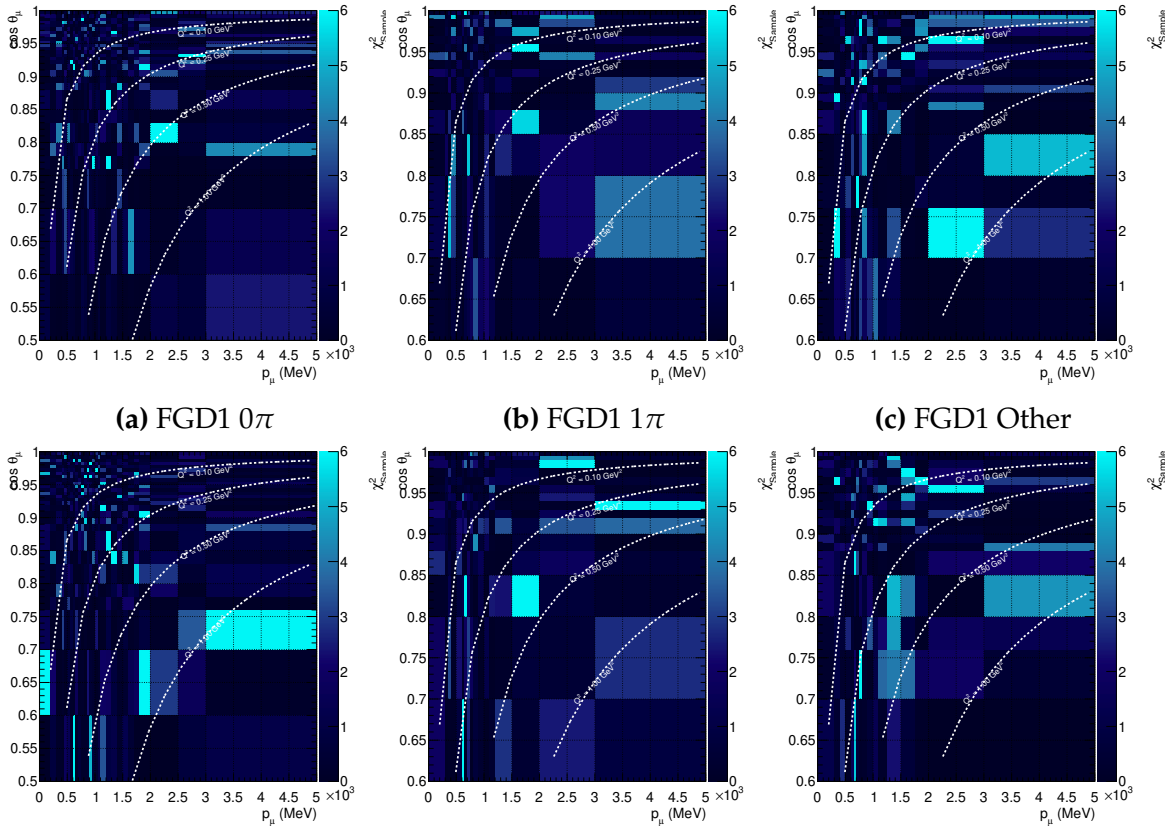
(b) Fluctuations of the posterior predictive distributions from the overall MCMC, calculating the χ^2 of the posterior predictive distribution and its statistical fluctuations: $p_{full} = 0.428$ (in solid red), $p_{red} = 0.000$ (in dashed red)

Figure 6.56.: One dimensional p-values using drawn MCMC steps and the posterior predictive, for the full (solid red) and reduced (dashed red) ND280 parameterisations. A notably worse p-value is obtained from using the reduced ND280 parameterisation, whereas the full ND280 parameterisation shows a good p-value

$\cos\theta_\mu$ underestimation is present until $\cos\theta_\mu \sim 0.76$ which is then overestimated until $\cos\theta_\mu \sim 0.9$. However, the over and underestimation is still within statistical 1σ of the data. The prefit distributions adequately describe $\cos\theta_\mu$ but fail the p_μ distributions.

The 1π selections p_μ are consistently overestimated between 400-550 MeV, which continues up to 900 MeV for FGD1. As p_μ increases, the disagreements sit within 1σ of the data. The $\cos\theta_\mu$ distributions are less well-described, with many points outside the 1σ range. Both FGDs are over-predicted until $\cos\theta_\mu \sim 0.9$ and there is inconsistent behaviour above that with the data points sitting at roughly $\pm 1\sigma$ of the prediction. The prefit distribution appears to roughly cover the data in both p_μ and $\cos\theta_\mu$.

Figure 6.61 shows the FHC Other distributions. The prefit clearly underestimates the data in both p_μ and $\cos\theta_\mu$ by 1-3σ in every bin, compatible with the 2017 results. The post-fit distributions in p_μ pulls the distributions more in-line with data although mis-models the high p_μ behaviour. The $\cos\theta_\mu$ distributions are similarly under-predicted much of the time, and poorly modelled above $\cos\theta_\mu \sim 0.9$. The uncertainties clearly decrease for the prediction, although contrary to the 0π and 1π selections the post-fit description does not lay close to the data.



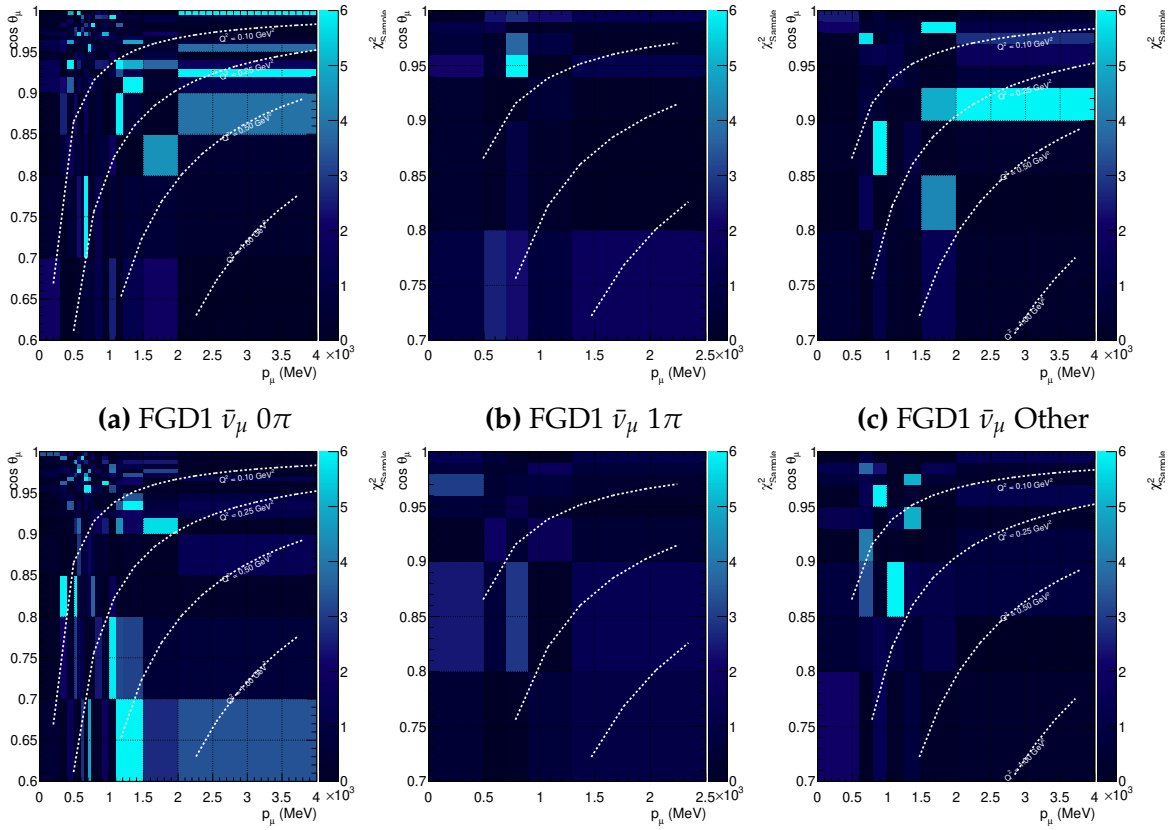
(a) FGD1 0π **(b)** FGD1 1π **(c)** FGD1 Other

(d) FGD2 0π **(e)** FGD2 1π **(f)** FGD2 Other

Figure 6.57.: Likelihood contributions from the posterior predictive spectrum to data for FHC selections

Figure 6.62 shows the new RHC 0π and 1π selections. Contrary to the FHC selections, the prior prediction covers the p_μ behaviour well, and the $\cos\theta_\mu$ distribution is well described for FGD2 except in the forward region, and FGD1 appears overestimated. The postfit distributions mostly lay within 1σ of the data: where FGD1 appears consistently overestimated between 500-1 GeV, FGD2 moves between over and underestimation. The $\cos\theta_\mu$ distributions appear mostly well modelled, with few points outside the 1σ of the data. For FGD2 the most forward-going bin is the largest deviation, which the model overestimates. For FGD1 the largest outlier appears to be statistical fluctuations.

It is more difficult to draw conclusions about the 1π distributions due to the lacking statistical power of the samples. The prior model appears to describe the data well, which is also the case for the post-fit. The $\cos\theta_\mu$ distributions are similar to the 0π, where behaviour is mostly good and the worst modelling happens in the most forward-going bin in FGD2.



(a) FGD1 $\bar{\nu}_\mu$ 0π **(b)** FGD1 $\bar{\nu}_\mu$ 1π **(c)** FGD1 $\bar{\nu}_\mu$ Other

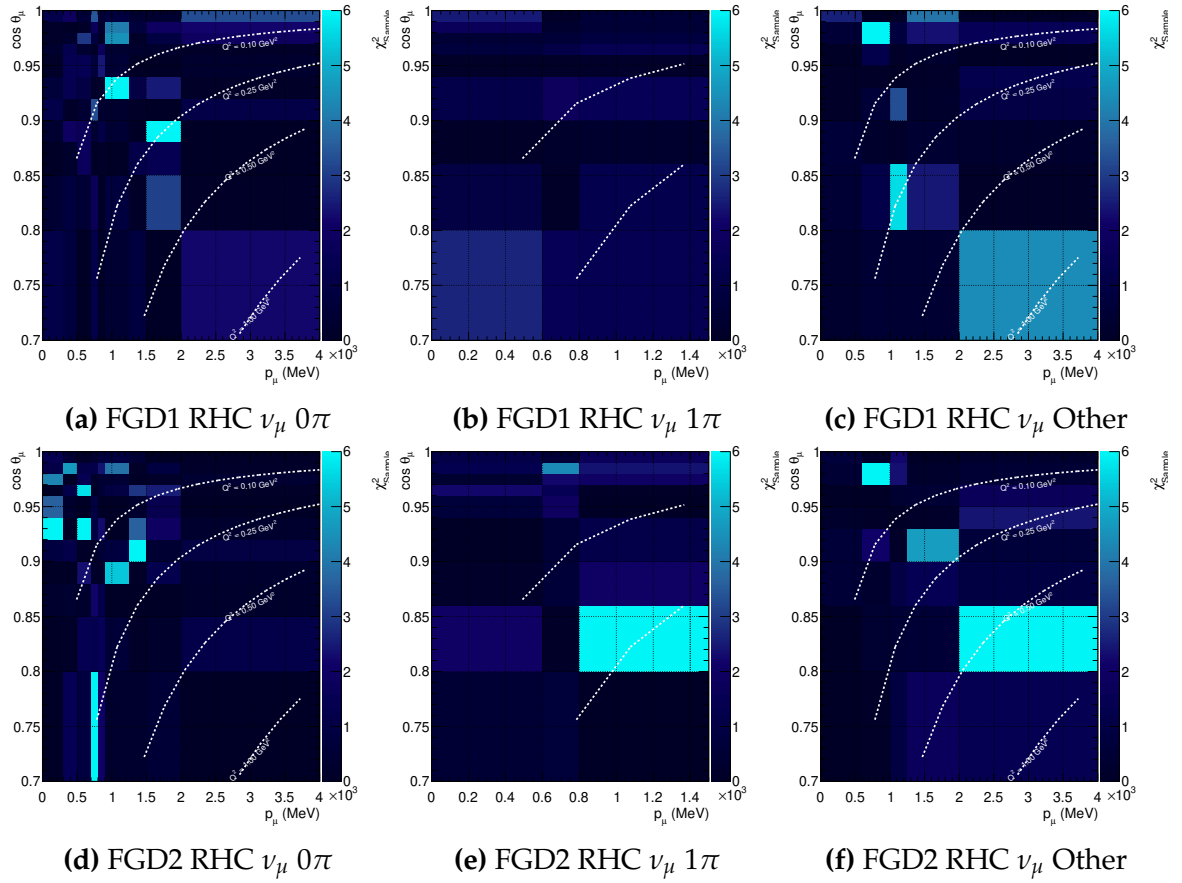
(d) FGD2 $\bar{\nu}_\mu$ 0π **(e)** FGD2 $\bar{\nu}_\mu$ 1π **(f)** FGD2 $\bar{\nu}_\mu$ Other

Figure 6.58.: Likelihood contributions from the posterior predictive spectrum to data for RHC $\bar{\nu}_\mu$ selections

The RHC $\bar{\nu}_\mu$ CCOther selections are shown in Figure 6.63. In contrast to the FHC distributions, the description of the Other selections appear adequate both before and after the fit. FGD1 and FGD2 are inconsistent, with FGD1 underestimated in all p_μ bins and in the first five $\cos\theta_\mu$ bins, whereas FGD2 oscillates more. The fit does little to move the parameters but reduce the uncertainties significantly.

Figure 6.64 shows the 0π and 1π distributions for ν_μ in RHC. The largest difference in p_μ is the first bin of FGD2, which moves from 3σ to 2σ after the fit. Generally, FGD1 is better described than FGD2. The $\cos\theta_\mu$ distributions are adequate prefit except for the most forward-going bins which are heavily underestimated. The post-fit adjusts to match this to some extent, but is still 2σ outside the statistical error of the data.

The 1π distributions in p_μ are compatible before and after the fit, but due to low bin numbers not many more conclusions can be made. The $\cos\theta_\mu$ distributions show the familiar behaviour of underestimating the most forward-going bins in both FGDs.



(a) FGD1 RHC ν_μ 0π **(b)** FGD1 RHC ν_μ 1π **(c)** FGD1 RHC ν_μ Other

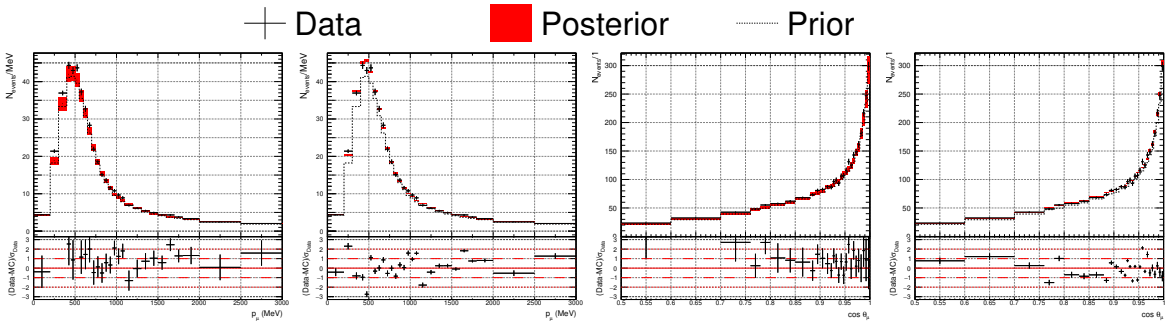
(d) FGD2 RHC ν_μ 0π **(e)** FGD2 RHC ν_μ 1π **(f)** FGD2 RHC ν_μ Other

Figure 6.59.: Likelihood contributions from the posterior predictive spectrum to data for RHC ν_μ selections

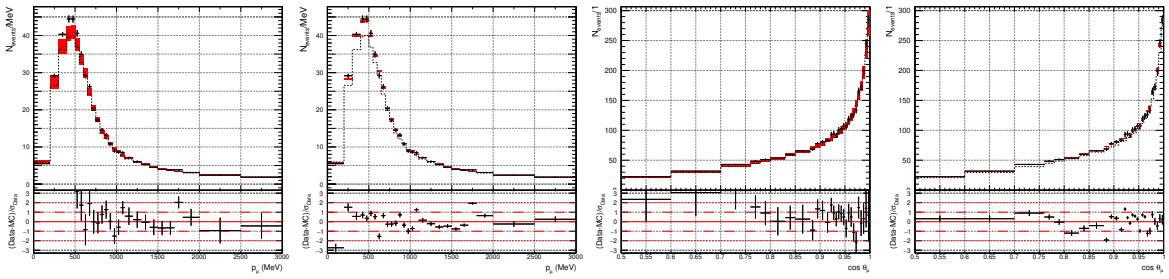
The post-fit corrects for this, and the worst bins postfit are $\cos\theta_\mu \sim 0.95$ for FGD1 and $\cos\theta_\mu \sim 0.75$ for FGD2.

Figure 6.65 finally shows the ν_μ RHC CC Other selections, where predictions are poor in the first and last bins of p_μ for FGD1, and some underestimation in FGD1 and FGD2 for the prefit. The postfit seems to follow FGD2 data better than FGD1, although the shapes are similar with more extreme differences seen in FGD1. For $\cos\theta_\mu$ there is the familiar trend of underestimating, especially at higher $\cos\theta_\mu$. Above $\cos\theta_\mu = 0.9$ the prefit prediction is off by 2σ of the data. The post-fit corrects most of this behaviour and sits predominantly within 1σ of the data.

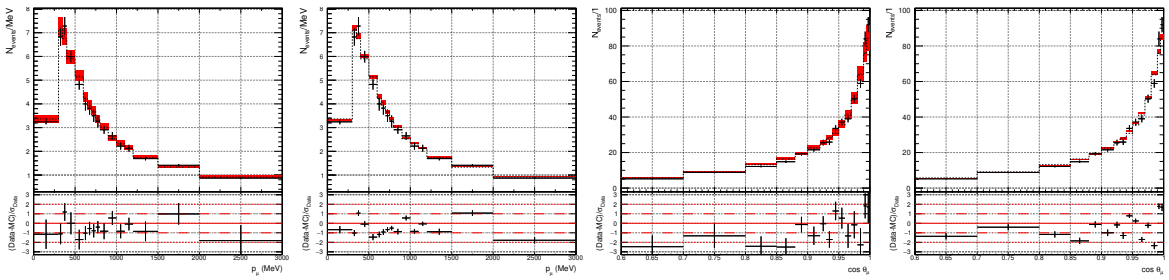
In summary, the likelihood contributions and projected p_μ, $\cos\theta_\mu$ distributions confirm much of what was seen in the 2017 analysis. An inadequate description of the CCOther sample, especially in the very forward region, is present in the prior model and the post-fit model attempts to correct for this as much as possible. The 0π and 1π post-fit model struggles in the most forward going region, and improves significantly



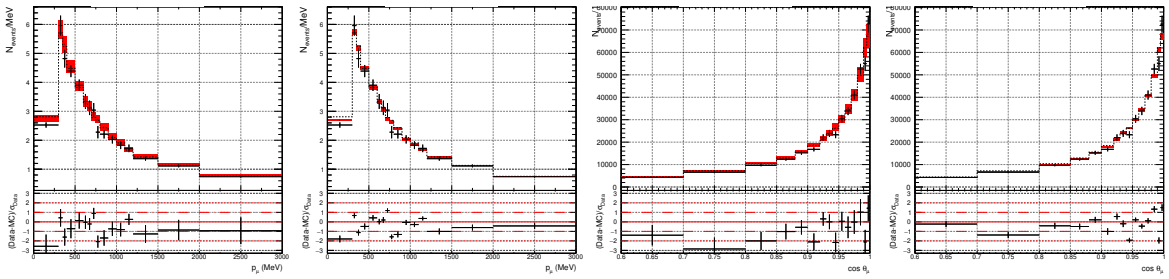
(a) FGD1 0π



(b) FGD2 0π



(c) FGD1 1π

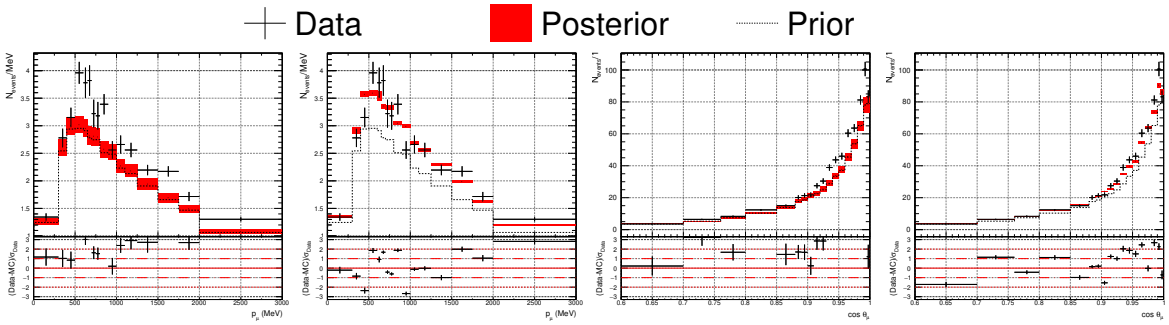


(d) FGD2 1π

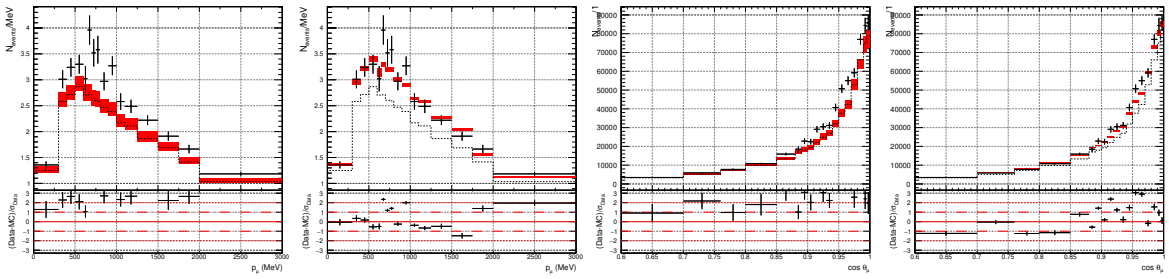
Figure 6.60.: FHC selections p_μ and $\cos\theta_\mu$ projections before and after fit

after the fit, especially in p_μ. The FGDs are mostly consistent, although FGD2 0π has more scattered likelihood contributions than FGD1, and vice versa for the 1π selection.

The new RHC samples, the prior model does an adequate job of describing the projections and the deficiency present in FHC selections at $p_\mu = 0.6$ GeV is not observed. This may hint at different neutrino/anti-neutrino modelling being required. The shape of the post-fit distributions is not necessarily improved after the fit, although it



(a) FGD1 Other



(b) FGD2 Other

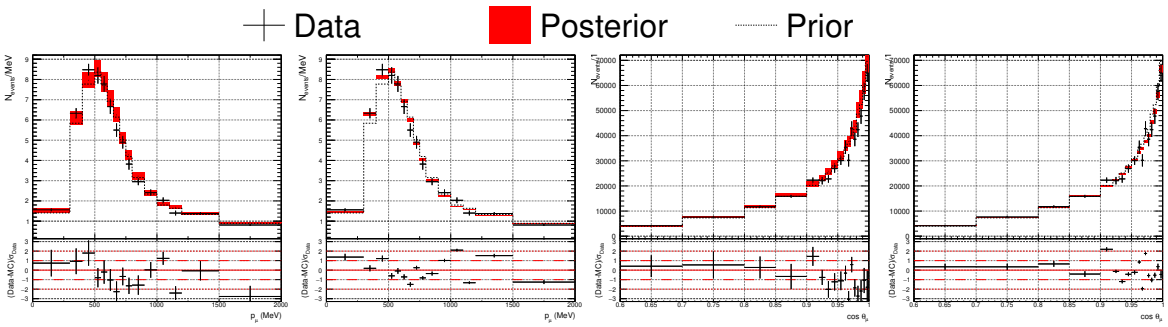
Figure 6.61.: FHC selections p_μ and $\cos\theta_\mu$ projections before and after fit

lessens the errors significantly. The most forward-going bin is problematic for both 0π and 1π selections in FGD2, whereas it is well modelled in FGD1. Interestingly, the underestimation of the FHC Other samples is not present in the RHC equivalent selections, and the data is well described by the prior and posterior model.

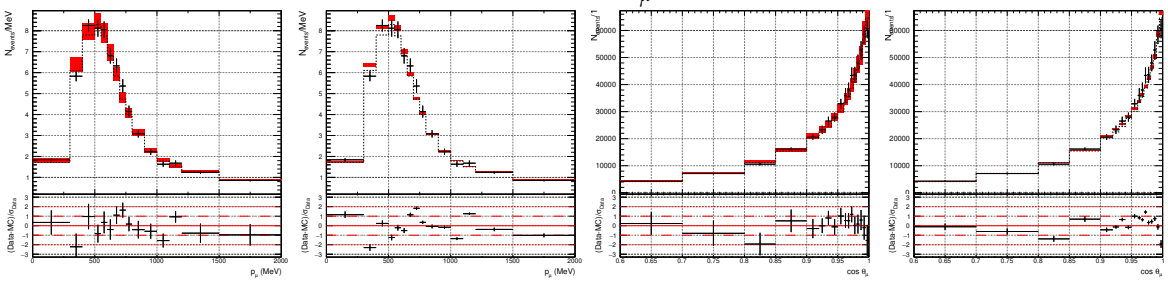
The ν_μ RHC samples appear to uniformly suffer from poor description in the very forward region for all selections. This is somewhat corrected in the post-fit, although still 1-2σ of the statistical error on the data. The p_μ distributions are binned fairly coarsely, but the prefit and postfit model appear to describe the data adequately.

6.7.8. Alternate Model and Compatibility Studies

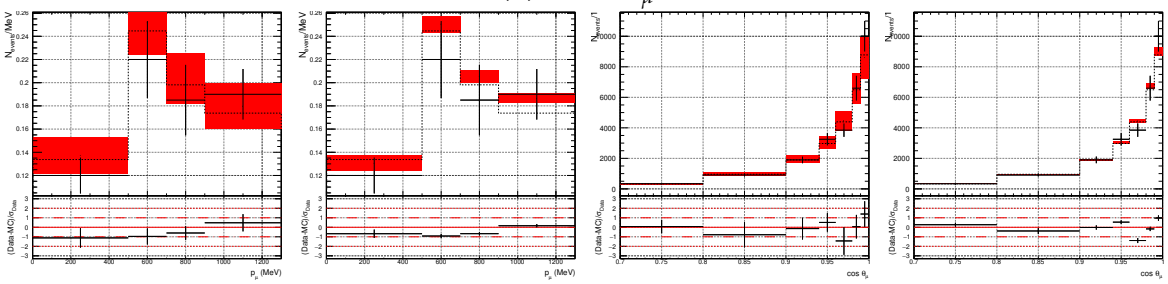
Two alternative studies were performed for this analysis: neutrino versus anti-neutrino runs, and FGD1 versus FGD2 selections. The results largely agreed with those of 2017 in Appendix G, where parameter pulls were sometimes 2σ from the full fit to data. Once the parameter values were propagated to SK the effect was small ($< 1\sigma$), although stronger than in 2017 due to the increased statistics and smaller parameter uncertainties. The studies are found in Appendix G.



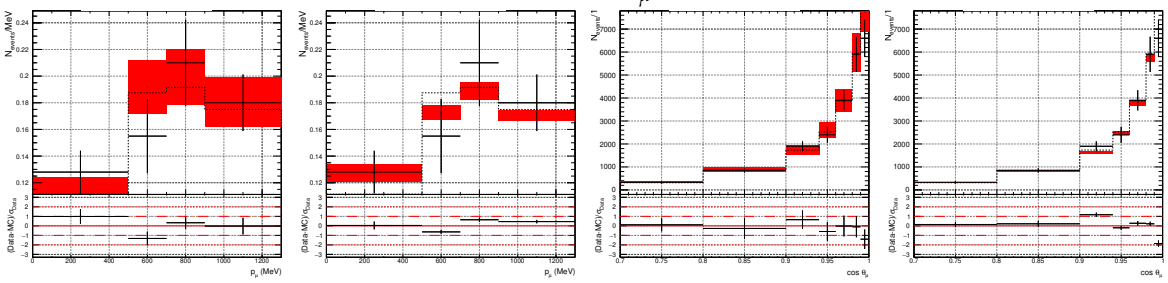
(a) FGD1 $\bar{\nu}_\mu$ 0π



(b) FGD2 $\bar{\nu}_\mu$ 0π



(c) FGD1 $\bar{\nu}_\mu$ 1π

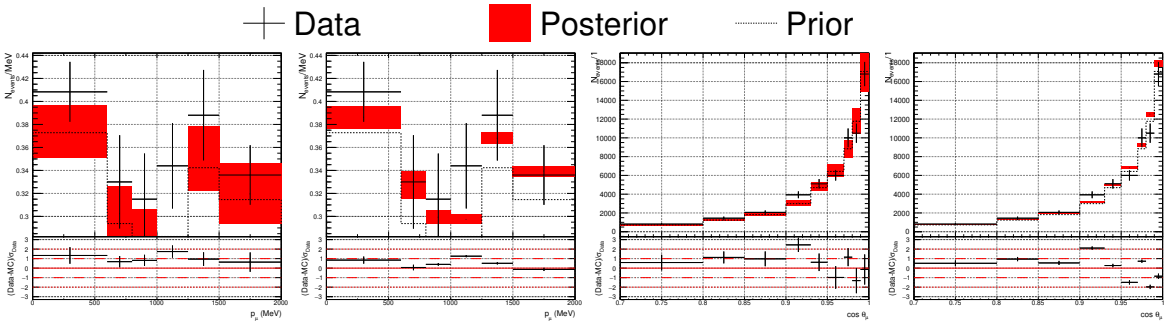


(d) FGD2 $\bar{\nu}_\mu$ 1π

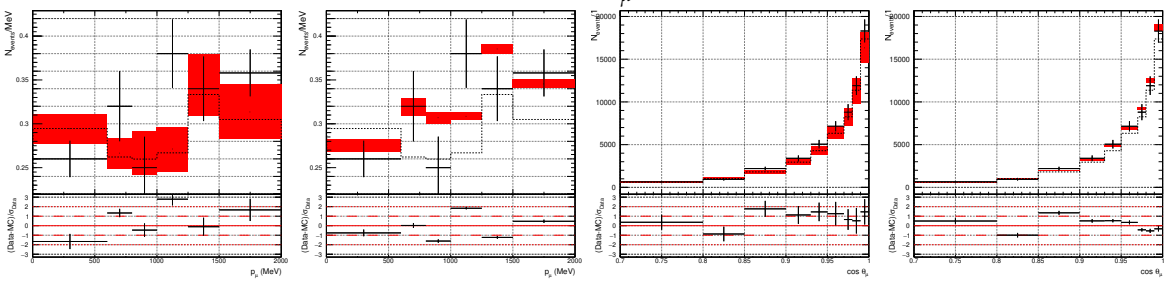
Figure 6.62.: RHC selections p_μ and $\cos\theta_\mu$ projections before and after fit

6.8. Impact on T2K Oscillation Analyses

Using the oscillation parameters listed in Table 5.20, the posterior predictive spectrum at SK is calculated with the 2018 analysis, and is compared to the results using only the prior information and the 2017 posterior.



(a) FGD1 $\bar{\nu}_\mu$ Other



(b) FGD2 $\bar{\nu}_\mu$ Other

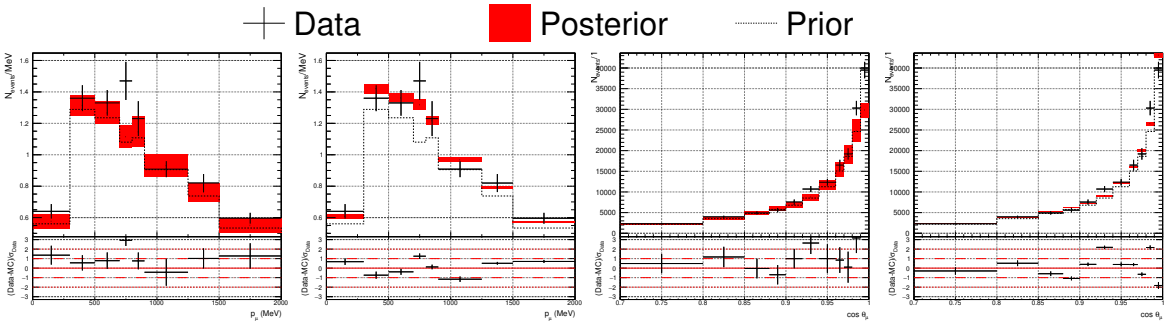
Figure 6.63.: RHC selections p_μ and $\cos\theta_\mu$ projections before and after fit

Sample	Event rate		$\delta N/N$ (%)		
	Pre-fit	Post-fit	Pre-fit	Post-fit	2017
1Rμ	249.86 ± 34.96	270.30 ± 6.46	13.99	2.39	3.06
1Re	65.62 ± 9.95	73.33 ± 2.89	15.16	3.94	3.99
1Re 1de	7.70 ± 0.93	7.02 ± 0.33	12.08	4.64	4.75
1Rμ RHC	61.50 ± 7.21	65.09 ± 1.55	11.72	2.38	2.76
1Re RHC	7.64 ± 0.95	8.12 ± 0.29	12.43	3.57	4.15

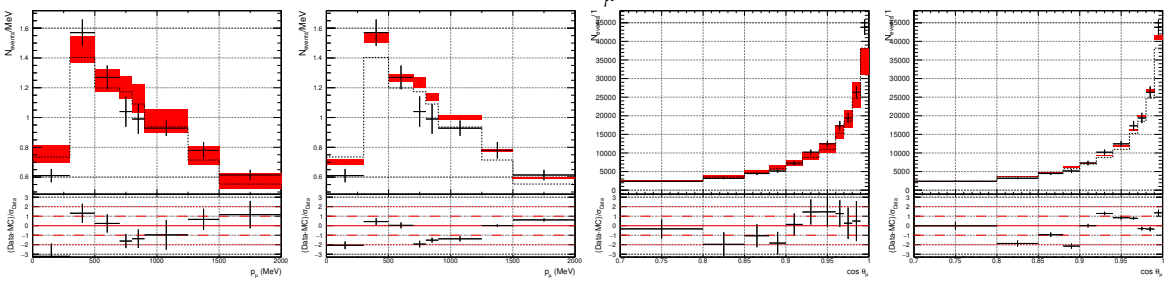
Table 6.19.: T2K-SK event rates and uncertainties from flux and interaction systematics with and without near-detector constraints from the 2018 analysis (not including SK and oscillation parameter errors)

Table 6.19 shows the integrated event rates for the SK selections before and after the fit, with a comparison to the 2017 result in Table 5.21. The 1Rμ selection sees the largest improvement, reducing the uncertainty by another 22% compared to the 2017 fit. For the statistics limited 1Re selections the effect is between 0-12%, and barely has an impact on reducing the uncertainty of the 1Re FHC selection. The central value also shifts considerably for the 1Rμ selections.

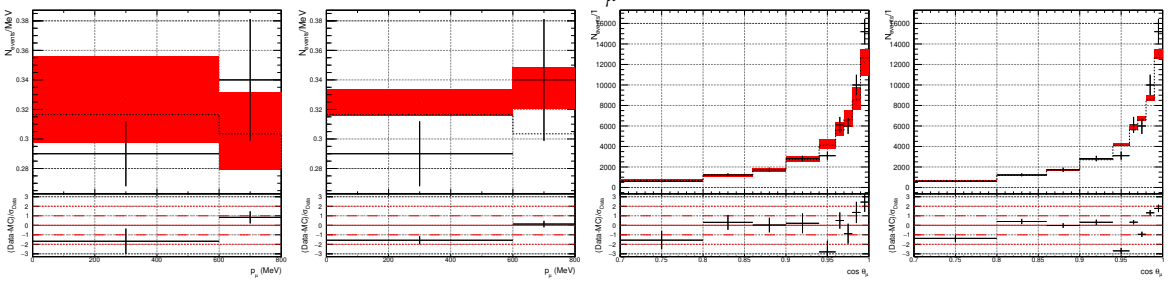
The SK E_{rec} distributions are shown in Figure 6.66 where the increase in events for the 1Rμ selections comes from the low E_{rec} region. Whereas the 1Re FHC selection



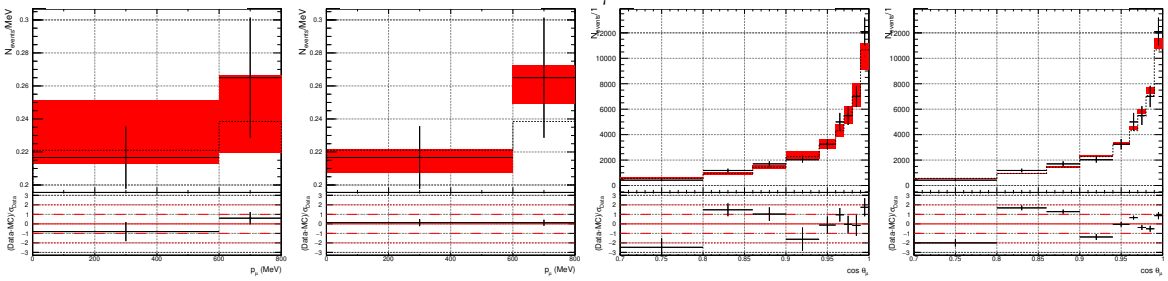
(a) FGD1 ν_μ RHC 0π



(b) FGD2 ν_μ RHC 0π



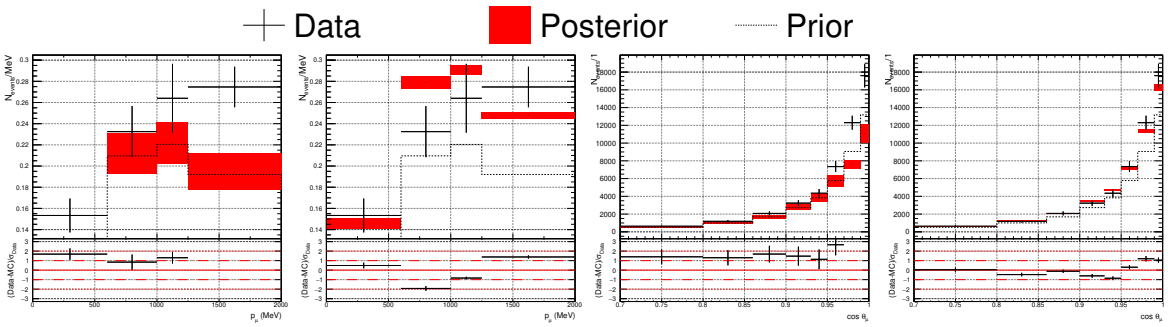
(c) FGD1 ν_μ RHC 1π



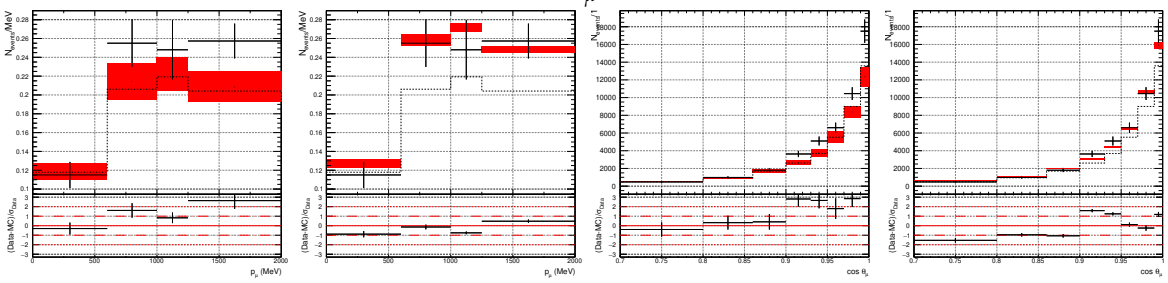
(d) FGD2 ν_μ RHC 1π

Figure 6.64.: RHC ν_μ selections p_μ and $\cos\theta_\mu$ projections before and after fit

appears much the same, 1Re RHC is enhanced throughout, especially in the oscillation peak. The 1Re1de is marginally enhanced, moving closer to the prior. The most extreme shifts sit at 1σ of the 2017 result and slightly outside the 1σ of the prior.



(a) FGD1 ν_μ RHC Other



(b) FGD2 ν_μ RHC Other

Figure 6.65.: RHC ν_μ selections p_μ and $\cos\theta_\mu$ projections before and after fit

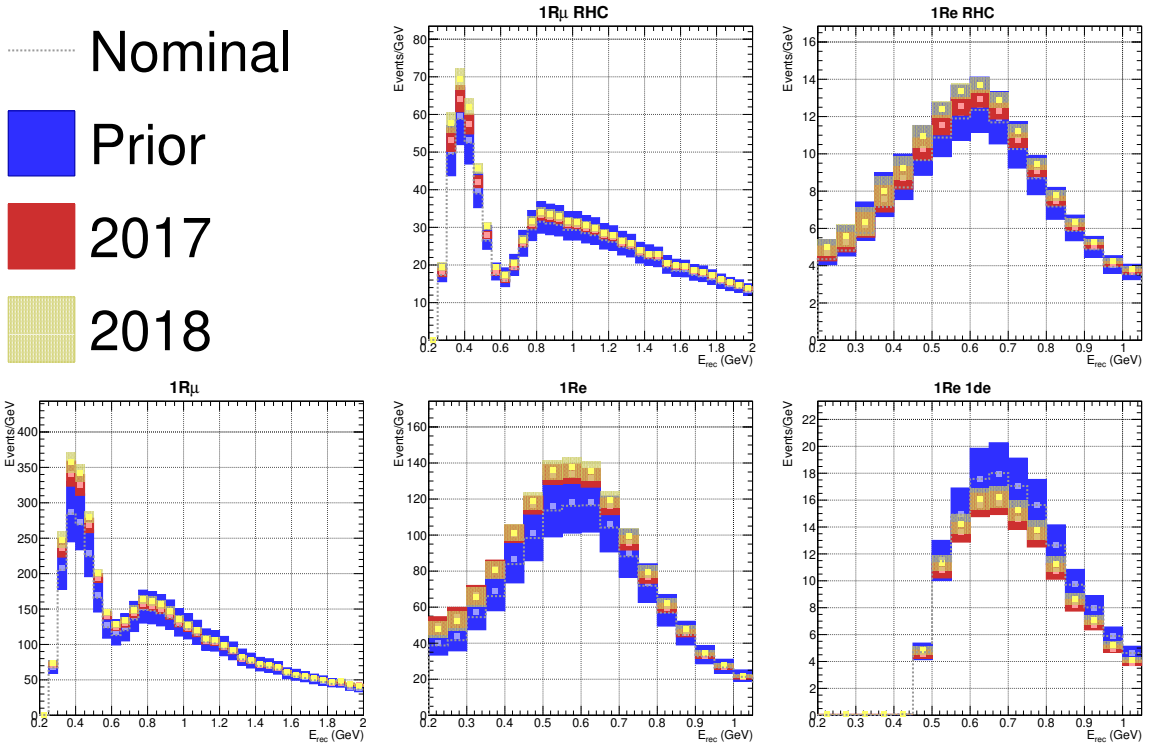


Figure 6.66.: Impact of the full 2018 fit on SK spectra compared to the prior and 2017 results

Sample	Event rate			$\delta N/N$ (%)		
	Full	Reduced	Multi-track	Full	Reduced	Multi-track
1Rμ	270.30 ± 6.46	260.91 ± 7.47	272.34 ± 8.43	2.39	2.86	3.10
1Re	73.33 ± 2.89	71.49 ± 2.70	72.63 ± 3.12	3.94	3.78	4.30
1Re 1de	7.02 ± 0.33	7.02 ± 0.33	7.16 ± 0.38	4.64	4.70	5.31
1Rμ RHC	65.09 ± 1.55	61.48 ± 1.83	66.52 ± 2.01	2.38	2.98	3.02
1Re RHC	8.12 ± 0.29	7.80 ± 0.32	8.15 ± 0.33	3.57	4.10	4.05

Table 6.20.: T2K-SK event rates and uncertainties from flux and interaction systematics comparing the impact of ND280 systematics parameterisations

6.8.1. Impact of ND280 Systematics Parameterisation

The full fit to data in section 6.7 highlighted the importance of the parameterisation of ND280 systematics, in which the resulting parameters sometimes lie outside 1σ of each other. The effect of the reduced and full parameterisation on the integrated event rate at SK is shown in Table 6.20. There is noticeable shift in the central value for the reduced parameterisation for 1Rμ, although the 1Re samples all agree well. The full parameterisation delivers the smallest uncertainties, with the multi-track being worse than the 2017 result.

The propagation to SK in E_{rec} is shown in Figure 6.67, where the largest difference comes from the multi-π versus multi-track. The two multi-π selections are compatible for the 1Re FHC and RHC selections, and differ in the 1Rμ selections when $E_{rec} > 0.6$ GeV although consistently well within uncertainty.

In conclusion, the parameterisation of ND280 detector systematics has an effect on the extraction of oscillation parameters at T2K. The effect is mostly contained within the 1σ uncertainty from systematics, but should be improved upon in future studies outside of this work.

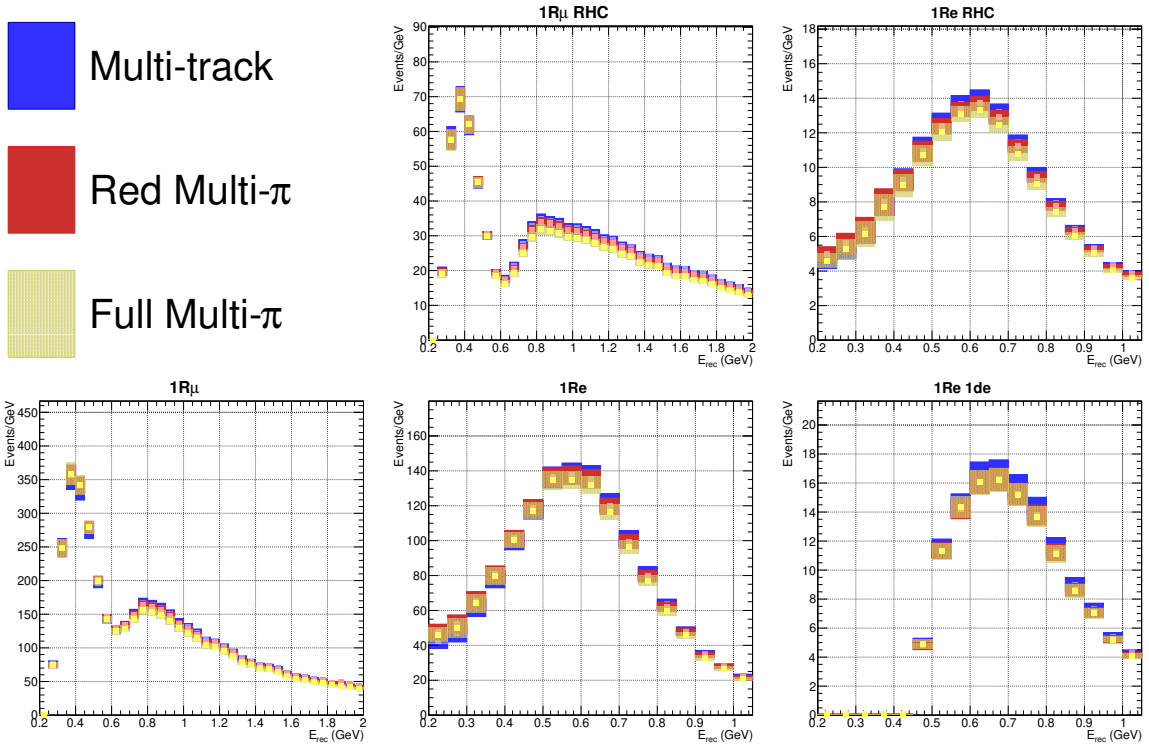


Figure 6.67.: Impact of the full 2018 fit on SK spectra compared to the using the reduced multi-π parameterisation and the multi-track selection

Chapter 7

Future Sensitivity of the Framework

This brief chapter reports on the achievable constraints from ND280 data using the full 7.8E21 POT. For this study, the run 2-8 POT (1.78E21 total with 1.15E21 FHC, 6.28E20 RHC) is scaled up to the full T2K POT of 7.8E21, which roughly assumes 75% FHC and 25% RHC running. The model is identical to that of the 2018 analysis in chapter 6 and the posterior predictive method is used to create the event rates and their uncertainties. The data is the Asimov data, defined as being the prediction of the model when its parameters are set to their priors.

Comparing the ND280 post-fit event rates in Table 7.1 to the 2017 equivalent in Table 6.12, the uncertainties halve from approximately quadrupling the POT. The total event rates for ν_μ are $\sim 280,000$ 0π events, $\sim 60,000$ 1π events, and $\sim 60,000$ Other events. For the RHC selection there are $\sim 60,000$ 0π events, $\sim 5,000$ 1π events, and $8,000$ Other events. The uncertainties on ν_μ selections fall below 0.6% and the 0π $\bar{\nu}_\mu$ selection sits at half a percent. The largest uncertainties, owing to the lowest event rates, are in the RHC 1π and Other selections.

The constraints from the fit to ND280 Asimov data is propagated to SK using either the run 2-8 POT or the full POT at SK. The oscillation parameters in Table 5.20 are applied and the posterior predictive method is used to make the central values and their uncertainties.

The results in Table 7.2 show uncertainties of 1.6-1.85% for the 1Rμ and 3-3.25% for the 1Re selections at SK. This is a reduction of almost a percent over the 2018 analysis, coming purely from the statistical update. In reality, model developments between now and when the full T2K POT is achieved[a] will undoubtedly change the central values and their uncertainties. However, it is encouraging to see that the required DUNE and Hyper-K sensitivities of 1-2% are achievable with this framework.

[a] Likely in 2021 [13].

Sample	Event rate	$\delta N/N$ (%)
FGD1 0π	138182.4 ± 363.0	0.26
FGD1 1π	35054.6 ± 171.7	0.49
FGD1 Other	29785.4 ± 151.9	0.51
FGD2 0π	138917.4 ± 364.2	0.26
FGD2 1π	28137.5 ± 153.5	0.55
FGD2 Other	28712.2 ± 157.3	0.55
FGD1 $\bar{\nu}_\mu$ 0π	27909.1 ± 153.1	0.55
FGD1 $\bar{\nu}_\mu$ 1π	2344.1 ± 44.9	1.92
FGD1 $\bar{\nu}_\mu$ Other	4478.1 ± 63.3	1.41
FGD2 $\bar{\nu}_\mu$ 0π	27508.1 ± 154.0	0.56
FGD2 $\bar{\nu}_\mu$ 1π	2118.2 ± 43.7	2.06
FGD2 $\bar{\nu}_\mu$ Other	4131.7 ± 60.7	1.47
FGD1 ν_μ RHC 0π	10957.9 ± 94.8	0.87
FGD1 ν_μ RHC 1π	3719.1 ± 58.8	1.58
FGD1 ν_μ RHC Other	3502.9 ± 56.6	1.62
FGD2 ν_μ RHC 0π	11187.7 ± 95.2	0.85
FGD2 ν_μ RHC 1π	2935.3 ± 52.1	1.77
FGD2 ν_μ RHC Other	3427.9 ± 55.4	1.62
Total	503010.7 ± 705.4	0.14

Table 7.1.: ND280 posterior predictive event rates using Asimov data for the full projected 7.8E21 POT

Sample	Event rate		$\delta N/N$ (%)	
	ND280	ND280+SK	ND280	ND280+SK
1Rμ	235.20 ± 3.86	530.23 ± 9.73	1.64	1.84
1Re	62.54 ± 2.09	142.40 ± 4.40	3.34	3.09
1Re 1de	7.26 ± 0.27	16.41 ± 0.45	3.72	2.74
1Rμ RHC	56.60 ± 1.04	252.02 ± 4.26	1.84	1.69
1Re RHC	6.98 ± 0.24	31.39 ± 1.01	3.44	3.22

Table 7.2.: T2K-SK event rates and uncertainties from flux and interaction systematics, using Asimov data for the full projected 7.8E21 POT. "ND280" refers to using full POT at ND280 and run 2-8 POT at SK, and "ND280+SK" refers to using the full POT at both ND280 and SK

Chapter 8

Conclusion and Remarks

The work in this thesis presented a Bayesian Markov Chain Monte Carlo method for reducing systematic uncertainties in the long baseline neutrino oscillation experiment Tokai to Kamioka (T2K). The constraints on systematics came from external sources, such as hadron production and neutrino cross-section experiments, as well as internal near detector data from the INGRID and ND280 detectors. The framework uses a simultaneous near and far detector analysis, expanding on using a frequentist ND280-only fit to constrain the systematics in analyses with far detector data. The MCMC method avoids making assumptions on the probability density functions, scales considerably better with parameter dimensionality, and does not suffer when there are local minima or discontinuities in the likelihoods. Compared to conventional frequentist likelihood minimisation, the MCMC method is much more suitable for future large neutrino oscillation fits requiring complex treatment of systematics, such as the proposed T2K–Super-Kamiokande, T2K-NOνA, Hyper-Kamiokande and DUNE analyses.

The uncertainties on the predicted event rates at Super-Kamiokande were reduced from 12-14% to 2-4%, enabling world-leading constraints on multiple oscillation parameters to be extracted at T2K, unachievable without using near detector data. The predictability of the model with the new constraints delivered p-values above 5%, with one exception out of fourteen selections. The first analysis presented herein introduced a more robust systematics parameterisation, notably on the description of neutrino interactions in the nuclear medium. It was used in the official 2017 T2K publications and the results presented at Neutrino 2018. The second analysis expanded on the first, using new ND280 selections and almost double the near detector data, and will be used in official analyses after 2018. A future sensitivity study—using the full projected 7.8E21 protons-on-target—found uncertainties on the far detector event rates at the 1-2% scale, fulfilling the requirements of the future long baseline neutrino oscillation experiments DUNE and Hyper-Kamiokande.

The analyses uncovered unsatisfactory interaction modelling and parameter values disagreeing with the priors from external data. The multi-particle “CC Other” selec-

tion at ND280 was found to be poorly modelled with p-values of 0.000. Numerous compatibility studies—comparing subsets of data with each other and variations of the interaction model and priors—were found to have a negligible effect on the event spectra at Super-Kamiokande. In the second analysis, the parameterisation of ND280 detector systematics and correct neutrino/anti-neutrino modelling had the largest effect on the predicted event spectra at Super-Kamiokande.

Looking ahead, the analyses will in the near future benefit from updated modelling. There are new models for CCQE (z-expansion [224], multi-nucleon [47]), and single pion production [225] being implemented. The priors on the flux uncertainties are expected to be reduced by $\sim$50% [226] using hadron production data from the NA61/SHINE T2K replica target run. Efforts are also underway to include high and backwards going FGD selections which include ECal PID information, expanding kinematic phase space and increasing statistics by 20-30%. Furthermore, dedicated ν_e selections are being developed to help inform ν_e/ν_μ differences. Lastly, exclusive P0D selections will contribute a 100-150% increase in statistics, especially on neutrino-water interactions.

Finally, the ND280 upgrades [227] for the T2K-II project [13] will add another fine grained detector with a cuboid design, physically replacing the P0D. This will enable finer vertex measurements, improving the understanding of multi-nucleon and final state interaction processes. The TPC coverage is designed to be near-hermetic, matching the kinematic phase space of the far detector.

Appendix A

Selection Efficiencies and Purities, 2017

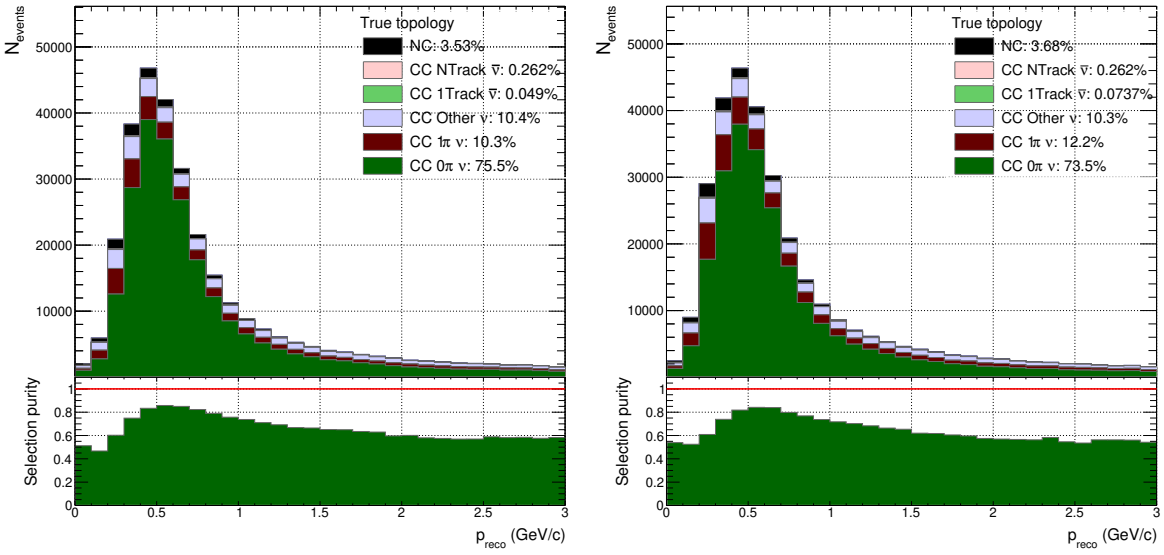
This chapter presents on the efficiencies and purities of the 2017 ND280 analysis, summarised in Table 5.2.

A.1. ν_μ in FHC

Using the aforementioned cuts we can study the lepton tagging efficiency and purity of each ND280 ν_μ selection. The number of events are the raw number of generated Monte-Carlo events without any weighting applied. We show the efficiencies as a function of reconstructed lepton candidate momentum, p_{reco}.

Figure A.1 shows the topology purity for the CC0π selection in FGD1 and FGD2. The purity peak coincides with the event peak with $\sim$85% efficiency and falls off in both directions. The true CC1π and CCOther topology constitute the selection very similarly, at about 10% across the momentum range. As we move up in momentum the CC DIS cross-sections—the largest contribution to the CCOther final state—increase whilst the CCQE and 2p2h cross-sections—the largest contributions to the CC0π final state—decrease. CC1π and CC DIS interactions can produce low-momentum pions which aren't reconstructed in the detector and CC DIS can produce a π^- which may be mistaken for the lepton candidate. Furthermore, the pions can undergo secondary interactions after exiting the nucleus, causing them to be undetected. The NC contribution comes primarily from the NC1π^- via resonance interaction, in which the π^- is identified as the lepton candidate and there are no other particles in the final state. We also note barely any anti-neutrino contamination, owing to the sign selection from the magnet, the low $\bar{\nu}_\mu$ flux in FHC and the smaller cross-section. Averaging over the entire range we have purity of 75.5% for FGD1 and 73.5% for FGD2.

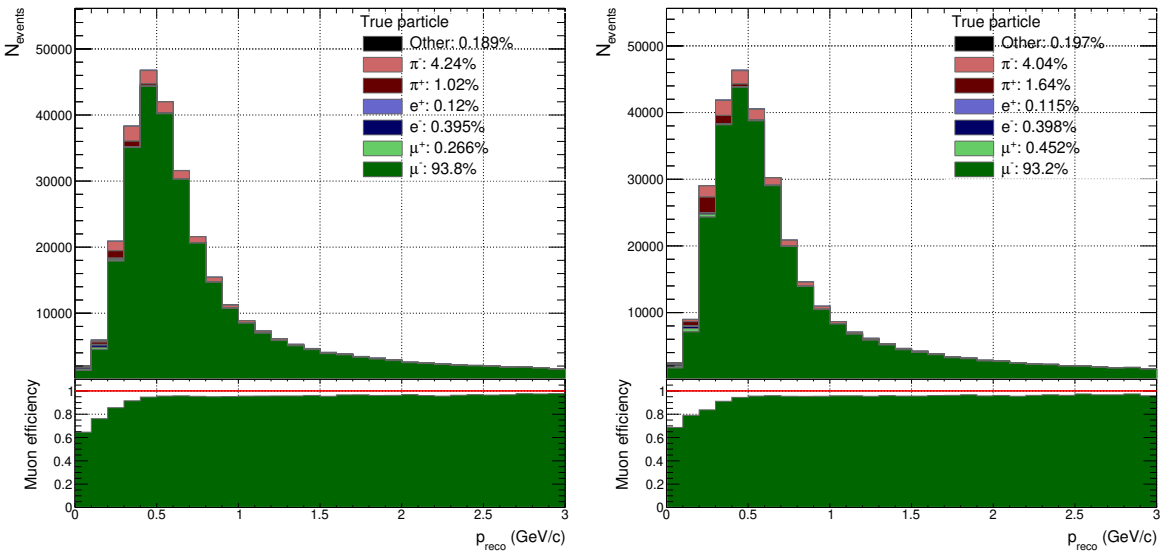
Figure A.2 shows the muon tagging efficiency. We observe good performance over the range of muon momentum, starting with $\sim$65% at low momentum, plateauing at $\sim$95% above 500 MeV/c for both FGD1 and FGD2, which is where the majority of events reside. Averaging over the entire range, the muon tagging performance is 93.8%



(a) FGD1 **(b)** FGD2

Figure A.1.: Breakdown of CC0π selection events' true event topology for FGD1 and FGD2

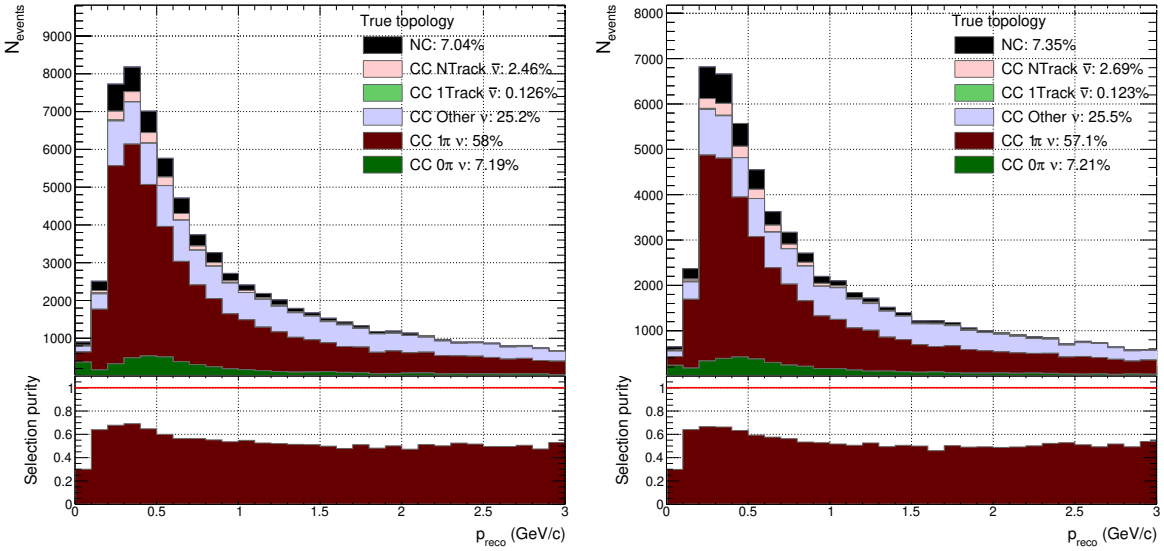
for FGD1 and 93.2% for FGD2. The largest background is π^- from CC1π, CCOther and NC interactions, in which the π^- is either created at the interaction vertex—e.g. an NC1π^- via a resonance where there is no μ^-, or a CCOther interaction creating multiple pions in which one π^- has a higher reconstructed momentum than the μ^-—or through final-state-interactions (FSI) in which a nucleon, π^0 or π^+ undergoes scattering on nucleons in the nucleus to produce the π^-.



(a) FGD1 **(b)** FGD2

Figure A.2.: Breakdown of selection CC0π events' true lepton candidate for FGD1 and FGD2

Figure A.3 shows the topology purity for the CC1π selection. Owing to identifying one single π^+ and one single π^-, the purity is notably worse for CC1π compared to CC0π and peaks at $p_{reco} \sim 0.4\,\text{GeV/c}$ with $\sim$70% purity. The purity takes the biggest hit from CCOther feed-down at 25%, in which either the lepton candidate is identified as a π^- with an accompanying π^+, or events with a $\{\mu^-, \pi^+, \pi^{-,0}\}$ have the latter pion unreconstructed from high-angle and low momentum tracks, leading to a poorly determined PID. The CCOther feed-down increases with p_{reco} as the CC DIS cross-section increases: the main cause of the decreasing purity with increase p_{reco}. The CC0π contributions comes from the outgoing proton being reconstructed as a π^+, or when the nucleon rescatters after exiting the nucleus, producing a pion-like track that gets associated with the primary vertex by mistake. The CC0π contribution is concentrated in the first momentum bin, in which it makes up $\sim$50%. The NC topology contributes 7% by producing a $\{\pi^-, \pi^+\}$ state through NC DIS or NC1π with FSI which gets reconstructed as the $\{\mu^-, \pi^+\}$ final state. The purity for CC1π across the full momentum range is 58% and very similar for FGD1 and FGD2.

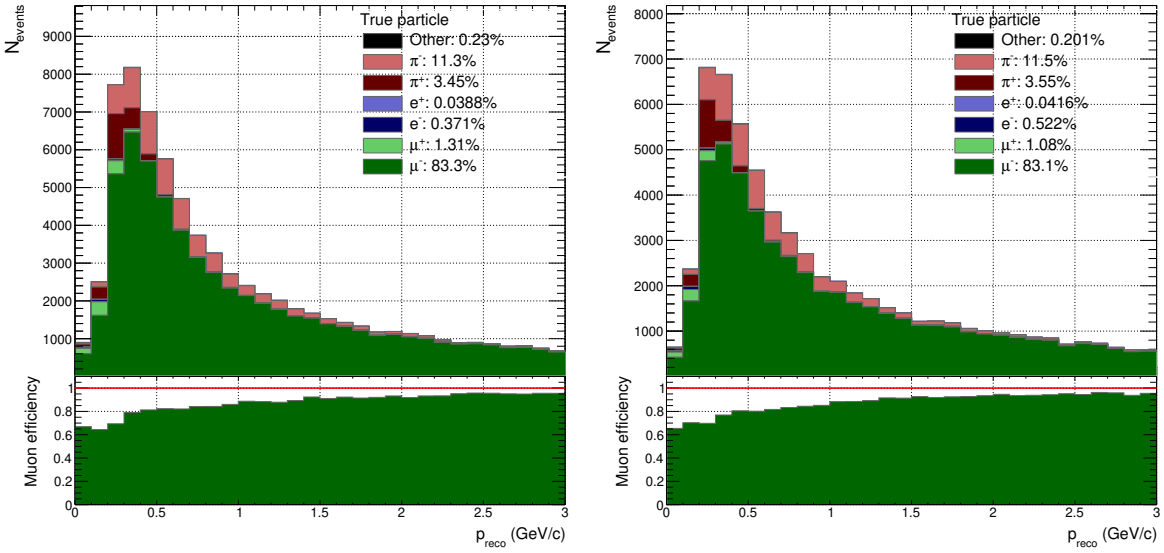


(a) FGD1 **(b)** FGD2

Figure A.3.: Breakdown of selection CC1π events' true event topology for FGD1 and FGD2

Figure A.4 shows the muon tagging efficiency for CC1π^+. As for the purity, the efficiency is comparably worse due to the additional pion requirement, averaging at 83% across the p_{reco} range. Common with the CC0π efficiency in Figure A.2 the major background is π^-, which now constitutes 11% instead of 4%. The π^- comes primarily from DIS interactions and CC1π^0 via resonances in which the π^0 undergoes

a charge-exchange FSI, as discussed earlier. The π^+ contribution comes from high momentum pions which do not bend sufficiently to get a good PID: since the initial CC-inclusive search is done based on highest-momentum track this track is selected as the lepton candidate which curves similarly to a high-momentum μ^-. As p_{reco} increases the efficiency tends to similar values as the CC0π selection.

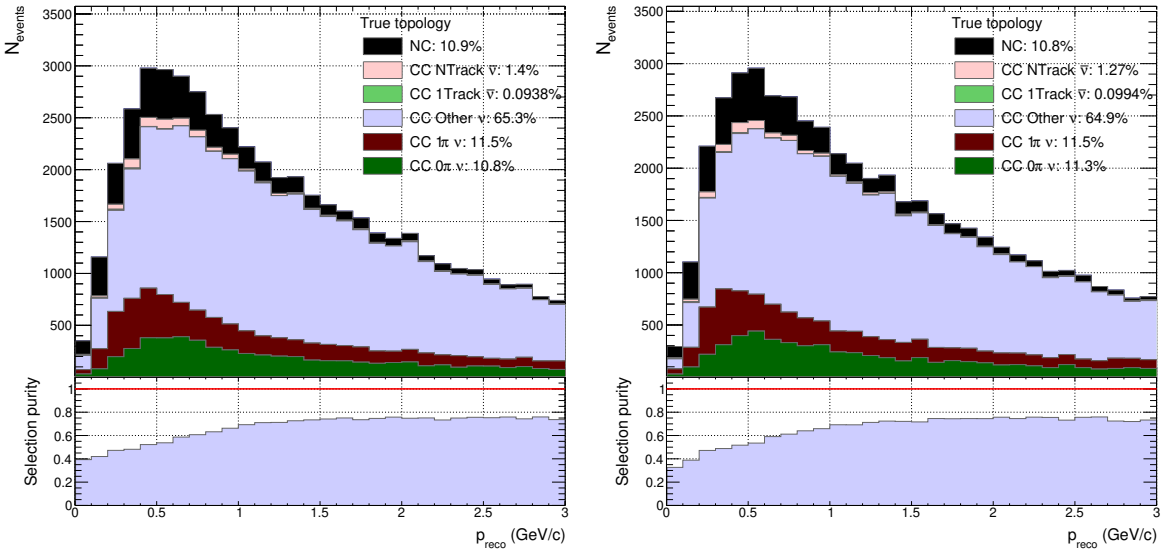


(a) FGD1 **(b)** FGD2

Figure A.4.: Breakdown of selection CC1π events' true lepton candidate for FGD1 and FGD2

Figure A.5 shows the topology purity for the CCOther selection. As expected from the other purities, CCOther increases with p_{reco} due to the increase in the CC DIS cross-section. The CC0π and CC1π topology seeps in to this selection by creating $\pi^{\pm,0}$ after exiting the nucleus that become(s) wrongly associated with the primary interaction vertex, or by the reconstruction falsely identifying an outgoing proton as a π^+. In the NC case, it is enough to have a interaction which produces a $\{\pi^-, \pi^0\}$ combination, which can happen directly through NC DIS or indirectly with NC1π via resonances, in which a secondary interaction occurs and the new track gets wrongly associated to the primary vertex. Initially, the purity starts at 35-40% and plateaus around 80% at $p_{reco} \sim 1.5$ GeV/c. Overall, the purity of the selection is 65% for both FGDs over the entire range, with $\sim 10\%$ each from NC, CC0π and CC1π.

Figure A.6 shows the muon tag efficiency for the CC Other sample, which is notably worse than for previous selections: on average 73%. This is expected, since the CCOther selection has at least 2 tracks ($1\mu^-$, $1\pi^\pm$, $1\pi^0$) but often even more. It is sufficient to have an interaction in which $N_\pi \geq 2$ and $p_\pi > p_\mu$ to wrongly identify



(a) FGD1 **(b)** FGD2

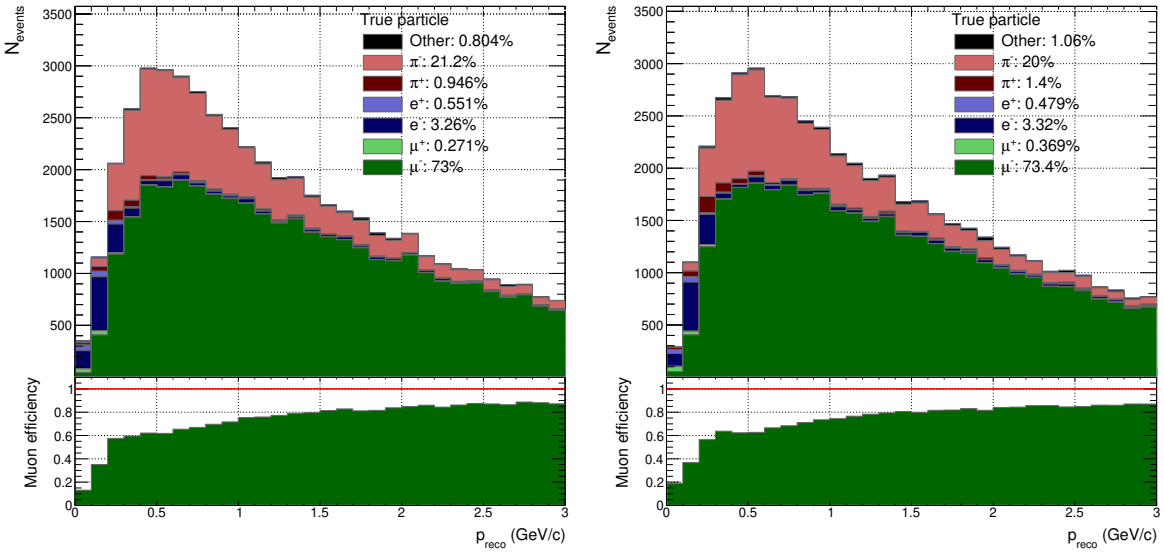
Figure A.5.: Breakdown of selection CCOther events' true event topology for FGD1 and FGD2

the lepton candidate. Owing to the many tracks in this topology due to the CC and NC DIS interaction, it is no surprise to see a 20% contribution from π^-. Furthermore, at low p_{reco} electrons are selected a majority of the time, coming from two sources: 1) relatively high threshold of ν_μ CC DIS compared to ν_e CC DIS due to the muon mass (even though the ν_e flux is much lower than ν_μ) and 2) since the CCOther topology is the only topology to allow for π^0, these are likely to produce $e^\pm$ pairs in the TPC. If the π^0 shower occurs early in the TPC and the interaction vertex is traced to a downstream layer of the FGD, the electron may be falsely associated with the interaction vertex, and if $p_e > p_\mu$ is picked as the highest momentum candidate. To pass the TPC μ PID cut we would require a low momentum e^- to match the dE/dx of a μ^-, which in Figure 5.1 happens at $p \sim 100$ MeV/c.

A.2. $\bar{\nu}_\mu$ in RHC

As for the ν_μ case, we study the efficiency and purities of the anti-neutrino CC1Track and CCNTrack samples.

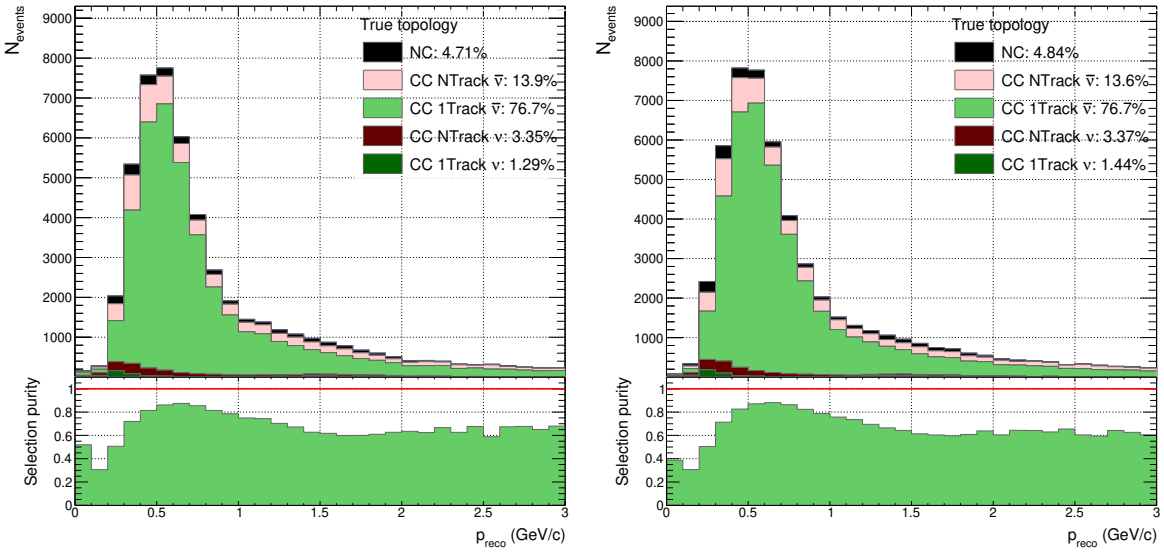
Figure A.7 shows the $\bar{\nu}_\mu$ CC 1 track topology purity. The purity peaks at 85% with the event distribution peak ($p_{reco} \sim 0.6$ GeV) and decreases to $\sim 60\%$ at higher momentum. The wrong-sign ν_μ equivalent selection have a small effect, notable only at low momentum. The largest background is the $\bar{\nu}_\mu$ N tracks topology, in which one



(a) FGD1 **(b)** FGD2

Figure A.6.: Breakdown of selection CCOther events' true lepton candidate for FGD1 and FGD2

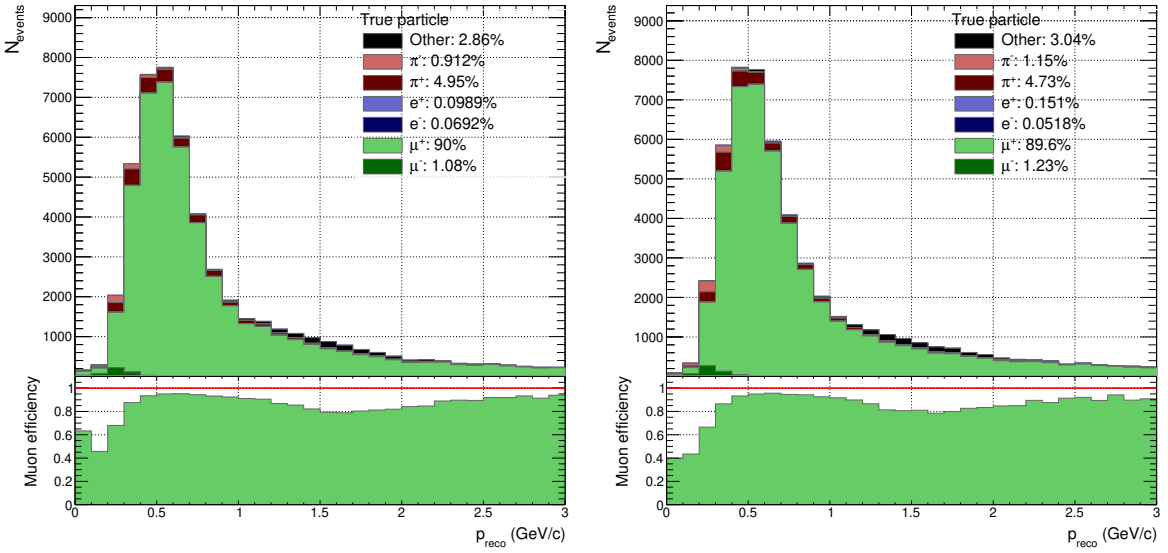
of the pions are unreconstructed. The NC topology enters primarily by the NC1π^+ via resonance interaction, in which the π^+ is reconstructed as a μ^+. Over the whole range the purity is 76.7% where the analogous ν_μ selection in Figure A.1 had a purity of 75.5%, so are very similar in performance.



(a) FGD1 **(b)** FGD2

Figure A.7.: Breakdown of $\bar{\nu}_\mu$ CC 1Trk selection events' true event topology for FGD1 and FGD2

Figure A.8 shows the muon efficiency for $\bar{\nu}_\mu$ CC 1 track selection. As with the ν_μ selections, both FGDs have efficiencies of 90% over the whole range, peaking at 95% at the event distribution maximum around $p_{reco} \sim 0.5$ GeV/c. The wrong-sign ($\bar{\nu}_\mu$) background makes up 1% of selected lepton candidates, and NC 5%.

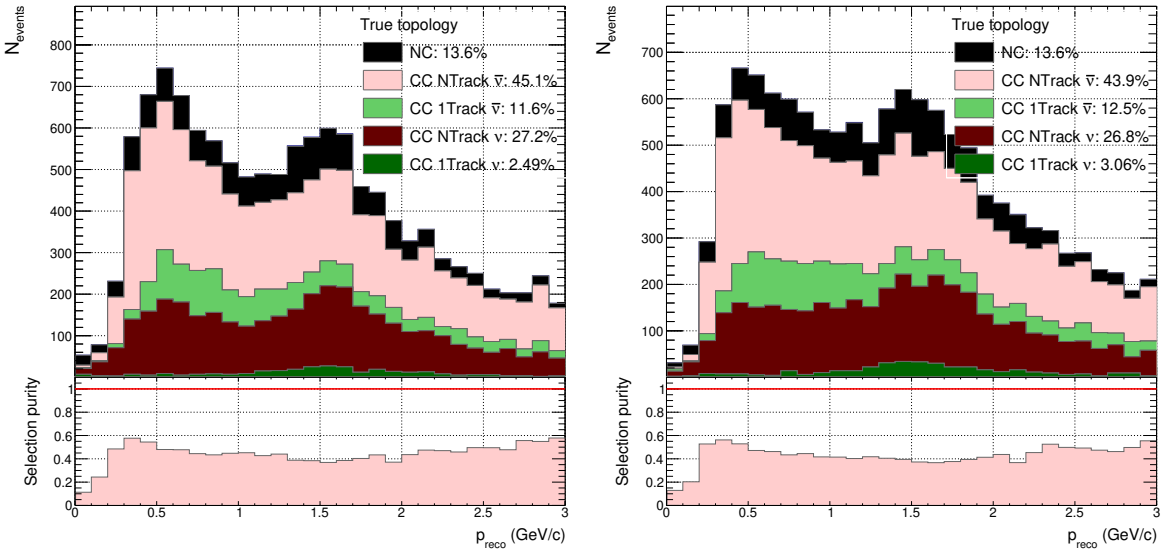


(a) FGD1 **(b)** FGD2

Figure A.8.: Breakdown of $\bar{\nu}_\mu$ CC 1Trk selection events' true lepton candidate for FGD1 and FGD2

As with the ν_μ samples, the CCNTrack selection purity in Figure A.9 is much lower than for the 1 track. It peaks at 60% and decreases to 40% at intermediate p_{reco} to increase to 60% at higher momentum. Overall, the wrong-sign CC N track topology (ν_μ) is the largest contribution at 27%, the NC contribution is 14% and $\bar{\nu}_\mu$ CC 1 track is non-negligable at 12% for both FGDs. Since the CCNtrack selection only requires $N > 1$, the ν_μ CC N track enters by the $\{\mu^-,\pi^+\}$ being identified as $\{\pi^-, \mu^+\}$, on top of the usual possibilities of broken tracks and missed secondary pions. The CC1Track contamination comes from energetic protons being reconstructed as a μ^+ or π^+, and lesser so producing secondary pions and/or nucleons, leading to more particles associated with the primary vertex.

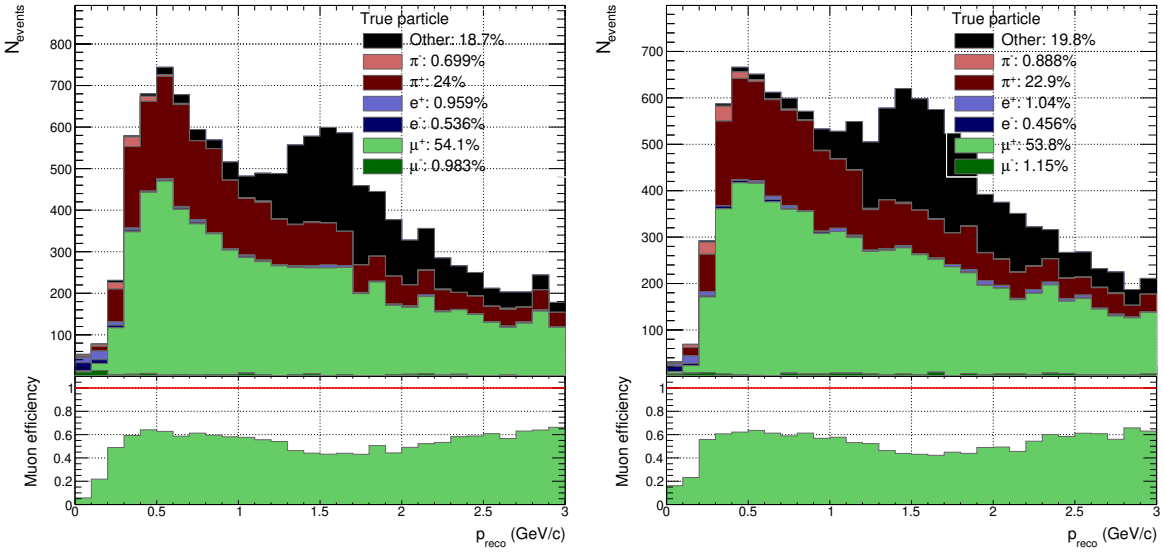
Figure A.10 shows the selected lepton candidate, which over the entire momentum range is 54%. At low momentum the efficiency is very poor but peaks at $p_{reco} \sim 0.5$ GeV/c at about 60%. π^+ make up $\sim$24% of the lepton candidates, having the largest impact between 0.5-1.5 GeV/c. At $\sim$1.5 GeV/c the "Other" category rises sharply, making up 30% of the lepton candidates. This population is predominantly



(a) FGD1 **(b)** FGD2

Figure A.9.: Breakdown of $\bar{\nu}_\mu$ CC NTrk selection events' true event topology for FGD1 and FGD2

protons being identified as μ^+ in the TPC PID algorithm due to the dE/dx, which happens when $1.3 < p < 1.7\,\text{GeV/c}$ as seen in Figure 5.1. Looking ahead at the ν_μ in RHC muon efficiency in Figure A.14, the "Other" population contributes a mere 0.6% over the entire momentum range since the lepton candidate is required to be of negative charge. After the proton "bump" the efficiency rises to 60% again.



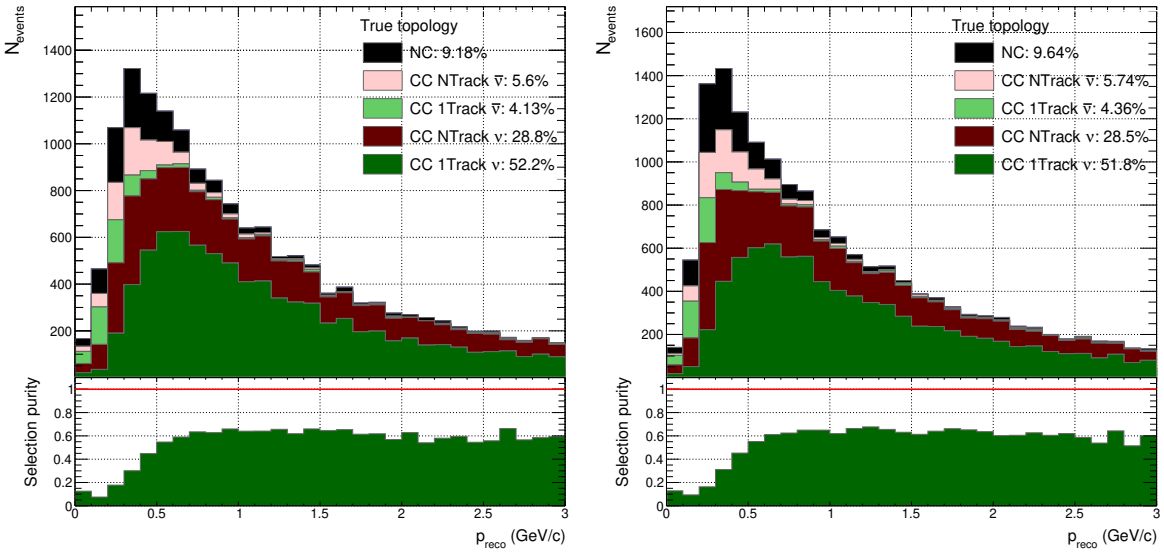
(a) FGD1 **(b)** FGD2

Figure A.10.: Breakdown of $\bar{\nu}_\mu$ CC NTrk selection events' true lepton candidate for FGD1 and FGD2

A.3. ν_μ in RHC

As in previous sections we now study the muon tagging efficiency and topological purity of the final ν_μ in RHC selections.

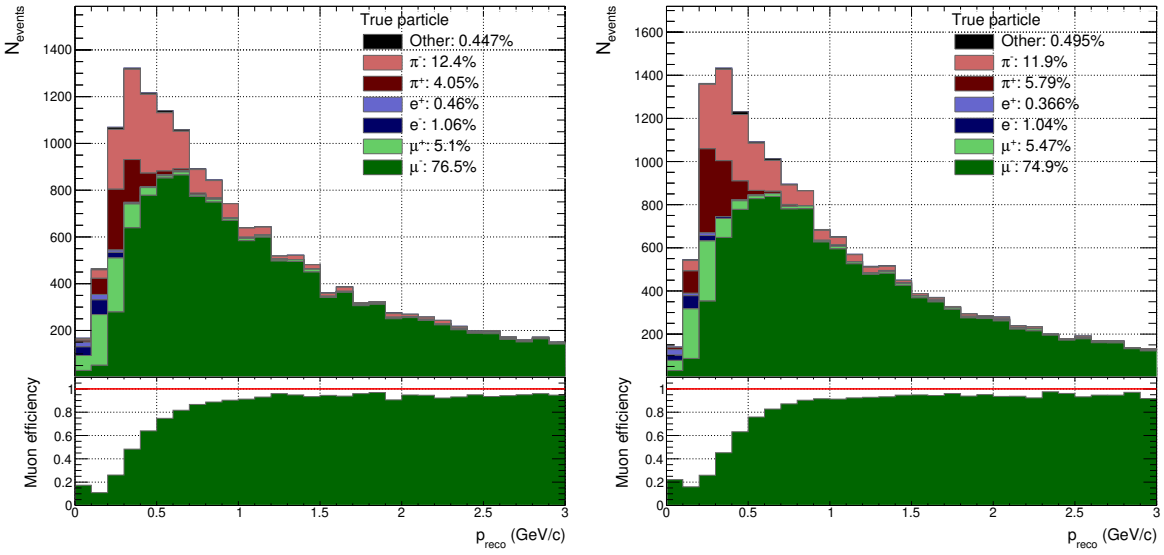
Figure A.11 shows the purity of the CC1Track selection, where we note a poor purity at low momentum, plateauing at 60% at 0.8 GeV/c, averaging at 52%. The overall ν_μ CCNtrack contribution is 29%, 10 percentage units larger than for $\bar{\nu}_\mu$ and is roughly constant over the full range. The wrong-sign contribution is 10% in total, and NC is 10%. The wrong-sign and NC contributions happen primarily at low momentum and vanish above $p_{reco} = 1$ GeV/c.



(a) FGD1 **(b)** FGD2

Figure A.11.: Breakdown of ν_μ in RHC CC 1Trk selection events' true event topology for FGD1 and FGD2

Figure A.12 shows the muon efficiency which closely follows the pattern of the purity. The efficiency is very poor up until 300 MeV/c and then sharply rises to plateau at 95% at 1 GeV/c. However the event distribution peaks in region of low efficiency, causing the average to be 75%. The background varies significantly in the low momentum range: at lowest momentum it's composed of $e^\pm$ since the TPC has similar energy loss for e and μ in this region. Around the event peak, 1/2-2/3 of the selected leptons are background, almost equally wrong-sign muons and both signs of pions. The π^- background comes primarily from NC1π^- and $\bar{\nu}_\mu$ CC1π^- where the muon is missed.

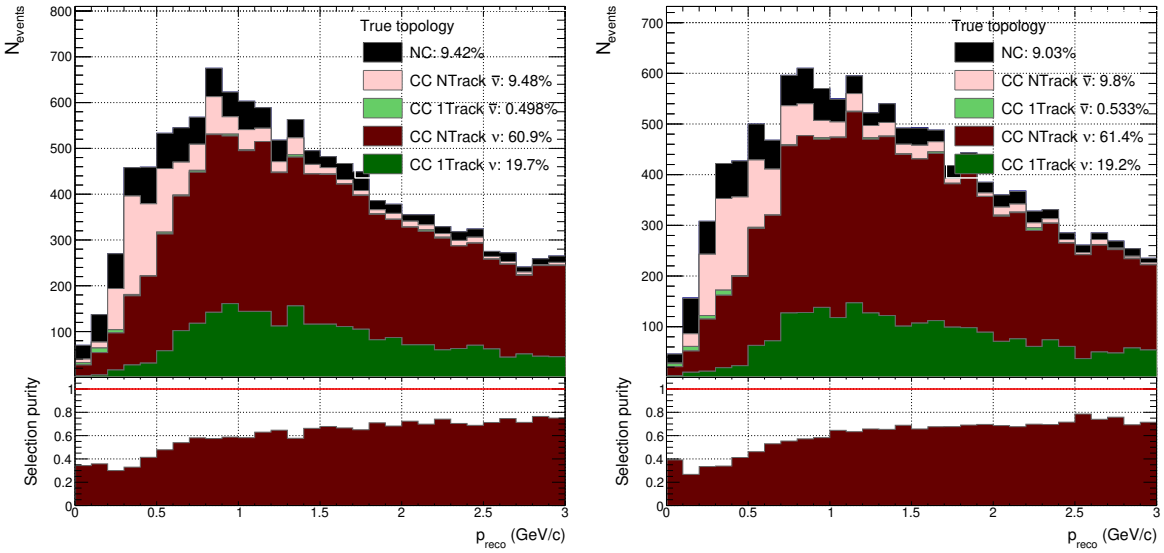


(a) FGD1 **(b)** FGD2

Figure A.12.: Breakdown of ν_μ in RHC CC 1Trk selection events' true lepton candidate for FGD1 and FGD2

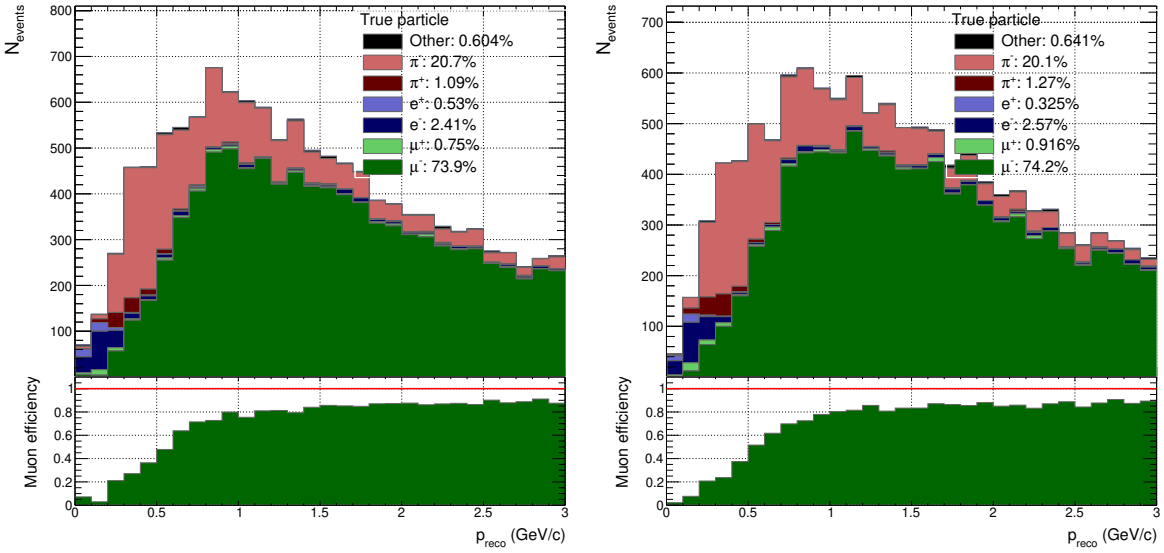
For the NTrack distribution purity in Figure A.13 we approximately 60% purity, which plateaus at 70% above 1.5 GeV/c. The largest background is from same-sign CC1Trk where a broken track or secondary interaction creates a false second (or more) track associated with the vertex. The NC contribution is approximately same size and shape to the wrong-sign CCNTrk, at 10%, largest below $p_{reco} \sim 1$ GeV/c. For NC the contribution comes from reconstructing a $\{\pi^-, \pi^{0,+}\}$ pair as a $\{\mu^-, \pi^{0,+}\}$ pair, and the wrong sign NTrack comes from $\{\mu^+, \pi^-\}$ as $\{\pi^+, \mu^-\}$, just as in the $\bar{\nu}_\mu$ RHC selection, for which the NC contamination was 14% and wrong-sign NTrack was 27%. However, the $\bar{\nu}_\mu$ RHC CCNtrack selection additionally stands the risk of a proton being reconstructed as the μ^+, which is why the purity is worse.

Inspecting the muon tagging efficiency in Figure A.14, we observe several traits common with the $\bar{\nu}_\mu$ CCNTrack and ν_μ CCOther selections: at low momentum the lepton tag is primarily from e^- due to the similar energy loss of e, μ and π in this region; as we increase lepton candidate momentum we create μ,π systems in which the π^- has the higher momentum and is assumed the μ^- candidate. The efficiency rises sharply at 0.3 GeV/c and plateaus at 80% above 1 GeV/c, coinciding with the event distribution peak. Over the entire range the efficiency is 74% and the π^- background is 20%.



(a) FGD1 **(b)** FGD2

Figure A.13.: Breakdown of ν_μ in RHC CC NTrk selection events' true event topology for FGD1 and FGD2



(a) FGD1 **(b)** FGD2

Figure A.14.: Breakdown of ν_μ in RHC CC NTrk selection events' true lepton candidate for FGD1 and FGD2

Appendix B

Selection Efficiencies and Purities, 2018

This chapter presents on the efficiencies and purities of the 2017 ND280 analysis, summarised in Table 6.2.

B.1. ν_μ in FHC

Since the FHC selection is unchanged to the 2017 analysis, the efficiency and purities are very similar and here we only compare FGD1 CC0π in 2018 to 2017 in Figure B.1, and refer to Table 6.2 for the summary.

B.2. $\bar{\nu}_\mu$ in RHC

The CC0π RHC selections are much the same as the 2017 1Track selection, and as such the purity in Figure B.2 is almost identical to Figure A.7. Both FGDs have similar purities and the largest contamination is right-sign single events in which the pion is missed, at about 9.5%. The NC contribution is 5% and the total wrong-sign contribution is $\sim$5%, similar to the 1Trk selection. The difference comes primarily from the removed upper bound on the muon likelihood cut. Moving up in momentum, the purity reduces to about 60% from 85% at the event peak.

Figure B.3 shows the muon tagging efficiency, which again is comparable to the 1Trk equivalent in Figure A.8. The largest mis-id comes from π^+ being reconstructed as the muon, and the wrong-sign contamination is small. At $\sim$1.5 GeV we see the characteristic proton bump—which makes up 3% of the total—in which the dE/dx of a proton is very similar to that of a muon, causing it to be the selected highest momentum positive track with a muon likelihood.

Moving to the 1π $\bar{\nu}_\mu$ selection, the purity is shown in Figure B.4. The $\bar{\nu}_\mu$ RHC selection sees similar performance to the ν_μ FHC equivalent in Figure A.3, reaching an overall purity of $50-54$%, with FGD1 being the higher. The wrong-sign contamination

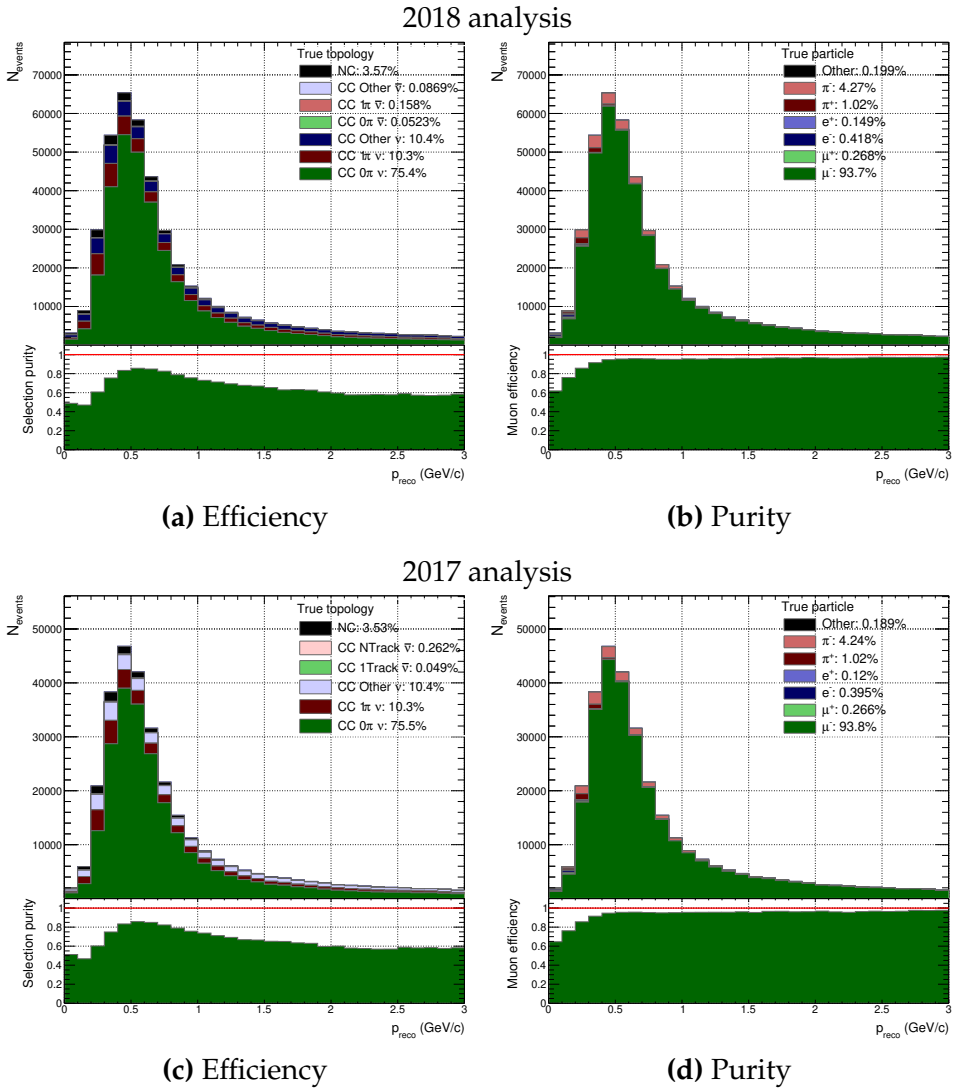
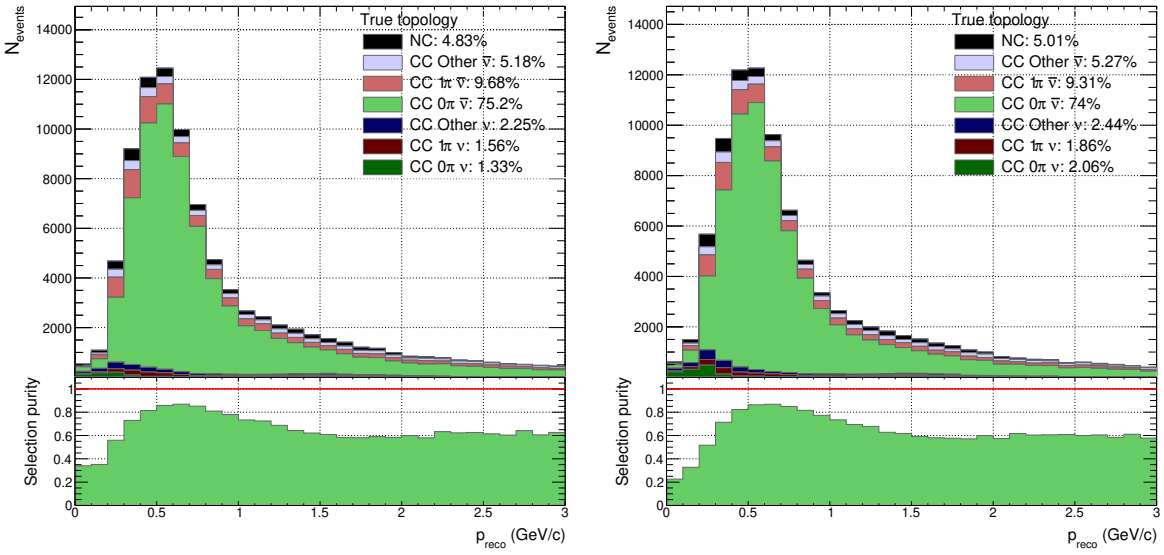


Figure B.1.: FGD1 0π efficiencies and purities for 2017 and 2018 analyses

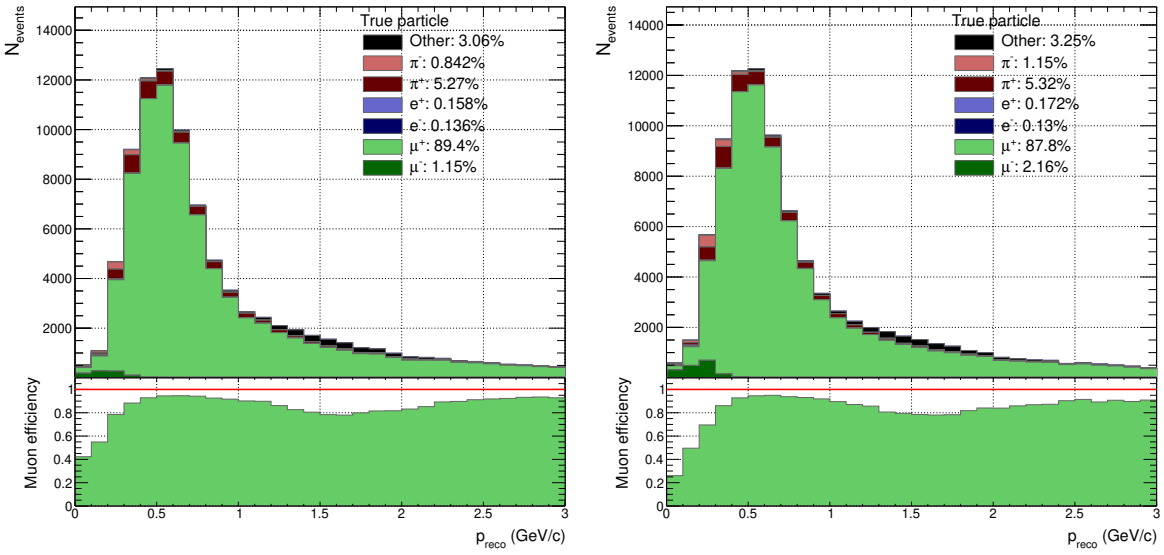
is significant at 25-30%, coming from a π^+ in a CC1π^+ or CC DIS event being reconstructed as the muon candidate, and the μ^- reconstructed as the π^-. The right-sign CCOther contamination is about 10%, owing mostly to one or several missed π. We again see the CC0π ν_μ peak at 1.5 GeV, where the proton (likely from a CCQE or 2p2h interaction) is reconstructed as the μ^+.

Figure B.5 shows the muon tagging efficiency, which in the event peak sits at 80% and decreases to 50% with increasing p_{reco}. Tagging the right-sign pion as the lepton candidate happens 20% of the time, and protons at $p \sim 1.5$ GeV about 15% of the time, making up a large fraction of mis-identification. However, the 1π selection is still almost 10% more efficient and pure than the old CCNTrack selection.



(a) FGD1 **(b)** FGD2

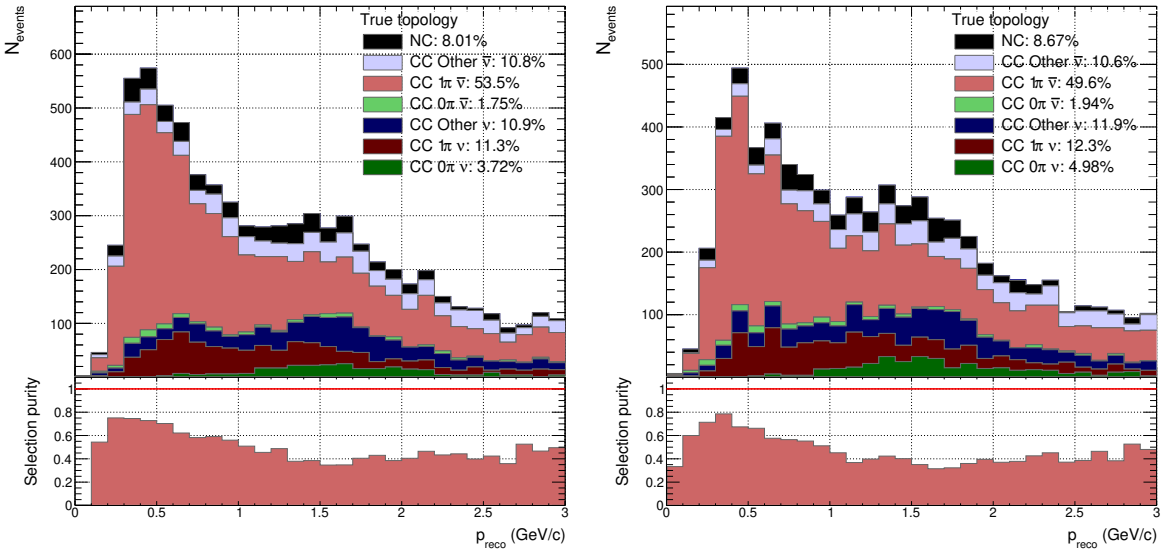
Figure B.2.: Breakdown of $\bar{\nu}_\mu$ CC0π selection events' true event topology for FGD1 and FGD2



(a) FGD1 **(b)** FGD2

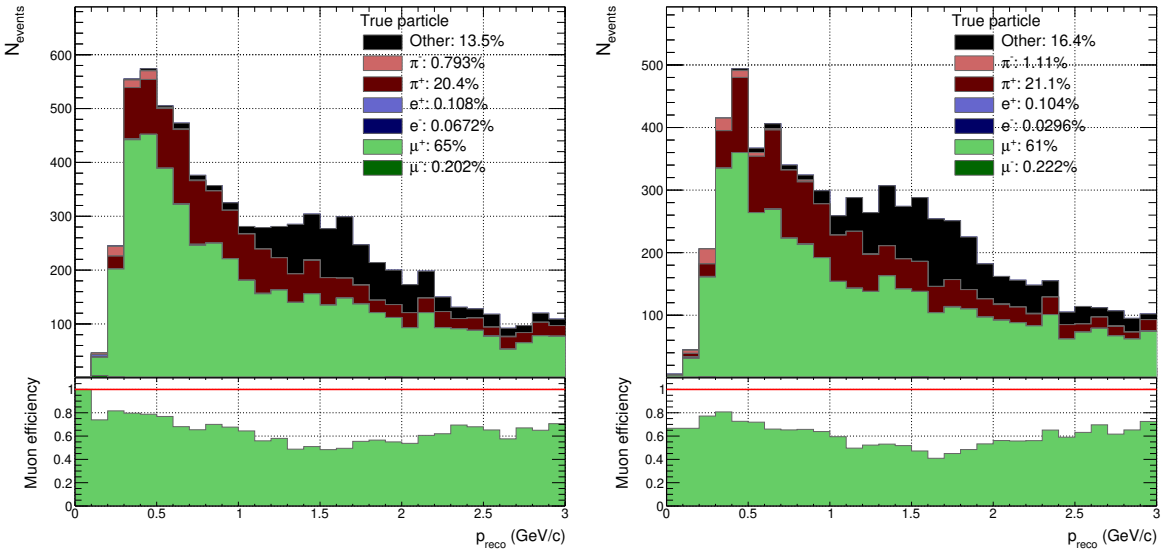
Figure B.3.: Breakdown of $\bar{\nu}_\mu$ CC0π selection events' true lepton candidate for FGD1 and FGD2

Finally Figure B.6 shows the purity of the $\bar{\nu}_\mu$ CCOther selection, which collects all $\bar{\nu}_\mu$ CC candidates that weren't classified as CC0π or CC1π. As with the equivalent ν_μ selection, the sample suffers from low purity due to broken tracks and secondary interactions, leading to a mis-reconstructed number of pions in the event. The selection has an almost equal efficiency for $\bar{\nu}_\mu$ CCOther events as it does for ν_μ CCOther events, and in FGD2 it's indeed more pure of wrong-sign events. It has a high NC contamination



(a) FGD1 **(b)** FGD2

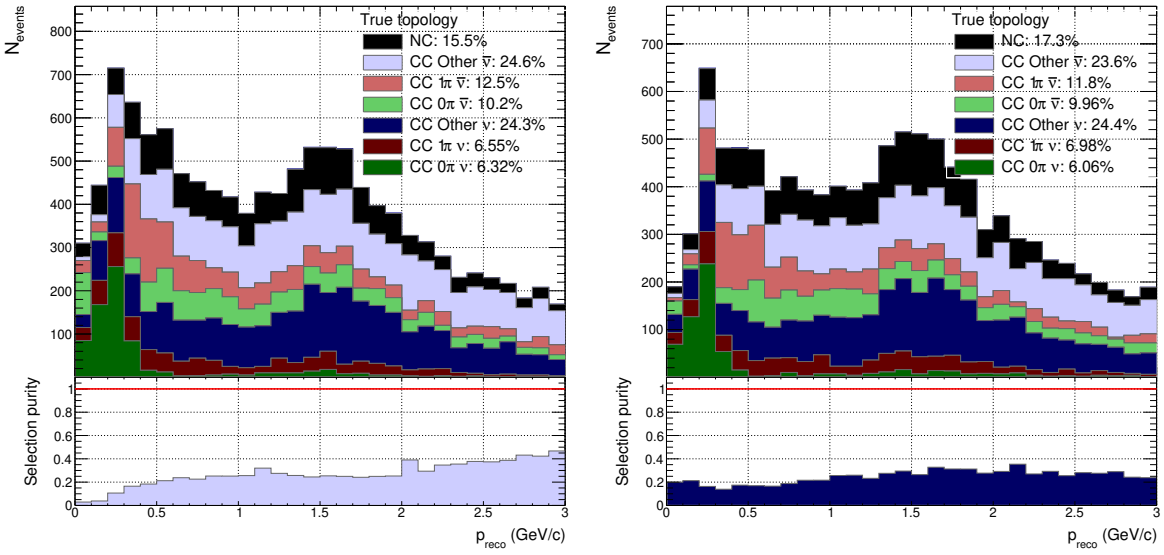
Figure B.4.: Breakdown of $\bar{\nu}_\mu$ CC1π selection events' true event topology for FGD1 and FGD2



(a) FGD1 **(b)** FGD2

Figure B.5.: Breakdown of $\bar{\nu}_\mu$ CC1π selection events' true lepton candidate for FGD1 and FGD2

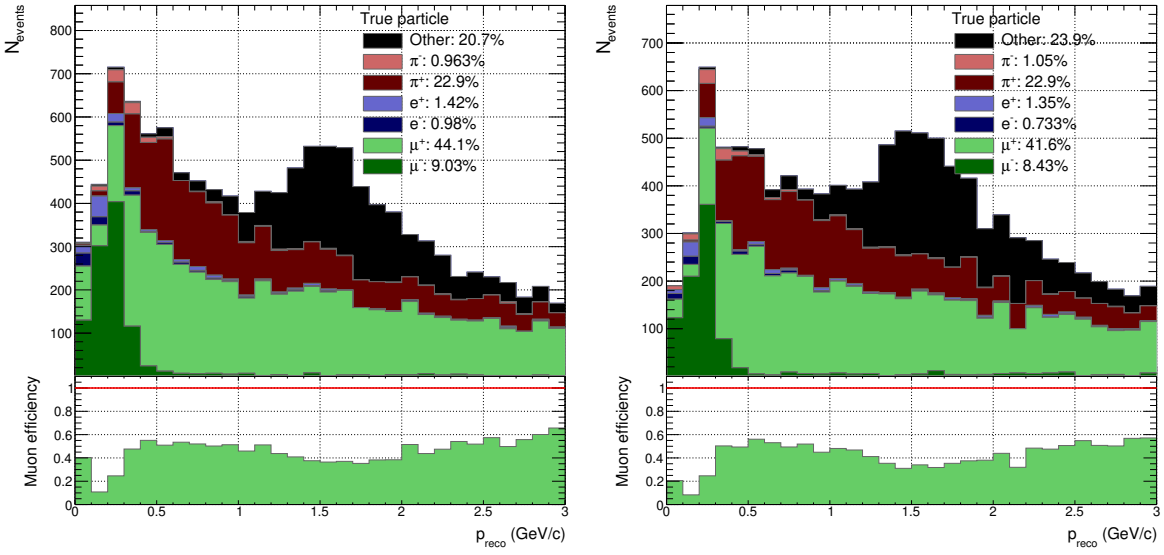
due to collecting high pion multiplicity events, causing a pion to be reconstructed as a muon in the TPC. At low momentum, the purity is close to zero, being swamped by wrong-sign 0π events in which the low momentum μ^- is identified as a μ^+, owing to the changed likelihood cut which in the 2017 analysis was present to remove such events. The wrong-sign 0π and 1π contributions largely vanish above 500 MeV and the wrong sign component is almost exclusively ν_μ CCOther.



(a) FGD1 **(b)** FGD2

Figure B.6.: Breakdown of $\bar{\nu}_\mu$ CCOther selection events' true event topology for FGD1 and FGD2

The muon tagging efficiency of the $\bar{\nu}_\mu$ CCOther selection is shown in Figure B.7, which echoes the conclusions above. The efficiency is below 50% and has almost equal parts proton tagging and π^+ tagging as contaminants. The wrong sign tag happens primarily at low momentum, in which the charge is reconstructed in the magnetic field. The proton bump at 1.5 GeV is especially present in this selection.

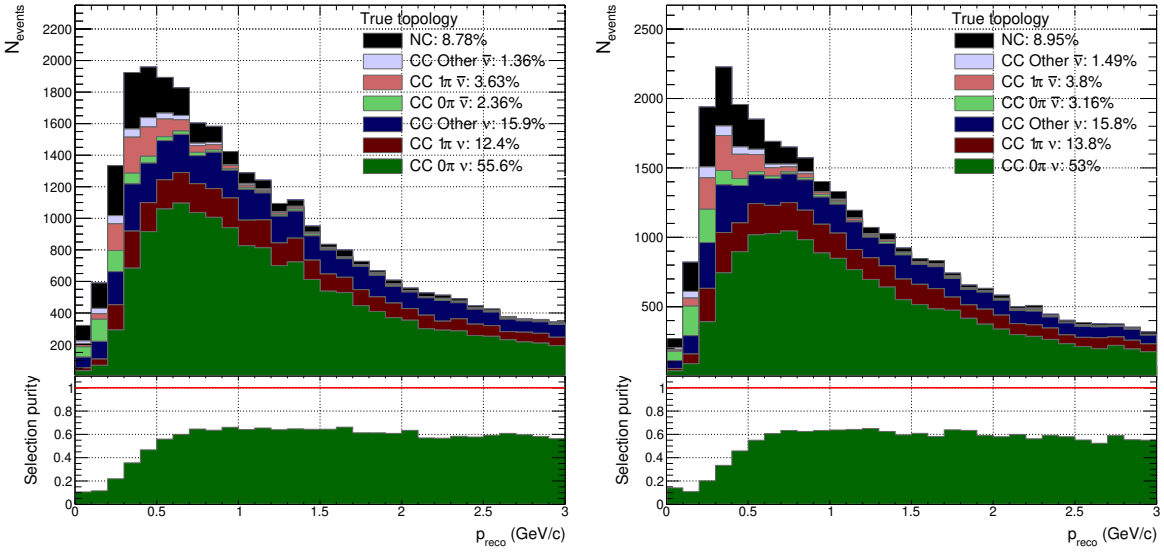


(a) FGD1 **(b)** FGD2

Figure B.7.: Breakdown of $\bar{\nu}_\mu$ CCOther selection events' true lepton candidate for FGD1 and FGD2

B.3. ν_μ in RHC

As with the $\bar{\nu}_\mu$ CC0π selection, the ν_μ RHC CC0π selection is largely identical to the 1Track equivalent in the 2017 analysis. The purity in Figure B.8 is above 53%, with large contamination from right-sign 1π and Other interactions, and the wrong-sign background making up 8%, slightly less than the 1Track case. The NC contamination is almost identical to the 1Track selection at 9%. We note the purity above 600 MeV stabilises at about 60%.

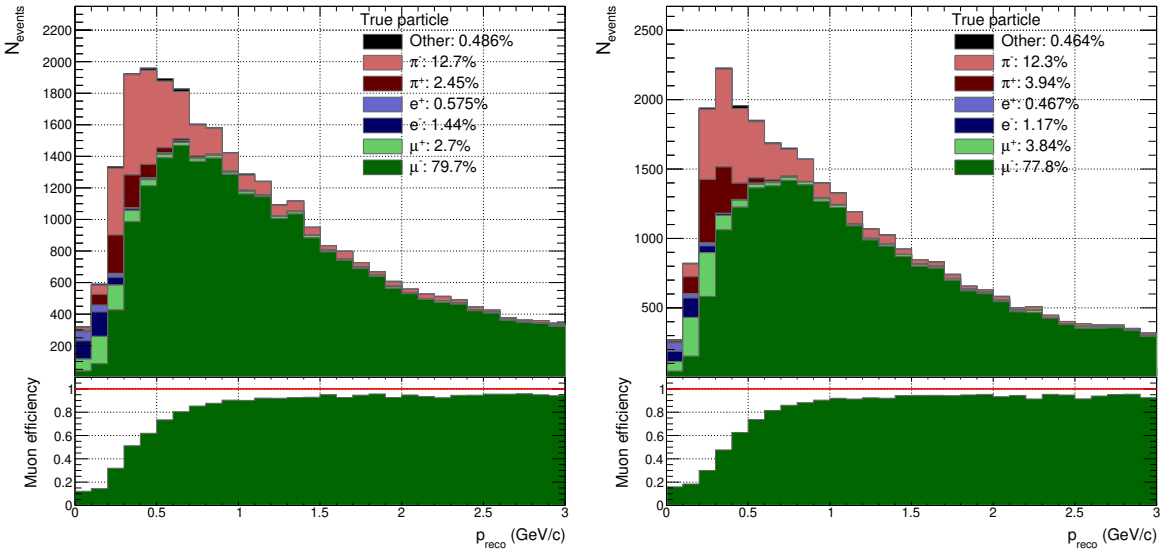


(a) FGD1 **(b)** FGD2

Figure B.8.: Breakdown of ν_μ RHC CC0π selection events' true event topology for FGD1 and FGD2

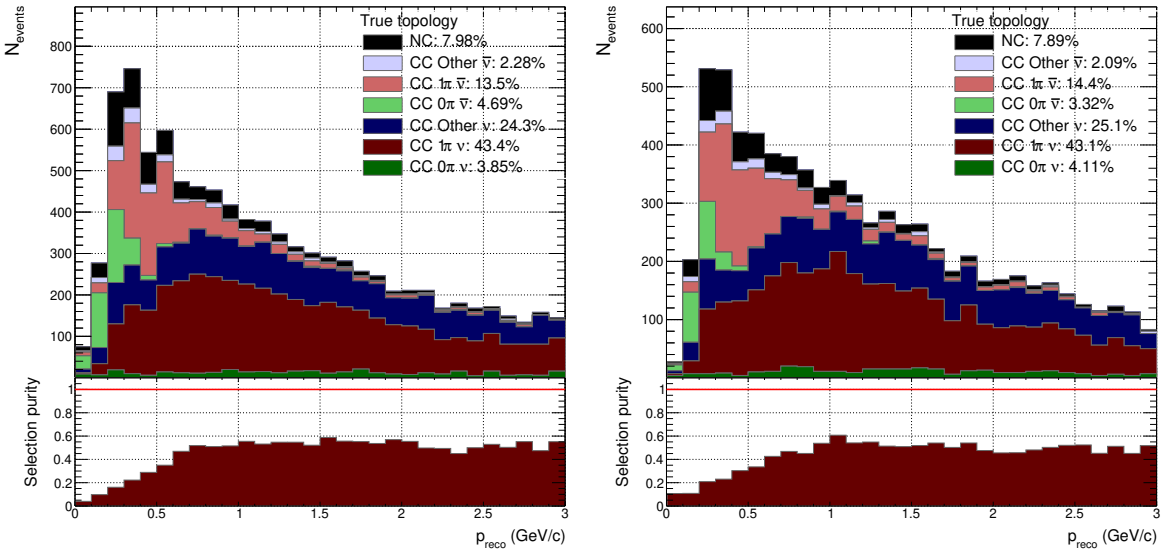
The muon tagging efficiency is shown in Figure B.9, where we note 90% above 1 GeV. At the event peak the efficiency sits at 55%, leading to overall 78%. In and below the event peak the main contamination is from π^- (13%) and as we go down in momentum the wrong-sign contributions increase due to wrongly reconstructing the charge in the magnet. At low momentum the wrong-sign component is $\times 10$ larger than the right-sign.

The CC1π purity is shown in Figure B.10, where we again see a large wrong-sign contribution at low momentum, primarily from $\bar{\nu}_\mu$ 1π events. The purity is 43% overall, and a meagre 20% in the event peak. The right-sign CCOther amount is constant with momentum, making up almost 1/3 at higher momentum.



(a) FGD1 **(b)** FGD2

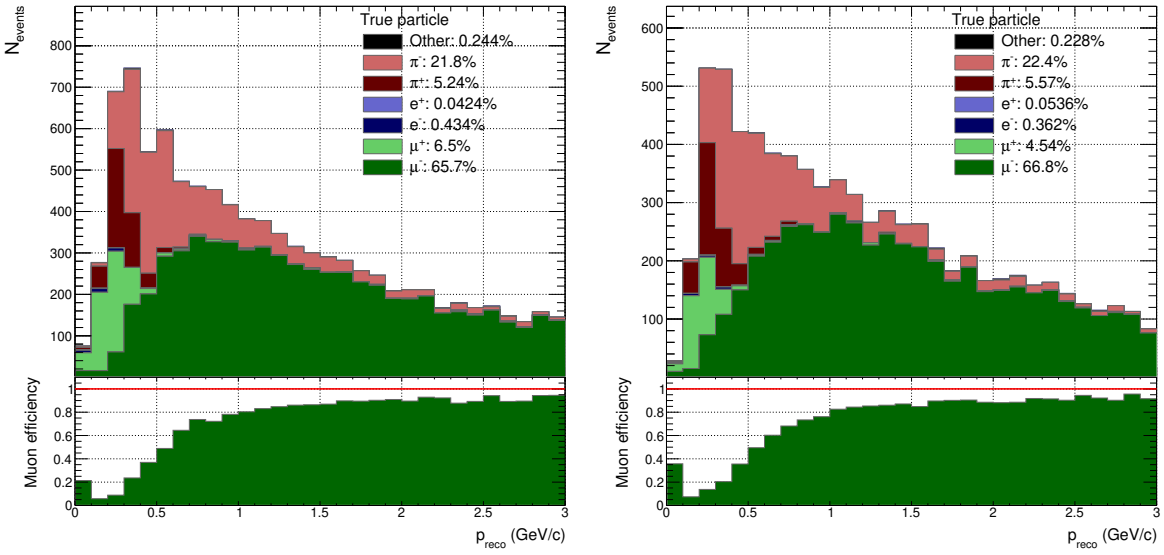
Figure B.9.: Breakdown of ν_μ RHC CC0π selection events' true lepton candidate for FGD1 and FGD2



(a) FGD1 **(b)** FGD2

Figure B.10.: Breakdown of ν_μ RHC CC1π selection events' true event topology for FGD1 and FGD2

The muon tagging efficiency in Figure B.11 performs similarly to the NTrack selection at 65%. At the event peak the efficiency is barely 20%–the rest split almost equally amongst π^-, π^+ and μ^+—but increases steadily to 85% at higher momentum, where to wrong-sign component vanishes. The total wrong-sign contribution is 12% but is dominant at low momentum. The π^- contribution is sizeable at 22%, which dies off at higher momentum.

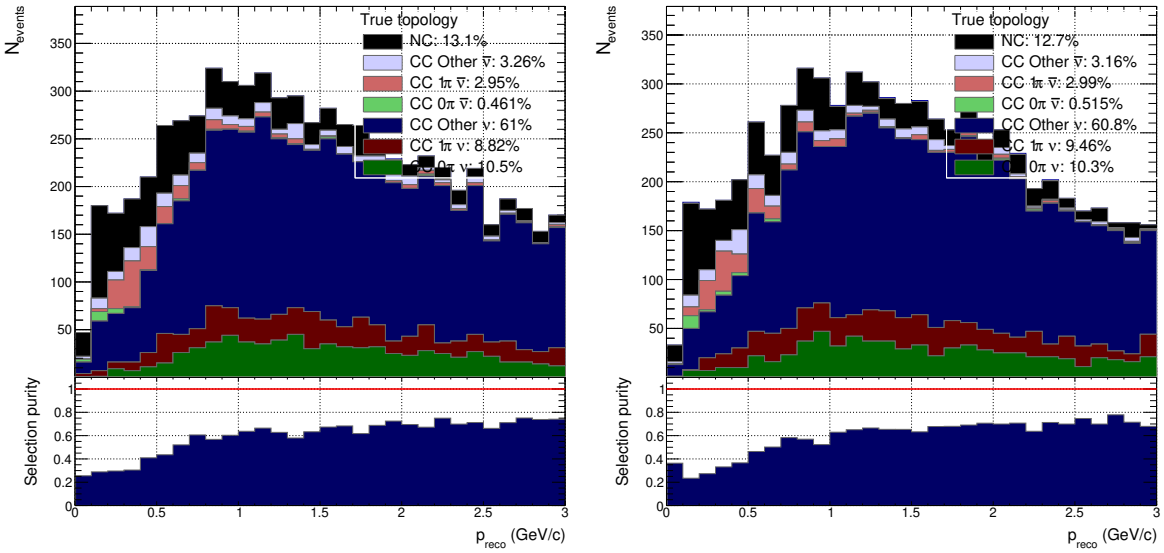


(a) FGD1 **(b)** FGD2

Figure B.11.: Breakdown of ν_μ RHC CC1π selection events' true lepton candidate for FGD1 and FGD2

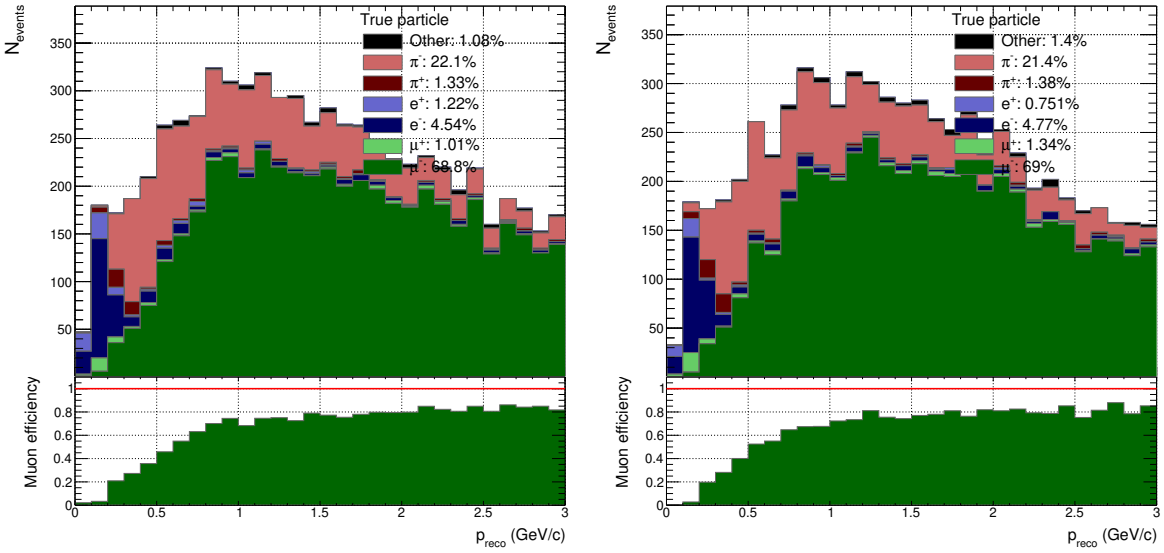
The ν_μ RHC CCOther selection's purity seen in Figure B.12 is relatively high compared to other CC Other selection; overall 61%. At low momentum the NC contribution is the largest, which is also the largest background overall at 13%. Interestingly, the right-sign 0π selection contaminates the sample 10% and is the second largest contamination. Since the sign selection looks for a negative track for ν_μ selections, the CC0π contribution can not come from a proton track being the muon candidate, and must be broken tracks being reconstructed as multiple pions.

The corresponding muon efficiency is shown in Figure B.13, where we see close to zero efficiency at low momentum. In this region the electron is the principal muon candidate, but dies down above 200 MeV. After that the π^- is the only competing background at $\sim$20%. The overall efficiency is 68% and stabilises at 1 GeV, coinciding with the event distribution peak.



(a) FGD1 **(b)** FGD2

Figure B.12.: Breakdown of ν_μ RHC CCOther selection events' true event topology for FGD1 and FGD2



(a) FGD1 **(b)** FGD2

Figure B.13.: Breakdown of ν_μ RHC CCOther selection events' true lepton candidate for FGD1 and FGD2

Appendix C

2D Nominal Monte-Carlo Distributions, 2017

This chapter presents on the two dimensional $p_\mu \cos\theta_\mu$ distributions for each selection present in the 2017 analysis. The simulated distributions use the prior central value and is referred to as the "nominal".

C.1. FGD1 ν_μ FHC

Figure C.1 shows the $p_\mu \cos\theta_\mu$ distributions and their ratios for FGD1, normalised to bin width. For FGD1 CC0π, the largest Data/MC discrepancies are in the very forward region around 500-1000 MeV/c, with some areas of low cross-section (e.g. p_μ 2-5 GeV/c, $\cos\theta_\mu$ 0.85-0.9) mismodelled. The lines of constant Q^2 suggest high Q^2 behaviour is over-estimated in Monte-Carlo, whereas for $0.05 < Q^2 < 0.15\,\text{GeV}^2$ it is under-estimated. For CC1π the most forward-goign bins are almost consistently under-estimated. As for CC0π, the $Q^2 > 0.1\text{GeV}^2$ region is over-estimated, but it is less clear at lower Q^2. For CCOther there is a band-like behaviour in Q^2 going from over estimation to underestimation up until $Q^2 \sim 0.1\,\text{GeV}^2$. The high-momentum areas are mostly under-estimated in Monte-Carlo. It is also clear that ND280 are dominated by $0.05 < Q^2 < 0.30\,\text{GeV}^2$ events.

C.2. FGD2 ν_μ FHC

Figure C.2 shows the same ν_μ selections for FGD2. The CC0π selection is very similar to the FGD1 CC0π selection, whereas the CC1π selection appears better modelled for FGD2 than FGD1, although the opposite is true for CCOther.

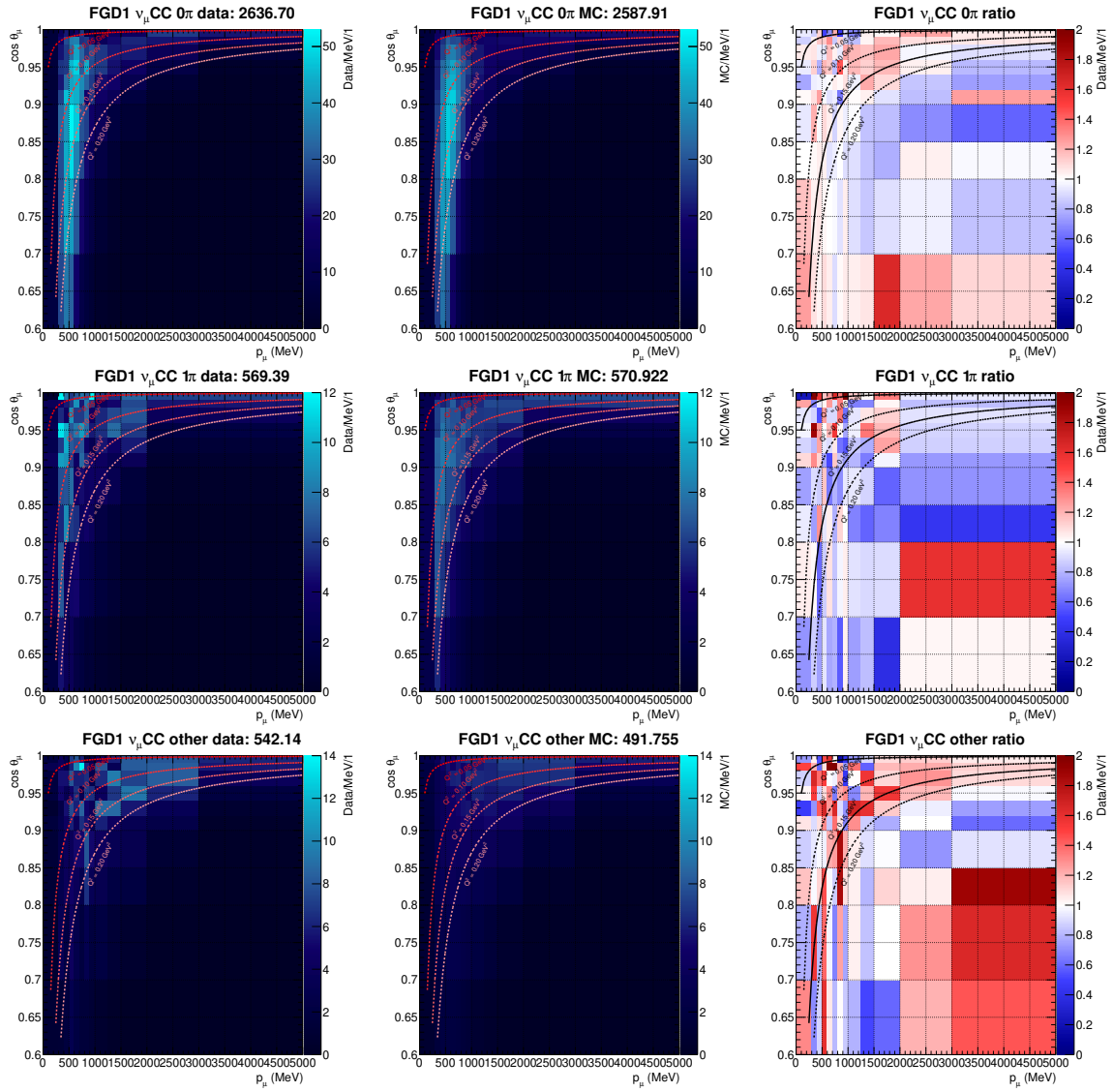


Figure C.1.: Data and nominal MC distributions and the Data/MC ratio for FGD1 ν_μ selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

C.3. FGD1+2 $\bar{\nu}_\mu$ RHC

Figure C.3 shows the $p_\mu \cos\theta_\mu$ for the $\bar{\nu}_\mu$ selection, which again sees mostly consistent behaviour for the two FGDs for both the 1Track and NTracks selection. The event are mostly underestimated at low p_μ and become overestimated as we go up in $\cos\theta_\mu$. The Q^2 bands appear present, notably in the 1 Track selections for $0.05 < Q^2 < 0.10\ \text{GeV}^2$.

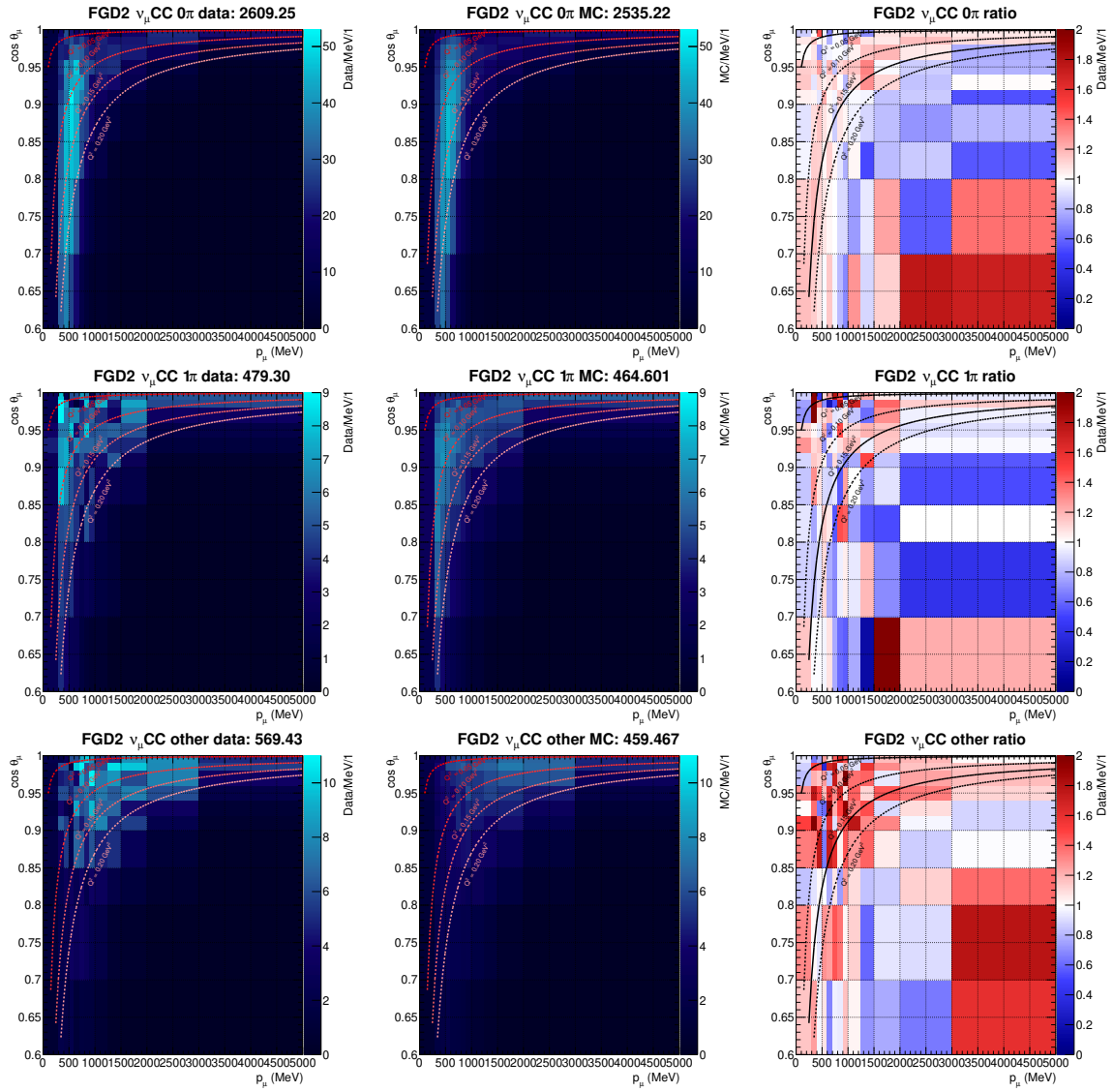


Figure C.2.: Data and nominal MC distributions and the Data/MC ratio for FGD2 ν_μ selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

C.4. FGD1+2 ν_μ RHC

Figure C.3 shows the ν_μ in RHC selections, which generally populate higher p_μ due to the ν_μ flux in RHC mode. The samples are also statistically small so are more prone to larger statistical fluctuations. The CC1Track selection appears to have a pattern of underestimation at low p_μ, following the Q^2 band to high p_μ and high $\cos\theta_\mu$.

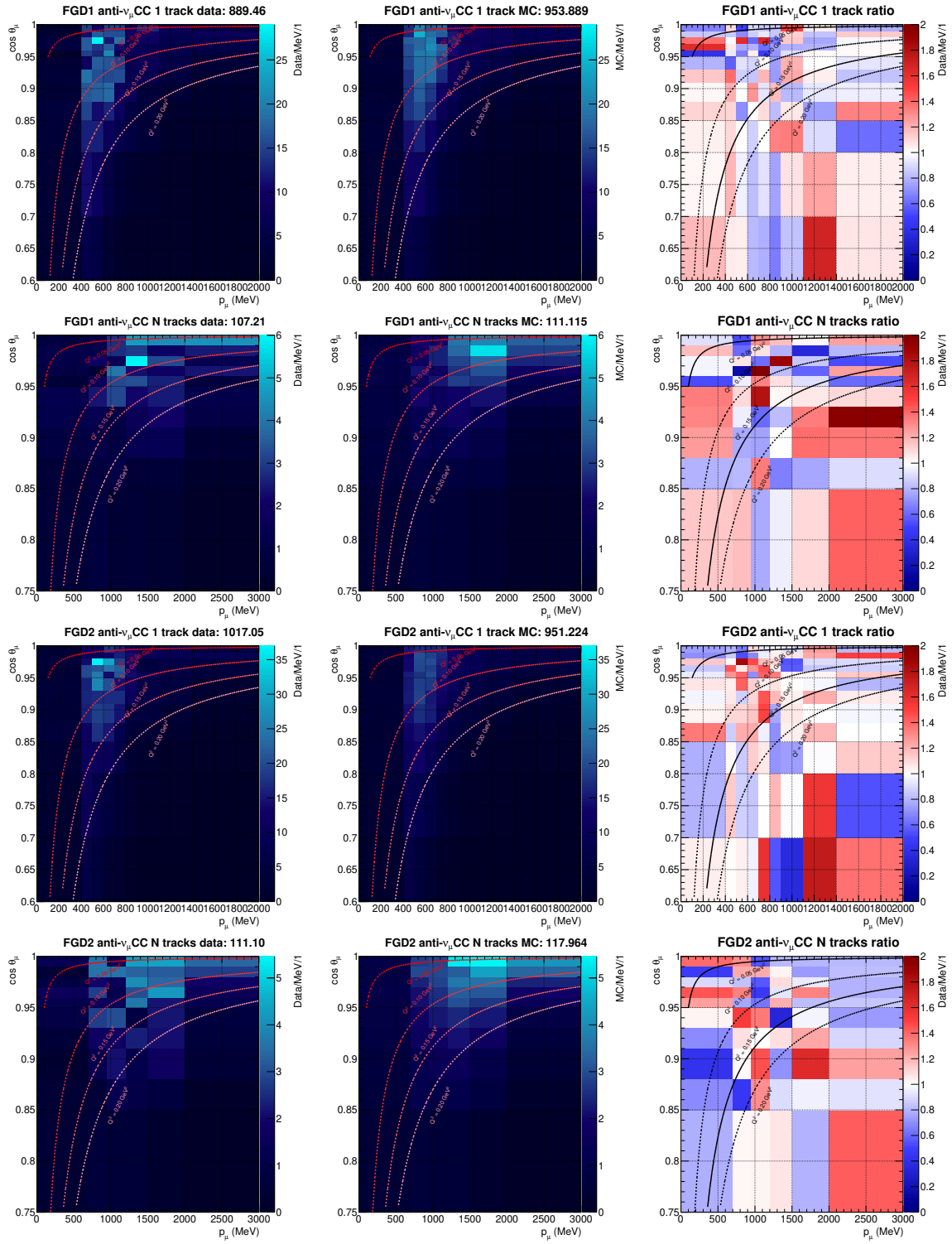


Figure C.3.: Data and nominal MC distributions and the Data/MC ratio for FGD1 and FGD2 $\bar{\nu}_\mu$ selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

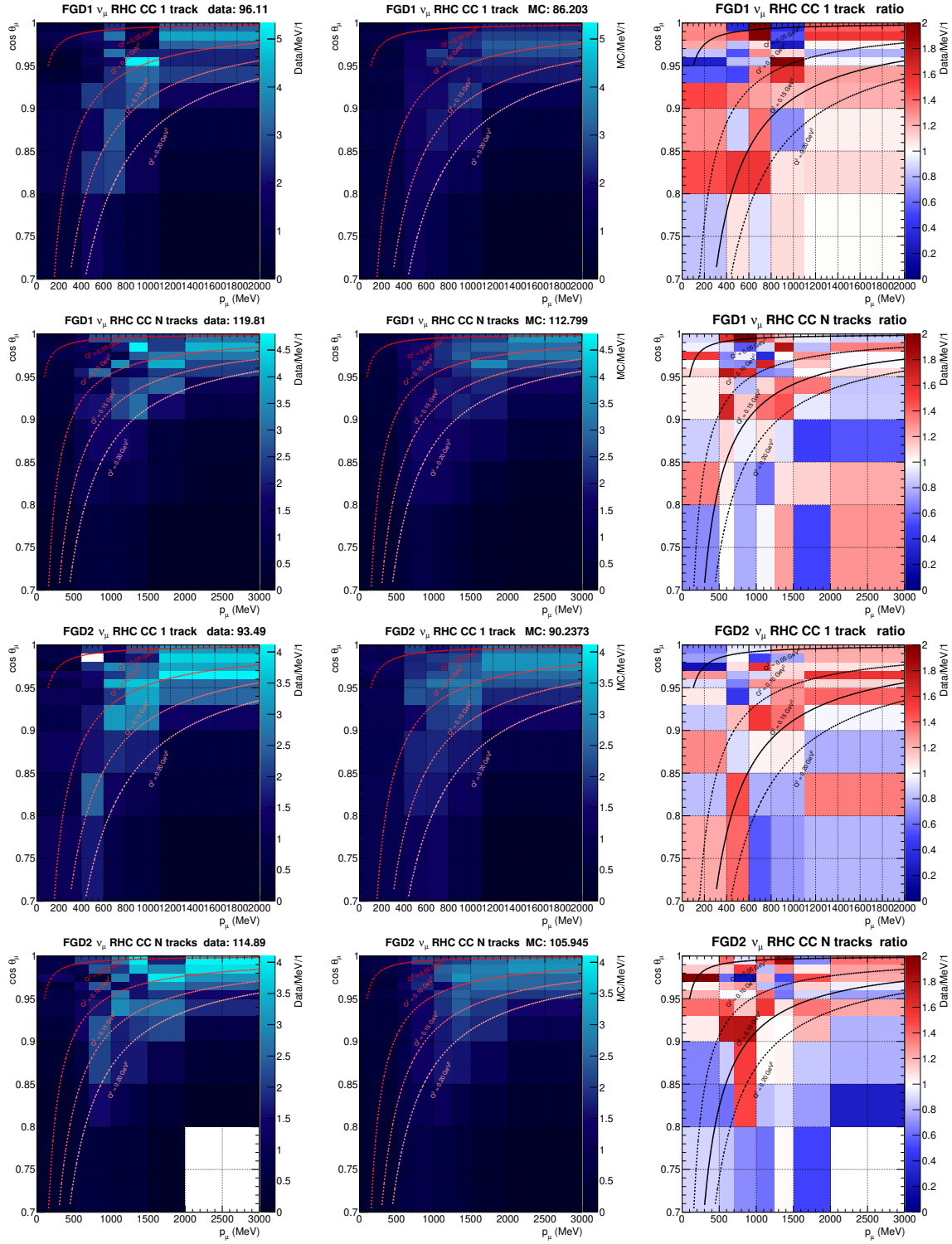


Figure C.4.: Data and nominal MC distributions and the Data/MC ratio for FGD1 and FGD2 ν_μ in RHC selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

Appendix D

2D Nominal Monte-Carlo Distributions, 2018

This chapter presents on the two dimensional $p_\mu \cos\theta_\mu$ distributions for each selection present in the 2018 analysis. The simulated distributions use the prior central value and is referred to as the "nominal".

D.1. FGD1 ν_μ FHC

Figure D.1 shows the data and nominal model prediction for the FGD1 FHC selections. We see in the restricted plotting region (excluding highest momentum and most backward bins, normalising to bin width), the data is consistently higher than the prediction. The CC0π selection looks overestimated at higher momentum between $\cos\theta_\mu$ =0.8-0.95. We also see some clear underestimates along lines of constant Q^2, between 0.07 and 0.15 GeV^2. For the CC1π selection we see a similar overestimate at high p_μ and $\cos\theta_\mu$ =0.8-0.95. For the CCOther selection we see a clear underestimate in almost all bins below $Q^2 = 0.2$ GeV^2. In general, the distributions are compatible and similar to the 2017 equivalents in Figure C.1.

D.2. FGD2 ν_μ FHC

Figure D.2 shows the nominal FHC ν_μ distributions for FGD2, with very similar behaviour to the FGD1 distributions.

D.3. FGD1 $\bar{\nu}_\mu$ RHC

Figure D.3 shows the first light for the new RHC CC0π, CC1π and CCNTrack selections. The CC0π selection is slightly over-estimated, but the nominal prediction

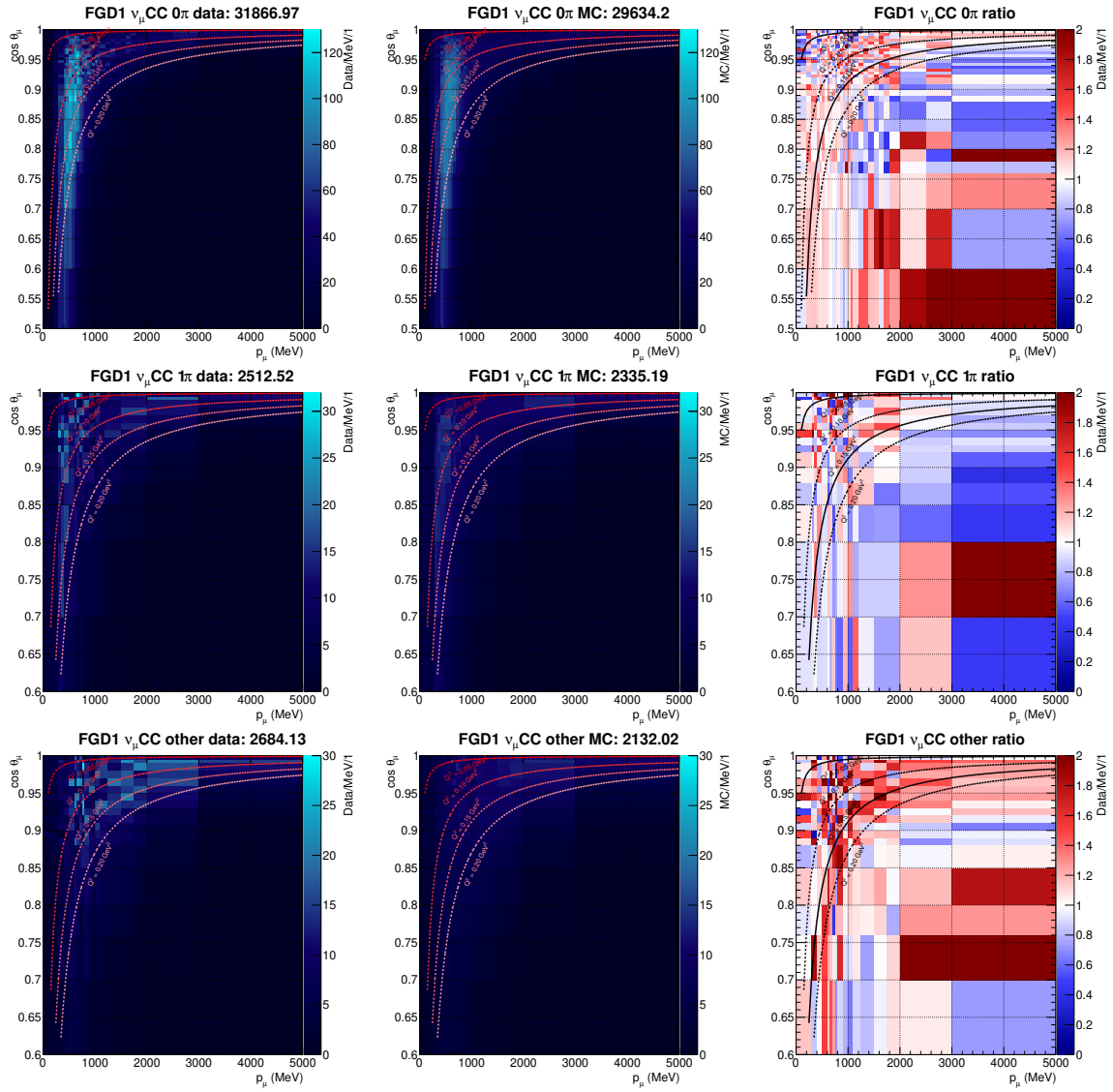


Figure D.1.: Data and nominal MC distributions and the Data/MC ratio for FGD1 FHC selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

looks more compatible with data than the FHC ν_μ distributions. The Data/MC ratio also doesn't appear to contain the same deficiency in Q^2. The CC1π distribution is consistently underestimated in the most forward bin, and hints at an overestimation at low Q^2. The CCOther distribution looks similar to the FHC equivalents in that it is almost consistently underestimated.

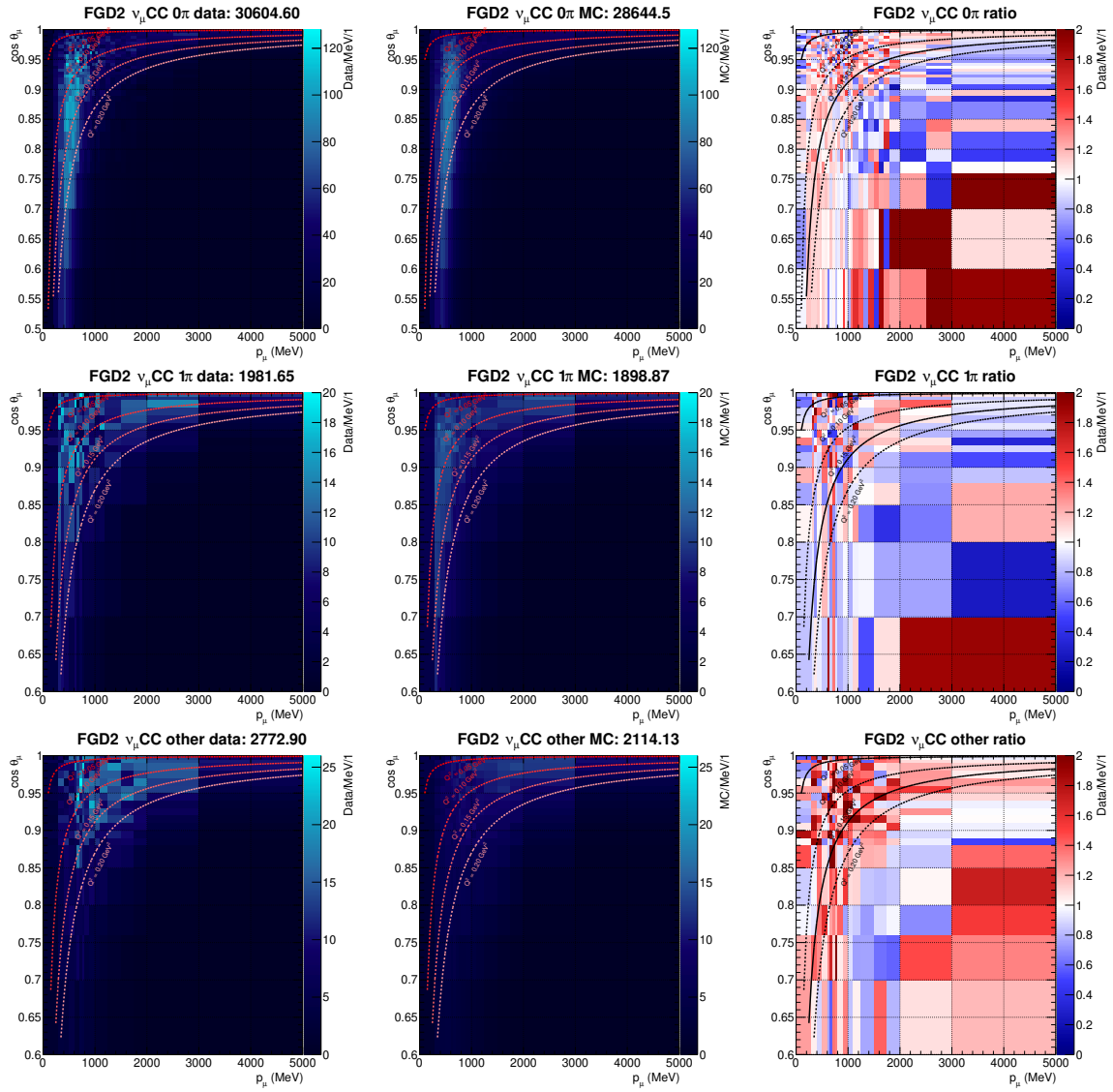


Figure D.2.: Data and nominal MC distributions and the Data/MC ratio for FGD2 FHC selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

D.4. FGD2 $\bar{\nu}_\mu$ RHC

Figure D.4 shows the new RHC selections for FGD2. The CC0π distribution is underestimated, in contrast to FGD1, and looks more similar to the FHC selections with patters of underestimation looking roughly constant in Q^2. The CC1π distribution appears to underestimate in Q^2 rather than overestimate as was the case for FGD1. The CCOther distribution however largely looks compatible with the FGD1 case and is underestimated in general.

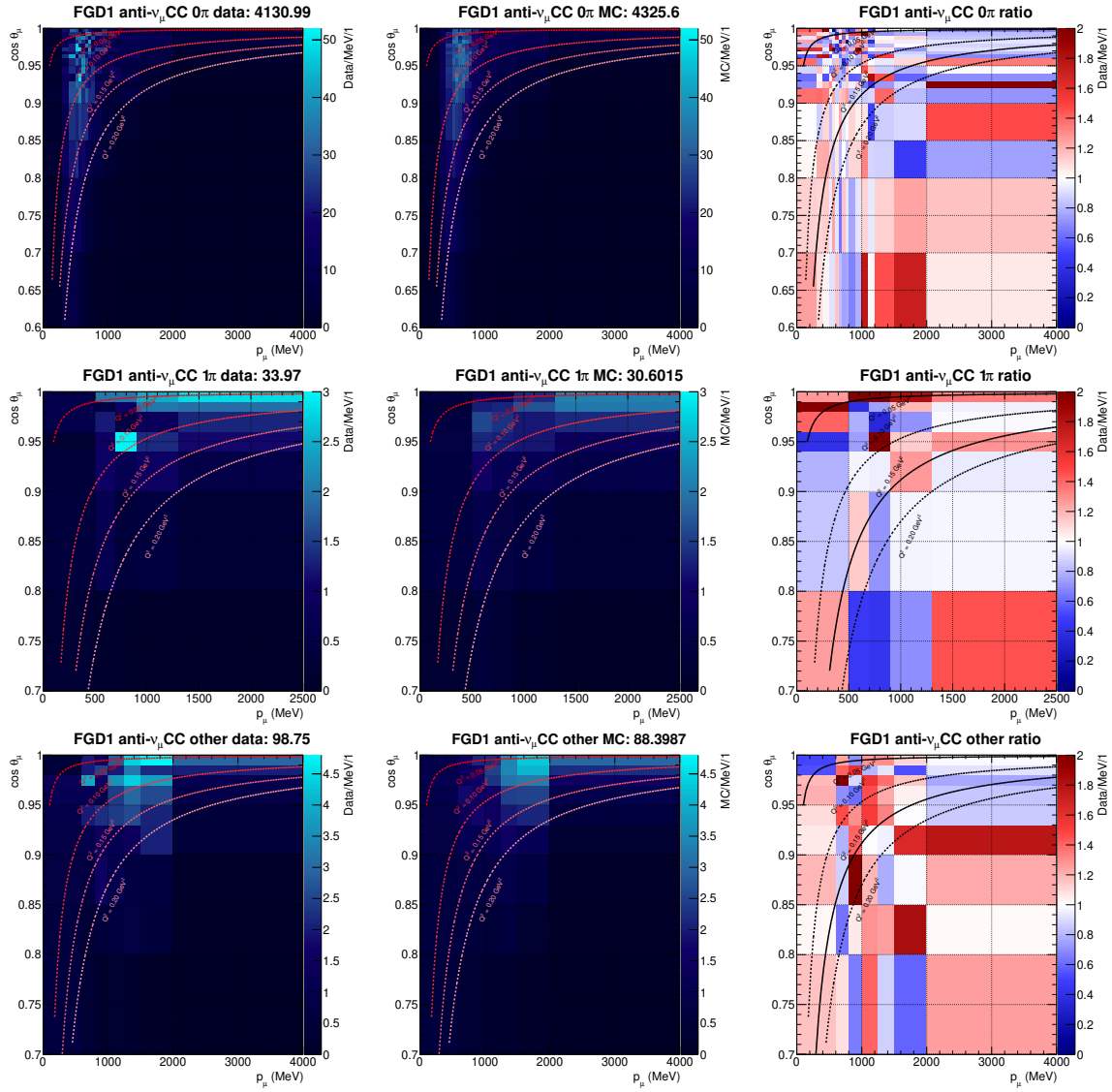


Figure D.3.: Data and nominal MC distributions and the Data/MC ratio for FGD1 $\bar{\nu}_\mu$ selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

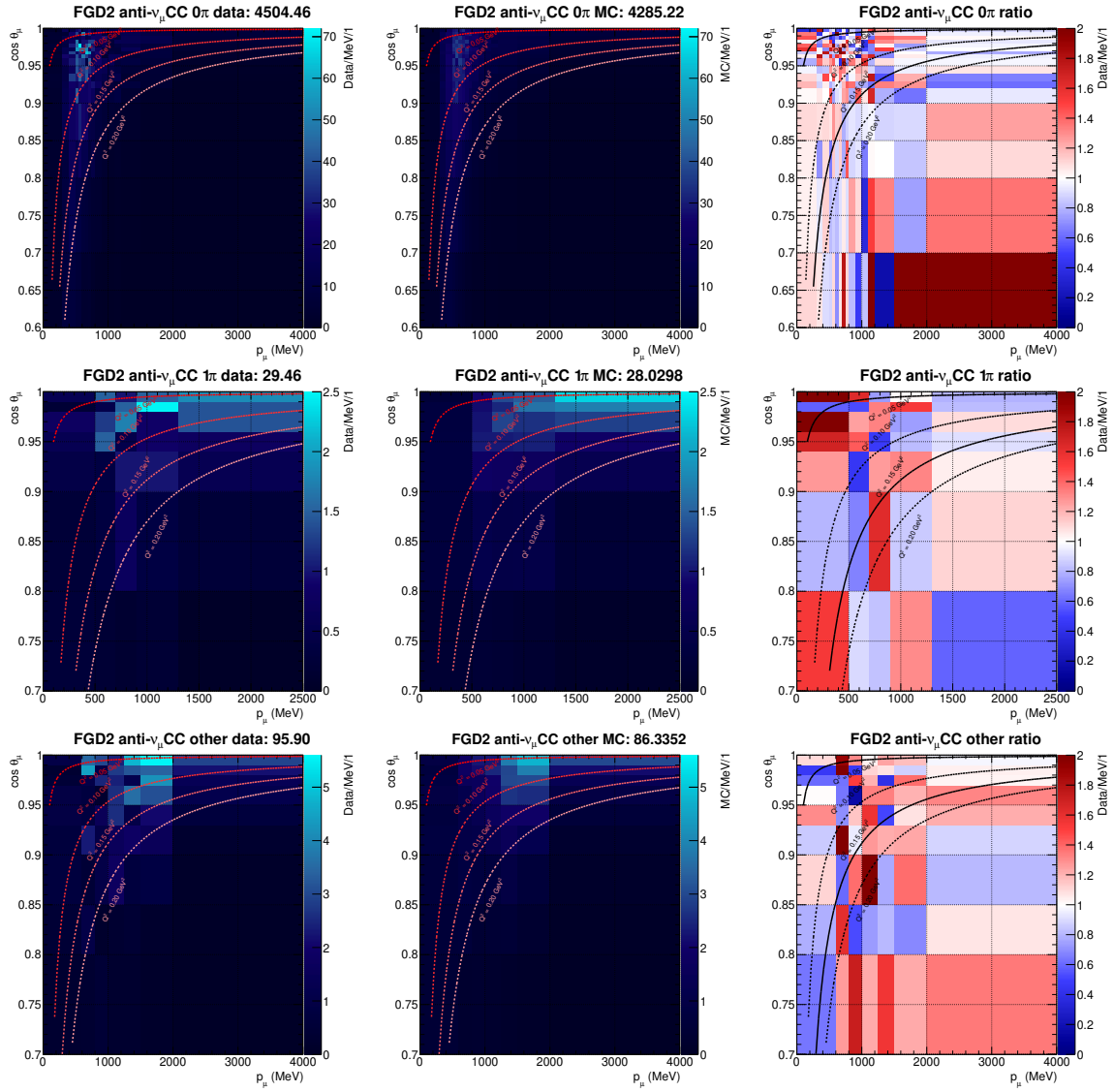


Figure D.4.: Data and nominal MC distributions and the Data/MC ratio for FGD2 $\bar{\nu}_\mu$ selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

D.5. FGD1 ν_μ RHC

Figure D.5 shows the new RHC ν_μ selections for FGD1. As expected, they are concentrated at higher p_μ, owing to the neutrino parents producing neutrinos of higher E_ν. The distributions are consistently underestimated, although the shapes are fairly well reproduced. The CC0π distribution again looks underestimated in constant Q^2, similar to the FHC ν_μ and RHC $\bar{\nu}_\mu$ selections. The CC1π selection is also similar to the RHC $\bar{\nu}_\mu$ equivalent and is underestimated at high $\cos\theta_\mu$ and high p_μ. The CCOther selection agrees well with the previous CCOther selections, being even more underestimated than for the others.

D.6. FGD2 ν_μ RHC

The FGD2 ν_μ RHC distributions in Figure D.6 are similarly underestimated throughout. For CC0π there is consistent underestimation at high p_μ and $\cos\theta_\mu$, which agrees with the FGD1 distribution. The data appears shifted towards higher momentum to the prediction, and have a larger spread. The CC1π selection is similarly patchy and is difficult to draw conclusions from. CCOther is consistent with FGD1, being mostly underestimated, although overestimated at low momentum.

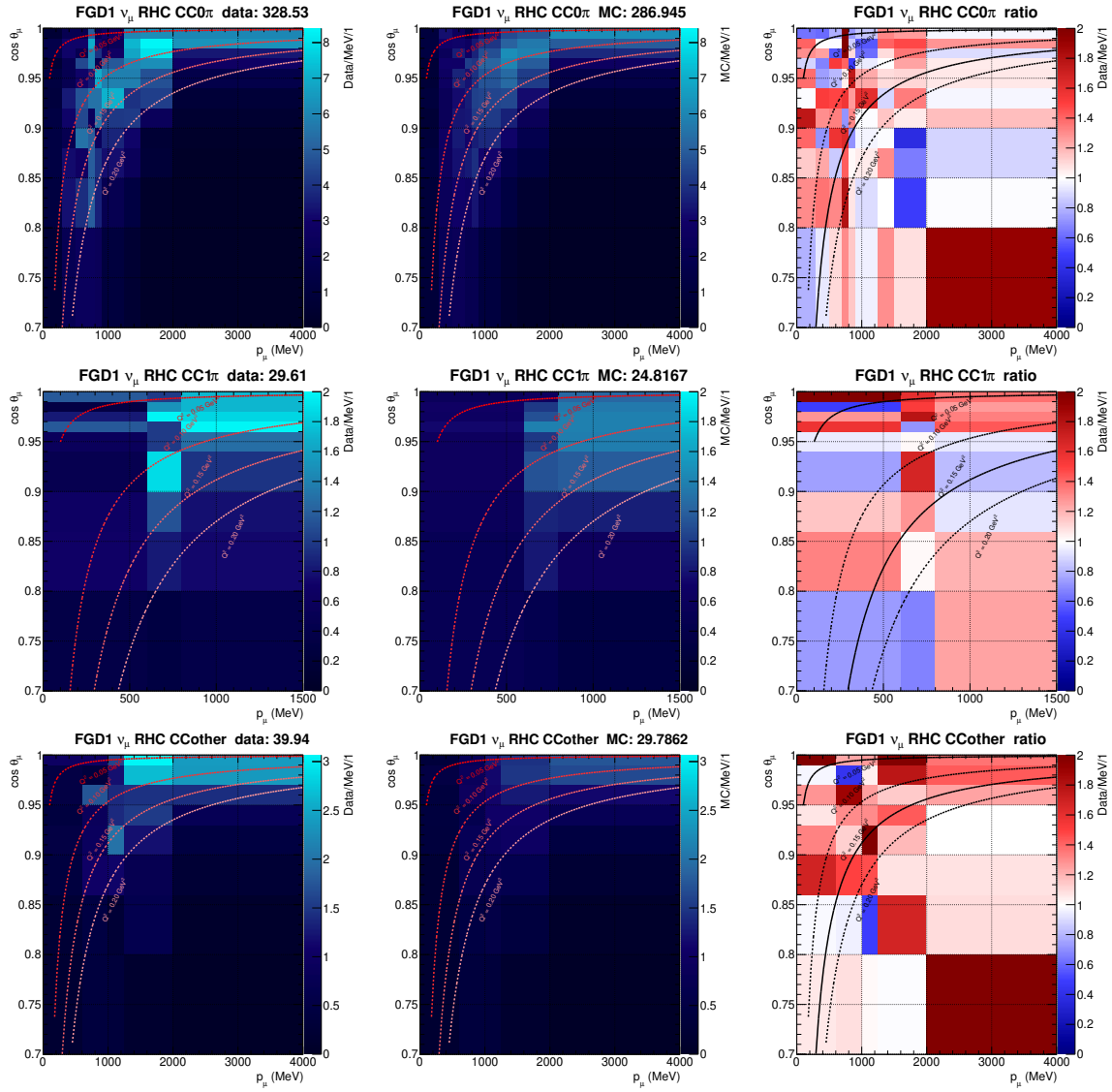


Figure D.5.: Data and nominal MC distributions and the Data/MC ratio for FGD1 ν_μ RHC selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

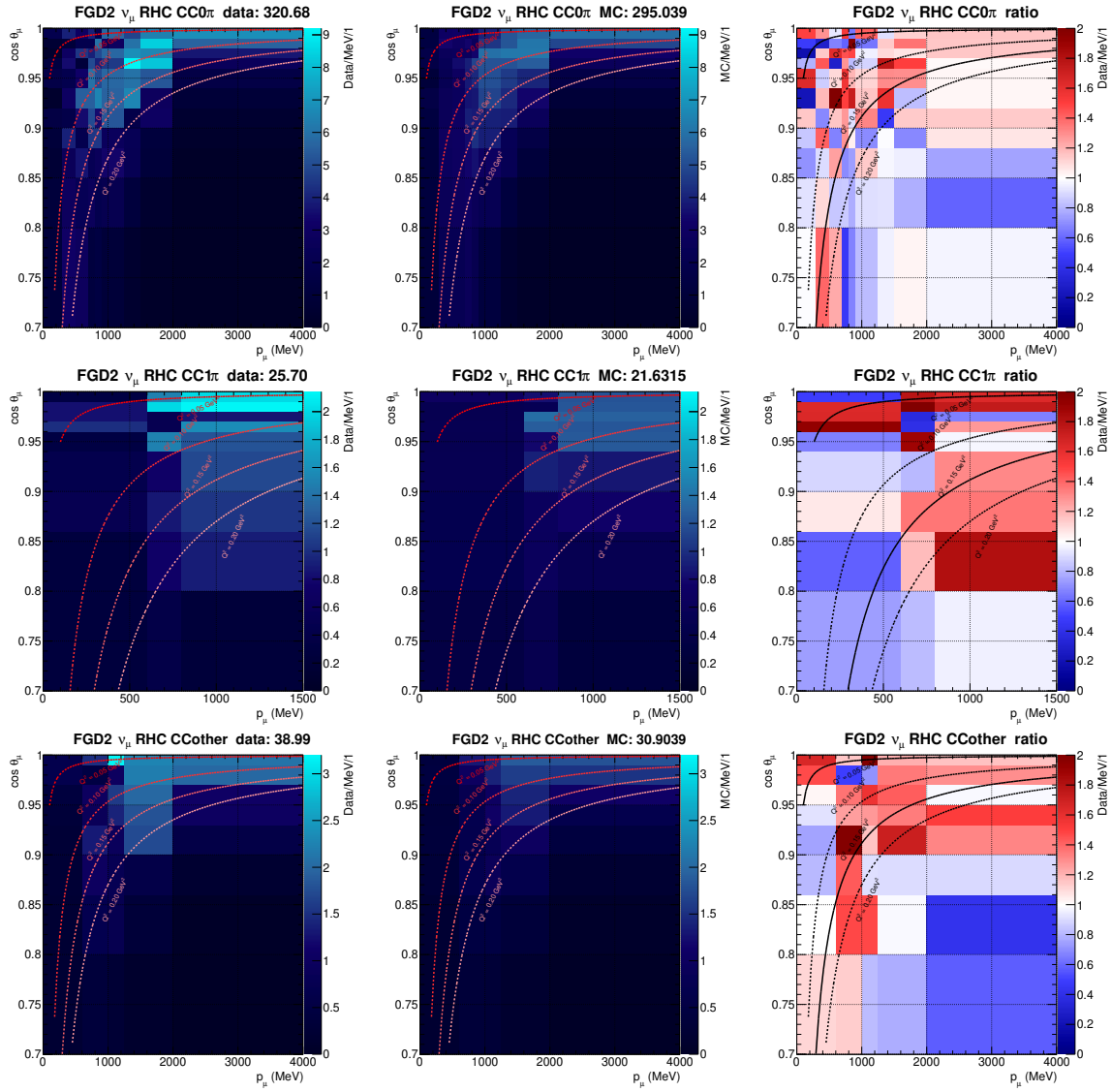


Figure D.6.: Data and nominal MC distributions and the Data/MC ratio for FGD2 ν_μ RHC selections. Lines of constant Q^2_{reco} are shown. Bin content is normalised to bin width

Appendix E

Validating to the BANFF Framework

As outlined in chapter 5, T2K has two near-detector fitters: one Markov Chain Monte Carlo fitter ("MaCh3") and one frequentist gradient descent fitter ("BANFF"). The results of the 2017 fits has been presented using the former, and this section shows the validation tests throughout the fitting process between the two fitters.

The validation procedure is not expected to entirely agree at all stages due to numerous technical and statistical differences. As covered earlier, a gradient descent minimiser is designed to find a global minimum of the test-statistic whereas a MCMC samples to build up a posterior distribution in high dimension. Comparing parameter values between the two is not a trivial exercise, and differences are expected.

An example of a technical difference is MaCh3 evaluates the weights from cross-section parameter variations on GPGPUs[a], which involves floating point operations in 32 (float) rather than 64 bits (double). The effect of this on 35,000 events from run 3c is shown in Figure E.1, where we observe weight differences on the 1E-7 scale per event, which agrees well with the float vs double precision. Hence we expect small differences in event rates, although too small to notice in a likelihood evaluation.

Weights from cross-section parameters were also compared for a selected few events and was found to be accurate to 1E-6.

E.1. Nominal Model Prediction

The nominal model and selection are run through both the MaCh3 and BANFF frameworks and the nominal event rates are compared in Table E.1. The largest difference is seen in the FGD1 1Trk selection $\mathcal{O}(10^{-3})$ percent and the total difference is $\mathcal{O}(10^{-4})$ percent, which were deemed acceptable.

Figure E.2 and Figure E.3 show the 2D $p_\mu \cos\theta_\mu$ distributions for FGD1 and FGD2 respectively, with BANFF-MaCh3 on the z-axis. We see the differences in Table E.1

[a]General Purpose Graphics Processing Units

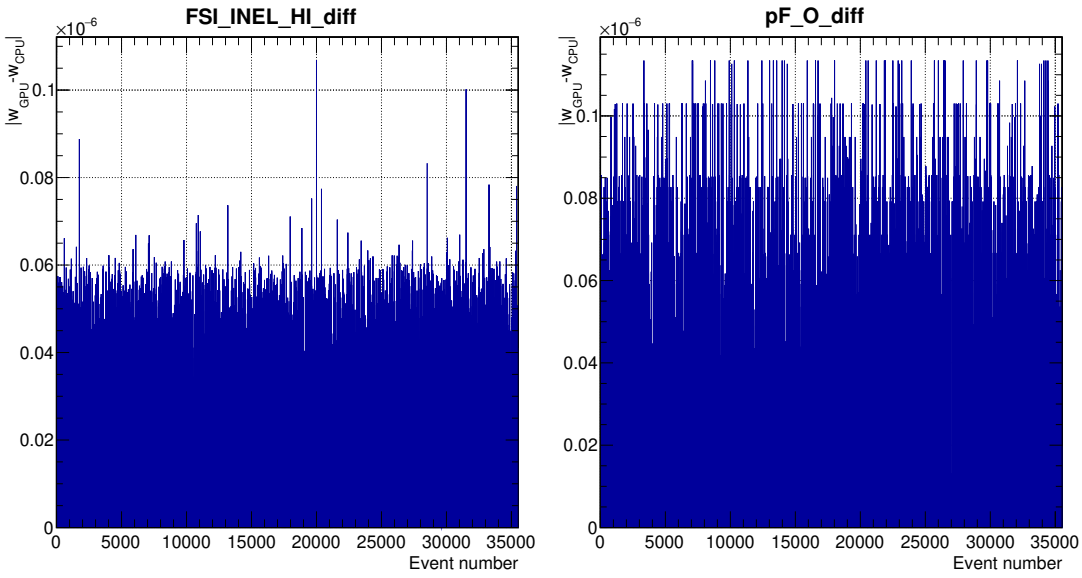
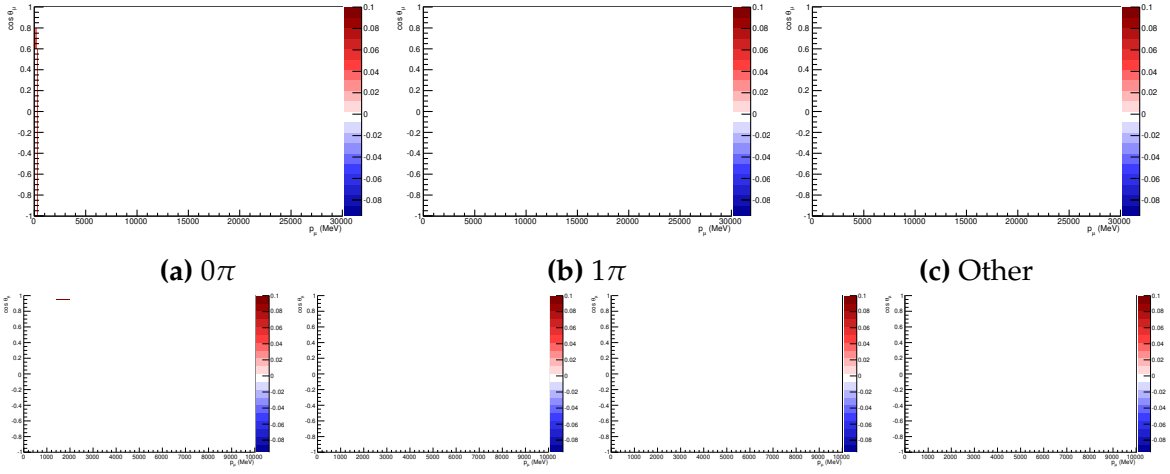


Figure E.1.: Absolute weight differences using a GPU versus CPU for a random parameter variation in MaCh3 for Run 3c Monte-Carlo

Sample	Data	BANFF	MaCh3	$\frac{\text{BANFF-MaCh3}}{\text{BANFF}}$
FGD1 0π	17136	16723.69	16723.8	-6.60E-6
FGD1 1π	3954	4381.48	4381.47	2.28E-6
FGD1 Other	4149	3943.95	3943.95	0E0
FGD2 0π	17443	16959.19	16959.3	-6.49E-6
FGD2 1π	3366	3564.23	3564.23	0E0
FGD2 Other	4075	3570.95	3570.94	2.80E-6
FGD1 1Trk	3527	3587.65	3587.77	-3.34E-5
FGD1 NTrk	1054	1066.91	1066.91	0E0
FGD2 1Trk	3732	3618.27	3618.29	-5.53E-6
FGD2 NTrk	1026	1077.24	1077.24	0E0
FGD1 ν_μ 1Trk	1363	1272.17	1272.17	0E0
FGD1 ν_μ NTrk	1370	1357.45	1357.45	0E0
FGD2 ν_μ 1Trk	1320	1262.63	1262.63	0E0
FGD2 ν_μ NTrk	1253	1246.71	1246.71	0E0
Total	64768	63632.53	63632.9	-5.81E-6

Table E.1.: BANFF and MaCh3 comparison of final event rates for the nominal model and data

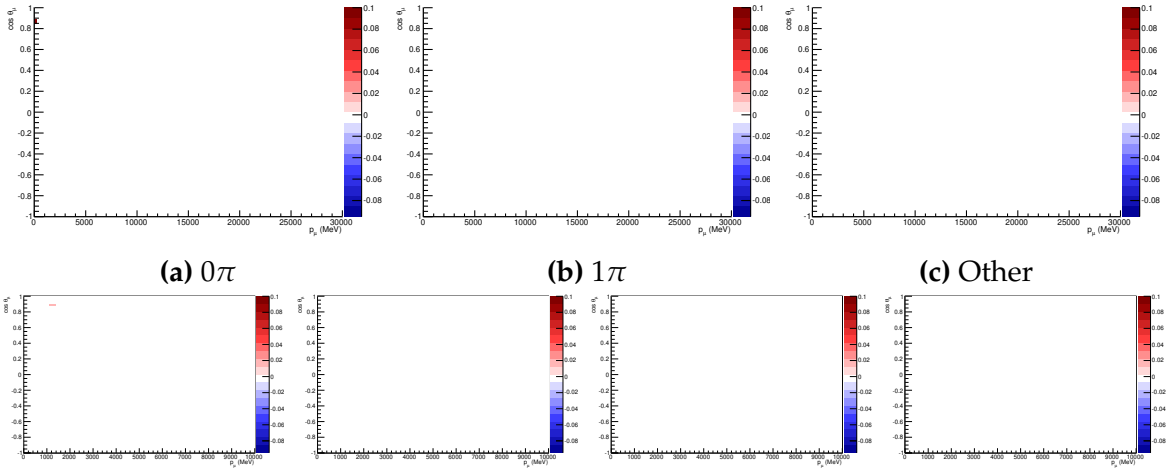
come primarily from one or two bins, and it was found these differences were due to the interaction parameter evaluations mentioned earlier.



(a) 0π **(b)** 1π **(c)** Other

(d) 1Trk **(e)** NTrk **(f)** ν_μ 1Trk **(g)** ν_μ NTrk

Figure E.2.: FGD1 selections showing the nominal MaCh3-BANFF events



(a) 0π **(b)** 1π **(c)** Other

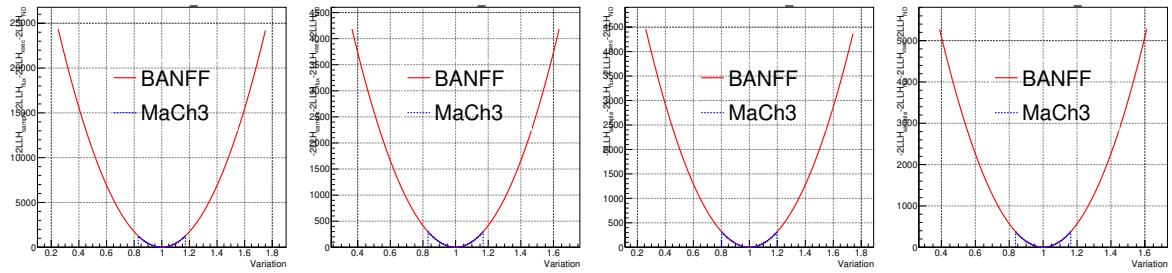
(d) 1Trk **(e)** NTrk **(f)** ν_μ 1Trk **(g)** ν_μ NTrk

Figure E.3.: FGD2 selections showing the nominal MaCh3-BANFF events

E.2. Log-Likelihood Scans

The likelihood scans in subsection 5.6.1 are also validated across the two groups. As then, likelihood response is scanned across a parameter one at a time, and all other parameters are fixed to their nominal values. This validation has two main purposes: checking the likelihood response of the prior term and the sample term. The plots shown below are for the total likelihood.

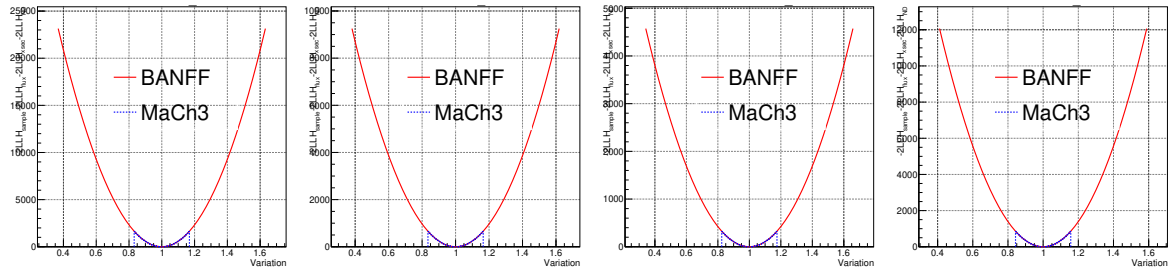
A selection of ND280 flux parameters are shown in Figure E.4. We see very good agreement for the different groups of parameters, and all other flux parameters display this property.



(a) FHC ν_μ 0.6-0.7 GeV **(b)** FHC ν_e 0.7-0.8 GeV **(c)** RHC ν_μ 1.0-1.5 GeV **(d)** RHC ν_μ 0.8-1.5 GeV

Figure E.4.: Likelihood scan comparison between BANFF and MaCh3 for ND280 flux parameters

Figure E.5 shows the same parameters as in Figure E.4 but for SK. We again note perfect agreement for all parameters.

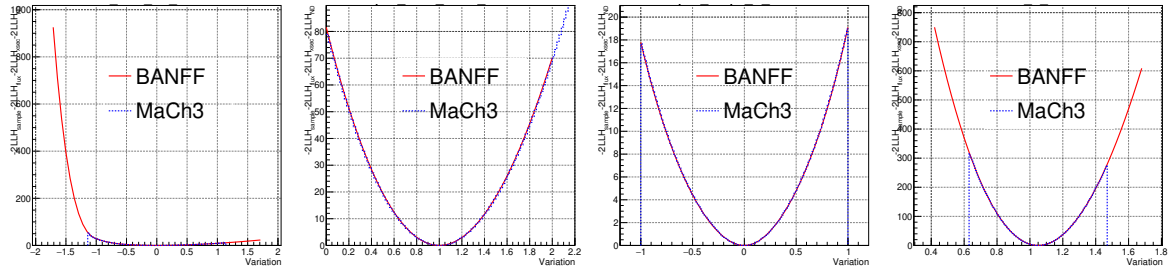


(a) FHC ν_μ 0.6-0.7 GeV **(b)** FHC ν_e 0.7-0.8 GeV **(c)** RHC ν_μ 1.0-1.5 GeV **(d)** RHC ν_μ 0.8-1.5 GeV

Figure E.5.: Likelihood scan comparison between BANFF and MaCh3 for SK flux parameters

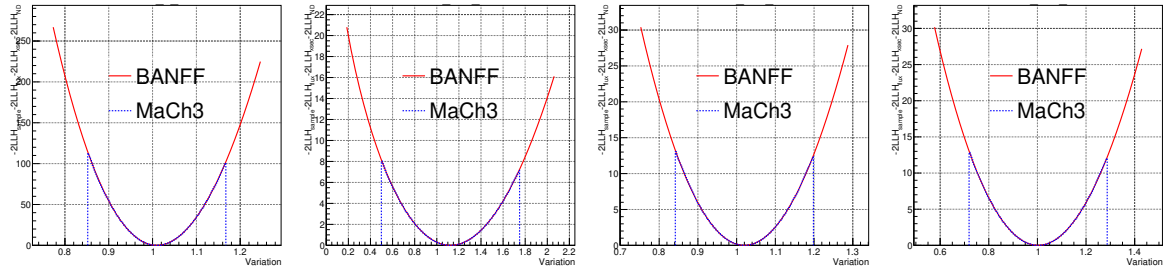
Figure E.6 shows the likelihood responses for a selection of interaction parameters, which also agree well.

Figure E.7 shows the likelihood response for a selection of ND280 detector parameters and again perfect agreement is seen.



(a) FSI Cex lo **(b)** 2p2h norm $\bar{\nu}$ **(c)** 2p2h shape C **(d)** BeRPA B

Figure E.6.: Likelihood scan comparison between BANFF and MaCh3 for interaction parameters



(a) FGD1 CC0π 0-1 GeV, 0.85-0.94 **(b)** FGD1 $\bar{\nu}_\mu$ CC1Trk 3-5 GeV, 0.99-1.00 **(c)** FGD2 CCOther 3-5 GeV, 0.85-0.90 **(d)** FGD2 $\bar{\nu}_\mu$ CCNTrk 1.5-2 GeV, 0.93-0.90

Figure E.7.: Likelihood scan comparison between BANFF and MaCh3 for ND280 detector parameters

E.3. Parameter Variations

We also validate the two frameworks on a parameter-by-parameter basis by varying all the parameters by the pre-fit 1σ. This validation is meant to purely look at the sample response to specific parameters and catch erroneous implementation and p_μ $\cos\theta_\mu$ responses—e.g. CCOther selections have strong responses to 2p2h variations, or a FGD1 CC0π event having a response to an FGD2 CC1π ND280 parameter, or a FHC ν_μ event having a response to a variation of the RHC flux parameters.

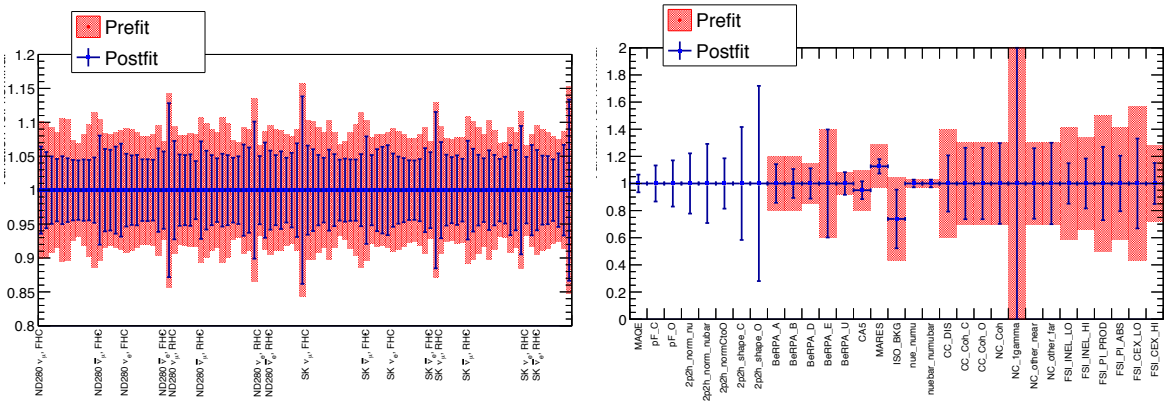
To demonstrate, Table E.2 shows the responses of all the samples to a variation in the BeRPA B interaction parameter. The largest difference is seen in the FGD1 CC0π, followed by the FGD2 CC0π sample, much in accordance with the nominal event rate differences in Table E.1.

Sample	-1σ		Nominal		+1σ		BANFF-MaCh3
	BANFF	MaCh3	BANFF	MaCh3	BANFF	MaCh3	BANFF
FGD1 0π	16081.46	16081.60	16723.69	16723.80	17365.92	17366.10	-1E-5
FGD1 1π	4372.10	4372.09	4381.48	4381.47	4390.85	4390.84	3E-6
FGD1 Other	3934.03	3934.03	3943.95	3943.95	3953.88	3953.87	2E-6
FGD2 0π	16315.62	16315.70	16959.19	16959.30	17602.77	17602.90	-7E-6
FGD2 1π	3556.52	3556.52	3564.23	3564.23	3571.94	3571.94	1E-6
FGD2 Other	3561.12	3561.12	3570.95	3570.94	3580.77	3580.77	0E-6
FGD1 1Trk	3478.23	3478.35	3587.65	3587.77	3697.08	3697.20	-3.3E-6
FGD1 NTrk	1063.27	1063.26	1066.91	1066.91	1070.55	1070.55	1E-6
FGD2 1Trk	3509.99	3510.01	3618.27	3618.29	3726.55	3726.58	-7E-6
FGD2 NTrk	1073.57	1073.57	1077.24	1077.24	1080.91	1080.91	3E-6
FGD1 ν_μ 1Trk	1235.74	1235.74	1272.17	1272.17	1308.60	1308.60	1E-6
FGD1 ν_μ NTrk	1349.19	1349.18	1357.45	1357.45	1365.72	1365.72	3E-6
FGD2 ν_μ 1Trk	1226.51	1226.51	1262.63	1262.63	1298.76	1298.76	1E-6
FGD2 ν_μ NTrk	1239.87	1239.87	1246.71	1246.71	1253.54	1253.54	4E-6

Table E.2.: 1σ variations for BeRPA B in MaCh3 and BANFF. N.B. the last column is calculated with event rates to four decimal places

E.4. Fitting Asimov Data

BANFF performs the same Asimov fit as MaCh3 in Figure E.8. However it does not suffer from marginalisation issues and additionally starts the fit at the local minimum, so has no biases. Because of this, direct comparisons are not particularly enlightening. Comparing to the MaCh3 result in subsection 5.6.4 and Figure 5.34, Figure 5.35 and Figure 5.36 we see compatible uncertainties on all parameters.



(a) Flux parameters **(b)** Interaction parameters

Figure E.8.: BANFF post-fit parameter when fitting the Asimov data set

E.5. Fitting Real Data

We compare the global minimum from BANFF to the 1D marginalised posteriors from MaCh3, keeping the marginalisation issues detailed in Figure 5.6.4 in mind.

The flux parameters are compared in Figure E.9 and we see a consistent shift in the parameters. The shapes are near identical, with BANFF seeing consistently higher flux parameters. This is largely expected due to the correlations of the flux with interaction parameters that showed marginalisation biases in the Asimov fit. The uncertainties on each of the parameters is consistent.

Figure E.10 shows the post-fit interaction parameters where differences aren't as pronounced. The uncertainties on p_F are slightly different due to the method of evaluating the error (arithmetic mean) and the non-Gaussian shape of the one dimensional posterior. 2p2h normalisations and BeRPA A and B are different, as

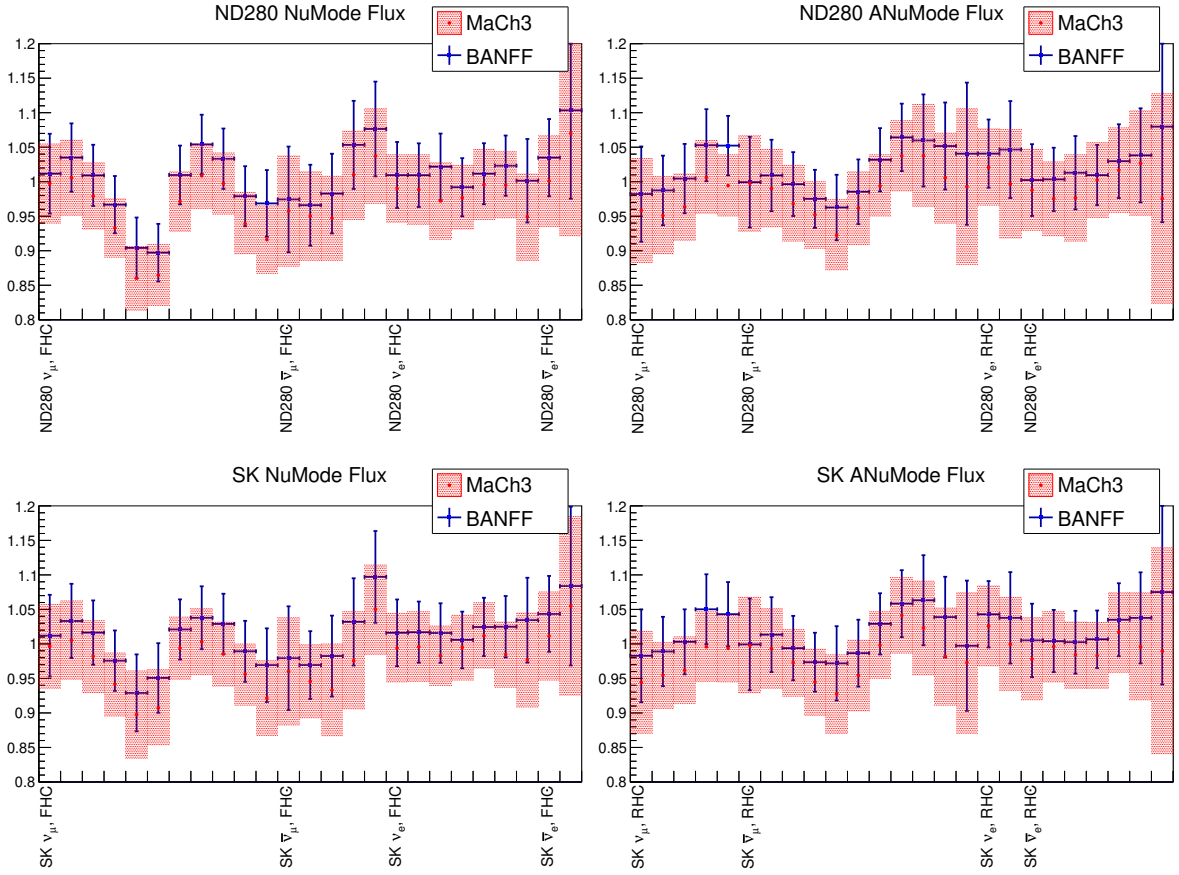


Figure E.9.: Flux parameters post-fit for BANFF and MaCh3

expected from the marginalisation bias present in the Asimov. The CC DIS parameter is also "fit lower" in MaCh3, likely due to the correlations with the flux parameters which are also fit higher.

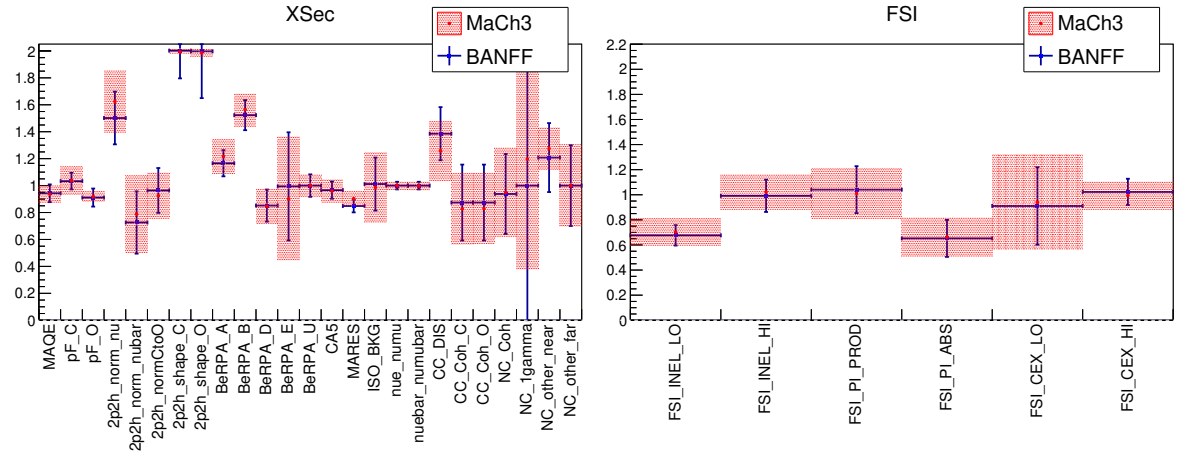


Figure E.10.: Interaction parameters post-fit for BANFF and MaCh3

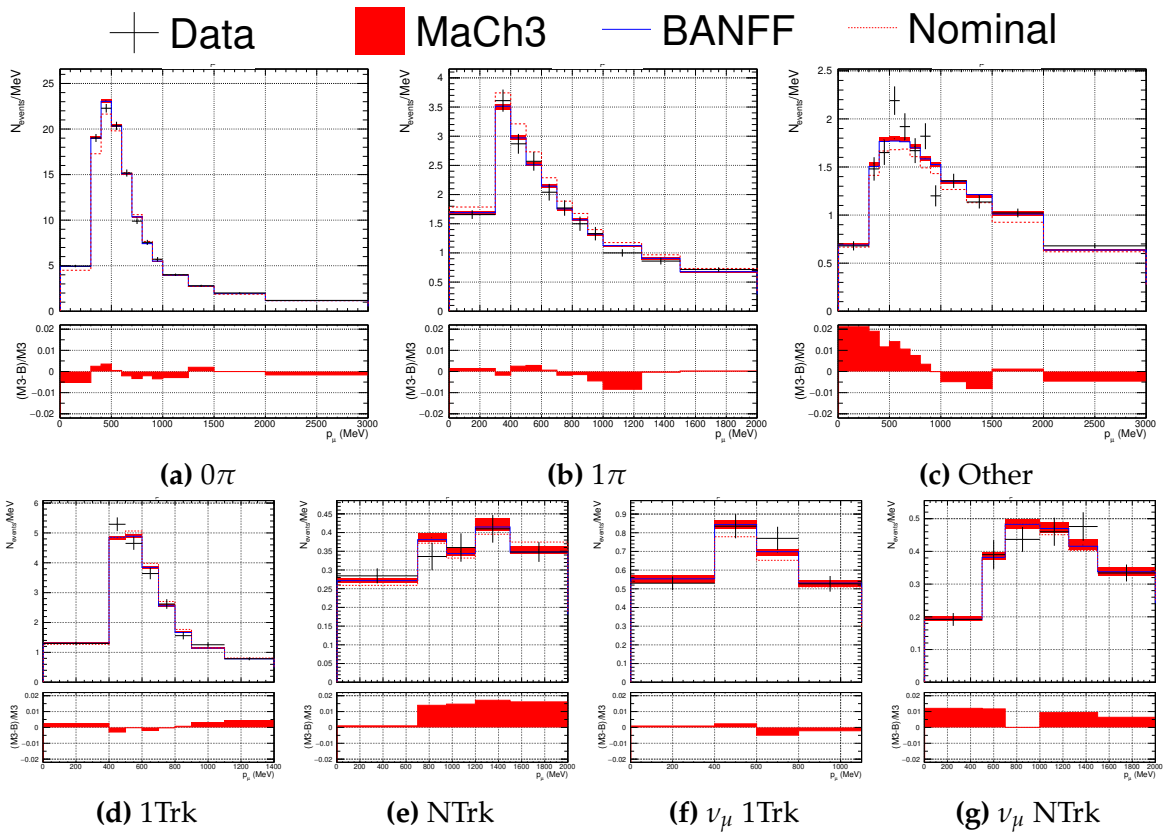
Finally we can also compare the final event rates and test-statistics in Table E.3. The difference in number of events is -15.38 (MaCh3 higher), corresponding to -0.02%. The test-statistic from the samples differs by -11.96 and as expected BANFF is lower. Comparing to the total number of fit bins (1624) and degrees of freedom (938), the $\Delta\left(-2\ln\mathcal{L}_s/\text{nDOF}\right) \sim -0.01$. The sample with the largest difference in test-statistic are the 0π selections and FGD2 CCOther, where the difference is approximately 2.

Sample	Event rate		-2 ln $\mathcal{L}_s$	
	BANFF	MaCh3	BANFF	MaCh3
FGD1 0π	17122.22	17123.80	169.97	172.21
FGD1 1π	4061.65	4054.18	164.06	164.04
FGD1 Other	4095.58	4103.96	223.49	224.03
FGD2 0π	17494.56	17500.70	164.40	166.15
FGD2 1π	3416.28	3409.63	162.62	162.71
FGD2 Other	3915.36	3914.40	168.82	171.17
FGD1 1Trk	3503.79	3509.37	117.79	117.80
FGD1 NTrk	1052.69	1062.70	74.98	76.50
FGD2 1Trk	3685.46	3678.66	129.09	129.84
FGD2 NTrk	1097.38	1108.52	78.95	80.34
FGD1 1Trk	1353.44	1347.50	66.99	66.51
FGD1 NTrk	1354.02	1358.99	62.00	61.75
FGD2 1Trk	1330.49	1323.30	62.79	64.34
FGD2 NTrk	1263.12	1265.69	75.17	75.70
Total	64746.02	64761.40	1721.12	1733.08

Table E.3.: Comparison of the event rates for data post-fit MC and -2 ln $\mathcal{L}_s$ contributions broken by sample for MaCh3 and BANFF

E.6. Post fit Distributions

We now compare the post-fit p_μ distributions from BANFF and MaCh3. The BANFF post-fit uses the parameter set which finds the global minimum, and MaCh3 uses the posterior predictive spectrum. The error on the MaCh3 prediction is the total uncertainty.



(a) 0π **(b)** 1π **(c)** Other

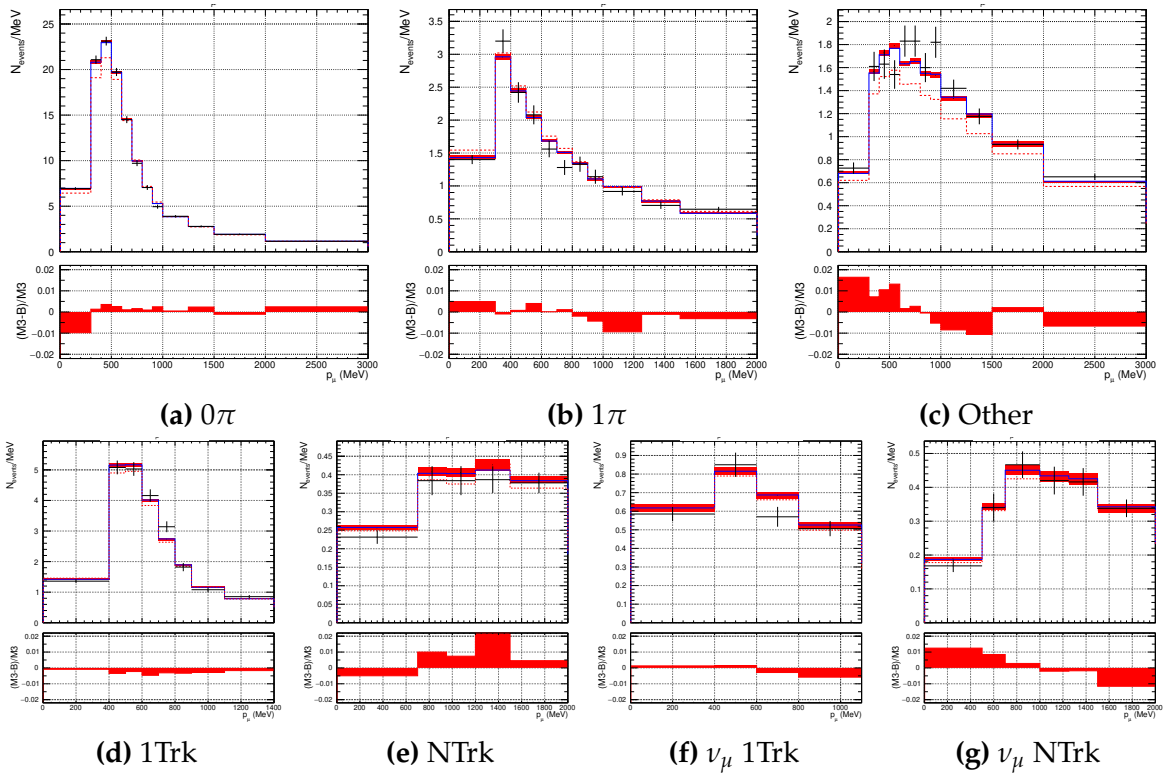
(d) 1Trk **(e)** NTrk **(f)** ν_μ 1Trk **(g)** ν_μ NTrk

Figure E.11.: FGD1 selections in p_μ with data, prefit, BANFF postfit and MaCh3 postfit

Figure E.11 shows the post-fit predictions for FGD1 and Figure E.12 for FGD2. The 0π, 1π and 1 track distributions are identical to within 1%. The Other and NTrk distributions show the largest differences, just above 2% for FGD1 CCOther.

E.7. SK prediction

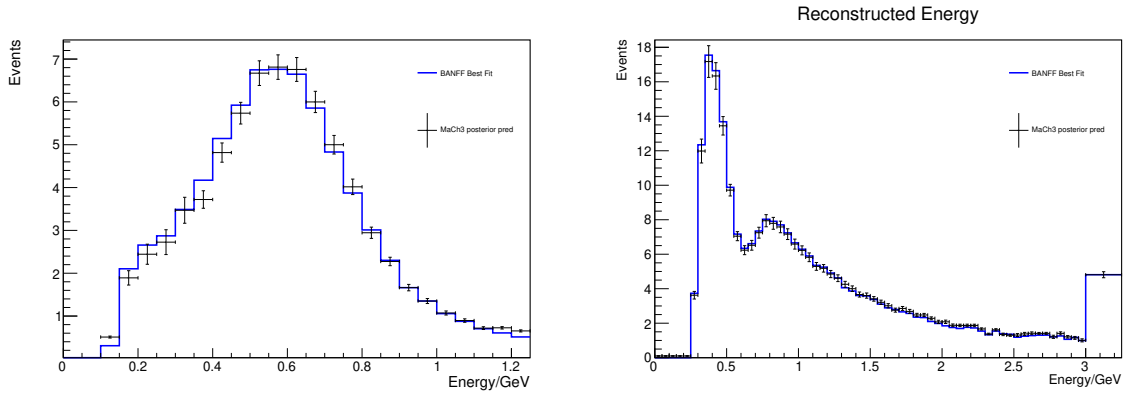
The BANFF central values and MaCh3 MCMC were used to make predictions in reconstructed neutrino energy, shown in Figure E.13. The curves agree well within 1σ (considering the BANFF line is missing the uncertainty band), and were deemed consistent.



(a) 0π **(b)** 1π **(c)** Other

(d) 1Trk **(e)** NTrk **(f)** ν_μ 1Trk **(g)** ν_μ NTrk

Figure E.12.: FGD2 selections in p_μ with data, prefit, BANFF postfit and MaCh3 postfit



(a) 1Re **(b)** 1Rμ

Figure E.13.: Super-Kamiokande event predictions using BANFF and MaCh3

Appendix F

Alternative Studies, 2017 Analysis

The largest concern for the 2017 analysis was the poor posterior predictive p-value, driven primarily by the modelling of the FGD1 CCOther selection. A number of compatibility studies were made to study the effects of this and various other model choices, e.g. using BeRPA and 2p2h shape.

For the alternative studies that had the largest effect on the ND280 fit, we propagate the parameter values results to SK and inspect the difference relative the reference model using the full data set. For the oscillation parameters we fix them to those in Table 5.20 and the SK detector parameters are set to their prior central values and are not varied.

The procedure mimics that of the posterior predictive method, in which MCMC steps are randomly selected after burn-in to give a parameter set $\vec{x}$, used to reweight the Monte-Carlo prediction which is filled in a two-dimensional histogram (N_{toys} vs E_{rec} here). A Gaussian is then fit to the bin-by-bin distributions and the mean and rms is extracted and taken as the central value and 1σ.

F.1. Neutrino vs Anti-Neutrino

The data-sets used for the full fit (runs 2 to 6) spanned neutrino (runs 2 to 4) and anti-neutrino runs (runs 5 to 6), as shown in Table 5.8. As seen in the table, the FHC/RHC ratio is roughly 1.5 and looking at Table 5.10 the FHC/RHC event rate ratio is 3.4. Hence the full run 2 to 6 fit is dominated by FHC running and hence by neutrino rather than anti-neutrino interactions.

To investigate the compatibility between neutrino and anti-neutrino data at T2K, we perform data fits to run 2 to 4 data and to run 5 to 6 data separately, comparing results to the full fit. Due to the different event rates we expect larger constraints on parameters from the FHC data (run 2 to 4) than the RHC data (run 5 to 6), and the overall constraint from the full fit to come primarily from the FHC data.

Figure F.1 shows the ND280 flux parameters after the fit with the prefit uncertainty band in red. As expected, the anti-neutrino fit barely moves the central values for the FHC parameters, and the small movements are due to the covariance matrix penalty tying the FHC and RHC flux together. The full fit often settles between the two separate fits' central values, but in cases of extreme pulls the full fit sits closer to the neutrino fit. The neutrino-only fit pulls the extreme cases a little further (e.g. e.g. FHC ν_μ $\sim$1 GeV, high E_ν ν_μ and $\bar{\nu}_\mu$), highlighting the slight tension with the prior covariance matrix. Generally the flux parameters are compatible throughout and often well within 1σ of each other.

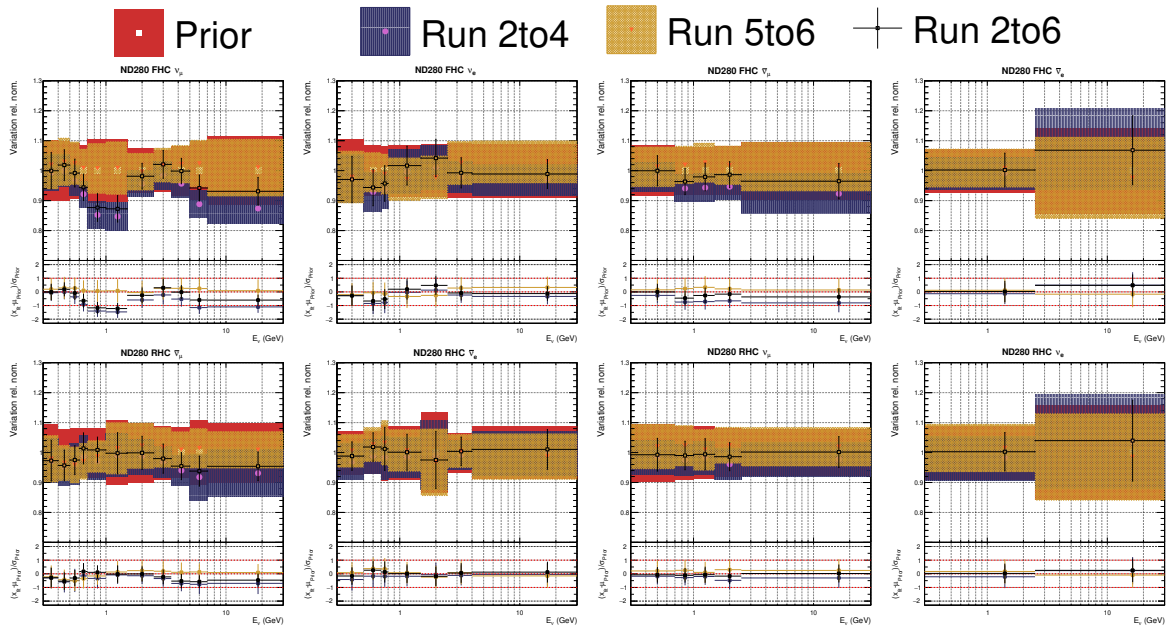


Figure F.1.: ND280 flux parameters after the data fit for different run periods

Figure F.2 shows the interaction parameters where we see more movement than for the flux parameters. As expected, the FHC fit doesn't constrain the $\bar{\nu}$ parameters, such as 2p2h norm $\bar{\nu}$, and the RHC fit barely constrains the ν parameters (only through the ν_μ in RHC samples). Generally the RHC fit results are closer to the prior values, largely due to the smaller statistics. Where the FHC fit pulls strongly (e.g. 2p2h shape at boundary, very high BeRPA B, $I_{1/2}$ background), the RHC fit generally sits inside the 1σ band instead.

Interestingly, the 2p2h norm $\bar{\nu}$ parameter is fitted higher in the RHC fit than it is in the joint fit, likely due to correlations with the flux parameters. We also note a smaller M_A^{QE} value for the anti-neutrino fit ($M_A^{QE} = 1.04 \pm 0.12$ GeV), almost identical to the

bubble chamber value [220]. Larger tensions are present in the single-pion parameters, where the two fits seem to pull M_A^{RES} and $I_{1/2}$ in opposite directions from the central value. Again, the RHC fit barely has enough statistics to pull 1σ from the prior, and more data is likely needed to draw conclusions.

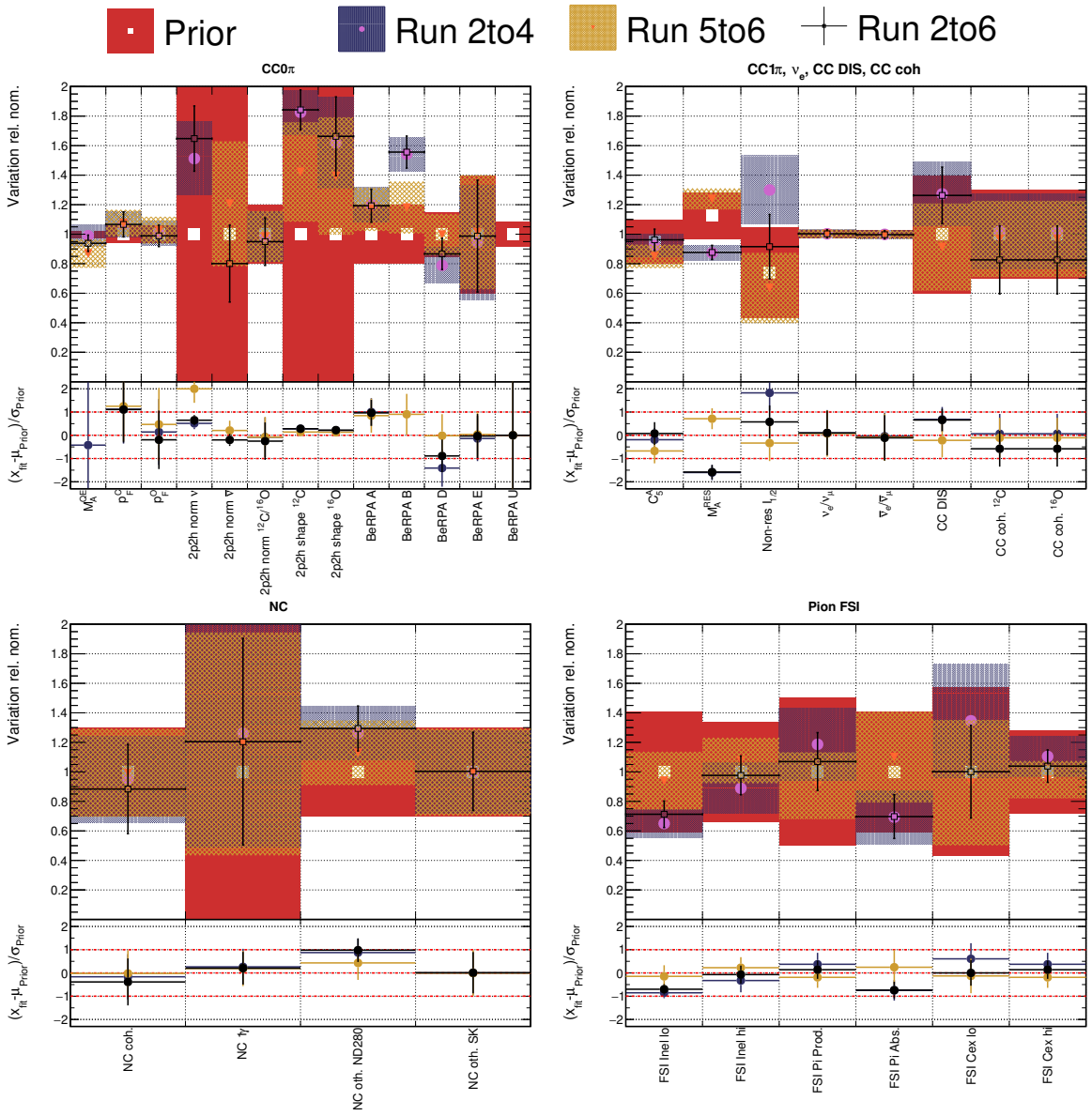


Figure F.2.: Interaction parameters after the data fit for different run periods

In conclusion, the comparisons suggests there are some tensions between neutrino and anti-neutrino data fits, with the anti-neutrino fits often favouring a best-fit closer to the central value. However this is largely due to the weaker constraint from the

sample likelihood due to the lower statistics, making the prior likelihood contribution more important.

The FHC (run 2 to 4) vs RHC (run 5 to 6) fits generally showed the RHC samples agreeing better with the priors through the parameter space. For the interaction parameters we saw significantly different values of M_A^{QE}, 2p2h shape, BeRPA, M_A^{RES}, non-resonant $I_{1/2}$ and some pion FSI parameters, and the uncertainties were much larger on the RHC data due to low statistics. This was worrying because unmodelled neutrino/anti-neutrino differences at Super-Kamiokande can potentially be soaked up in δ_{CP}, exaggerating its constraints.

Figure F.3 shows the three fits for the SK selections, where the differences are large albeit for FHC within 1σ of each other and the full fit. The FHC fit agrees well with the full fit for the FHC selections and the RHC fit agrees slightly worse with the full fit for the RHC selections.

The FHC fit prediction for the RHC samples has very large associated uncertainties and has a much larger normalisation than the full and RHC-only prediction, which would primarily affect the θ_{13} and θ_{23} mixing angles. The RHC-only prediction for FHC is consistently higher: for 1Rμ the effect is most noticeable in the above the oscillation dip at $E_{rec} \sim 0.6$ GeV, the 1Re shows most difference in the low E_{rec} range, and again the 1Re1de selection instead shows the largest effect at higher E_{rec}.

This study highlights the importance of collecting similar numbers of ν and $\bar{\nu}$ at ND280 and SK.

F.2. FGD1 vs FGD2

Reading off Table 5.10 the ratio of FGD1 to FGD2 data is roughly 1, as expected from the detector design. The target in FGD1 is plastic scintillator (C_8H_8), whereas FGD2 is alternating layers plastic scintillator and passive water layers. Hence the constraints on H_2O related parameters (e.g. 2p2h shape O) comes purely from FGD2. The location of the TPCs relative each FGD is also important from the view of reconstruction and related systematics: the PID for forward-going tracks is much better in FGD1 due to the two downstream TPCs, whereas FGD2 generally has better backwards going tracking capabilities due to the two upstream TPCs.

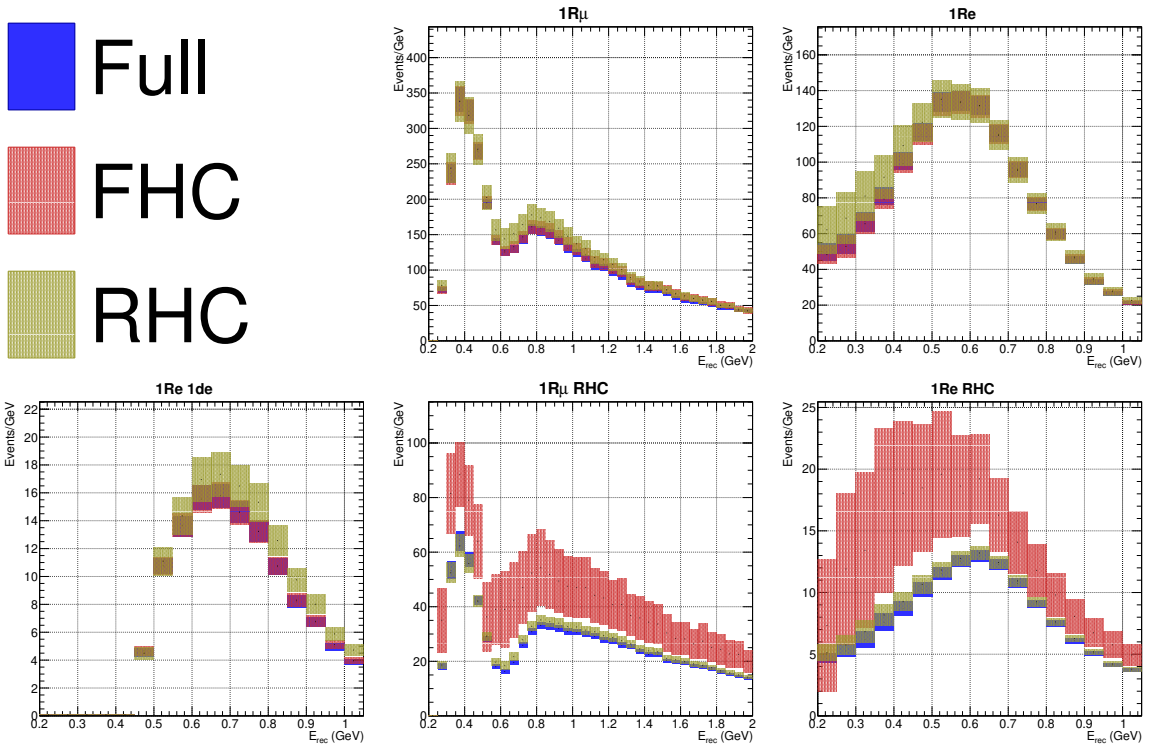


Figure F.3.: Impact of FHC vs RHC fit on SK spectra compared to full fit

Figure F.4 shows the ND280 flux parameters after the fit, in which we mostly similar results for the dominant parameters (correct-sign ν_μ and ν_e). The largest differences happen above $E_\nu = 3$ GeV, where the two fits diverge: FGD2 favouring a much lower flux normalisation at higher energies. The full fit then settles roughtly in between the two fits with an uncertainty covering each individual FGD fit's extreme 1σ uncertainty. Interestingly, FGD2 prefers a nominal flux weight at higher E_ν, whereas FGD1 pulls to 0.84. This pattern is repeated for all flux parameters.

Figure F.5 shows the interaction parameters for the fits to individual FGDs. The CC0π parameters are mostly in agreement, although there are differences in M_A^{QE} and BeRPA B. For FGD1 $M_A^{QE} = 1.21 \pm 0.08$ GeV , FGD2 $M_A^{QE} = 1.06 \pm 0.05$ GeV, and for both $M_A^{QE} = 1.13 \pm 0.07$ GeV, so once again the full fit settles between the FGD1 and FGD2 fit. Interestingly, FGD1 fits a similar value of M_A^{QE} to the notorious "MiniBooNE M_A^{QE} puzzle" at $M_A^{QE} = 1.35 \pm 0.17$ GeV [179] (also seen at K2K on oxygen $M_A^{QE} = 1.20 \pm 0.12$ GeV [218], at K2K on carbon $M_A^{QE} = 1.14 \pm 0.11$ GeV [217] and MINOS $M_A^{QE} = 1.23 \pm 0.18$ GeV [219]) , whereas FGD2 fits close to bubble chamber values at $M_A^{QE} = 1.069 \pm 0.016$ GeV [220]. A slightly lower value of BeRPA B is preferred for the FGD1 only fit, which we may expect from the strong correlation with

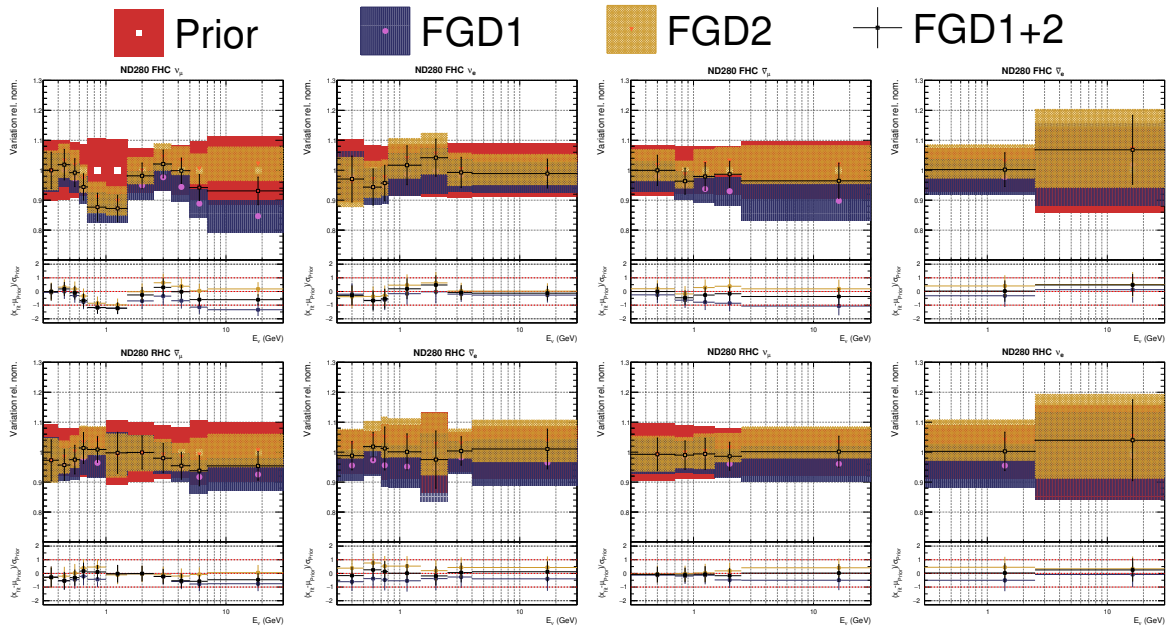


Figure F.4.: ND280 flux parameters after the data fit for FGD1 vs FGD2

M_A^{QE}. However the full FGD1+2 fit favours a value almost identical to FGD2 alone, with a small decrease in the parameter error.

The CC1π parameters are mostly the same with small differences in C_5^A and the coherent normalisation parameters, although they overlap well within eachother's 1σ uncertainty. For the FSI parameters, the full data fit tends towards the FGD2 result, and we note slightly larger errors on the FGD1, e.g. for FSI Inel lo.

In summary, the two FGDs seem to produce mostly compatible post-fit parameters. The only large difference is the flux normalisation parameters at high E_ν, where FGD1 prefers a much smaller normalisation than FGD2 ($\sim$0.85 vs $\sim$1.00). However, this region is very low statistics and barely contributes to the overall event rate at ND280 or SK, so is a small effect in practice. For the interaction parameters we observe slightly different parameters for M_A^{QE} and BeRPA B, in which the fit settles somewhere in between. The FGD1 vs FGD2 fit saw numerous difference in parameters: the high energy flux parameters, M_A^{QE}, BeRPA, 2p2h shape and pion absorption parameters all moved within 1σ of each other, sometimes on the border.

Figure F.6 shows the predicted oscillated SK event distributions for the three fits, in which all the bins agree well inside each other's 1σ. The largest deviations are again for the 1Rμ FHC and RHC samples with $0.2 < E_{rec} < 0.5$ GeV, in which the full fit sits higher than the individual fits. The 1Re selection do not show much change.

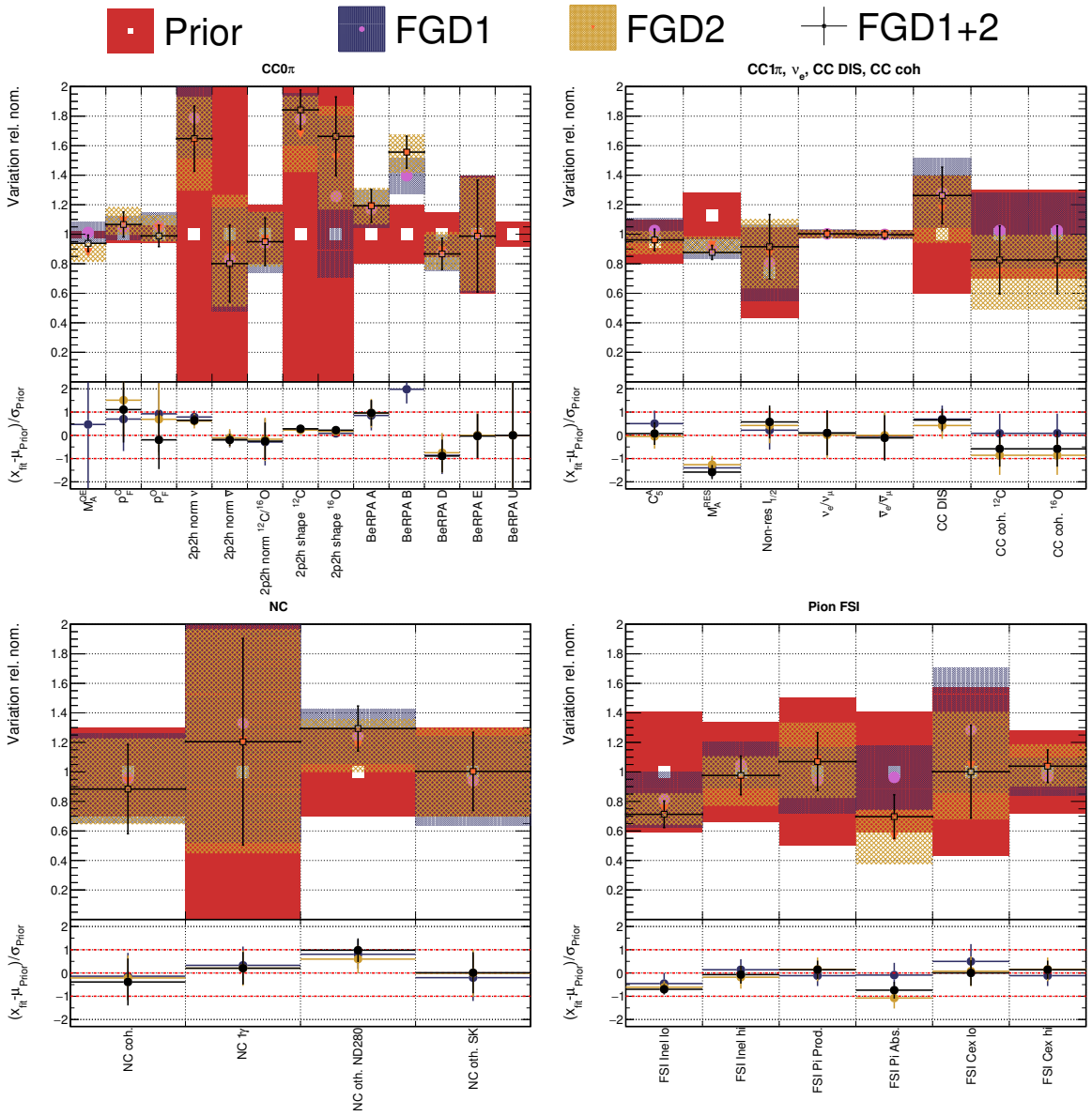


Figure F.5.: Interaction parameters after the data fit for FGD1 vs FGD2

F.3. Excluding the FGD1 CCOther Selection

The data fit showed a generally good description of the ND280 selections with acceptable p-values and best-fit test-statistics for all but FGD1 CCOther. Here we investigate the impact of the sample on the overall fit with the concern that if some parameterisation of systematics has gone awry it may produce erroneous parameter results propagated to SK.

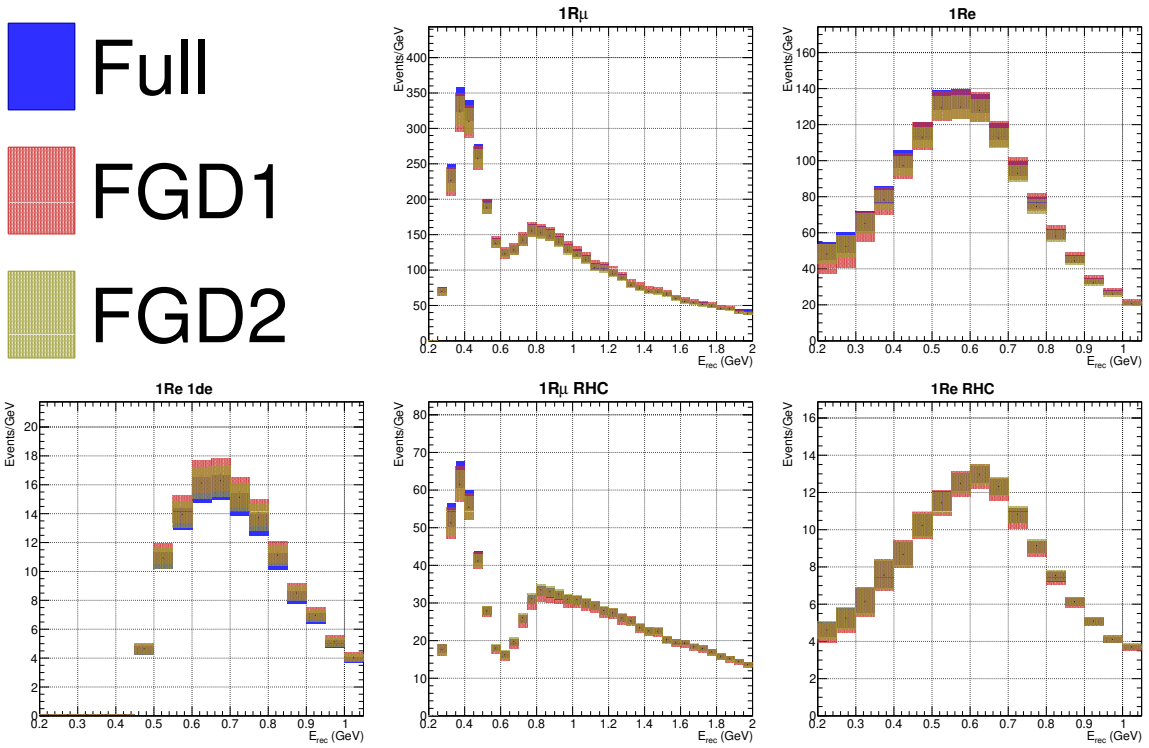


Figure F.6.: Impact of FGD1 vs FGD2 fit on SK spectra compared to full fit

The ND280 flux parameters in Figure F.7 are almost identical to the full 2017 fit and we note barely any differences: the largest change is 0.015 in the highest E_ν normalisation of FHC ν_μ.

The interaction parameters in Figure F.8 shows slightly more change in the single and multi-pion related parameters, although the CC0π parameters remain unchanged. The changes happen in the poorly constrained non-resonant $I_{1/2}$ background parameter, the CC DIS parameter, and some of the FSI parameters. All the differences are however well within the full data fit result.

The event rate and $-2\ln\mathcal{L}_s$ for including and excluding FGD1 CCOther is shown in Table F.1 and as expected there is no greater difference between the two fits in central values or uncertainties: the total difference across all samples is 50.5 events and for the test-statistic it's 2.25.

Finally we recalculate the sample-by-sample p-values in Table F.2 for direct comparison to the full data fit values in Table 5.19. The difference between the two fits' posterior predictive p-values is maximally 7% for FGD2 NTrk (from 58.7% to 51.7%) and all p-values are deemed acceptable.

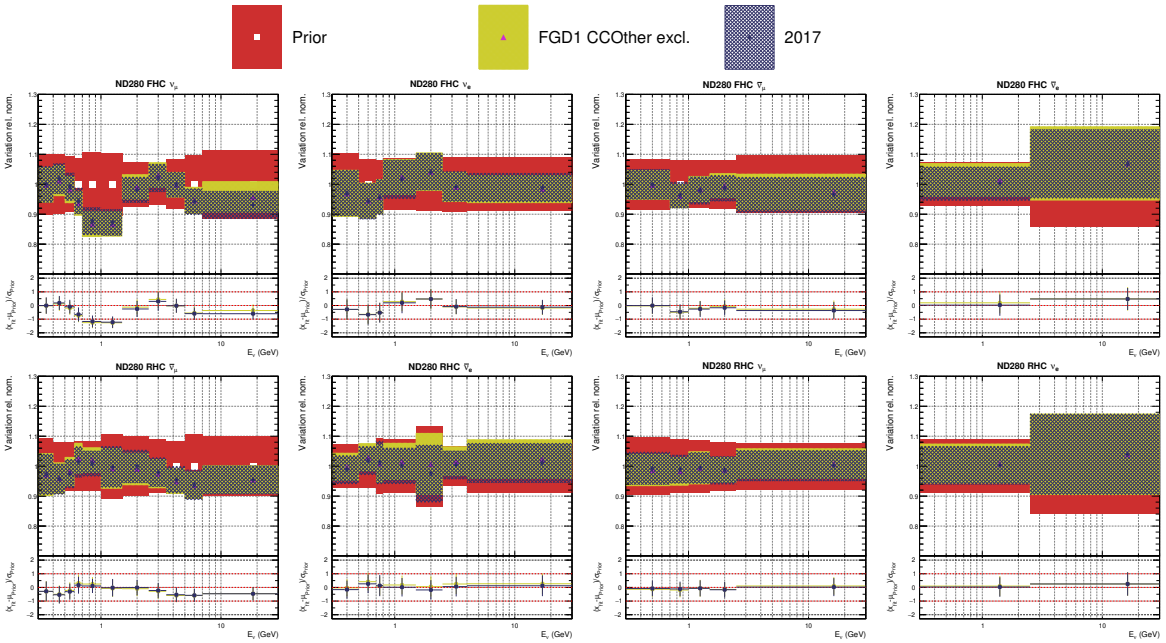


Figure F.7.: ND280 flux parameters after the data fit for excluding FGD1 CCOther

Sample	Event rate		$-2\ln\mathcal{L}_s$	
	2017	FGD1 CCOth	2017	FGD1 CCOth
FGD1 0π	17122.9 ± 120.0	17121.6 ± 120.9	172.21	170.75
FGD1 1π	4053.8 ± 54.3	4032.4 ± 54.0	164.04	162.14
FGD2 0π	17501.5 ± 122.5	17487.9 ± 121.4	166.15	166.28
FGD2 1π	3409.4 ± 48.2	3408.7 ± 46.1	162.71	162.26
FGD2 Other	3917.8 ± 50.8	3906.6 ± 55.8	171.17	172.05
FGD1 1Trk	3509.6 ± 50.1	3513.1 ± 50.1	117.80	117.15
FGD1 NTrk	1062.7 ± 21.9	1046.7 ± 23.6	76.50	76.36
FGD2 1Trk	3678.7 ± 51.3	3679.0 ± 53.0	129.84	128.44
FGD2 NTrk	1108.4 ± 23.5	1105.1 ± 19.8	80.34	80.72
FGD1 ν_μ 1Trk	1348.3 ± 23.1	1351.6 ± 24.5	66.51	66.38
FGD1 ν_μ NTrk	1359.1 ± 26.9	1351.4 ± 26.0	61.75	61.17
FGD2 ν_μ 1Trk	1323.1 ± 23.8	1327.6 ± 23.3	64.34	65.09
FGD2 ν_μ NTrk	1265.9 ± 24.2	1272.3 ± 25.1	75.70	78.00
Total	60658.1	60607.6 ± 244.1	1509.05	1506.81

Table F.1.: Event rate and test-statistic for 2017 with and without FGD1 CCOther. N.B. the total rate and test-statistic for 2017 has the values for FGD1 CCOther subtracted

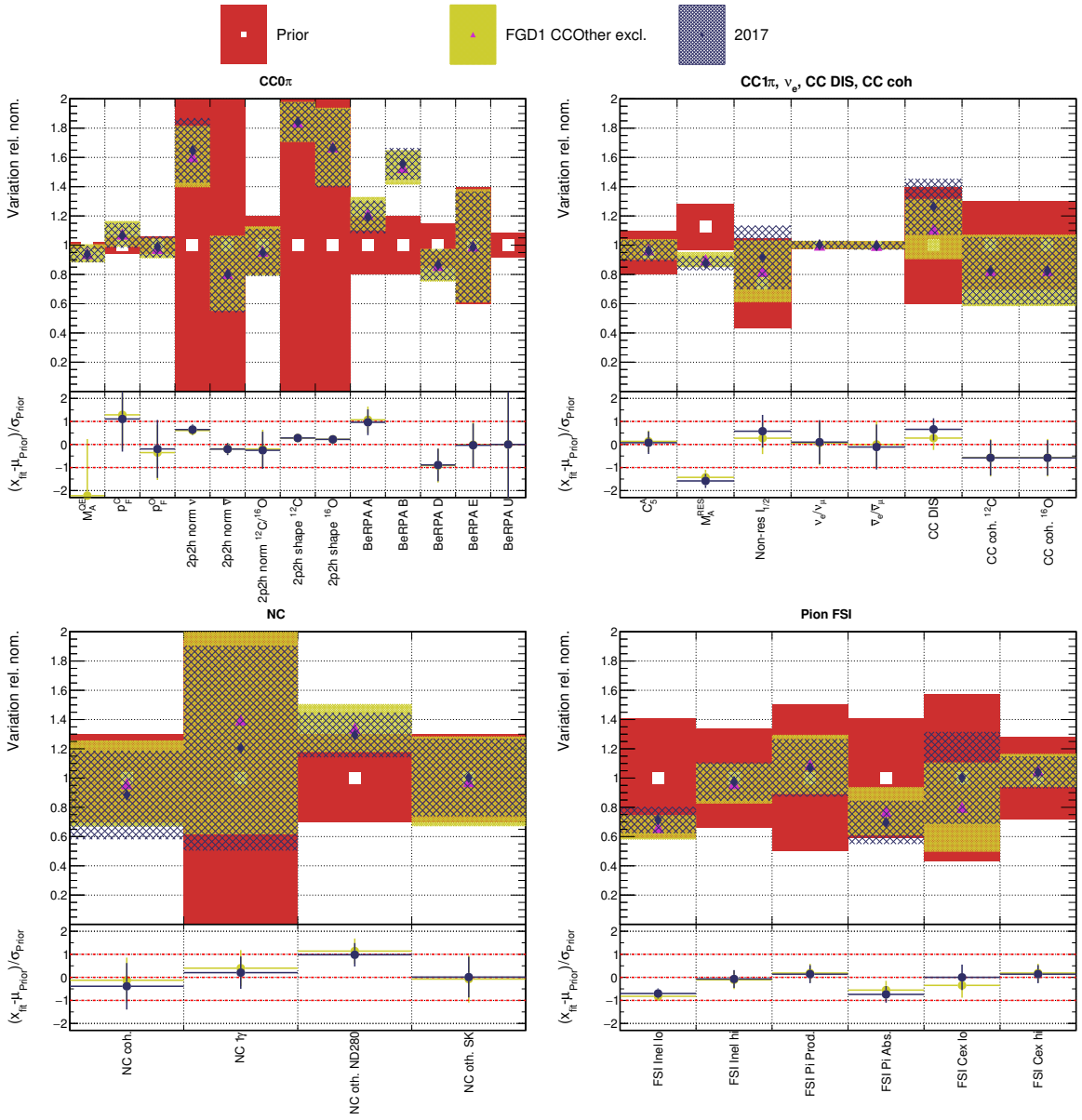


Figure F.8.: Interaction parameters after the data fit for excluding FGD1 CCOther

F.4. Using a 2015-like Model

The 2017 oscillation analysis saw numerous updates to the cross-section model which attempted to provide freedom to nuclear effects such as multi-nucleon interactions and RPA corrections. Ideally, the flux and interaction model is moving closer towards reality than being an "effective" model, being able to only describe ND280-like detectors in a flux similar to that at ND280.

Sample	Draw Fluc.	Pred. Fluc.
FGD1 0π	0.075	0.071
FGD1 1π	0.089	0.086
FGD2 0π	0.113	0.115
FGD2 1π	0.087	0.096
FGD2 Other	0.090	0.087
FGD1 1Trk	0.526	0.529
FGD1 NTrk	0.293	0.281
FGD2 1Trk	0.293	0.288
FGD2 NTrk	0.208	0.206
FGD1 ν_μ 1Trk	0.283	0.287
FGD1 ν_μ NTrk	0.860	0.859
FGD2 ν_μ 1Trk	0.313	0.307
FGD2 ν_μ NTrk	0.517	0.520

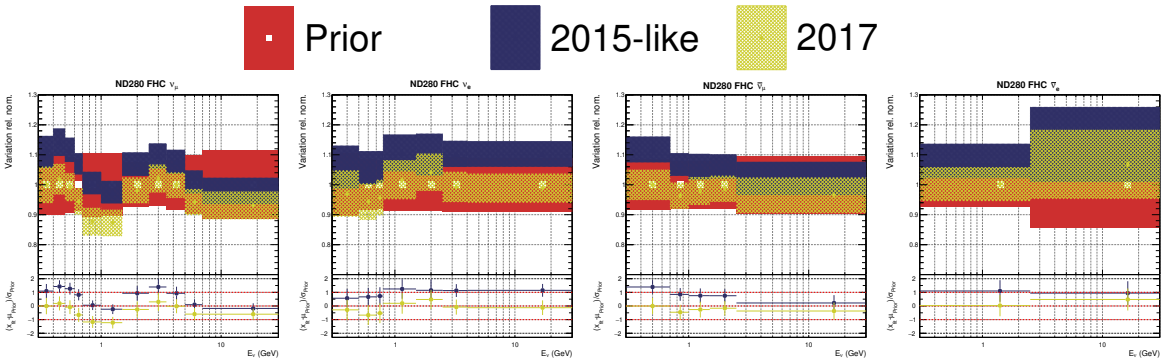
Table F.2.: Posterior predictive p-values for each sample after the data fit excluding FGD1 CCOther

In this section we compare the parameter values from the 2017 analysis to a modified 2017 model made to resemble that of 2015 [108]. In 2015 there was no freedom to change the 2p2h shape or the RPA correction, which was at the time thought to be a weakness in the analysis. The 2p2h normalisation were applied, albeit in a different fashion where 2p2h norm $\bar{\nu}$ was applied multiplicatively to the 2p2h norm ν weight for $\bar{\nu}$ 2p2h events, effectively correlating the parameters. The 2p2h normalisation was also separated for Carbon and Oxygen, but were found to be the same after the fit. For the 2017 analysis these parameters are uncorrelated. The priors on the single pion parameters changed slightly too, but overall only shifted central values by small amounts and gave larger uncertainties on the prior.

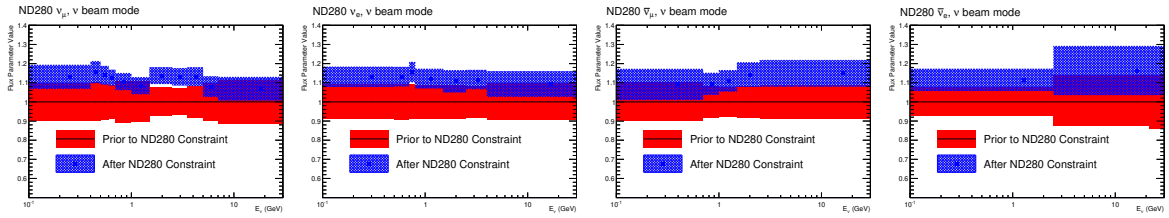
A straight-forward implementation of the 2015 cross-section model is therefore to fix 2p2h shape C and O and the BeRPA parameters to their nominal values and not vary them during the fit. Here we compare such a model to the results of the full data fit, drawing parallels with the 2015 results.

Figure F.9 and Figure F.10 shows the ND280 flux parameters for the two fits, including the result from 2015. [108]. The 2015-like fit result replicates the main feature of the 2015 fit: the notably high normalisation of $\sim 10-20\%$ throughout the flux parameters. However, it does not perfectly replicate the parameter results, as the

dip for the 2017 result for FHC ν_μ around $E_\nu = 0.8 - 1.0$ GeV is also present in the 2015-like fit, but not in the actual 2015 fit. The RHC parameters for the 2015-like fit in Figure F.10 agree more with the 2015 result with the key features replicated.



(a) 2017 and 2015-like

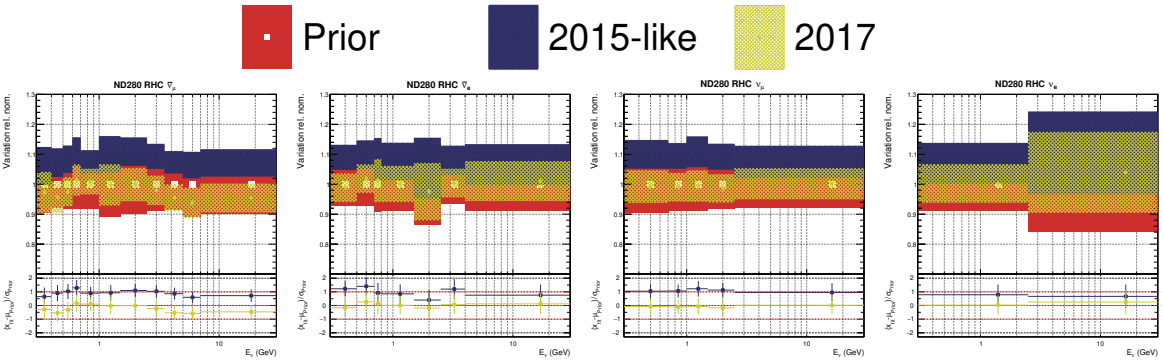


(b) 2015 results

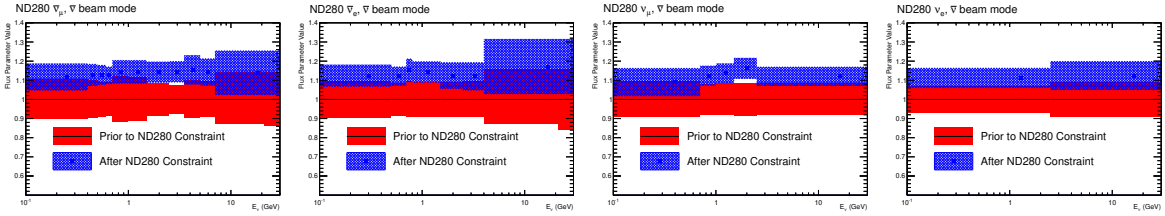
Figure F.9.: ND280 FHC flux parameters for 2017, 2015-like and 2015 analyses

Figure F.11 show the interaction parameters, in which we note a similar values for the directly comparable parameters, e.g. M_A^{QE}, p_F, C_5^A, CC DIS, NC other. We also note the difference in the CC DIS parameter, which moves from nominal (2015 and 2015-like) to roughly its upper 1σ. Since the CC DIS parameter is heavily connected to an event's E_ν (since the parameter is $0.4/E_\nu$), this is likely a reflection of the central values of the fluxes shifting from $\sim$1.1 to $\sim$1.0 in moving from the 2015 to 2017 analysis.

A Model with BeRPA and 2p2h Shape Fixed From the initial likelihood scans and 1σ variations to the Asimov data set in subsection 5.6.1 and subsection 5.6.2, it is expected that the "2015-like" quality of keeping BeRPA and 2p2h shape parameters fixed comes from the BeRPA variation. But both BeRPA and 2p2h shape are pulled away from the nominal correction, which the scans and parameter variations were centralised on. Two additional MCMCs were run on the data with variations of the



(a) 2017 and 2015-like



(b) 2015 results

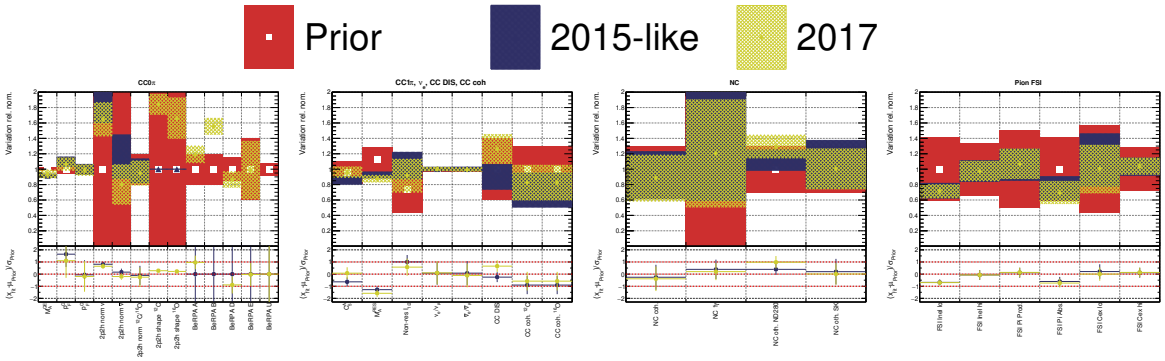
Figure F.10.: ND280 RHC flux parameters for 2017, 2015-like and 2015 analyses

2015-like cross-section model: fixing the BeRPA parameters at nominal and leaving 2p2h shape free, and vice versa.

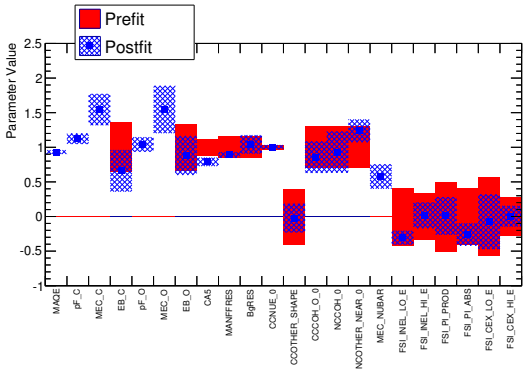
Figure F.12 shows the flux parameters after the fits for the different variations and confirms suspicions that the BeRPA parameters are causing the shift in flux parameters. Fixing 2p2h shape to their nominal values has a minor impact on the flux parameters and narrowly replicates the 2015-like fit.

Figure F.13 shows the cross-section parameters for the 2015-like variations and we notice much the same pattern. All the observable differences are in 2p2h normalisations, C_5^A and CC DIS all come from the difference in the flux parameter's central values.

In summary, it appears that the 2017 analysis agreeing with the flux parameter is down to moving freedom from E_ν normalisations (the flux parameters) to Q_2 normalisations (the BeRPA parameters). Although the BeRPA parametrisation's post-fit model is a strong distortion of the nominal RPA model—implying the parameterisation of the error should be reconsidered—it is more justified assigning error to BeRPA (which has no external data constraints) than the flux model (which has multitudes of external data and monitoring systems).



(a) 2017 and 2015-like



(b) 2015 results

Figure F.11.: Interaction parameters for 2017, 2015-like and 2015 analyses

The 2015-like parameterisation fixed 2p2h shape and BeRPA to their prior values and did not vary them in the fit. This roughly reproduced the results from the ND280 fit from 2015 [108].

Figure F.14 shows the E_{reco} distributions for the five samples at Super-Kamiokande. We note the largest difference in spectrum is in the 1Rμ sample between $0.2 < E_{rec} < 0.5$ GeV: the region which is sensitive to oscillation parameters (mostly θ_{23} and Δm^2_{23} or Δm^2_{32}). The shift is less than 1σ of the full data fit. The 1Re selection has a consistently lower event rate for the 2015-like fit, which has an impact on θ_{13} primarily, and the effect on the 1Re1de selection is opposite. The RHC selections are entirely consistent and see no deviations from each other.

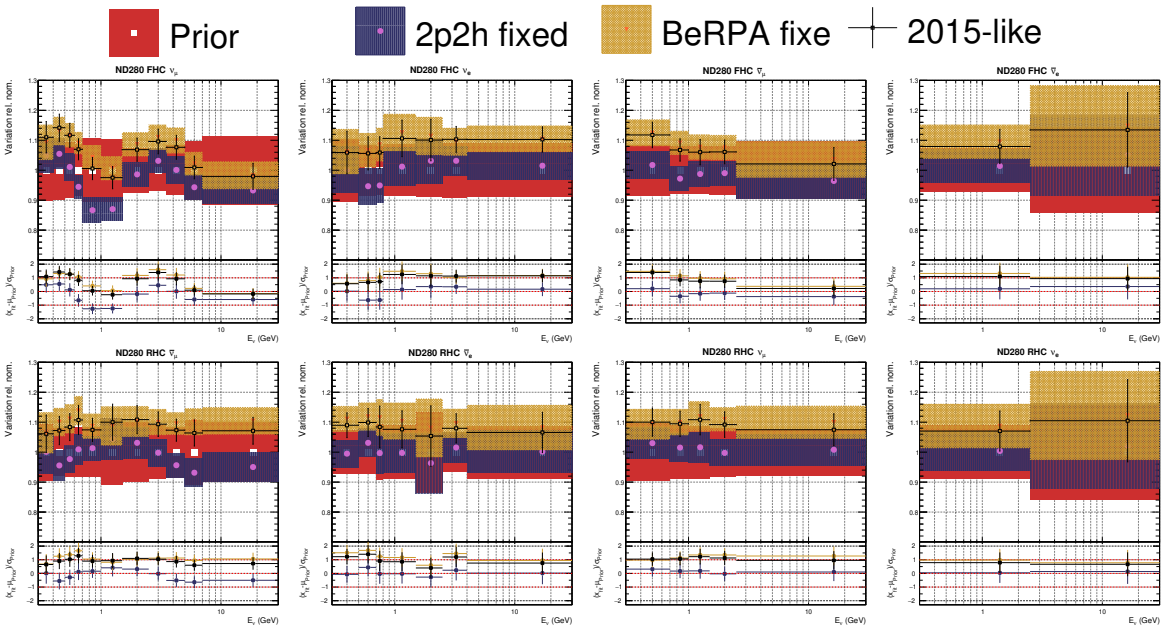


Figure F.12.: Flux parameters for 2015-like, 2p2h shape fixed and BeRPA fixed models

F.5. Invoking a Prior on M_A^{QE}

In the 2017 data fit it was decided to keep the flat prior on M_A^{QE}. This was justified by wanting a conservative CCQE model with large degrees of freedom. There is also the historical "MiniBooNE M_A^{QE} puzzle", in which fits to MiniBooNE CCQE data inflated M_A^{QE} to 1.3-1.6 GeV from the bubble chamber value of $\sim$1.0 GeV. Fits to MiniBooNE and MINERvA "CCQE" data are further complicated by different signal definitions at the two experiments (e.g. subtraction of pionless delta decay) and a questionable model dependence [175]. However, now that a detailed nuclear model has been implemented for the CCQE/CC0π model, M_A^{QE} is feasibly a neutrino-nucleon parameter (i.e. not an effective parameter in the nuclear environment), and can be constrained from neutrino-nucleon scattering data.

Fits were done of the NEUT generator's CCQE model with the NUISANCE package to ANL [228, 229] and BNL [230] CCQE data in E_ν and Q^2 using NUISANCE [1], finding $M_A^{QE} = 1.03 \pm 0.04$ GeV in agreement with literature [220]. This result was then propagated into the analysis in the interaction parameter covariance matrix. In light of the data fit presented in section 5.7—in which $M_A^{QE} = 1.12 \pm 0.07$ GeV, right between the bubble chamber prior and the nominal value in NEUT—a moderate effect is expected.

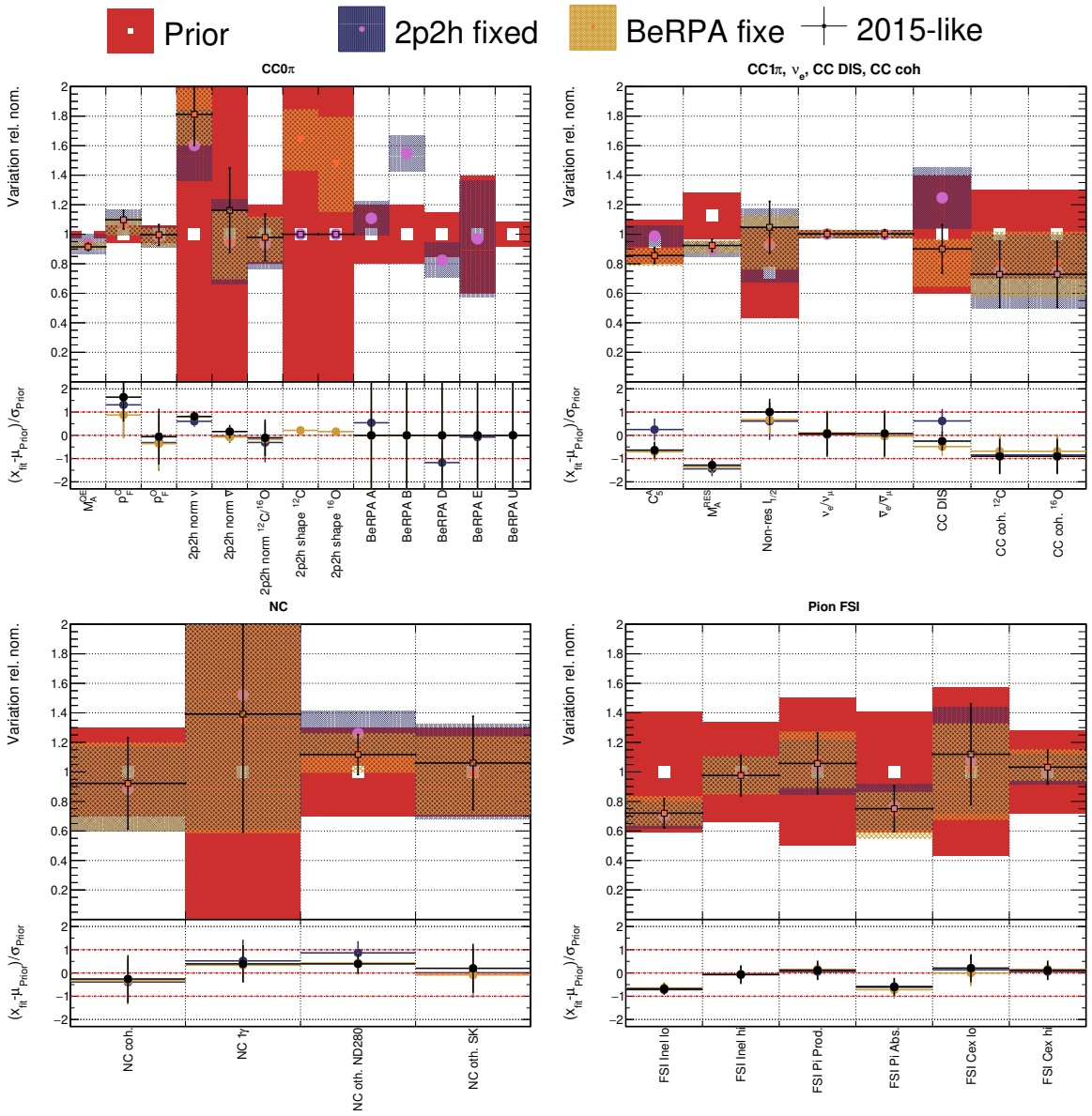


Figure F.13.: Interaction parameters for 2015-like, 2p2h shape fixed and BeRPA fixed models

Figure F.15 shows the ND280 flux parameters for the two models. A moderately higher flux normalisation is seen for all flux parameters: most noticeable for FHC ν_μ at $E_\nu < 0.6$ GeV. and RHC $\bar{\nu}_\mu$ between $0.8 < E_\nu < 5$ GeV. However, the differences are relatively small and maximally 0.05 in normalisation.

The interaction parameters for the two M_A^{QE} priors are found in Figure F.16, where very little difference is observed. Since M_A^{QE} is fit to 1.12 ± 0.07 GeV (equivalent to 0.93 relative the nominal 1.2 GeV) and the prior is $M_A^{QE} = 1.03 \pm 0.05$, we expect the ND280 data to pull M_A^{QE} up from the bubble chamber prior, and roughly land

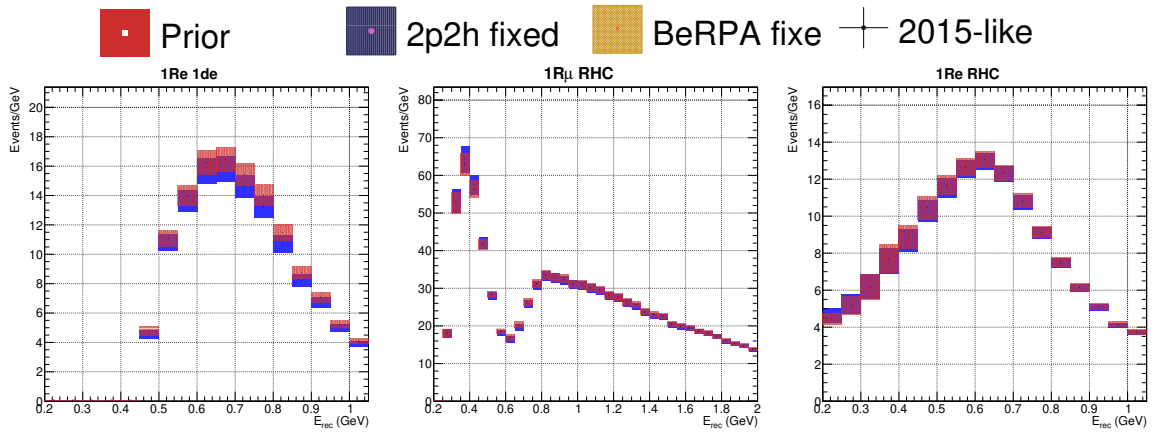


Figure F.14.: Impact of 2015-like fit on SK spectra compared to full fit

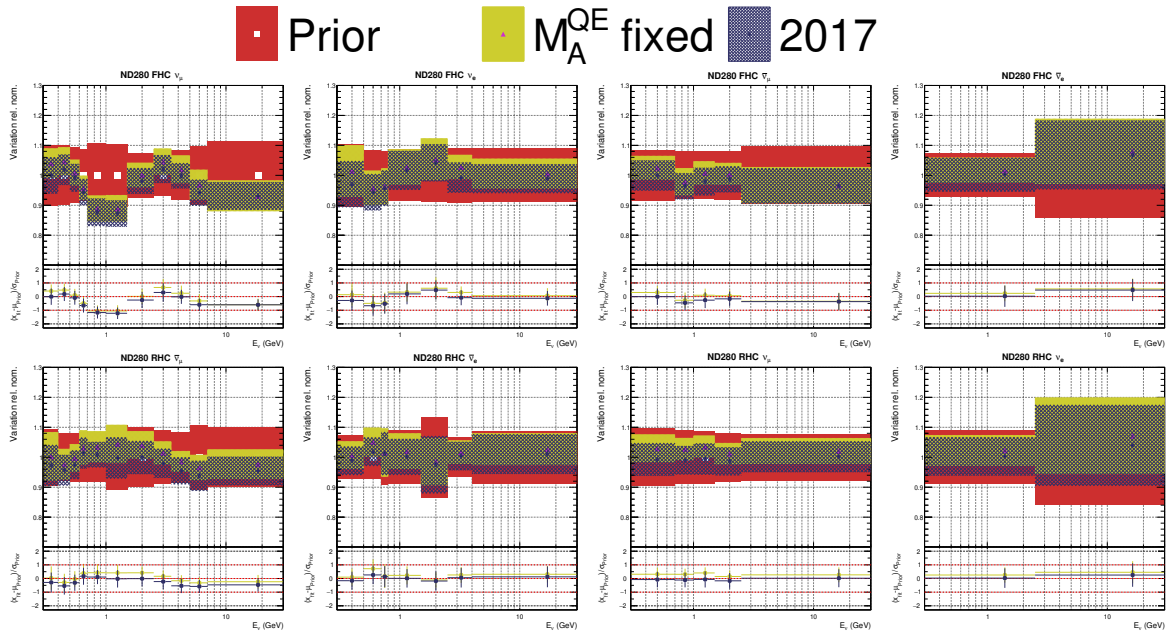


Figure F.15.: ND280 flux parameters for interaction model with M_A^{QE} prior and the 2017 fit

in between, which indeed is observed: the post-fit value for the fit with the prior is $M_A^{QE} = 1.07 \pm 0.04$ GeV. We note small shifts in the CC0π parameters (most in BeRPA D, which controls high Q^2 and M_A^{QE} strongly constrains), and the largest shifts happen in pion-related parameters: the C_5^A parameter almost halves in uncertainty, and M_A^{RES} shifts slightly. This is expected from the single pion contamination in the CC0π samples. The CC DIS parameter also shifts closer to nominal, likely due to the correlations with the flux again.

In summary, the prior on M_A^{QE} has a small effect on the parameter values propagated to the oscillation analysers.

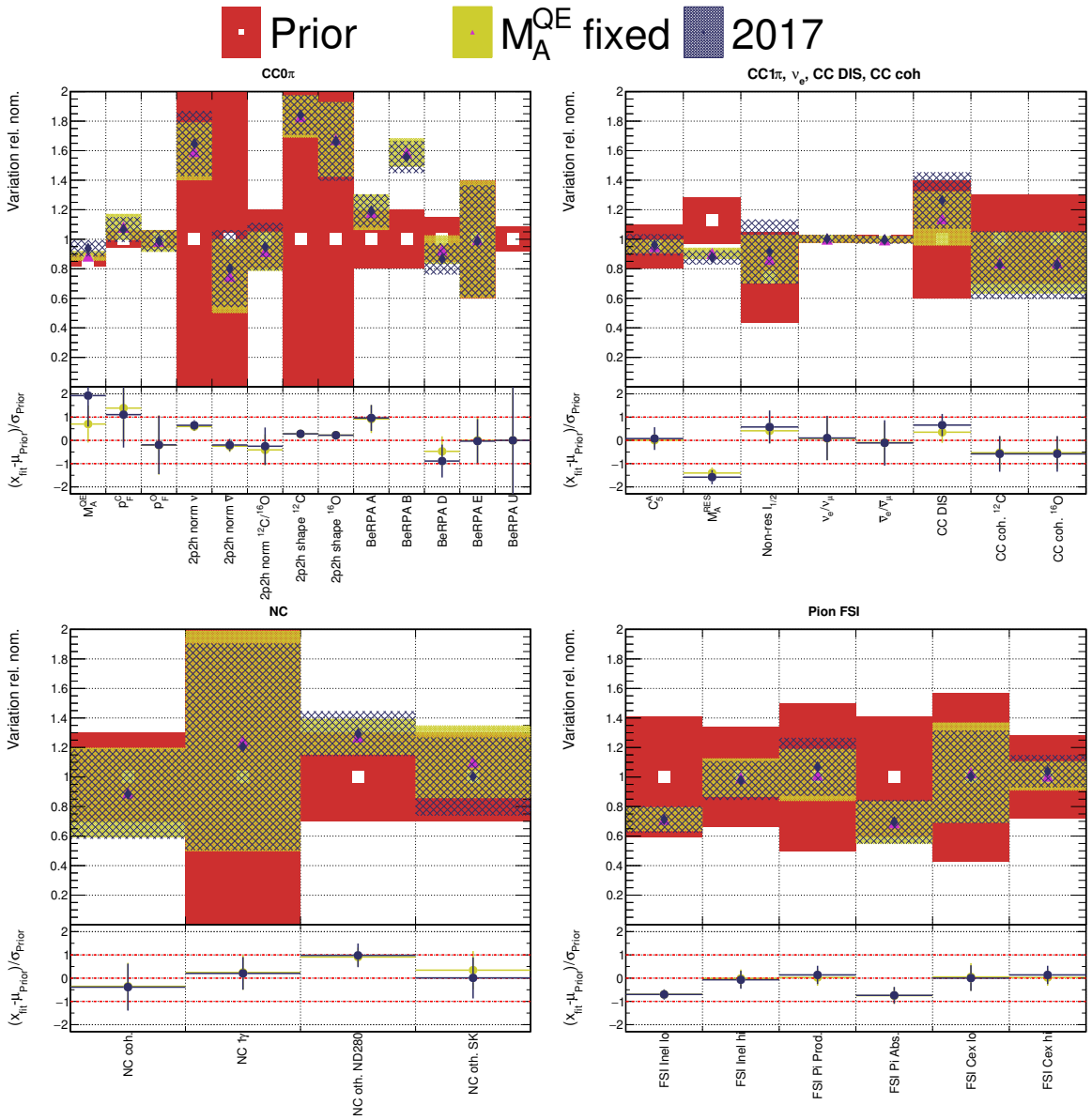


Figure F.16.: Interaction parameters for interaction model with M_A^{QE} prior and the 2017 fit

F.6. A Model with Flat Single Pion Production Priors

Similar to the CCQE/CC0π cross-section modelling, there are concerns that the neutrino-nucleon parameters in single pion production are absorbing unmodelled nuclear effects, effectively turning the model "unnatural": the post-fit model describes ND280 data well but no longer does so for bubble chamber data.

The CC0π modelling has improved drastically over recent years with multi-nucleon effects [47, 48], spectral function calculations [50], initial state models [174] being

developed. However, few models have extended into the "delta-region", in which the Δ baryon decays to produce a pion-nucleon state, a large contributor to the 1π topology at ND280.

The data fit in section 5.7 showed M_A^{RES} and the non-resonant $I_{1/2}$ background of the single pion parameters pulled far from their priors, and studies of neutrino vs anti-neutrino selections showed different preferences for the parameters. Similar to the study of the impact of the M_A^{QE} prior, we here ignore the single pion parameter priors, leaving them entirely free in the fit to ND280 data.

Figure F.17 shows the ND280 flux parameters for the two fits and the change is minimal, similar to the effect of the M_A^{QE} prior. The shape of the 2017 data fit is largely maintained with a slightly larger normalisation on a 0.01 scale (1%).

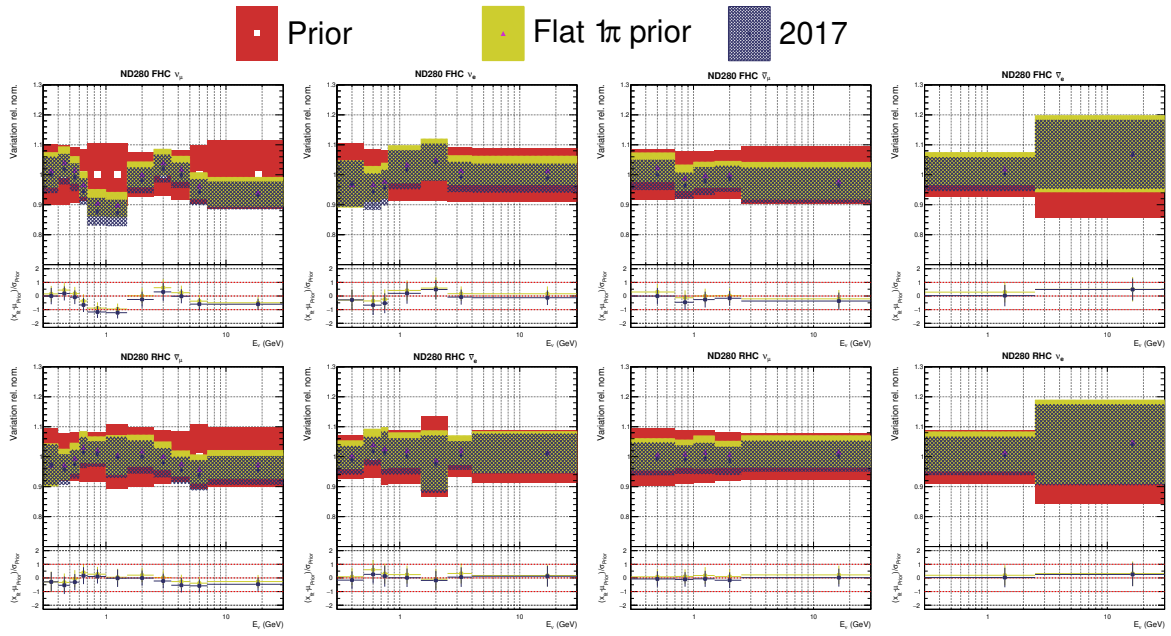


Figure F.17.: ND280 flux parameters for interaction model without 1π priors and the 2017 fit

The CC0π interaction parameters in Figure F.18 also show a minimal difference to the full fit. The only substantial shift is in the single pion parameters whose priors are now ignored. C_5^A shifts from the nominal value down by half a sigma of the original prior, but maintains the same size for the uncertainty. M_A^{RES} shifts slightly downward (away from the prior), although it is minimal—indicating the prior is already largely over-run by the ND280 data in the 2017 reference fit. We also note a slightly larger uncertainty due to the likelihood contribution now only coming from the samples. The non-resonant $I_{1/2}$ parameter also moves further from the prior, although just on

the top of the 1σ uncertainty that the prior would provide had it been applied. We see small shifts in the coherent normalisations, due to correlations coming from the 1π and Other selections[a].

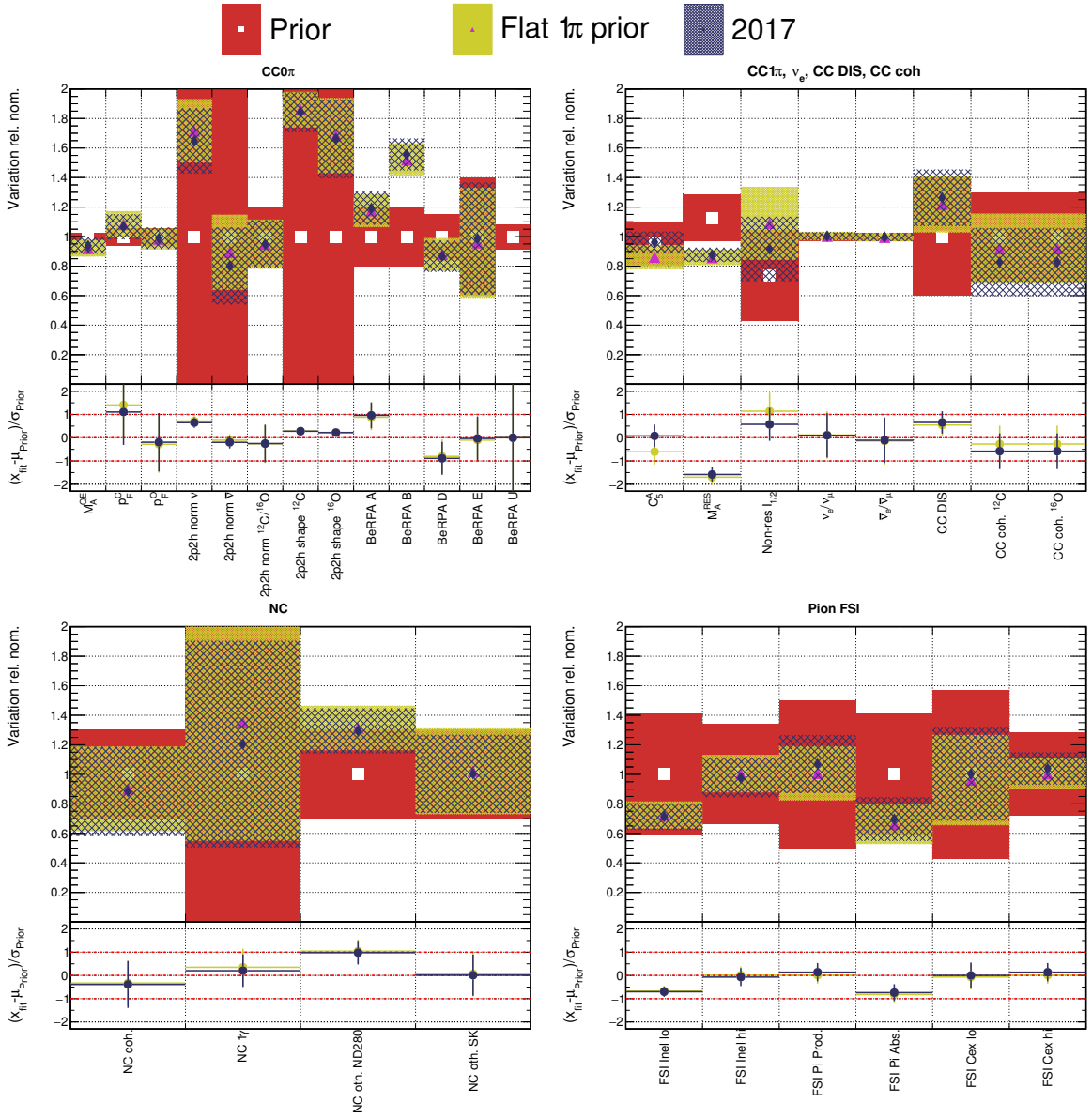


Figure F.18.: Interaction parameters for interaction model without 1π priors and the 2017 fit

In summary it appears releasing the single pion priors has little effect on the parameter values in the fit. The large excursion of M_A^{RES} from the prior is maintained

[a]CC coherent has an identical final state to CC1π but without the nucleon—however no nucleons are required in the ND280 selection.

at a similar post-fit value, and the two other parameters move away from the prior, although not significantly. The differences between the prior and post-fit does still indicate tension between the neutrino-nucleon and neutrino-nucleus single pion model.

Appendix G

Alternative Studies, 2018 Analysis

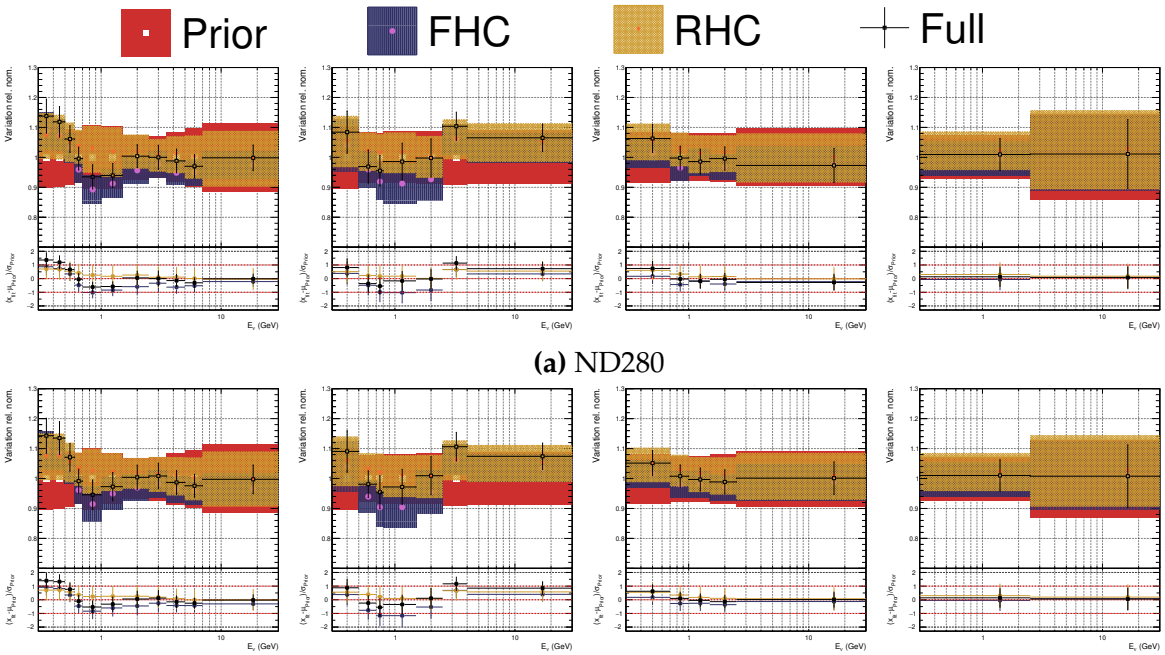
This section fits different subsets of the data and compares it to the full result in 2018, similar to what was done for the 2017 analysis in Appendix F. Since the model only marginally changed and the new parameters are entirely compatible with last year, we here focus on compatibility between FGD1 and FGD2, and FHC and RHC runs rather than variations in the systematics.

G.1. Neutrino vs Anti-Neutrino

In the 2017 analysis we found BeRPA to be marginally different between FHC and RHC runs, with the full data settling near the FHC result. With 1.99x more collected POT in FHC and 1.63x in RHC mode, it's interesting to see if the data subsets prefer different BeRPA models. RHC events still only make up 20% of the total data (24795/121432), so a larger relative effect of the prior is expected.

Figure G.1 shows the FHC flux parameters after the fit. We see good compatibility at low energies throughout, and the two FGDs disagreeing around 1 GeV for the ν_μ and ν_e fluxes, returning to agreeing at 1.5 GeV. Interestingly, the FHC+RHC fit favours a higher flux normalisation at low energies than either of the separate fits (115% vs 109%), but still sit within the parameter error. The difference at 1 GeV manifests by the FHC+RHC fit sitting between the two, slightly towards the FHC data fit.

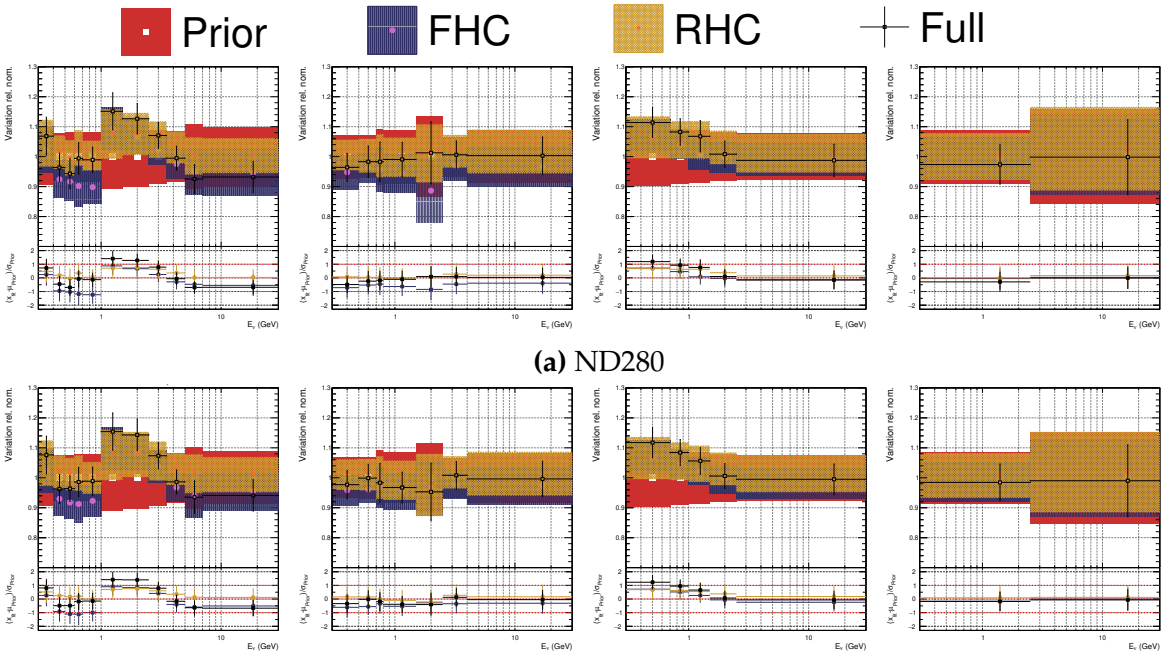
The RHC parameters in Figure G.2 show a similar pattern: when the separate FHC and RHC fits prefer a high flux normalisation, the full fit prefers an even higher flux normalisation, and when the two pull away from eachother, the full fit settles in between. The only differences are for the $\bar{\nu}_\mu$ and $\bar{\nu}_e$ between 0.4-1.0 GeV, where the RHC fit (which has sample sensitivity to these parameters) sits slightly above the prior and the FHC fit (which only constrains via correlations with the FHC parameters) sits below.



(a) ND280

(b) SK

Figure G.1.: FHC flux parameters, fitting to data with different horn configurations



(a) ND280

(b) SK

Figure G.2.: RHC flux parameters, fitting to data with different horn configurations

The interaction parameters in Figure G.3 show the largest differences between the two horn configurations. Most CC0π-related parameters agree up to BeRPA B, which sees a large inflation from the prior, which the RHC data seems to prefer. The full data

fit sits right on the FHC-only results, demonstrating the power of the high statistics samples over the prior. The CC0π parameters are very similar to the 2017 equivalent results in Figure F.2.

As in 2017, the single pion parameters are also significantly different for FHC and RHC runs, with the patterns similar but pulls more extreme. Generally the RHC results are closer to the prior. There have already been indications in data [202] and modelling [225] that differences in neutrino and anti-neutrino single pion production may be unaccounted for, which here is supported by T2K data.

The remaining differences are found in the pion final state interaction probabilities, where we see weaker constraints from the RHC samples, many times agreeing within 1σ with the FHC. As for other parameters where FHC and RHC parameter are in tension, the full fit settles in between with errors to cover the individual fits' 1σ. The parameters in tension appear to be the inelastic and absorption probabilities, where the RHC selections prefer a higher than nominal value, and FHC the opposite.

The different BeRPA parameterisations after fitting FGD1 and FGD2 data are shown in Figure G.4. The RHC data clearly prefers a more RPA-like prescription and has larger errors (as expected from the small data set), whereas the FHC data drives the data fit, with an even more extreme enhancement than in 2017 around $Q^2 \sim 0.5$ GeV^2.

As with the 2017 analysis, comparing neutrino and anti-neutrino fits to the full fit has highlighted differences in BeRPA, single pion production and pion FSI probabilities. The tensions from 2017 seem to remain in the larger 2018 fit and should be addressed in the future.

G.2. FGD1 vs FGD2

We now compare using FGD1 and FGD2 selections to using both, identical to in 2017. The 2017 analysis saw relatively large differences in the flux parameters, notably at high E_ν, and the interaction parameters had different M_A^{QE} and BeRPA B values.

The FHC flux parameter are shown in Figure G.5 where we mostly see compatibility. Similar to the FHC vs RHC case, the full FGD1+FGD2 fit favours a higher flux at low E_ν than the separate FGD1 and FGD2 fits, which is repeated whenever the flux parameter values are high. The results are generally compatible.

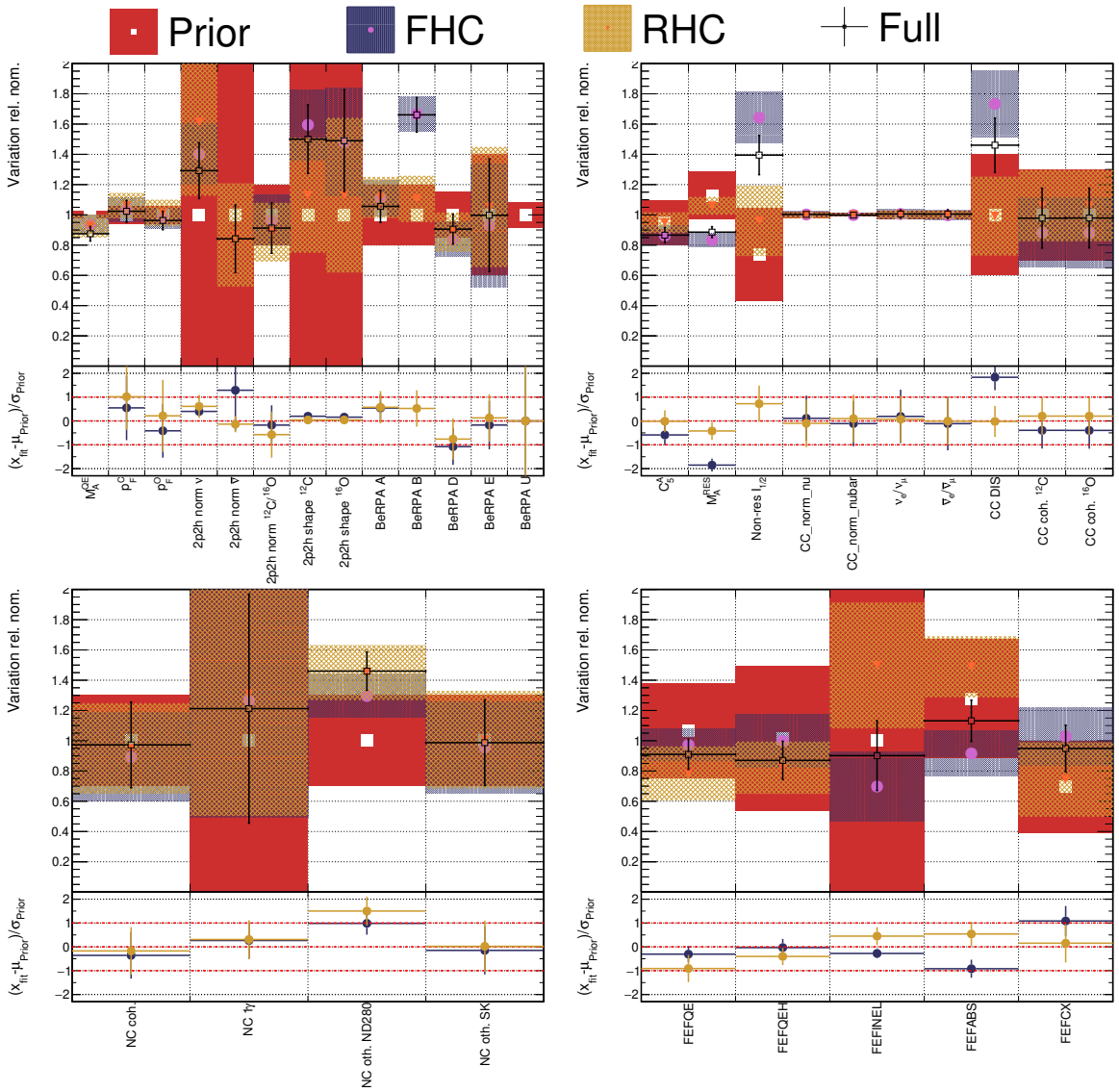
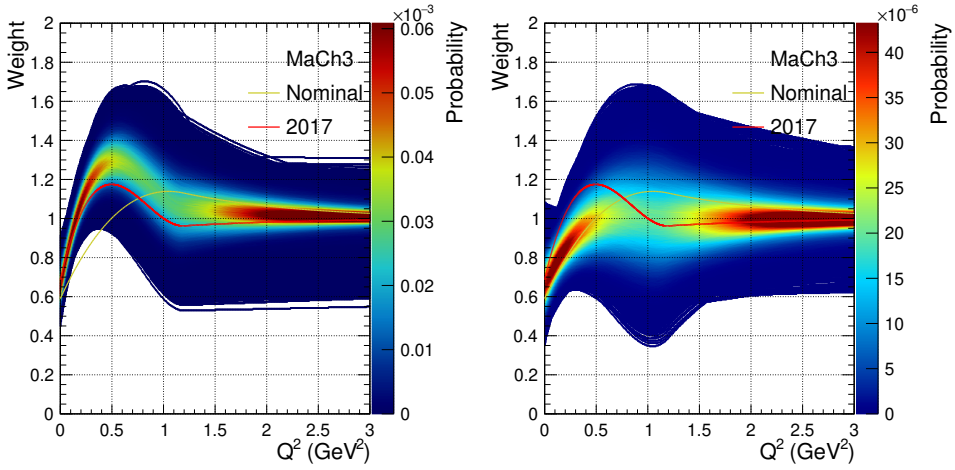


Figure G.3.: Interaction parameters, fitting to data with different horn configurations

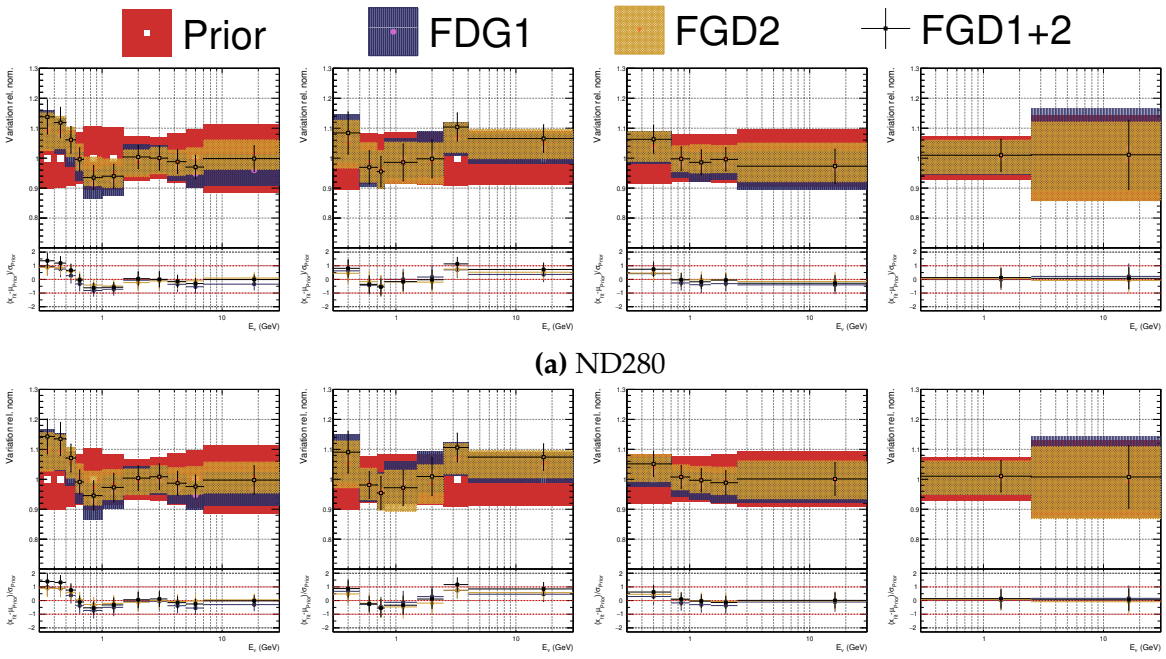
For the RHC flux parameters in Figure G.5 the compatibility is also good with the same patterns repeated.

The interaction parameters in Figure G.7 are also more compatible than in 2017, but show some similarities. Notably M_A^{QE} is again different for FGD1 and FGD2, with FGD1 favouring an inflated "MiniBooNE-like" value ($\sim$1.2 GeV) and FGD2 a more bubble-chamber like value ($\sim$1.0 GeV). The 2p2h shape C parameter is slightly different (although within error) and the full fit settles closer to the FGD2 fitted value than FGD1. The BeRPA parameters are compatible, with the largerst difference in BeRPA D, which controls the behaviour around 0.8-1.0 GeV^2.



(a) FHC **(b)** RHC

Figure G.4.: BeRPA parameterisations for fitting to data with different horn configurations

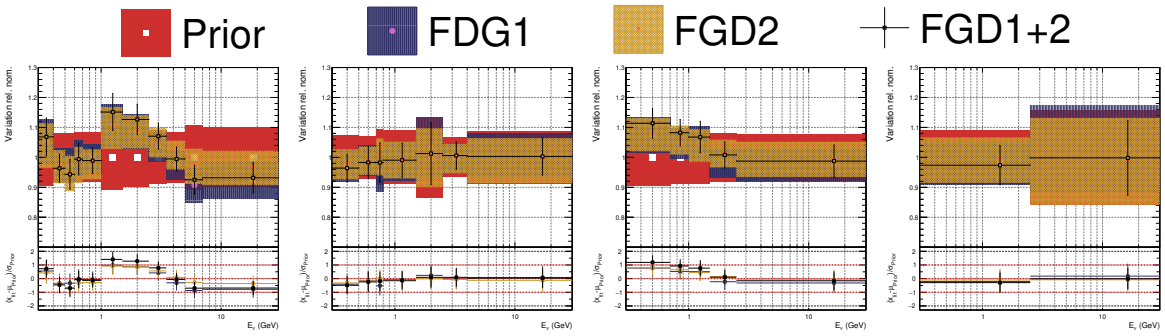


(a) ND280

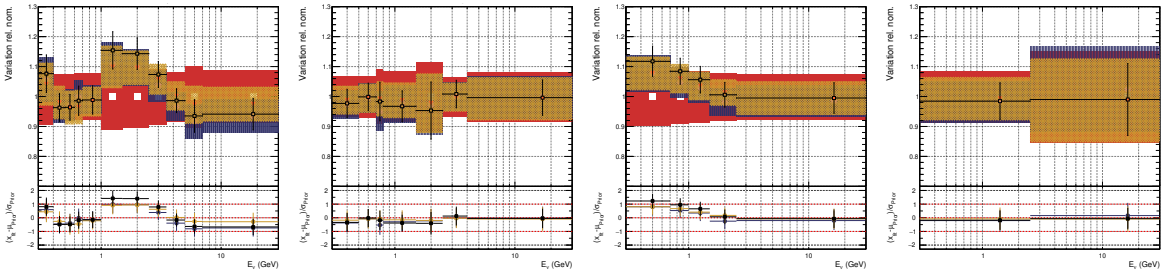
(b) SK

Figure G.5.: FHC flux parameters, fitting to data with FGD1 and FGD2

There is some tension in the C_5^A single pion production parameter, where FGD1 favours a value agreeing with the prior and FGD2 below that. This was seen in 2017 also, and was correlated with the differences in single pion production parameters. Both FGDs prefer a non-resonant $I_{1/2}$ background higher than nominal, with FGD2 inflating it by 200% (or 140% of nominal), again indicating insufficient single pion modelling.



(a) ND280



(b) SK

Figure G.6.: RHC flux parameters, fitting to data with FGD1 and FGD2

The pion final state parameters are also mostly compatible with the 2017 fit, where FGD1 often prefers a larger parameter value than FGD2. The pion charge exchange parameter is the first time the new FSI parameter priors are pulled outside the 1σ, although FGD2 prefers a value much in-line with the prior.

In conclusion, the FGD1 vs FGD2 compatibility largely agrees with the 2017 conclusions, with better agreement in the flux parameters. M_A^{QE} and BeRPA D are the largest CC0π differences, and the single pion parameters and pion FSI are pulled differently too. Throughout, all the parameters are within 1σ of the full fit (and often of each other).

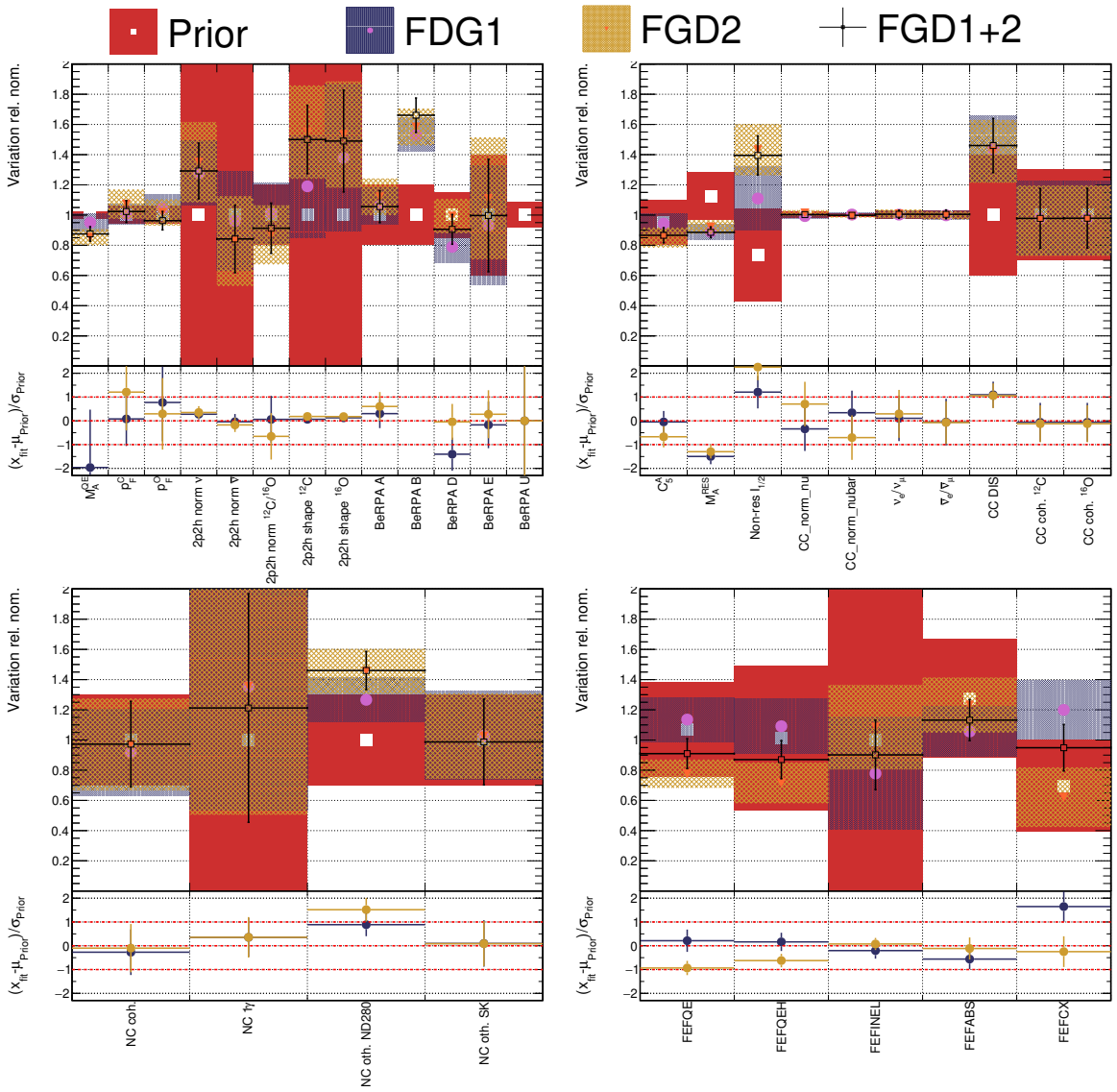


Figure G.7.: Interaction parameters, fitting to data with FGD1 and FGD2

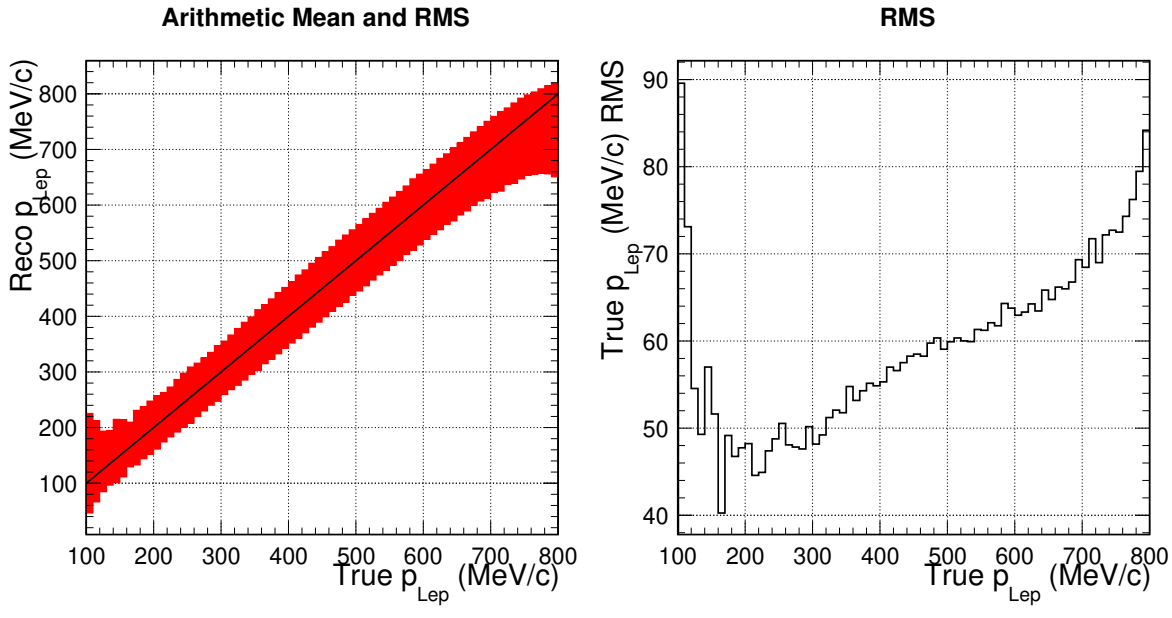
Appendix H

ND280 detector resolution

This appendix outlines the detector resolution of ND280 and exists as a reference for the binning studies. The resolution plots presented here were made by:

- Use all generated ND280 Monte-Carlo from run 2 to 8
- Run the event selections and study events that pass the ND280 selections
- Require an event to have a lepton candidate, placing it in a CC selection
- Select the lepton candidate and plot its true and reconstructed kinematics

Figure H.1 shows the resolution of events included in the analysis projected on the lepton momentum. In the peak region of $p_{Lep} = 200 - 300$ MeV/c the resolution is better than 50 MeV. Above 250 MeV it roughly linearly increases to $\Delta p_{Lep} = 80$ MeV/c at $p_{Lep} = 800$ MeV/c.

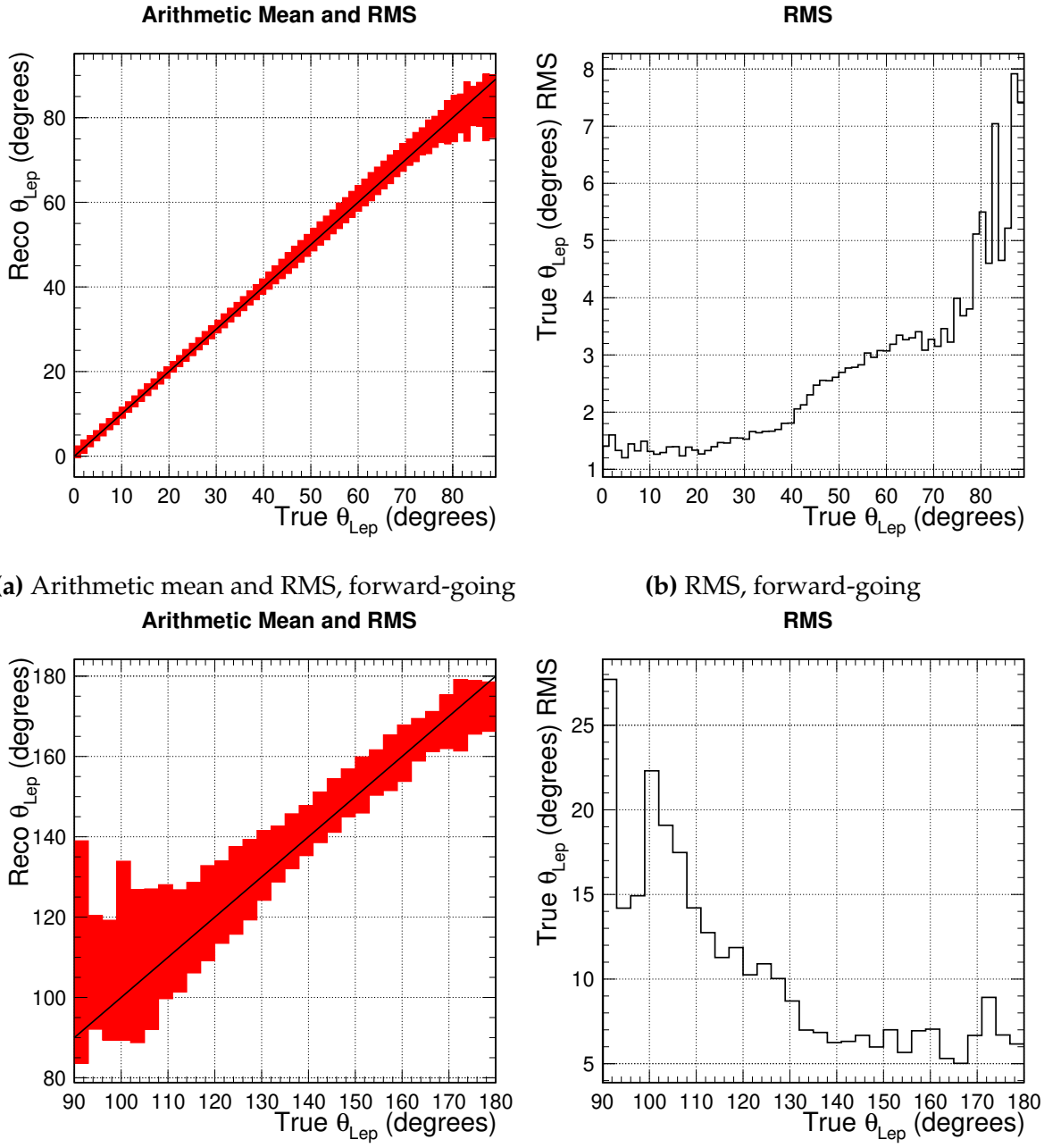


(a) Arithmetic mean and RMS

(b) RMS

Figure H.1.: ND280 momentum resolution for CC-inclusive event's lepton candidates

Figure H.2 shows the θ_{Lep} resolution of ND280 for selected events. The resolution is largely flat between $\theta_{Lep} = 0 - 30^{\circ}$ at 1.5°, which linearly increases up to 3°at $\theta_{Lep} = 70^{\circ}$. Higher angles are poorly reconstructed ($\Delta\theta_{Lep} = 6 - 25^{\circ}$) due to the FGD bar orientation, and returns to 7°below $\theta_{Lep} = 140^{\circ}$.



(a) Arithmetic mean and RMS, forward-going

(b) RMS, forward-going

(c) Arithmetic mean and RMS, backward-going

(d) RMS, backward-going

Figure H.2.: ND280 θ_{Lep} resolution for CC-inclusive event's lepton candidates for forward-going (upper panel) and downward-going (lower panel)

Bibliography

[1] P. Stowell *et al.*, Journal of Instrumentation **12**, P01016 (2017), http://stacks.iop.org/1748-0221/12/i=01/a=P01016.

[2] The T2K Collaboration, M. Hartz *et al.*, T2K Neutrino Oscillation Results with Data up to 2017 Summer, KEK Seminar, Summer, 2017.

[3] The NOνA Collaboration, A. Radovic *et al.*, Latest Oscillation Results from NOvA, Fermilab JETP seminar, 2018.

[4] Hyper-Kamiokande Proto-Collaboration, K. Abe *et al.*, arXiv (2018), https://arxiv.org/abs/1805.04163.

[5] DUNE Collaboration, R. Acciarri *et al.*, arXiv (2016), https://arxiv.org/abs/1601.05471.

[6] B. Pontecorvo, Sov. Phys. JETP **26**, 984 (1968), [Zh. Eksp. Teor. Fiz.53,1717(1967)].

[7] B. Pontecorvo, Sov. Phys. JETP **7**, 172 (1958), [Zh. Eksp. Teor. Fiz.34,247(1957)].

[8] Z. Maki, M. Nakagawa, and S. Sakata, Progress of Theoretical Physics **28**, 870 (1962).

[9] T2K Collaboration, M. Wascko, T2K Status, Results, and Plans, Neutrino 2018, 2018.

[10] NOνA Collaboration, M. Sanchez, NOνA Results and Prospects, Neutrino 2018, 2018.

[11] The T2K and NOνA Collaborations, T2K and NOvA collaborations to produce joint neutrino oscillation analysis, T2K and NOνA news, 2018.

[12] The T2K and NOνA Collaborations, NOvA-T2K Joint Workshop on Neutrino Interaction Uncertainties in Oscillation Measurements, T2K and NOνA news, 2017.

[13] The T2K Collaboration, K. Abe *et al.*, Progress of Theoretical and Experimental Physics **2015**, 043C01 (2015).

[14] LBNE, C. Adams *et al.*, (2013), 1307.7335.

[15] Daya Bay Collaboration, F. P. An *et al.*, Phys. Rev. D **95**, 072006 (2017), https://link.aps.org/doi/10.1103/PhysRevD.95.072006.

[16] IceCube Collaboration, M. Aartsen *et al.*, Phys. Rev. Lett. **120**, 071801 (2018), https://link.aps.org/doi/10.1103/PhysRevLett.120.071801.

[17] Super-Kamiokande Collaboration, K. Abe *et al.*, Phys. Rev. D **97**, 072001 (2018), https://link.aps.org/doi/10.1103/PhysRevD.97.072001.

[18] Particle Data Group, C. Patrignani *et al.*, Chin. Phys. **C** (2016).

[19] J. Chadwick, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **142**, 1 (1933), http://rspa.royalsocietypublishing.org/content/142/846/1.full.pdf.

[20] J. Chadwick, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **136**, 692 (1932), http://rspa.royalsocietypublishing.org/content/136/830/692.full.pdf.

[21] W. Pauli, Rapports du SeptiÃĺme Conseil de Physique Solvay, GauthierâĂŘVillars, Paris, 1933.

[22] E. Fermi, Zeitschrift für Physik **88**, 161 (1934).

[23] F. Reines and C. L. Cowan, Phys. Rev. **92**, 830 (1953).

[24] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, Science **124**, 103 (1956), http://science.sciencemag.org/content/124/3212/103.full.pdf.

[25] R. Davis, Phys. Rev. **97**, 766 (1955).

[26] G. Danby *et al.*, Phys. Rev. Lett. **9**, 36 (1962).

[27] J. Bienlein *et al.*, Physics Letters **13**, 80 (1964).

[28] M. Block *et al.*, Physics Letters **12**, 281 (1964).

[29] M. L. Perl *et al.*, Phys. Rev. Lett. **35**, 1489 (1975).

[30] K. Kodama *et al.*, Physics Letters B **504**, 218 (2001).

[31] The ALEPH Collaboration, The Delphi Collaboration, The L3 Collaboration, The SLD Collaboration, The LEP Electroweak Working Group, The SLD Electroweak and Heavy Flavour Groups, J. A. Bagger *et al.*, Physics Reports **427**, 257 (2006).

[32] M. Lattanzi, Journal of Physics: Conference Series **718**, 032008 (2016).

[33] V. Gribov and B. Pontecorvo, Physics Letters B **28**, 493 (1969).

[34] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[35] M. Kobayashi and T. Maskawa, Progress of Theoretical Physics **49**, 652 (1973).

[36] B. Kayser, eConf **C040802**, L004 (2004), hep-ph/0506165.

[37] V. Barger, K. Whisnant, S. Pakvasa, and R. J. N. Phillips, Phys. Rev. D **22**, 2718 (1980).

[38] S. J. Parke, Phys. Rev. Lett. **57**, 1275 (1986).

[39] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978).

[40] S. Mikheyev and A. Smirnov, Soviet Journal of Nuclear Physics **42**, 913 (1985).

[41] A. Y. Smirnov, Physica Scripta **2005**, 57 (2005).

[42] T. Katori and M. Martini, J. Phys. **G45**, 013001 (2018), 1611.07770.

[43] U. Mosel, Annual Review of Nuclear and Particle Science **66**, 171 (2016), https://doi.org/10.1146/annurev-nucl-102115-044720.

[44] L. Alvarez-Ruso, Y. Hayato, and J. Nieves, New Journal of Physics **16**, 075015 (2014).

[45] C. Berger and L. M. Sehgal, Phys. Rev. D **79**, 053003 (2009), https://link.aps.org/doi/10.1103/PhysRevD.79.053003.

[46] A. Botrugno and G. Co', Nuclear Physics A **761**, 200 (2005).

[47] J. Nieves, I. R. Simo, and M. J. V. Vacas, Phys. Rev. C **83**, 045501 (2011), https://link.aps.org/doi/10.1103/PhysRevC.83.045501.

[48] R. Gran, J. Nieves, F. Sanchez, and M. J. V. Vacas, Phys. Rev. D **88**, 113007 (2013), https://link.aps.org/doi/10.1103/PhysRevD.88.113007.

[49] J. T. Sobczyk and J. Żmuda, Phys. Rev. C **87**, 065503 (2013).

[50] O. Benhar and A. Fabrocini, Phys. Rev. C **62**, 034304 (2000), https://link.aps.org/doi/10.1103/PhysRevC.62.034304.

[51] J. E. Sobczyk, Phys. Rev. C **96**, 045501 (2017).

[52] P. Martel *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **612**, 46 (2009).

[53] Y. Hayato, Acta Physica Polonica B **40** (2009).

[54] L. Alvarez-Ruso *et al.*, Progress in Particle and Nuclear Physics **100**, 1 (2018).

[55] C. Andreopoulos *et al.*, Nucl. Instrum. Meth. **A** (2010).

[56] T. Golan, J. Sobczyk, and J. Zmuda, Nucl. Phys. **B**, 499 (2012).

[57] E. P. Guerra, PhD thesis, York University, 2018.

[58] K. Gallmeister, U. Mosel, and J. Weil, Phys. Rev. C **94**, 035502 (2016).

[59] IceCube Collaboration, T. DeYoung, Latest Neutrino Physics Results from IceCube and ANTARES, Neutrino 2018, 2018.

[60] A. BELLERIVE, International Journal of Modern Physics A **19**, 1167 (2004), https://www.worldscientific.com/doi/pdf/10.1142/S0217751X04019093.

[61] A. Bellerive, J. Klein, A. McDonald, A. Noble, and A. Poon, Nuclear Physics B **908**, 30 (2016), Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015.

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

[63] K. S. Hirata *et al.*, Phys. Rev. Lett. **63**, 16 (1989).

[64] B. T. Cleveland *et al.*, The Astrophysical Journal **496**, 505 (1998).

[65] SAGE Collaboration, J. N. Abdurashitov *et al.*, Phys. Rev. C **60**, 055801 (1999).

[66] W. Hampel *et al.*, Physics Letters B **447**, 127 (1999).

[67] SNO Collaboration, Q. R. Ahmad *et al.*, Phys. Rev. Lett. **89**, 011301 (2002).

[68] G. Ranucci *et al.*, Journal of Physics: Conference Series **675**, 012036 (2016).

[69] C. Kraus and S. J. Peeters, Progress in Particle and Nuclear Physics **64**, 273 (2010), Neutrinos in Cosmology, in Astro, Particle and Nuclear Physics.

[70] S. Moriyama and S.-K. Collaboration, Journal of Physics: Conference Series **888**, 012005 (2017).

[71] M. Honda, T. Kajita, K. Kasahara, and S. Midorikawa, Phys. Rev. D **70**, 043008 (2004).

[72] F. Reines *et al.*, Phys. Rev. Lett. **15**, 429 (1965).

[73] C. Achar *et al.*, Physics Letters **18**, 196 (1965).

[74] T. J. Haines *et al.*, Phys. Rev. Lett. **57**, 1986 (1986).

[75] K. Hirata *et al.*, Physics Letters B **205**, 416 (1988).

[76] W. Allison *et al.*, Physics Letters B **391**, 491 (1997).

[77] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, Phys. Rev. Lett. **81**, 1562 (1998), https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[78] T. Kajita, New Journal of Physics **6**, 194 (2004).

[79] ANTARES, A. Albert *et al.*, (2017), 1711.01496.

[80] SNO Collaboration, B. Aharmim *et al.*, Phys. Rev. D **80**, 012001 (2009).

[81] Super-Kamiokande, Z. Li *et al.*, (2017), 1711.09436.

[82] G. J. Feldman, J. Hartnell, and T. Kobayashi, Advances in High Energy Physics **2013**.

[83] A. Aguilar-Arevalo *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **599**, 28 (2009).

[84] L. Aliaga *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **743**, 130 (2014).

[85] SBND, MicroBooNE, S. Tufanli, PoS **EPS-HEP2017**, 141 (2017).

[86] MINOS Collaboration, D. G. Michael *et al.*, Phys. Rev. Lett. **97**, 191801 (2006), https://link.aps.org/doi/10.1103/PhysRevLett.97.191801.

[87] K2K Collaboration, M. H. Ahn *et al.*, Phys. Rev. D **74**, 072003 (2006), https://link.aps.org/doi/10.1103/PhysRevD.74.072003.

[88] K2K Collaboration, S. Yamamoto *et al.*, Phys. Rev. Lett. **96**, 181801 (2006).

[89] MINOS Collaboration, P. Adamson *et al.*, Phys. Rev. Lett. **110**, 171801 (2013), https://link.aps.org/doi/10.1103/PhysRevLett.110.171801.

[90] T2K Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **112**, 061802 (2014), https://link.aps.org/doi/10.1103/PhysRevLett.112.061802.

[91] NOvA Collaboration, P. Adamson *et al.*, Phys. Rev. Lett. **116**, 151806 (2016), https://link.aps.org/doi/10.1103/PhysRevLett.116.151806.

[92] R. Acquafredda *et al.*, Journal of Instrumentation **4**, P04018 (2009).

[93] S. Amerio *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **527**, 329 (2004).

[94] LSND Collaboration, A. Aguilar *et al.*, Phys. Rev. D **64**, 112007 (2001).

[95] MiniBooNE Collaboration, A. A. Aguilar-Arevalo *et al.*, Phys. Rev. Lett. **110**, 161801 (2013).

[96] HARP–CDP Group, A. Bolshakova *et al.*, Phys. Rev. D **85**, 092008 (2012).

[97] A. Pastore, Nuclear Physics B - Proceedings Supplements **217**, 163 (2011), Proceedings of the Neutrino Oscillation Workshop (NOW 2010).

[98] OPERA Collaboration, N. Agafonova *et al.*, Phys. Rev. Lett. **120**, 211801 (2018).

[99] M. Antonello *et al.*, The European Physical Journal C **73**, 2345 (2013).

[100] The OPERA Collaboration, N. Agafonova *et al.*, Prepared for JHEP (2018).

[101] H. Furuta *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **662**, 90 (2012).

[102] RENO Collaboration, J. H. Choi *et al.*, Phys. Rev. Lett. **116**, 211801 (2016), https://link.aps.org/doi/10.1103/PhysRevLett.116.211801.

[103] F. P. An *et al.*, Phys. Rev. Lett. **108**, 171803 (2012), https://link.aps.org/doi/10.1103/PhysRevLett.108.171803.

[104] RENO Collaboration, J. K. Ahn *et al.*, Phys. Rev. Lett. **108**, 191802 (2012), https://link.aps.org/doi/10.1103/PhysRevLett.108.191802.

[105] Y. Abe *et al.*, Journal of High Energy Physics **2016**, 163 (2016).

[106] The RENO Collaboration, G. Bak *et al.*, arxiv **1806.00248**, https://arxiv.org/abs/1806.00248.

[107] Y. Abe *et al.*, Journal of High Energy Physics **2014**, 86 (2014).

[108] The T2K Collaboration, K. Abe *et al.*, Phys. Rev. **D** (2017).

[109] NOvA Collaboration, P. Adamson *et al.*, Phys. Rev. Lett. **118**, 151802 (2017).

[110] P. Adamson *et al.*, Phys. Rev. Lett. **112**, 191801 (2014).

[111] Y.-F. Li, J. Cao, Y. Wang, and L. Zhan, Phys. Rev. D **88**, 013008 (2013).

[112] The KamLAND Collaboration, A. Gando *et al.*, Phys. Rev. D **83**, 052002 (2011).

[113] P. Huber, Phys. Rev. Lett. **118**, 042502 (2017).

[114] M. Dentler, Á. Hernández-Cabezudo, J. Kopp, M. Maltoni, and T. Schwetz, Journal of High Energy Physics **2017**, 99 (2017).

[115] NEOS Collaboration, Y. J. Ko *et al.*, Phys. Rev. Lett. **118**, 121802 (2017).

[116] I. Alekseev *et al.*, International Journal of Modern Physics: Conference Series **46**, 1860044 (2018), https://www.worldscientific.com/doi/pdf/10.1142/S2010194518600443.

[117] PROSPECT, J. Ashenfelter *et al.*, (2018), 1806.02784.

[118] STEREO, H. AlmazÃąn *et al.*, (2018), 1806.02096.

[119] Y. Abreu *et al.*, Journal of Instrumentation **12**, P04024 (2017).

[120] The T2K Collaboration, KEK Proposal (2001), http://neutrino.kek.jp/jhfnu/loi/loi.v2.030528.pdf.

[121] The T2K Collaboration, KEK Proposal (2006), http://j-parc.jp/researcher/Hadron/en/pac_0606/pdf/p11-Nishikawa.pdf.

[122] NINJA Collaboration, T. Fukuda *et al.*, Proposal for precise measurement of neutrinop-water cross-section in NINJA physics run, Proposal for J-PARC and KEK, 2017.

[123] T. Ovsiannikova *et al.*, Physics of Particles and Nuclei **48**, 1014 (2017),

https://doi.org/10.1134/S1063779617060478.

[124] M. Antonova *et al.*, Journal of Instrumentation **12**, C07028 (2017), http://stacks.iop.org/1748-0221/12/i=07/a=C07028.

[125] Y. Yamazaki *et al.*, Accelerator technical design report for J-PARC, 2003.

[126] The T2K Collaboration, K. Abe *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **659**, 106 (2011), http://www.sciencedirect.com/science/article/pii/S0168900211011910.

[127] T. Sekiguchi *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **789**, 57 (2015), http://www.sciencedirect.com/science/article/pii/S0168900215004672.

[128] T2K Collaboration, K. Abe *et al.*, Phys. Rev. D **87**, 012001 (2013), https://link.aps.org/doi/10.1103/PhysRevD.87.012001.

[129] K. Matsuoka *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **624**, 591 (2010), http://www.sciencedirect.com/science/article/pii/S016890021002098X.

[130] K. Suzuki *et al.*, Progress of Theoretical and Experimental Physics **2015**, 053C01 (2015).

[131] D. Beavis, C. A., and I. Chiang, (1995), Physics design report, Report Number BNL–52459.

[132] T2K Collaboration, K. Abe *et al.*, Phys. Rev. D **93**, 072002 (2016).

[133] K. Abe *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **694**, 211 (2012), http://www.sciencedirect.com/science/article/pii/S0168900212002987.

[134] T. Kikawa, Nuclear Physics B - Proceedings Supplements **229-232**, 451 (2012), Neutrino 2010.

[135] P.-A. Amaudruz *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **696**, 1 (2012), http://www.sciencedirect.com/science/article/pii/S0168900212008789.

[136] C. Nielsen, *Constraining the flux and cross section models using carbon and oxygen*

targets in the off-axis near detector for the 2016 joint oscillation analysis at T2K, PhD thesis, University of British Columbia, 2017.

[137] N. Abgrall *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **637**, 25 (2011), http://www.sciencedirect.com/science/article/pii/S0168900211003421.

[138] I. Giomataris *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **560**, 405 (2006).

[139] Y. Giomataris, P. Rebourgeard, J. Robert, and G. Charpak, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **376**, 29 (1996).

[140] C. Giganti, *Particle Identification in the T2K TPCs and study of the electron neutrino component in the T2K neutrino beam*, PhD thesis, IRFU, Saclay, 2010.

[141] S. Assylbekov *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **686**, 48 (2012), http://www.sciencedirect.com/science/article/pii/S0168900212005153.

[142] D. Allan *et al.*, Journal of Instrumentation **8**, P10019 (2013), http://stacks.iop.org/1748-0221/8/i=10/a=P10019.

[143] S. Aoki *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **698**, 135 (2013), http://www.sciencedirect.com/science/article/pii/S0168900212011242.

[144] S. Fukuda *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **501**, 418 (2003), http://www.sciencedirect.com/science/article/pii/S016890020300425X.

[145] H. Nishino *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **610**, 710 (2009), http://www.sciencedirect.com/science/article/pii/S0168900209017495.

[146] K. Abe *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **737**, 253 (2014), http://www.sciencedirect.com/science/article/pii/S016890021301646X.

[147] S. Ahn *et al.*, Physics Letters B **511**, 178 (2001).

[148] S.-K. Collaboration, Outline of the start of refurbishment work for Super-Kamiokande, 2018.

[149] GEANT4, S. Agostinelli *et al.*, Nucl. Instrum. Meth. **A506**, 250 (2003).

[150] G. Battistoni *et al.*, AIP Conference Proceedings **896**, 31 (2007), https://aip.scitation.org/doi/pdf/10.1063/1.2720455.

[151] A. Ferrari, P. R. Sala, A. FassÃš, and J. Ranft, *FLUKA: A multi-particle transport code (program version 2005)*CERN Yellow Reports: Monographs (CERN, Geneva, 2005), http://cds.cern.ch/record/898301.

[152] T. BÃűhlen *et al.*, Nuclear Data Sheets **120**, 211 (2014), http://www.sciencedirect.com/science/article/pii/S0090375214005018.

[153] R. Brun *et al.*, *GEANT: Detector Description and Simulation Tool; Oct 1994*CERN Program Library (CERN, Geneva, 1993), http://cds.cern.ch/record/1082634, Long Writeup W5013.

[154] C. Zeitnitz and T. Gabriel, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **349**, 106 (1994), http://www.sciencedirect.com/science/article/pii/0168900294906130.

[155] A. Fiorentini *et al.*, T2K Technical Note **217** (2017).

[156] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (Chapman & Hall/CRC Interdisciplinary Statistics, 1995).

[157] S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, *Handbook of Markov Chain Monte Carlo* (CRC Press, 2011).

[158] S. Sharma, Annual Review of Astronomy and Astrophysics **55**, 213 (2017), https://doi.org/10.1146/annurev-astro-082214-122339.

[159] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial* (Oxford Science Publications, 2006).

[160] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Journal of Chemical Physics **21** (1970).

[161] W. K. Hastings, Biometrika **57** (1970).

[162] G. O. Roberts, A. Gelman, and W. R. Gilks, The Annals of Applied Probability **7** (1997), http://faculty.wcas.northwestern.edu/ lchrist/course/Gerzensee_2011/The%20Annals%20of%

[163] A. Gelman and D. B. Rubin, Statist. Sci. **7**, 457 (1992), https://doi.org/10.1214/ss/1177011136.

[164] A. Gelman, X.-L. Meng, and H. Stern, Statistica Sinica **6** (1996).

[165] X.-L. Meng, The Annals of Statistics **22** (1994).

[166] A. Gelman, Y. Goegebeur, F. Tuerlinckx, and I. V. Mechelen, Appl. Statist. **49** (2000).

[167] G. E. P. Box, J. R. Statist. Soc. **A** (1980).

[168] F. James, Reference Manual, Version 94.1, CERN Program Library Long Writeup D506 **D** (1994), https://root.cern.ch/sites/d35c7d8c.web.cern.ch/files/minuit.pdf.

[169] L. Haegel, *Measurement Of Neutrino Oscillation Parameters Using Neutrino And Antineutrino Data Of The T2K Experiment*, PhD thesis, University of Geneva, 2017.

[170] K. E. Duffy, *Measurement of the Neutrino Oscillation Parameters* $\sin^2\theta_{23}$, Δm^2_{32}, $\sin^2\theta_{13}$, *and* δ_{CP} *in Neutrino and Antineutrino Oscillation at T2K*, PhD thesis, Oriel College, University of Oxford, 2016.

[171] R. G. Calland, *A 3 Flavour Joint Near and Far Detector Neutrino Oscillation Analysis at T2K*, PhD thesis, University of Liverpool, 2014.

[172] N. Abgrall *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **701**, 99 (2013), http://www.sciencedirect.com/science/article/pii/S016890021201234X.

[173] HARP Collaboration, M. Apollonio *et al.*, Phys. Rev. C **80**, 035208 (2009), https://link.aps.org/doi/10.1103/PhysRevC.80.035208.

[174] J. E. Amaro, C. Maieron, J. Nieves, and M. Valverde, The European Physical Journal A - Hadrons and Nuclei **24**, 343 (2005), https://doi.org/10.1140/epja/i2005-10034-2.

[175] C. Wilkinson *et al.*, Phys. Rev. D **93**, 072010 (2016), https://link.aps.org/doi/10.1103/PhysRevD.93.072010.

[176] C. L. Smith, Physics Reports **3**, 261 (1972), http://www.sciencedirect.com/science/article/pii/0370157372900105.

[177] R. Bradford, A. Bodek, H. Budd, and J. Arrington, Nuclear Physics B - Proceedings Supplements **159**, 127 (2006), http://www.sciencedirect.com/science/article/pii/S0920563206005184, Proceedings of the 4th International Workshop on Neutrino-Nucleus Interactions in the Few-GeV Region.

[178] R. Smith and E. Moniz, Nuclear Physics B **43**, 605 (1972), http://www.sciencedirect.com/science/article/pii/0550321372900405.

[179] MiniBooNE Collaboration, A. A. Aguilar-Arevalo *et al.*, Phys. Rev. D **81**, 092005 (2010), https://link.aps.org/doi/10.1103/PhysRevD.81.092005.

[180] MiniBooNE Collaboration, A. A. Aguilar-Arevalo *et al.*, Phys. Rev. D **88**, 032001 (2013), https://link.aps.org/doi/10.1103/PhysRevD.88.032001.

[181] MINERvA Collaboration, G. A. Fiorentini *et al.*, Phys. Rev. Lett. **111**, 022502 (2013), https://link.aps.org/doi/10.1103/PhysRevLett.111.022502.

[182] MINERvA Collaboration, L. Fields *et al.*, Phys. Rev. Lett. **111**, 022501 (2013), https://link.aps.org/doi/10.1103/PhysRevLett.111.022501.

[183] D. Rein and L. M. Sehgal, Annals of Physics **133**, 79 (1981), http://www.sciencedirect.com/science/article/pii/0003491681902426.

[184] D. Rein, Zeitschrift für Physik C Particles and Fields **35**, 43 (1987), https://doi.org/10.1007/BF01561054.

[185] K. S. Kuzmin, V. V. Lyubushkin, and V. A. Naumov, Nuclear Physics B - Proceedings Supplements **139**, 158 (2005), http://www.sciencedirect.com/science/article/pii/S0920563204007820, Proceedings of the Third International Workshop on Neutrino-Nucleus Interactions in the Few-GeV Region.

[186] K. M. Graczyk and J. T. Sobczyk, Phys. Rev. D **77**, 053003 (2008), https://link.aps.org/doi/10.1103/PhysRevD.77.053003.

[187] C. Berger and L. M. Sehgal, Phys. Rev. D **76**, 113004 (2007), https://link.aps.org/doi/10.1103/PhysRevD.76.113004.

[188] K. M. Graczyk, J. Żmuda, and J. T. Sobczyk, Phys. Rev. D **90**, 093001 (2014), https://link.aps.org/doi/10.1103/PhysRevD.90.093001.

[189] K. M. Graczyk, D. Kiełczewska, P. Przewłocki, and J. T. Sobczyk, Phys. Rev. D **80**, 093001 (2009), https://link.aps.org/doi/10.1103/PhysRevD.80.093001.

[190] K. M. Graczyk and J. T. Sobczyk, Phys. Rev. D **77**, 053001 (2008), https://link.aps.org/doi/10.1103/PhysRevD.77.053001.

[191] G. M. Radecky *et al.*, Phys. Rev. D **25**, 1161 (1982), https://link.aps.org/doi/10.1103/PhysRevD.25.1161.

[192] M. Derrick *et al.*, Phys. Rev. D **23**, 569 (1981), https://link.aps.org/doi/10.1103/PhysRevD.23.569.

[193] T. Kitagaki *et al.*, Phys. Rev. D **34**, 2554 (1986), https://link.aps.org/doi/10.1103/PhysRevD.34.2554.

[194] N. J. Baker *et al.*, Phys. Rev. D **23**, 2495 (1981), https://link.aps.org/doi/10.1103/PhysRevD.23.2495.

[195] K. Furuno, NuInt02 proc. **KEK Preprint 2003-48** (2003).

[196] C. Wilkinson, P. Rodrigues, S. Cartwright, L. Thompson, and K. McFarland, Phys. Rev. D **90**, 112017 (2014), https://link.aps.org/doi/10.1103/PhysRevD.90.112017.

[197] MiniBooNE Collaboration, A. A. Aguilar-Arevalo *et al.*, Phys. Rev. D **83**, 052007 (2011), https://link.aps.org/doi/10.1103/PhysRevD.83.052007.

[198] MiniBooNE Collaboration, A. A. Aguilar-Arevalo *et al.*, Phys. Rev. D **83**, 052009 (2011), https://link.aps.org/doi/10.1103/PhysRevD.83.052009.

[199] The MiniBooNE Collaboration, A. A. Aguilar-Arevalo *et al.*, Phys. Rev. D **81**, 013005 (2010), https://link.aps.org/doi/10.1103/PhysRevD.81.013005.

[200] T. Le *et al.*, Physics Letters B **749**, 130 (2015), http://www.sciencedirect.com/science/article/pii/S0370269315005493.

[201] MINERvA Collaboration, B. Eberly *et al.*, Phys. Rev. D **92**, 092008 (2015), https://link.aps.org/doi/10.1103/PhysRevD.92.092008.

[202] MINERvA Collaboration, C. L. McGivern *et al.*, Phys. Rev. D **94**, 052005 (2016),

https://link.aps.org/doi/10.1103/PhysRevD.94.052005.

[203] K2K Collaboration, A. Rodriguez *et al.*, Phys. Rev. D **78**, 032003 (2008), https://link.aps.org/doi/10.1103/PhysRevD.78.032003.

[204] D. Rein and L. M. Sehgal, Nuclear Physics B **223**, 29 (1983), http://www.sciencedirect.com/science/article/pii/0550321383900901.

[205] MINERvA Collaboration, A. Higuera *et al.*, Phys. Rev. Lett. **113**, 261802 (2014), https://link.aps.org/doi/10.1103/PhysRevLett.113.261802.

[206] M. Glück, E. Reya, and A. Vogt, The European Physical Journal C - Particles and Fields **5**, 461 (1998), https://doi.org/10.1007/s100529800978.

[207] A. Bodek and U. K. Yang, AIP Conference Proceedings **670**, 110 (2003), https://aip.scitation.org/doi/pdf/10.1063/1.1594324.

[208] T. Sjostrand, Computer Physics Communications **82**, 74 (1994), http://www.sciencedirect.com/science/article/pii/0010465594901325.

[209] MINOS Collaboration, P. Adamson *et al.*, Phys. Rev. D **81**, 072002 (2010), https://link.aps.org/doi/10.1103/PhysRevD.81.072002.

[210] P. Lasorak, *A search for neutrino-induced single photons and measurement of oscillation analysis systematic errors with electron and anti-electron neutrino selections, using the off-axis near detector of the Tokai to Kamioka experiment*, PhD thesis, Queen Mary's University of London, 2018.

[211] E. Wang, L. Alvarez-Ruso, Y. Hayato, K. Mahn, and J. Nieves, Phys. Rev. D **92**, 053005 (2015), https://link.aps.org/doi/10.1103/PhysRevD.92.053005.

[212] M. Day and K. S. McFarland, Phys. Rev. D **86**, 053003 (2012), https://link.aps.org/doi/10.1103/PhysRevD.86.053003.

[213] L. Salcedo, E. Oset, M. Vicente-Vacas, and C. Garcia-Recio, Nuclear Physics A **484**, 557 (1988), http://www.sciencedirect.com/science/article/pii/0375947488903107.

[214] G. Rowe, M. Salomon, and R. H. Landau, Phys. Rev. C **18**, 584 (1978), https://link.aps.org/doi/10.1103/PhysRevC.18.584.

[215] R. Seki and K. Masutani, Phys. Rev. C **27**, 2799 (1983), https://link.aps.org/doi/10.1103/PhysRevC.27.2799.

[216] Particle Data Group, K. Nakamura *et al.*, J. Phys. **G** (2010).

[217] X. Espinal and F. Sanchez, AIP Conference Proceedings **967**, 117 (2007), https://aip.scitation.org/doi/pdf/10.1063/1.2834461.

[218] K2K Collaboration, R. Gran *et al.*, Phys. Rev. D **74**, 052002 (2006), https://link.aps.org/doi/10.1103/PhysRevD.74.052002.

[219] MINOS Collaboration, P. Adamson *et al.*, Phys. Rev. D **91**, 012005 (2015), https://link.aps.org/doi/10.1103/PhysRevD.91.012005.

[220] V. Bernard, L. Elouadrhiri, and U.-G. MeiÃ§ner, Journal of Physics G: Nuclear and Particle Physics **28**, R1 (2002), http://stacks.iop.org/0954-3899/28/i=1/a=201.

[221] DUET Collaboration, E. S. Pinzon Guerra *et al.*, Phys. Rev. C **95**, 045203 (2017), https://link.aps.org/doi/10.1103/PhysRevC.95.045203.

[222] J. Engel, Phys. Rev. C **57**, 2004 (1998), https://link.aps.org/doi/10.1103/PhysRevC.57.2004.

[223] K. S. Kim and M. K. Cheoun, AIP Conference Proceedings **1189**, 163 (2009).

[224] A. S. Meyer, M. Betancourt, R. Gran, and R. J. Hill, Phys. Rev. D **93**, 113015 (2016).

[225] M. Kabirnezhad, *Improvement of Single Pion Production for T2K experiment simulation tools*, PhD thesis, National Center for Nuclear Research, Warsaw, Poland, 2017.

[226] T. Vladisavljevic, (2018), 1804.00272.

[227] T2K, D. Sgalaberna, PoS **EPS-HEP2017**, 518 (2017).

[228] T. Kitagaki *et al.*, Phys. Rev. D **42**, 1331 (1990), https://link.aps.org/doi/10.1103/PhysRevD.42.1331.

[229] S. J. Barish *et al.*, Phys. Rev. D **16**, 3103 (1977), https://link.aps.org/doi/10.1103/PhysRevD.16.3103.

[230] N. J. Baker *et al.*, Phys. Rev. D **23**, 2499 (1981), https://link.aps.org/doi/10.1103/PhysRevD.23.2499.

List of Figures

List of Tables