ROYAL HOLLOWAY, UNIVERSITY OF LONDON

Constraining Systematic Uncertainties at T2K using Near Detector Data

Author:
Will Parker

Supervisor:
Dr. Asher Kaboth

October 23, 2020

Abstract

T2K is a long baseline neutrino oscillation experiment in Japan. It was designed to make precise measurements of the parameters governing neutrino oscillations, and has been taking data since 2011. A muon (anti-)neutrino beam is produced at the Japan Proton Accelerator Research Complex (J-PARC), and is aimed towards the Super-Kamiokande (SK) detector 295 km away. In this analysis, Markov Chain Monte Carlo is used to fit the Monte Carlo prediction to data from the near detector, ND280, which measures the neutrino flux and interaction cross-sections before oscillation. The flux and interaction models are parametrised using external data, T2K beam line monitoring measurements, and theoretical calculations to set the prior values and uncertainties. Several updates have been made to the data samples, cross-section model, and fitting framework used for the 2020 oscillation analysis to maximise the constraint on these systematics, and reduce the impact they have on oscillation results. The near detector fit is crucial for T2K to make world-leading oscillation parameter measurements. The analysis presented here uses ND280 data from T2K runs 2–9, corresponding to 1.99×10^{21} protons on target (POT), to reduce the uncertainty on the SK prediction from 12–14% to 2–3%. The reduction of systematic uncertainties in detectors for future long baseline neutrino oscillation experiments is also investigated.

Declaration of Authorship

I, William Parker, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Acknowledgements

Contents

Contents

List of Figures

List of Tables

1. Introduction

Sakharov proposed in 1967 [1] that the abundance of matter over anti-matter observed in the Universe could be explained by interactions which meet three conditions. These are that: baryon number is violated, such that a different number of baryons can be produced to anti-baryons; C (charge) and CP (charge-parity) symmetries are violated, so that processes producing more baryons aren't compensated by processes producing more anti-baryons; and the interactions occur outside of thermal equilibrium, so the net amount of generated baryons isn't cancelled out by the time reverse of the process.

CP violation was first observed in the quark sector in 1964 [2], but it is not sufficient to alone explain the matter dominated Universe we observe. The PMNS mechanism for neutrino oscillations allows a CP violating phase, δ_{CP}, which could be non-zero. It has been shown that CP violation in leptons could produce the matter anti-matter imbalance through leptogenesis [3].

T2K [4] is a long baseline neutrino oscillation experiment, originally designed to make precision measurements of the oscillation parameters $\sin^2\theta_{23}$, Δm_{32}^2, and $\sin^2\theta_{13}$. However, one of the main focuses is now on measuring δ_{CP} to determine if CP is violated in neutrino oscillations, and if so by how much. As the datasets for both neutrinos and anti-neutrinos increase, and with the prospect of future neutrino oscillation experiments such as Hyper-Kamiokande [5] and DUNE [6] on the horizon, we will soon be able to determine the value of δ_{CP}. However, with the increase in sample size the treatment of systematics becomes more important. To achieve the required sensitivity, improvements to the neutrino interaction model are required.

The T2K near detector prediction is fitted to data to reduce systematic uncertainties from the cross-section, flux, and detector models used to build the MC prediction. This process typically reduces the uncertainty on event rates at the far detector, Super-Kamiokande, from $\sim$15% to $\sim$5%, without which T2K would not be able to make world leading measurements of oscillation parameters.

This thesis presents the results of the near detector fit for the 2020 oscillation analysis, as

1

well as describing the updates to the models and fitting framework since the last analysis. Chapter 2 outlines the history of neutrino physics, from initial discovery to current open questions. It also describes the theory of both neutrino interactions, particularly those present in the T2K detectors, and neutrino oscillations.

The T2K experimental setup is detailed in Chapter 3, giving an overview of the beamline, the on-axis near detector INGRID, the off-axis ND280, and the far detector Super-Kamiokande. The simulation used to predict measurements at each of the detectors is also described.

The statistical treatment of data is discussed in Chapter 4. This work is a Bayesian analysis, which uses Markov Chain Monte Carlo (MCMC) for fitting systematics to data. The MCMC method, and Bayesian techniques for post-fit analysis and interpretation are introduced here.

This is followed by a description of the near detector fit and inputs for the 2020 oscillation analysis in Chapter 5. Various improvements were made to the cross-section, flux, and detector models, as well as the fitting framework itself, since the previous analysis. The implementation of these, and the potential impact on the fit are discussed here.

The results of the fit are described in Chapter 6. First the validations of the model and fitting framework are shown, before the final results and impact at the far detector are presented. The impact of these results on the sensitivity to oscillation parameter measurements is then studied.

In Chapter 7, the potential impact of a new technology for long baseline neutrino oscillation experiment near detectors is discussed. Finally, this thesis concludes with a summary of the results presented.

The three sets of systematics, flux, cross-section, and detector, were provided by working groups within T2K. The beam group produced a covariance matrix of the flux systematics. The Neutrino Interaction Working Group (NIWG) recommended the cross-section systematics. This was partly informed by the studies on the binding energy dial performed by the author of this thesis, presented in Section 5.4.1.1. The ND280 Selection, Systematics and Validations group recommended the detector systematics. I made correlated throws of these systematics to produce a covariance matrix, and performed the studies on detector binnings presented in Section 5.4.3.3.

The implementation of these systematics into the analysis framework, the ND280 fits, validations, and postfit studies, were all my work. The joint near and far detector fits were run in parallel by several members of the MaCh3 group, including me. I produced the

comparisons of the joint fit results for fixing the binding energy, and using uniform binning, presented in Section 6.8.1.

2. Neutrino Physics

Neutrinos are spin-1/2, colourless, fermions with extremely small mass and zero electric charge. They only interact via the weak and gravitational forces, allowing them to travel great distances through matter without ever being affected by it, and so making them very difficult to detect. This has made neutrinos one of the most elusive particles, despite being one of the most abundant in the Universe.

This chapter gives an overview of the history of neutrino physics, from Wolfgang Pauli's 'desperate remedy' of an undetectable particle, to the 2015 Nobel prize winning results of SK [7] and SNO [8], as well as the relevant theory involved. Section 2.1 describes the initial evidence of the existence of neutrinos, and the discovery of the different flavours. The evidence of neutrino oscillations is presented in Section 2.2, separated by neutrino source. The theory behind the oscillation mechanism is introduced in Section 2.3, along with a discussion of the different interaction types relevant for long baseline neutrino oscillation experiments. Finally, Section 2.4 gives an overview of the current experimental status of neutrino oscillations, and the questions yet to be answered.

2.1. Neutrino Discovery

In 1933, using a magnetic spectrometer and a Geiger counter, Chadwick [9] measured a continuous energy spectrum of e^- from β decay. This appeared to violate the conservation laws for energy, momentum, and spin. This is because, assuming a two-body process, the emitted particle is carrying away the energy difference between the initial and final nuclear states, which should be constant, as is the case for α and γ decays. However, the vast majority of the emitted β particles measured had energies much lower than the expected value, but none had energies higher.

Neils Bohr initially proposed a statistical formulation of the conservation laws, suggesting individual decays could violate them as long as the overall average resulted in no net change. However, the upper limit on the measured β energies, which was confirmed by Ellis and

Mott [10], contradicted this theory. If the conservation laws were invalid, any amount of energy would be available in at least a small fraction of decays.

To attempt to resolve the problem, Pauli [11] proposed another, undetected, fermion would need to also be produced by the decay, and that it would be light and have zero electric charge. This would explain the observed spectrum as there was a fixed limit on the available energy from the conservation laws, but the new particle would take a varying fraction of it, with the β taking the rest. The β can therefore have a varied energy spectrum up to the hard limit, as had been measured. Pauli originally called the new neutral particle a 'neutron'.

In 1932, Chadwick discovered a massive neutral particle in the nucleus of atoms, which he also called a 'neutron'. Initially, Pauli did not distinguish between the two namesake particles. Two years later, Fermi devised a framework by which the light chargeless fermions could account for the missing energy [12] in β decay, in which he coined the name 'neutrino' for them, meaning 'the little neutral one'.

In this theory of β decay, Chadwick's neutron could decay to a proton, emitting a β and Pauli's neutrino:

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

The neutrality and lightness of the neutrino meant they would be very difficult to detect, and it was not until 1956 that the first experimental observation took place, by Reines and Cowan [13]. Given the small chance of interaction, an extremely large flux of neutrinos was required, and so it was only with the advent of nuclear reactors that detection became viable.

Two tanks of water sandwiched between three tanks of liquid scintillator containing 110 photomultiplier tubes (PMTs) were placed near the Savannah River Plant. It was predicted that anti-neutrinos could interact with protons via inverse β decay:

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

The tanks of water provided a large amount of target protons for the anti-neutrinos to interact with, producing neutrons and positrons. Positron annihilation would then produce two γ-rays, causing a flash in the scintillator liquid which could be detected by the PMTs.

The water was doped with 40 kg of $CdCl_2$, so that the free neutron could be detected via neutron capture:

$$n + {}^{108}\text{Cd} \rightarrow {}^{109}\text{Cd}^* \rightarrow {}^{109}\text{Cd} + \gamma. \tag{2.3}$$
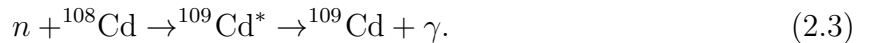
The coincidence of a neutron capture 5 μs after a positron annihilation provided an unambiguous signature of an anti-neutrino reaction. This was the first experimental evidence of the existence of neutrinos.

A similar experiment by Ray Davis [14] in 1962 used a tank of chlorine doped water placed near the Brookhaven nuclear reactor, to search for the interaction:

$$\bar{\nu} + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar}. \tag{2.4}$$

However, no excess of argon was detected. This along with the Reines and Cowan result lead to the theory of lepton number conservation. This meant that anti-neutrinos could not be produced from interactions involving leptons, and neutrinos could not be produced from interactions involving anti-leptons.

In 1962, Ledermen, Schwartz, and Steinberger [15] explored the possibility of the existence of two separate neutrino flavours. Protons were initially accelerated to 15 GeV, before being impinged on a beryllium target. This produced many pions, which decayed into muons and neutrinos:

$$\pi \rightarrow \mu + \nu_\mu. \tag{2.5}$$

A 13 m steel shield then stopped the muons and any surviving pions, leaving a beam of neutrinos. A spark chamber was used to detect when these neutrinos interacted with target aluminium sheets. An excess of muons compared to electrons was observed, showing that neutrinos produced with a muon, produce another muon when they interact. This proved that muon neutrinos were distinct from electron neutrinos. Measurements at CERN confirmed this result in 1963 [16].

When the tau lepton was discovered in 1975, it was expected that there would be a corresponding neutrino. However, as tau leptons have a half-life of only 300 fs, it is difficult to use them to detect tau neutrinos. This, coupled with the fact that tau neutrinos are extremely rare, meant that the first direct detection was not until 2000, at the DONUT experiment [17]. Protons were accelerated to 800 GeV and impinged on 36 m of Tungsten. This produced D_s mesons, which quickly decayed to tau anti-neutrinos and tau leptons, which would then decay to produce a tau neutrino. A kink in the tau lepton's path, detected with nuclear emulsion, was used to identify the tau decay, and prove the existence of the tau
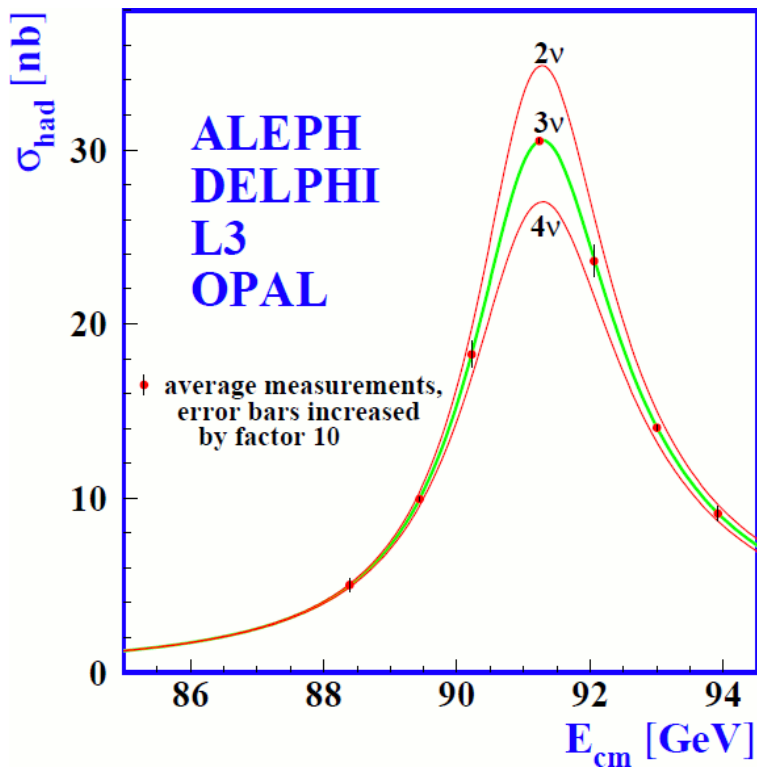
Figure 2.1.: The cross-section for Z production as a function of energy. The red and green lines show the prediction in the case of 2, 3, and 4 active neutrino flavours. The data points are from a combination of the ALEPH, DELPHI, L3, and OPAL experiments at LEP. Figure from [20].

neutrino.

The possibility of further active neutrino flavours has been ruled out by measurements of Z decays at the LEP (Large Electron Postitron) collider and the Stanford Linear Accelerator Center (SLAC) [18]. The width of the Z mass peak is the sum of the visible partial width (from decays to leptons and hadrons), and the invisible partial width (assumed to be from decays to N_ν light neutrino species). Assuming each flavour contributes equally, N_ν was measured to equal 2.9840±0.0082. The results are shown in Figure 2.1, clearly most consistent with the three active neutrinos case. This has been supported by measurements of the expansion rate of the early Universe, which is consistent with N_ν=3.04±0.18 [19].

However, this does not mean there cannot be further types of neutrino which do not interact via the weak force and are therefore unable to couple to the Z boson, known as sterile neutrinos. There are many experiments currently searching for sterile neutrinos, though no firm evidence has yet been detected. Furthermore, there could be more heavy neutrino flavours, which are just too massive to be produced in Z decays.

2.2. Neutrino Oscillations Evidence

Neutrino oscillations are now well established, with many experiments measuring various aspects of the phenomena in different regimes. This was not always the case, however. In this section, the early evidence for oscillations of all flavours, and potential steriles, is discussed for different neutrino sources.

2.2.1. Solar Neutrinos

Electron neutrinos are produced in the Sun by a number of different mechanisms. The largest flux comes from the nuclear fusion of four Hydrogen atoms into a Helium (the *pp* chain):

$$4p \rightarrow {}^4\text{He} + 2e^+ + 2\nu_e. \tag{2.6}$$
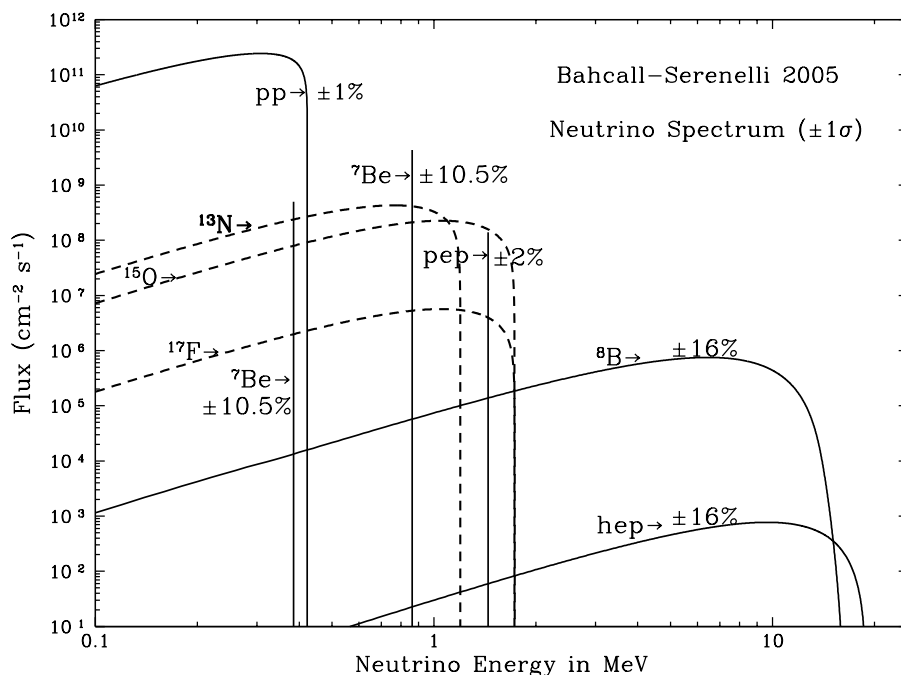
Figure 2.2.: Flux of solar neutrinos at Earth as a function of energy for different production mechanisms, according to Bahcall's solar model. Figure from [21].

However, the energy of the resulting neutrinos is below detection threshold for most experiments. Other mechanisms produce higher energy neutrinos but with a lower flux, as shown in Figure 2.2. Most solar neutrino experiments therefore measure the flux produced through boron decay:

$$_{5}^{8}\text{B} \rightarrow {}_{4}^{8}\text{Be} + e^+ + \nu_e. \tag{2.7}$$

2. Neutrino Physics

Ray Davis and John Bahcall devised an experiment to measure the flux of neutrinos produced via this mechanism in 1968 [22]. The design was similar to Davis' previous experiment, measuring the argon produced in the reaction in Equation 2.4 using a 380 m^3 underground tank of chlorine-based cleaning fluid (C_2Cl_4) in the Homestake Mine.

Bahcall calculated the predicted number of neutrinos from the Sun from the B decay chain using the Standard Solar Model, as well as the number of argon nuclei they would produce in the tank. However, consistently, only a third of the amount of neutrinos were detected than had been expected. The discrepancy between the predicted and measured number of neutrinos became known as The Solar Neutrino Problem.

Bruno Pontecorvo proposed a solution in 1968 [23], which involved neutrinos changing flavour as they propagated through space. This phenomenon, known as neutrino oscillations, was similar to the CKM [24] [25] formalism of quark mixing. Some of the electron neutrinos produced in the Sun would therefore change flavour before reaching the Earth, causing the flux deficits measured. However, this required neutrinos to have mass, which would be a significant modification to the Standard Model of particle physics. The success of the massless theory of the neutrino meant that this explanation initially did not gain much support.

Instead, initial efforts focused on modifying the solar model so that the prediction fit the data. The model depended on accurate knowledge of the pressure and temperature inside the Sun's core, and so it was thought that if the temperature was lower than had been assumed in the prediction, then fewer neutrinos would be expected to be produced. However, advances in helioseismology allowed improved measurements of the core temperature which were consistent with the original value.

No modification to the solar model itself could accommodate the measured fluxes either. The overall reduced flux required a lower core temperature, whereas the measured shape of the energy distribution of the neutrinos required a higher temperature. This is because the different nuclear processes producing different neutrino energies have different temperature dependencies. Modifying the solar model in any way would always result in at least one of these discrepancies increasing.

Pontecorvo's resolution was still not widely accepted though. It was also suspected that there was something wrong with the experimental setup, but later results were consistent with the deficit.

In 1989, the Kamiokande experiment [26] measured the solar neutrino flux using a large water Cherenkov detector. The electron recoil from elastic scattering was used to detect

electron neutrinos, but again measured a deficit to the predicted number.

The GALLEX [27] and SAGE [28] experiments in the early 1990s also used radiochemical detection, but with Gallium to measure:

$$\nu_e + {}^{71}\text{Ga} \rightarrow e^- + {}^{71}\text{Ge}. \tag{2.8}$$

This interaction had a lower energy threshold, allowing measurement of the flux from the *pp* chain. Both experiments confirmed the discrepancy between prediction and measurement.

Despite the *pp* chain being better understood than B-8 decay, the results of these experiments all relied upon the solar model and so were still not fully accepted.

The first solar-model independent measurement of the solar neutrino flux came from the SNO experiment in 2002 [29]. SNO detected both the ν_e flux via charged-current interactions:

$$\nu_e + d \rightarrow p + p + e^-, \tag{2.9}$$

and the total ν flux via neutral-current interactions:

$$\nu_\alpha + d \rightarrow p + n + \nu_\alpha, \tag{2.10}$$

and:

$$\nu_\alpha + e^- \rightarrow \nu_\alpha + e^-. \tag{2.11}$$

The electron neutrino flux measured was consistent with the previous results, but the total flavour independent flux was consistent with the predicted values, as shown in Figure 2.3. This was strong evidence that electron neutrinos were changing flavour into muon and tau neutrinos before they got to the Earth.

2.2.2. Atmospheric Neutrinos

Interactions between cosmic rays and nuclei in the Earth's atmosphere produce neutrinos via pion decay:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu; \ \mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \tag{2.12}$$

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu; \ \mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \tag{2.13}$$

Atmospheric neutrinos typically have a higher energy than solar neutrinos, with the flux
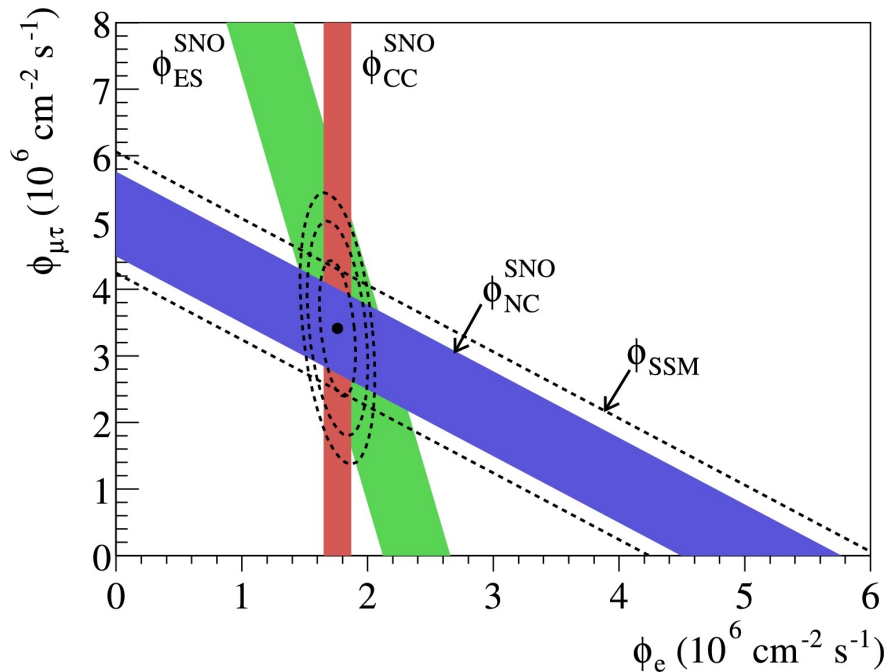
Figure 2.3.: The flux of $\mu + \tau$ vs e ^{8}B solar neutrinos measured at SNO. The dashed lines show the total flux predicted by the BP2000 solar model [30]. The blue, red, and green bands show the flux measured through NC, CC, and elastic scattering reactions respectively. The intersect of the bands is at the bestfit values for ϕ_e and $\phi_{\mu\tau}$, showing that the combined fluxes are consistent with the prediction. Figure from [29].

peaking between 1-10 GeV.

Reines [31] first measured the atmospheric neutrino flux at the Kolar Gold Fields mines in India in 1965. The measured ratio of ν_μ flux to ν_e flux was lower than had been predicted, a result which was confirmed by the IMB [32] and Kamiokande [33] experiments. The deficit was statistically convincing, but not significant enough to be considered compelling evidence of neutrino flavour change. These results became known as the Atmospheric Neutrino Anomaly.

The upgrade to the Kamiokande experiment, Super Kamiokande (SK), measured the flux of atmospheric ν_μ as a function of incoming angle in 1998. The data was split into upwards-going and downwards-going samples, allowing measurements at different distances from production. Upward-going neutrinos would have to travel through the Earth before reaching the detector, not just the distance from the atmosphere to the surface. In theory, since the neutrinos are produced isotropically in the atmosphere, there should be the same amount of upward-going and downward-going neutrinos. However, a large deficit in the ratio of ν_μ to ν_e was observed in the upward sample, as shown in Figure 2.4.

The dependence of the number of ν_μ observed on the distance travelled could be explained
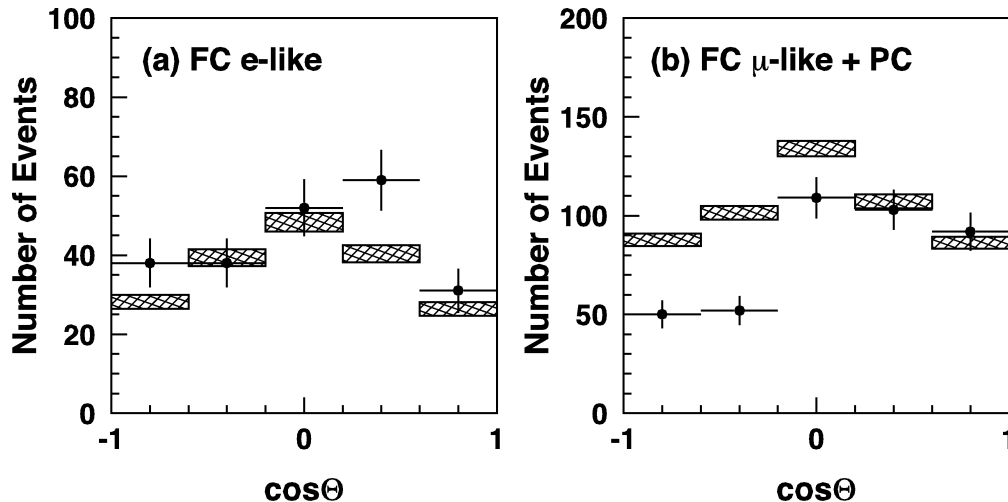
Figure 2.4.: The atmospheric neutrino flux as a function of angle from the first 414 days of Super-Kamiokande data. The boxes represent the prediction, the crosses represent the measured counts. Figure from [34].

in the context of neutrino flavour mixing. The upward-going neutrinos were changing flavour as they propagated the longer distance through the Earth to the detector. The SK result was strong evidence of ν_μ disappearance through oscillation.

2.2.3. Reactor Neutrinos

Reactor neutrinos have a similar energy flux to solar neutrinos, peaking between $1-10$ MeV. Like atmospheric neutrinos, experiments for detecting reactor neutrinos can measure oscillations across different baselines.

The KamLAND experiment [35] measured the $\bar{\nu}_\mu$ flux from 55 nuclear power reactors across Japan at a range of different distances, with a flux-weighted average baseline of 180 km. This gave sensitivity to Δm_{21}^2. The number of $\bar{\nu}_\mu$ detected was smaller than would be expected if neutrinos could not change flavour, and the measured probability of oscillation as a function of energy and distance was in agreement with Pontercorvo's theory of neutrino flavour mixing, as shown in Figure 2.5. This provided further evidence that neutrinos oscillate while propagating through space.

There are several reactor experiments, RENO [37], Daya Bay [38], Double Chooz [39], and its predecessor Chooz [40], with a shorter baseline of $\sim$1 km, sensitive to θ_{13}. Chooz initially set an upper limit on $\sin^2\theta_{13}$, before the others measured a large $\sin^2\theta_{13}$, which allows long baseline neutrino oscillation experiments to observe ν_e appearance. They have
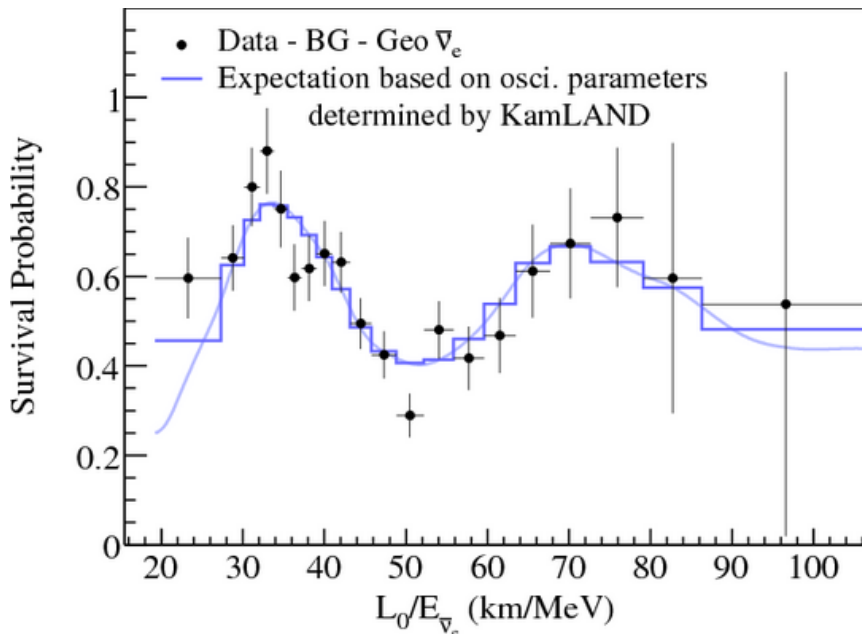
Figure 2.5.: Ratio of the KamLAND measured flux to expectation for no-oscillation as a function of L_0/E where L_0 is the flux-weighted average baseline ($L_0 = 180$ km). The blue histogram and curve show the expectation asuming PMNS oscillation parameters. Figure from [36].

all also measured an excess of neutrinos at $E_\nu \sim 5$ MeV compared to prediction, as shown in Figure 2.6. This could be due to poor flux modelling, or the existence of sterile neutrinos.

2.2.4. Accelerator Neutrinos

More recently, accelerators have been used to produce beams of neutrinos to study oscillations. This offers more control over the energies and baselines involved. Most long baseline ($\sim$100 km) accelerator experiments have two detectors, to measure the beam before and after oscillation. K2K [42], a long baseline (250 km) experiment in Japan, was the first to measure oscillations in such a way, using a beam of predominantly ν_μ. MINOS [43] also observed ν_μ disappearance consistent with the K2K result, with an even longer baseline (735 km). ν_e appearance in a ν_μ beam was discovered by the next generation experiments T2K and NOνA [44]. ν_τ appearance in a ν_μ beam was first observed by the OPERA [45] experiment in 2010, with a 730 km baseline.

Short baseline ($\sim$1 km) accelerator experiments are used to search for sterile neutrino oscillations, as well as measuring neutrino cross-sections. LSND [46] measured oscillation parameters in contradiction to other experiments in 2001, hinting at the existence of a sterile neutrino. However, subsequent experiments such as KARMEN [47] in 2001, ICARUS [48] in
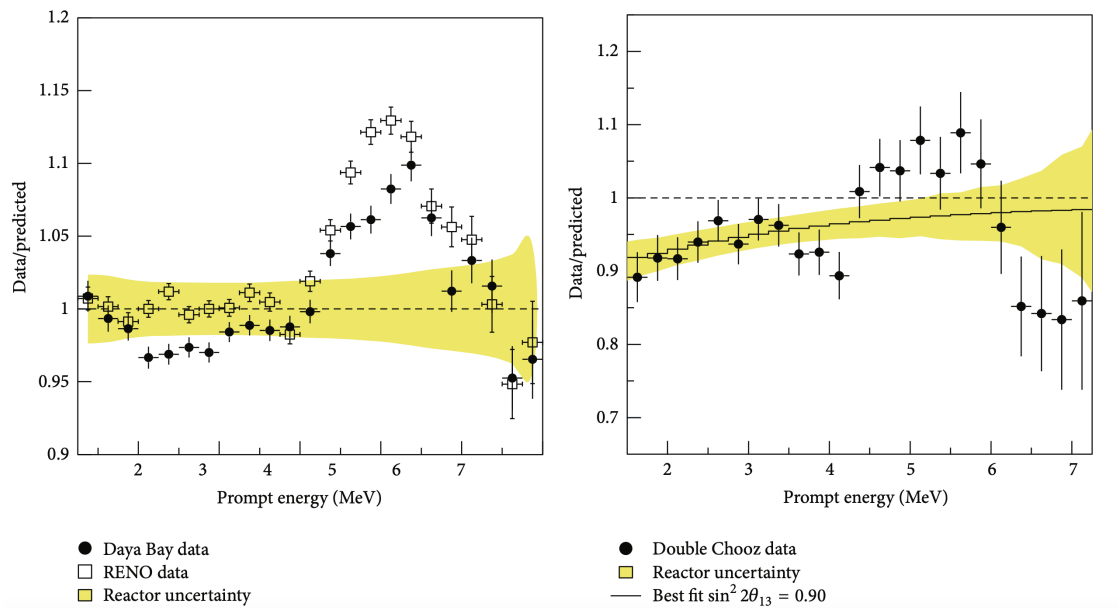
Figure 2.6.: Ratio of measured to expected $\bar{\nu}_e$ flux as a function of energy for Daya Bay and RENO (left) and Double Chooz (right). Figure from [41].

2004, and MiniBooNE [49][1] in 2007 did not agree with the LSND result. Currently there is no convincing evidence of the existence of any extra neutrino flavours.

2.3. Oscillation Theory

In this section, the theory of the mechanism behind neutrino oscillations and the interaction channels used to detect them are described. Firstly, the probabilities for a neutrino to oscillate are derived. The implications of the CP violating phase are then briefly discussed, before an overview of interactions within long baseline neutrino oscillation experiments is presented.

2.3.1. Oscillation Mechanism

Neutrinos are produced in charged current weak interactions, and therefore are produced in a definite flavour state: electron, muon or tau. This is defined by the massive lepton they are produced with, conserving lepton number in the interaction. The existence of flavour mixing is a direct consequence of neutrinos having mass. If neutrinos are not massless, there exists some set of mass eigenstates, each with definite mass m_i. There is no reason that

---

[1]Later MiniBooNE results were more compatible with LSND [50], though both are still considered controversial.

these should be equal to the flavour eigenstates, but as they both form a complete set the flavour states are each a linear combination of the mass states:

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i}^* |\nu_i\rangle. \tag{2.14}$$

Similarly, the mass states are each superpositions of the flavour states:

$$|\nu_i\rangle = \sum_\alpha U_{\alpha i} |\nu_\alpha\rangle, \tag{2.15}$$

where Roman subscripts are used to denote mass states, and Greek subscripts are used to refer to flavour states.

The lepton mixing matrix, U, relates the two sets of states in the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) formalism of neutrino oscillations [51]. If this were the identity matrix, the sets of eigenstates would be identical and neutrinos would not change flavour, but the experimental evidence described in Section 2.2 shows this is not the case. The PMNS matrix is often expressed in the form:

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3}. \end{pmatrix} \tag{2.16}$$

Each element, $U_{\alpha i}$, corresponds to the amplitude of the mass eigenstate i within the flavour eigenstate α. The flavour contents are shown in Figure 2.7, according to current best measurements. The left hand side is for the normal mass hierarchy, where $m_3^2 > m_2^2$, whereas the right handside is for the inverted mass hierarchy, where $m_2^2 > m_3^2$. It is known that $m_2^2 > m_1^2$ from measurements of neutrino oscillations in matter in the Sun, but the nature of the mass hierarchy is not known beyond this. As the atmospheric and reactor mass-splittings are measured through neutrino oscillations in a near-vacuum, which only depend on the absolute values, the signs of Δm_{32}^2 and Δm_{31}^2 are not yet known.

The propagation of the mass eigenstates can be described by solutions to the plane wave equation:

$$|\nu_i(t)\rangle = e^{-i(E_i t - \bar{p}_i \cdot \bar{x}_i)} |\nu_i(0)\rangle, \tag{2.17}$$

where t is time of propagation, E_i is the energy of the mass eigenstate i, $\bar{p}_i$ is the momentum, and $\bar{x}_i$ is the position. $|\nu_i(0)\rangle$ is the initial state of the mass eigenstate.

In the lab frame:
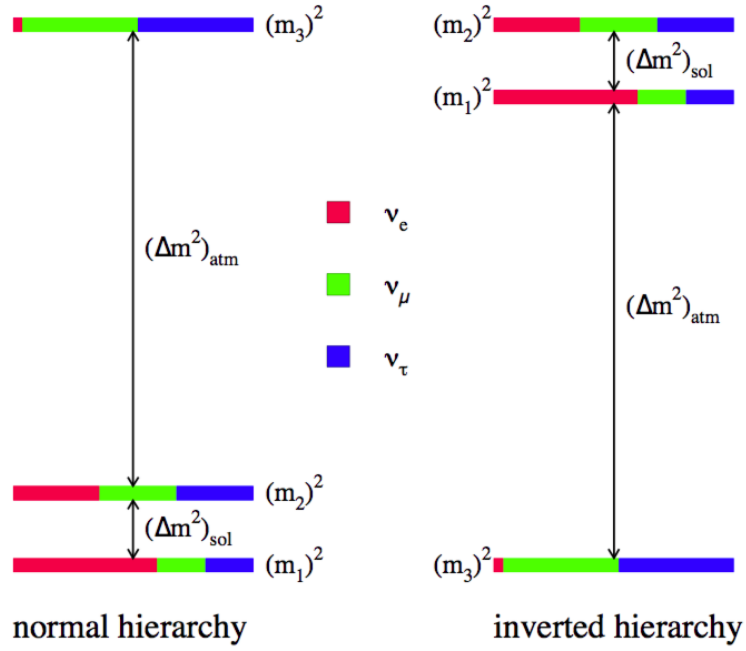
Figure 2.7.: The flavour content and mass differences of the three mass eigenstates, for both the normal and inverted hierarchys. Figure from [52].

$$p_i \cdot x_i = |\mathbf{p}_i| L, \qquad (2.18)$$

where L is the distance travelled.

In the ultrarelativistic limit: $|\mathbf{p}_i| = p_i \gg m_i$, so:

$$p_i \cdot x_i = EL, \qquad (2.19)$$

and the energy can be approximated as:

$$E_i = \sqrt{p_i^2 + m_i^2} \approx p_i + \frac{m_i^2}{2p_i} \approx E + \frac{m_i^2}{2E}, \qquad (2.20)$$

where E is the total energy of the neutrino. Therefore, taking the natural units, $c = 1$, $t = L$:

$$E_i t \approx L(\frac{m_i^2}{2E} + E), \qquad (2.21)$$

and so Equation 2.17 can be written as:

$$|\nu_i(t)\rangle = e^{-i\frac{m_i^2 L}{2E}} |\nu_i(0)\rangle . \qquad (2.22)$$

The different mass eigenstates therefore propagate differently as they have different masses. Equation 2.15 can be interpreted as the probability of a mass eigenstate i interaction producing a charged lepton α, as the fraction of flavour α in eigenstate ν_i can be calculated as $|U_{\alpha i}^*|$. Although produced in a definite flavour state, as a neutrino travels away from its source the mass states become out of phase. The resulting interference means that the neutrino's wavefunction evolves to contain components from all three flavour states. The amplitude for a neutrino initially having flavour α, being detected as having flavour β after propagating a certain distance is:

$$\mathcal{A}(\nu_\alpha \rightarrow \nu_\beta) = \sum_i \sum_j \langle \nu_j | U_{\beta j} e^{-i\frac{m_i^2}{2E}} U_{\alpha i}^* | \nu_i \rangle . \tag{2.23}$$

As charged leptons in the standard model only couple to neutrinos of the same flavour:

$$\begin{aligned}
\delta_{\alpha\beta} &= \langle \nu_\alpha | \nu_\beta \rangle \\
&= \langle \sum_i U_{\alpha i}^* \nu_i | \sum_j U_{\beta j}^* \nu_j \rangle \\
&= \sum_{i,j} U_{\alpha i} U_{\beta j}^* \langle \nu_i | \nu_j \rangle \\
&= \sum_i U_{\alpha i} U_{\beta j}^*,
\end{aligned} \tag{2.24}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta. The probability of a neutrino oscillating from flavour α to β is then found to be:

$$\begin{aligned}
P(\nu_\alpha \rightarrow \nu_\beta) &= |\mathcal{A}(\nu_\alpha \rightarrow \nu_\beta)|^2 \\
&= \delta_{\alpha\beta} - 4\Sigma_{i>j} \mathcal{R}(U_{\alpha i}^\star U_{\beta i} U_{\alpha j} U_{\beta j}^\star) \sin^2(\frac{\Delta m_{ij}^2 L}{4E}) \\
&\quad + (-)2\Sigma_{i>j} \mathcal{I}(U_{\alpha i}^\star U_{\beta i} U_{\alpha j} U_{\beta j}^\star) \sin^2(\frac{\Delta m_{ij}^2 L}{4E}),
\end{aligned} \tag{2.25}$$

where $\Delta m_{i,j}^2$ is the difference in mass of mass eigenstates i and j, and the negative sign is for anti-neutrinos.

The PMNS matrix is often parametrised as three matrices in terms of three mixing angles, θ_{12}, θ_{13}, θ_{23}, and a CP-violating phase, δ_{CP}:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\mathrm{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\mathrm{CP}}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{\mathrm{CP}}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{\mathrm{CP}}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{\mathrm{CP}}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{\mathrm{CP}}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{\mathrm{CP}}} & c_{23}c_{13} \end{pmatrix},$$

(2.26)

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. There are two extra complex phases which are only non-zero if neutrinos are Majorana particles [53]. Even if this is the case, they would be extremely difficult to measure experimentally, and they do not affect oscillation probabilities, so are not considered here.

This form of the unitary matrix is often used as it makes it easier to interpret the oscillation parameters, as they are separated by the different types of experiment in which they can be measured.

The first matrix contains terms only in θ_{23}, the mixing angle involved in most atmospheric neutrino oscillations. If ν_e are neglected, and atmospheric oscillations are considered as a two flavour process $\nu_\mu \to \nu_\tau$, then $\theta_{atm} \approx \theta_{23}$.

The second matrix contains terms only in θ_{13}, the mixing angle involved in most reactor neutrino oscillations, and the CP-violating phase, δ_{CP}. If $\delta_{\mathrm{CP}} \neq 0$, neutrinos oscillate differently to anti-neutrinos: $\nu_\alpha \to \nu_\beta \neq \bar{\nu}_\alpha \to \bar{\nu}_\beta$. However, when multiplied out δ_{CP} only appears in terms with other angles, and so can only be measured if these are all non-zero. This parametrisation of U emphasises the dependence on θ_{13}, as it is the smallest angle and was the last to be measured.

The third matrix contains terms only in θ_{12}, the mixing angle involved in most solar neutrino oscillations. If ν_τ are neglected, and solar oscillations are considered as a two flavour process $\nu_e \to \nu_\mu$, then $\theta_{sol} \approx \theta_{12}$.

This has all assumed the neutrinos have been propagating through a vacuum. In most circumstances, the neutrinos actually travel through matter, and so interactions with the medium must be accounted for. In the Earth, there are two types of interactions which can occur. These are charged current scattering of ν_e off an electron, and neutral current scattering of any flavour of neutrino off an electron, neutron, or proton. The neutral current scattering is flavour independent and so does not affect oscillation probabilities. The charged couple scattering however, only occurs for ν_e, and so needs to be considered in oscillation calculations. The effect increases with distance through matter travelled, and the electron

density of the medium. The full treatment of how this process affects oscillations is known as the MSW effect [54], but is beyond the scope of this thesis.

As the Earth is made predominantly of matter rather than anti-matter, the MSW effect has a different impact for neutrinos than it does for anti-neutrinos. This produces an effect that mimics CP violation, causing $\nu_\alpha \to \nu_\beta \neq \bar{\nu}_\alpha \to \bar{\nu}_\beta$, and so careful treatment is required for measurements of $\delta_{\rm CP}$. At T2K, the average matter density between the near and far detectors of 2.6 g/cm^3 [55] has very little effect on the oscillation probabilities, but the matter effects are fully taken into account nonetheless.

Putting all this together, Equation 2.25 can therefore be written for the T2K detection channels, ν_μ ($\bar{\nu}_\mu$) disappearance and ν_e ($\bar{\nu}_e$) appearance, as:

$$P(\nu_\mu \to \nu_\mu) \approx 1 - 4\cos^2\theta_{13}\sin^2\theta_{23}\,(1 - \cos^2\theta_{13}\sin^2\theta_{23})\sin^2\frac{\Delta m_{32}^2 L}{4E}, \qquad (2.27)$$

and

$$
\begin{aligned}
P(\nu_\mu \to \nu_e) \approx{} & \sin^2\theta_{23}\sin^2 2\theta_{13}\sin^2\frac{\Delta m_{31}^2 L}{4E} \\
& + \sin 2\theta_{23}\sin 2\theta_{23}\sin 2\theta_{13}\cos\theta_{13}\sin\frac{\Delta m_{21}^2 L}{4E}\sin\frac{\Delta m_{31}^2 L}{4E} \\
& \times (\cos\frac{\Delta m_{32}^2 L}{4E}\cos\delta_{\rm CP} - (+)\sin\frac{\Delta m_{32}^2 L}{4E}\sin\delta_{\rm CP}).
\end{aligned}
\qquad (2.28)
$$

Here, the positive sign in brackets is for anti-neutrinos. The neutrino and anti-neutrino disappearance probabilities in Equation 2.27 are identical, whereas the appearance probabilities in Equation 2.28 have opposing signs for the third term, allowing CP violation to be measured if $\delta_{\rm CP} \neq 0$.

In these probabilities, the matter effect terms have been neglected as they are small, and the solar oscillation terms (involving Δm_{21}^2) have been neglected as T2K is not sensitive to these processes. They are however, taken into account in all oscillation calculations in T2K analyses.

Equations 2.27 and 2.28 show that the probability for oscillation is dependent on the mass splittings rather than the absolute values of the masses. This is perhaps unsurprising, as it is the differences in the mass eigenstates that cause oscillations. Neutrino oscillation experiments can therefore only measure the difference between the masses, and not the values themselves. However, if all neutrino states were massless the mass differences would also be zero. The evidence of neutrino oscillations outlined in Section 2.1 is therefore evidence that at least two of the three mass eigenstates have non-zero mass.

As well as the oscillation parameters, the probability is also dependent on the experimental

parameters L and E. Therefore, by measuring the probability for oscillation at a known baseline and energy, the oscillation parameters can be determined. Using non-natural units:

$$\Delta m_{ij}\frac{L}{4E} = \Delta m_{ij}^2 \text{ (eV}^2)\frac{1.27L \text{ (km)}}{E \text{ (GeV)}},$$ (2.29)

meaning with a 295 km baseline and 0.6 GeV beam, T2K is sensitive to $\Delta m_{ij} \gtrsim 10^{-3}\text{eV}^2$.

This has all assumed that there are only three flavours of neutrino, in line with the measured values of the Z^0 decay width [18, 19]. However, there could exist sterile neutrinos, briefly discussed at the end of Section 2.1. In this case, the mixing matrix would need to be modified to accommodate oscillations involving these new neutrinos, and further mass states may be required.

2.3.2. CP Violation

In 1967, Sakharov proposed three conditions that would be required for an interaction to produce different amounts of matter and anti-matter [1]. These were:

- **Baryon Number Violation:** For a process to produce an unequal number of baryons and anti-baryons, there must be a net non-zero baryon number.

- **C and CP Violation:** Processes producing a net amount of baryons must not be balanced by processes producing a net amount of anti-baryons.

- **Departure from Thermal Equilibrium:** In thermal equilibrium, CPT symmetry would cause the time reversal of baryon producing interactions to balance the net amount.

The abundance of matter over anti-matter observed in the Universe therefore suggests that CP violating interactions must have occurred. These have been observed in the quark sector [2], but not by a sufficient amount to completely explain the asymmetry.

Equation 2.28 shows that neutrino oscillations could violate CP symmetry, if δ_{CP} is non-zero. Theories of leptogenesis propose that an abundance of leptons over anti-leptons in the Universe can be converted to a net number of baryons through sphaleron processes [56].

Leptogenesis models [57] [58] suggest that CP violation in the neutrino sector could be sufficient to explain the baryon asymmetry in the Universe.

2.3.3. Neutrino Interactions in Long Baseline Oscillation Experiments

To detect neutrino oscillations, it is inherently vital that the flavour composition of a flux of neutrinos is carefully analysed. As the the number of neutrinos measured is a convolution of the flux, cross-section, detector efficiency, and oscillation probability, it is important each of these components is well understood to make an accurate measurement. Due to their low interaction rate and subsequent lack of data, neutrino cross-sections contribute significantly to the total uncertainty, and this is only going to become more pronounced with increased statistical power for oscillation experiments. It is therefore essential to study neutrino interactions for precision measurements of oscillations to be made.

Equation 2.28 tells us that the probability of oscillation is dependent on the mixing angles, mass splittings, distance travelled, and energy of the neutrino. The baseline length and energy are experimental variables which can be chosen. The length is a fixed, known quantity for accelerator based neutrino experiments, but it is not feasible to produce a mono-energetic source of neutrinos, and so the energy spectrum for each flavour is used. The probability of oscillation is measured from the number of neutrinos of each flavour before and after oscillation. The oscillation parameters can then be calculated from the measured ($P(\nu_\alpha \to \nu_\beta)$) and experimental ($L$ and E) variables.

As neutrinos cannot be detected directly, the number of neutrinos is inferred from the secondary particles produced when they interact. This requires a thorough understanding of all the interactions that could take place. If the cross-section for a flavour is not known accurately, the number of neutrinos of that flavour, and hence the oscillation probability, will be determined incorrectly.

The energy spectrum is reconstructed from a set of observables in a detector. To do this, the kinematics of the event must be well understood, otherwise the energy spectrum will be distorted, and hence the oscillation parameters incorrectly calculated. Understanding these kinematics is highly dependent on accurately identifying the type of interaction.

As neutrinos are electrically neutral and colourless, they can only interact via the weak force. We divide these interactions into two types: Charged Current (CC), mediated by the $W^\pm$ boson, and Neutral Current (NC), mediated by the Z^0 boson.

The Feynman diagram for an example of an NC interaction, NC electron elastic scattering, is shown on the left hand side of Figure 2.8. A neutrino interacts with an electron, causing it to recoil, which can be detected. However, as this process could occur for any flavour of neutrino, detecting it is not useful for determining the flavour composition of a neutrino flux.
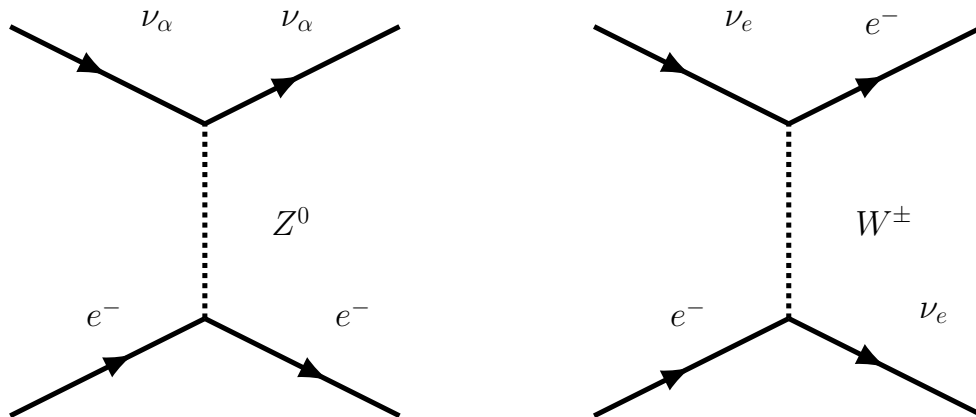
Figure 2.8.: Feynman diagrams for NC electron elastic scattering (left) and CC elastic scattering (right). The NC interaction can occur for any neutrino flavour $\alpha = e, \mu, \tau$, whereas the CC interaction can only occur for an incoming ν_e.

For this reason, NC interactions are not useful for oscillation measurements.

The Feynman diagram for CC elastic scattering is shown on the right hand side of Figure 2.8. Here, the only neutrino flavour this is possible for is ν_e, and so a measurement of this process does tell us about the flavour composition of the neutrino flux. These events are therefore selected to be part of oscillation analyses. However, the experimental signature is identical to that of NC scattering, which could be any flavour. In this way, NC electron elastic scattering forms an irreducible background to CC elastic scattering events. These arguments can be extended to all NC and CC interactions: NC interactions produce charged leptons with flavour uncorrelated to the incoming neutrino, whereas CC interactions produce charged leptons which match the flavour of the incoming neutrino, and so can be used to determine if a neutrino has oscillated. The NC event rate is therefore simulated in the MC and used in the predicted event rates.

At T2K energies, there are three main types of interaction, but with a range of complex nuclear effects involved. At low energies, <1 GeV, quasi-elastic (QE) -like scattering dominates, but nucleon-nucleon correlations and the distribution of nucleons within a nucleus need to be accounted for. QE events are called as such as the kinematics are similar to electron scattering, as the transfer of energy, Q^2, is small. At intermediate energies, around ~1 GeV, resonance production (RES) becomes important, where the nucleon is excited into a baryonic resonance, before decaying. At higher energies, deep inelastic scattering (DIS) dominates, with parton distribution functions becoming important. Feynman diagrams for these three processes are shown in Figure 2.9, and their cross-sections for neutrinos (rather
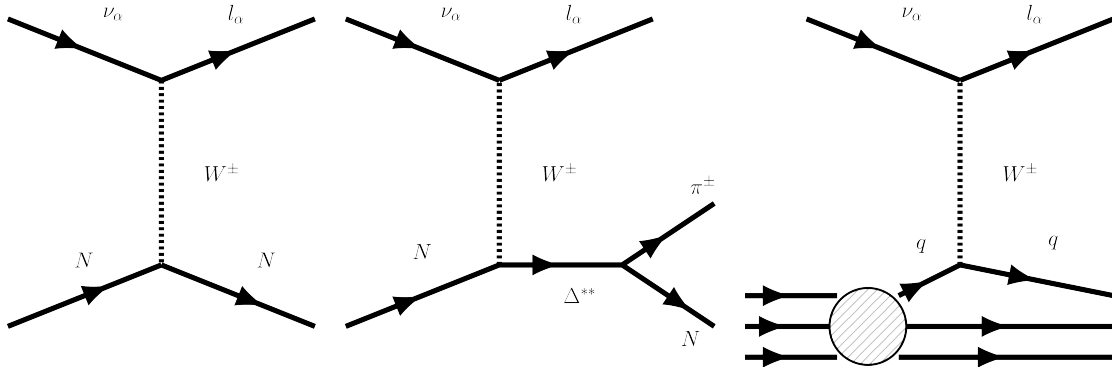
Figure 2.9.: Feynman diagram for CCQE (left), CC RES (center), and CC DIS (right) interactions.

than anti-neutrinos) as a function of energy are shown in Figure 2.10. The transitions between these energy regimes are currently poorly modelled [59].

The T2K neutrino beam peaks at 0.6 GeV, and so CCQE events dominate. CCQE interactions are well understood, and well constrained by data. Furthermore, as the they are two-body processes, and assuming the initial nucleon is at rest, the incoming neutrino energy, $E_{\rm rec}$, can be reconstructed from the final state lepton's momentum and angle:

$$E_{\rm rec} = \frac{m_p^2 - (m_n - E_{\rm b})^2 + m_\mu^2 + 2(m_n - E_{\rm b})E_\mu}{2(m_n - E_{\rm b} - E_\mu + p_\mu \cos\theta_\mu)}, \qquad (2.30)$$

where m_p is the mass of the proton, m_n is the mass of the neutron, $E_{\rm b}$ is the binding energy of the neutron, and m_μ, E_μ, p_μ and $\cos\theta_\mu$ are the mass, energy, momentum, and angle of the final state lepton.

Equation 2.30 shows that accurate neutrino energy reconstruction depends on having an accurate value of the binding energy. The binding energy itself is the energy required for the incoming neutrino to release a nucleon from the target nucleus. This manifests itself as missing energy in the interaction. If an incorrect value for $E_{\rm b}$ is used in the reconstruction, an incorrect value of $E_{\rm rec}$ will be obtained, biasing the measurement of oscillation parameters. As the average binding energy per nucleon is not well constrained by external data, this is one of the dominant systematic uncertainties in previous T2K oscillation analyses.

The assumption that the initial state nucleon is at rest, is also not strictly true, particularly for low Q^2 events. Nucleons are constantly moving around, and the initial momentum distribution within a nucleus is not well modelled. This causes an uncertainty in the reconstruction of the neutrino energy.
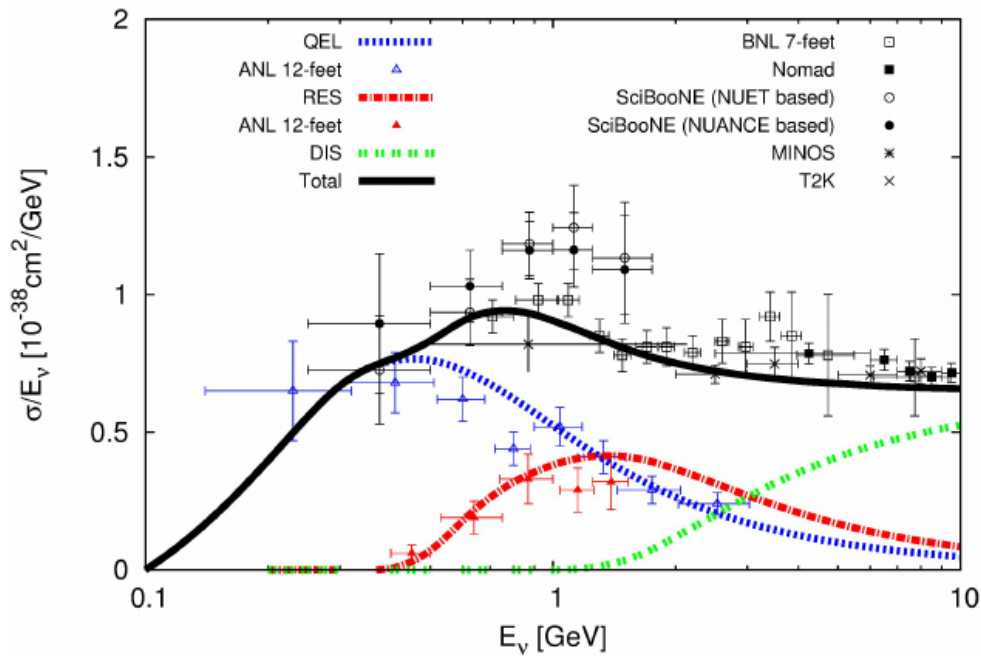
Figure 2.10.: Breakdown of the CC ν_μ cross-section for QE, RES, and DIS interactions, along with data from various experiments. Figure from [60].

As described in Section 5.2, T2K samples data by event topology as seen by the detector, whereby CCQE events are selected as CC 0π. There is a large background of CCQE-like events, which are also selected as CC 0π. These can be events with a pion(s) in the final state which is not detected, misreconstructed, or is absorbed before detection. Another contribution to this background comes from events where a neutrino interacts with a correlated pair of nucleons, known as 2p2h (two-particle-two-hole). An example 2p2h Feynman diagram is shown in Figure 2.11. This also leaves a 0π final state, meaning it is vitally important these interactions are well modelled. Equation 2.30 only applies for true CCQE events, so if 2p2h events are mistaken for CCQE, the neutrino energy will be incorrectly reconstructed, biasing the oscillation results.

The most dominant interaction type producing a 1π final state is CC RES. These interactions are not as well understood as CCQE, and as they are three-body processes the neutrino energy reconstruction is not as simple. If the pion produced is below detection threshold, the event will be classed as CC 0π, forming an irreducible background to CCQE events.

Coherent (COH) π production events, where the incoming neutrino interacts with an entire nucleus leaving it in the same final state as it was initially, also produce a 1π final state, as shown in Figure 2.12. Events on nuclei are less understood than events on nucleons, but a target of purely free neutrons would be impossible to construct in practice, and interac-
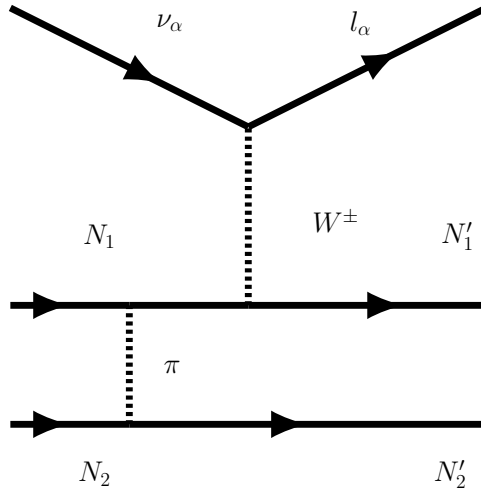
Figure 2.11.: Feynman diagram for a 2p2h interaction. As they both produce 0π final states, these events form a background to CCQE interactions.
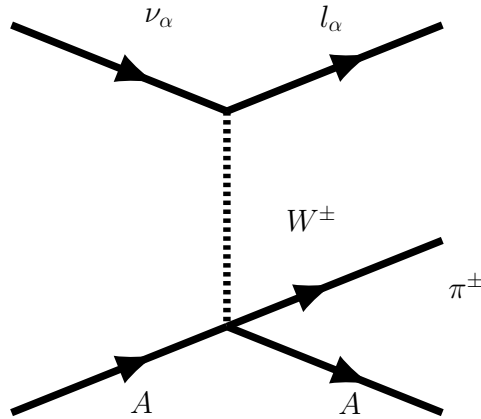
Figure 2.12.: Feynman diagram for a CC COH π production interaction.

tions on Hydrogen nuclei (purely protons) are only available to anti-neutrinos, which have a lower cross-section[2]. Nuclei targets are therefore used, and so coherent scattering events are inevitable.

As the energy increases, a larger Q^2 becomes available, and so inelastic events become accessible. In CC DIS interactions, the incoming neutrino scatters off an individual constituent quark rather than a nucleon or nucleus. As the hit quark recoils, the nucleon containing it fragments, producing multiple πs. These events are a significant fraction of the CC $N\pi$ (N>1) samples.

Final state interactions (FSI) can cause events to be miss-classified. The particles produced

---

[2]In fact, neutrino events on free electrons are even better understood than on free nucleons. However, the cross-sections for these interactions are much lower, and constructing a pure electron target is as infeasible as for neutrons.

in the event can interact with other nucleons as they propagate out of the nucleus. This can alter their momentum and direction significantly. The multiplicity of outgoing particles can also be changed, as they can be absorbed before leaving the nucleus, or produce more particles through collisions. FSI can therefore cause the particles leaving the nucleus to be different to those produced in the original interaction.

Although FSI doesn't generally affect the final state lepton, it can still have a big impact on oscillation results. For example, if the π produced in a CC RES event is absorbed before being detected, the event could be classified as CCQE. Equation 2.30 would then not be valid, and E_ν would not be reconstructed correctly.

2.4. Current Experimental Status

The current generation long baseline accelerator experiments, T2K and NOνA, are measuring the accelerator oscillation parameters to greater precision, as well as trying to determine $\delta_{\rm CP}$ and the mass hierarchy. The short baseline neutrino oscillation experiments, such as MicroBooNE [61] and ICARUS, are searching for sterile neutrinos, and trying to resolve tensions with the LSND result. These form the Short Baseline Neutrino (SBN) program at Fermilab [62], along with the SBN near and far detectors, due to come online in 2020. The main aim of the SBN program is to unambiguously confirm or disprove previous anomalous measurements, as well as performing detailed studies of neutrino-nucleus interactions at the GeV energy scale.

The solar oscillation parameters are well constrained, but Borexino [63], is measuring neutrinos produced via the ^{8}B, ^{7}Be, *pep*, and *pp* processes. SNO+ [64] will aim to confirm these results as well as performing detailed studies of the MSW effect.

The current atmospheric neutrino oscillation experiments, such as IceCube [65], ANTARES [66], and SK, are measuring the atmospheric parameters to increased precision, as well as studying specific zenith angles, and therefore baselines, to investigate the MSW effect.

The current reactor neutrino experiments, such as RENO, Double Chooz, Daya Bay, and KamLAND, are measuring $\sin\theta_{13}$ with increasing precision. They are also searching for steriles and trying to resolve tensions between the measured and predicted flux at $E_\nu \sim 5$ MeV. DANSS [67], NEOS [68], PROSPECT [69], STEREO [70] and SoLiδ [71] are all very short baseline reactor experiments (L $\sim$10 m), and have each confirmed the 5 MeV excess. However, currently none have found significant evidence of a sterile neutrino.

Many solar, atmospheric, reactor, and accelerator neutrino experiments have been per-

Parameter	Mass Hierarchy	$\sin\theta_{23}$ Octant	Value
Δm_{21}^2	-	-	(7.5 ± 0.18) x 10^{-5} eV^2
$\sin^2\theta_{12}$	-	-	$0.307^{+0.013}_{-0.012}$
Δm_{32}^2	NH	-	(2.444 ± 0.034) x 10^{-3} eV^2
	IH	-	(-2.55 ± 0.04) x 10^{-3} eV^2
$\sin^2\theta_{23}$	NH	Lower	$0.542^{+0.019}_{-0.022}$
		Upper	$0.512^{+0.019}_{-0.022}$
	IH	-	$0.536^{+0.023}_{-0028}$
$\sin^2\theta_{13}$	-	-	2.18 ± 0.07 x 10^{-2}
δ_{CP}	NH	Upper	$1.37^{+0.18}_{-0.17}$ rad

Table 2.1.: Current best measurements of oscillation parameters. Values from [72].

formed over many years, with the aim of measuring the key oscillation parameters to an increasing level of accuracy and precision. The Particle Data Group determine the current world-leading measurements of each of the parameters [72]. These are shown in Table 2.1.

The best measurement of the solar parameters are from a global fit of solar and KamLAND data, using a constraint on θ_{13} from reactor and accelerator experiments for Δm_{21}^2.

The atmospheric parameters are best measured using a fit of T2K, SK, NOvA, MINOS, and IceCube data. These depend on the mass hierarchy, and the θ_{23} octant.

A fit of data by reactor experiments Daya Bay, RENO, and Double Chooz is the most precise measurement of θ_{13}. The best measurement of δ_{CP} comes from a fit of T2K, SK, and NOvA data. The T2K and SK results assume the normal hierarchy, and use a constraint on θ_{13} from reactor experiments. The NOvA result assumes the normal hierarchy and the upper θ_{23} octant ($\sin^2\theta_{23} > 0.5$).

2.4.1. Open Questions

For several decades now, neutrino experiments have been uncovering more and more information about neutrinos, and their interactions. However, there are still several fundamental unknowns.

As mentioned in Section 2.4, current data suggests that δ_{CP} is non-zero. However, more data is needed to unambiguously conclude this, and if so, what its value actually is, and if it corresponds to sufficient CP violation to account for the matter – anti-matter asymmetry

we see in the Universe today. The next generation accelerator experiments, DUNE and HK, aim to obtain precise measurements of δ_{CP}.

The ordering of mass states is also still unknown. Although we know $m_2^2 > m_1^2$, we don't know if m_3^2 is higher or lower than the other two. Being such a fundamental property of neutrinos, not knowing the mass hierarchy limits our ability to measure many aspects of neutrino physics. Determining the hierarchy would allow more precise measurements of the other oscillation parameters, as well as having a big impact on our understanding of Supernovae. This is one of the aims of future accelerator experiments DUNE and HK, future atmospheric experiments SNO+ and IceCube, as well as the future reactor experiment JUNO [73].

Furthermore, the absolute scale of the masses are not known. Although we can measure the square of the mass differences, this tells us nothing about the actual values of the masses.

As well as this, the nature of the neutrino masses is unknown. Neutrinos could have Dirac masses, like all other fermions, or could have Majorana masses, meaning they are their own anti-particle. Knowing if neutrinos are Majorana or Dirac particles is very important for understanding the origin of small neutrino masses. Future neutrinoless double beta decay experiments, MAJORANA [74], nEXO [75], KamLAND-Zen [76], and SNO+ will be able to test the Majorana nature of neutrinos if the mass hierarchy is determined.

Finally, it is not known if the three flavours of neutrino that have already been detected are the only that exist. If there are more, it is not known how many more there are, and how their masses compare to the active flavours. If proved to exist, sterile neutrinos could even be the elusive Dark Matter we infer exists in the Universe, but don't currently know what it is [77]. The SNB program will improve constraints on sterile neutrinos.

2.5. Summary

This chapter has presented an overview of the history and theory of neutrino physics. Since the neutrino was first theorised in 1930, and first detected in 1956, many experiments have measured many of its properties, including the 2015 Nobel prize winning detection of neutrino oscillations. Three flavours have been discovered, and measurements of the Z decay width, along with cosmological constraints, show that any other flavours would have to be sterile.

Neutrino oscillations have been detected from a variety of different sources: solar, atmospheric, accelerator, and reactor. The parameters governing these oscillations are being measured with increasing precision, and T2K is now trying to determine the value of the

CP violating phase, which could explain the baryon asymmetry we observe in the Universe today.

The probability of a neutrino oscillating from one flavour to another depends on the distance travelled and the neutrino energy. The T2K beam energy and baseline is tuned to maximise the probability of oscillation. At T2K energies, there are three main CC interaction types: CCQE, CC Res, and CC DIS. CCQE is the best understood of these, and is the target signal interaction mode. However, the other CC modes, along with NC events, form a background to these interactions. It is therefore important to understand all their cross-sections, and final state interactions, to make accurate oscillation measurements.

There are still many unknowns within neutrino physics. The mixing angles and mass differences are being measured with increasing precision, but the absolute masses and mass hierarchy have not yet been determined. As well as this, anomalies in low energy flux measurements hint at the possibility of the existence of sterile neutrinos, but there is currently no concrete evidence that this is the cause yet.

3. The T2K Experiment
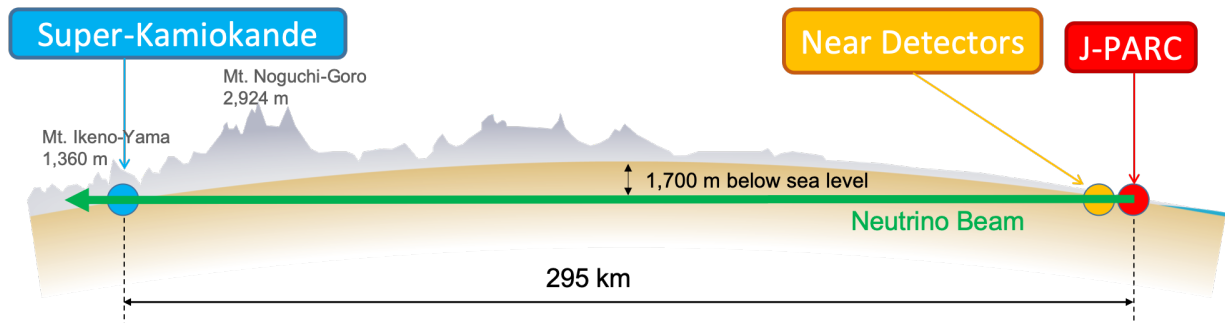
Figure 3.1.: The T2K experiment: Neutrinos are produced on the east coast of Japan, and are measured 280 m upstream by the near detectors, and 295 km away at the far detector, SK. Figure from [78].

T2K is a long baseline neutrino oscillation experiment in Japan, which has been taking data since 2010. It was designed to precisely measure $\sin^2\theta_{13}$ by observing the appearance of (anti)-electron neutrinos in a (anti)-muon neutrino beam, and $\sin^2 2\theta_{23}$ and Δm_{23}^2 by observing muon neutrino disappearance. In 2014, T2K first detected electron-neutrino appearance [79]. NOνA found consistent results in 2016 [80]. Since data taking began, T2K has made world leading measurements of $\sin^2\theta_{13}$, $\sin^2 2\theta_{23}$ and Δm_{23}^2. The current main aims of T2K focus on excluding possible values of δ_{CP} and determining the neutrino mass ordering.

In this chapter, the three main components of the experiment, the neutrino production beamline, the near detector suite, and the far detector, SK, are introduced. A schematic diagram of this setup is shown in Figure 3.1. The beam is produced on the east coast of Japan, at the Japanese Proton Accelerator Research Complex (J-PARC), and is described in Section 3.1. The near detector (ND) suite is located 280 m upstream of the beam source, where the ND280 and INGRID detectors measure the flux, cross-section, and direction of the beam. This is described in Section 3.2. Then, 295 km away in the west of Japan, the Super-Kamiokande water Cerenkov detector measures interactions after oscillation, and is described in Section 3.3.

At the ND suite, there are a number of different neutrino detectors. T2K was the first

experiment to use the off-axis technique, whereby SK and ND280 are located 2.5° away from the axis of the beam. This focuses the energy distribution of the beam into a narrower peak, as described in Section 3.2. This technique requires an accurate measurement of the beam direction, which is performed by a second, on-axis detector, INGRID.

As well as ND280 and INGRID, there are also detectors at the ND suite that are not directly used in the T2K oscillation analysis. The NINJA [81] experiment is designed to accurately measure the neutrino cross-section on water using nuclear emulsion techniques. The WAGASCI [82] and Baby MIND [83] experiments are designed to measure and constrain non-cancelling systematic uncertainties arising from the fact that ND280 and SK have different target materials. These will not be discussed any further, as data from these experiments is not used in this analysis.

3.1. Beamline

The neutrino beamline at J-PARC [84] was newly constructed for the T2K experiment, and is fed by a system of three accelerator facilities: a linear accelerator (LINAC), a rapid cycling proton synchrotron (RCS), and the main ring synchrotron (MR). The accelerator complex is shown in Figure 3.2. H^- ions are first accelerated to 400 MeV by the LINAC, before being converted to H^+ ions by charge stripping foils as they are injected into the RCS. Here, they are further accelerated to 3 GeV in 25 Hz cycles, with two bunches per cycle.

Approximately 5% of the bunches are injected into the MR, where they are accelerated to 30 GeV. The rest are supplied to other experiments at J-PARC. The MR holds eight bunches each of $\sim 3 \times 10^{14}$ protons. The beam is extracted from the MR by a set of five kicker magnets, which deflect the beam toward the neutrino beamline. The extraction happens in spills of eight proton bunches, separated by 560 ns. The spill cycle is ~ 0.5 Hz, and a single spill has a total duration of approximately 5 μs. A GPS system is used to link the timing of the information of the spills to the neutrino detector triggers, allowing better discrimination of backgrounds from the beam signal.

3.1.1. Neutrino Beamline

The neutrino beamline consists of two sections, the primary and secondary beamlines. These are shown in Figure 3.3. In the primary beamline, the proton beam is deflected towards the secondary beamline. It is vital the profile, position, intensity, and beam loss are known to be able to produce the stable and consistent neutrino beam required. Beam monitors in the
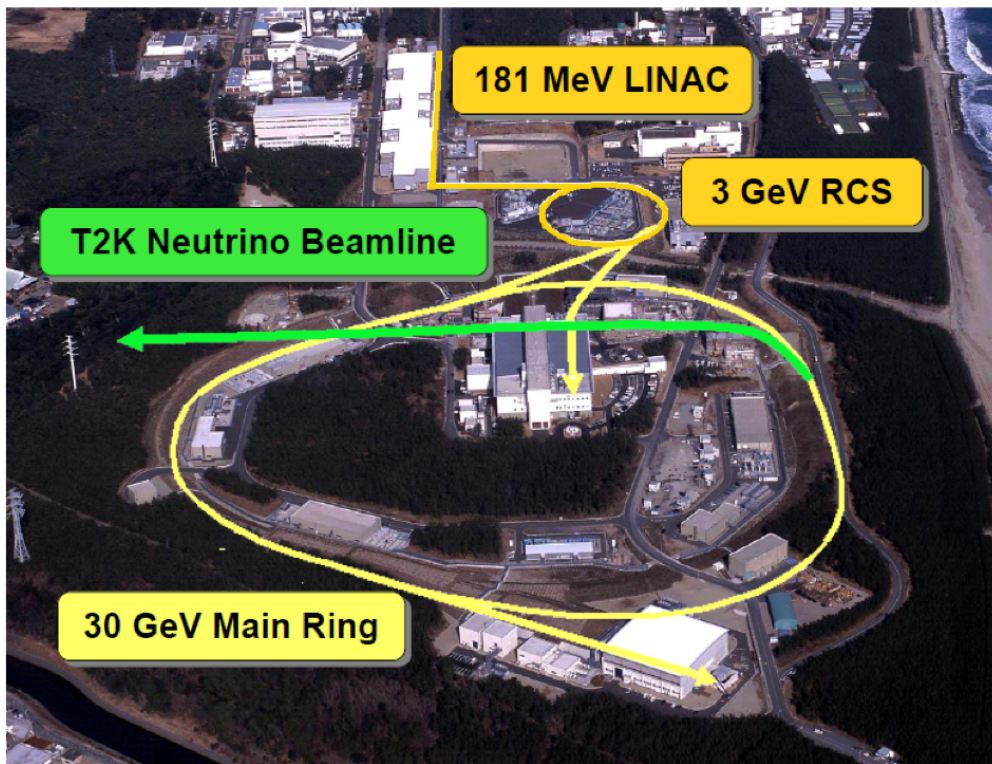
Figure 3.2.: The J-PARC accelerator complex, with the three main accelerators labelled. Figure from [85].

primary beamline perform these measurements.

The secondary beamline contains the proton target station, decay volume, and beam dump and muon monitor. This is shown in Figure 3.4. The proton target station, contains the graphite target, an optical transition radiation monitor (OTR), and the magnetic focussing horns. The protons are impinged on the target, producing secondary muons, pions, and kaons. The target is a 91.4 cm long graphite rod with a diameter of 2.6 cm. Helium gas is used to cool the target, to offset the heating effect of the beam. It is surrounded by a 2 mm thick graphite sleeve and a 0.3 mm titanium casing.

The target station sits within the first of the focusing horns. This collects the secondary mesons, and they are then focused by the second and third horns. In each horn, a toroidal magnetic field is produced by coaxial conductors. The strength of the field reduces with $1/r$ where r is the distance from the beam axis. The horns operate at 250 kA (1.7 T), which increases the neutrino flux at SK by a factor of 16 compared to 0 kA.

The current in the horns can be reversed to focus either positive or negatively charged mesons. The right sign particles are focused onto the beamline, while the wrong sign particles are deflected away. The current required to focus positive mesons is referred to as forward horn current (FHC), and the current required to focus negative sign particles is referred to
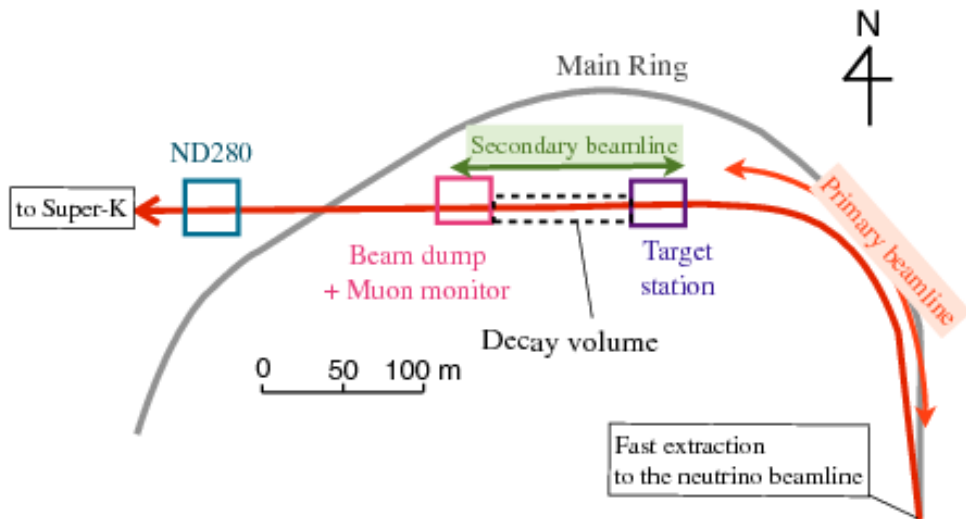
Figure 3.3.: The T2K neutrino beamline. Figure from [86].

as reverse horn current (RHC).

The focused mesons then travel down the 96 m long steel tunnel decay volume. Here, neutrinos are produced via the following decay processes:

$$\pi^+ \to \mu^+ + \nu_\mu$$
$$K^+ \to \mu^+ + \nu_\mu$$

(3.1)

in FHC, and:

$$\pi^- \to \mu^- + \bar{\nu}_\mu$$
$$K^- \to \mu^- + \bar{\nu}_\mu$$

(3.2)

in RHC.

There is also a small contribution of ν_e and $\bar{\nu}_e$ from decays such as:

$$\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$$
$$K^+ \to \pi^0 + e^+ + \nu_e$$

(3.3)

in FHC, and:

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$$
$$K^- \to \pi^0 + e^- + \bar{\nu}_e$$

(3.4)

in RHC.
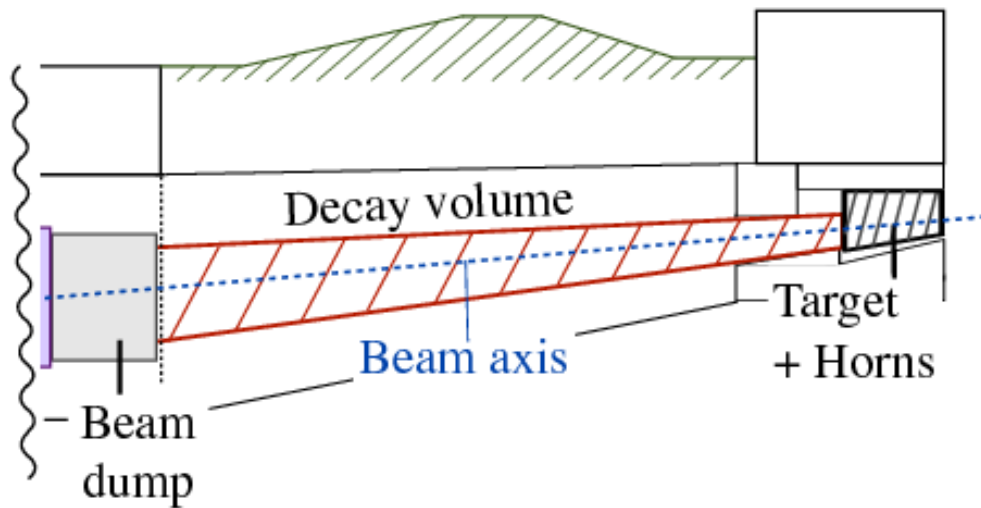
Figure 3.4.: Side view of the secondary beamline, showing the target station and focusing horns, decay volume, and beam dump. Figure from [86].

In each mode, there is some contamination of the wrong sign neutrino due to imperfect horn focusing, and the μ decays in Equations 3.3 and 3.4, as shown in Figure 3.5. As this is a larger effect for RHC mode, and as anti-neutrinos have a much smaller cross-section than neutrinos, there are many more neutrino interactions in RHC mode than anti-neutrino interactions in FHC mode.

After the decay channel, there is a beam dump made up of 3.17 m of graphite and 2.4 m of iron. This stops all surviving mesons, and all muons below 5 GeV. Neutrinos and higher momentum muons reach the muon monitor (MUMON) beyond the beam dump. The MU-MON consists of two independent detectors: Si PIN photodiodes, and ion chambers. These measure the muon profile on a bunch by bunch basis, which is a reliable measure of the beam direction and intensity as the majority of neutrinos in the beam are produced with a muon in a two-body decay. The MUMON measures the beam direction to an accuracy of 0.25 mrad, and the beam intensity to a precision of 3% [88].

It's not possible to count the total number of neutrinos produced in the beam, so the number of protons impinged on the target (POT) is used as a metric for the data collected by T2K. The total accumulated POT and beam power have been increasing since data taking began in 2010, as shown in Figure 3.6. Near detector data from runs 2–9 is used in this analysis.

(a) FHC Mode   (b) RHC Mode
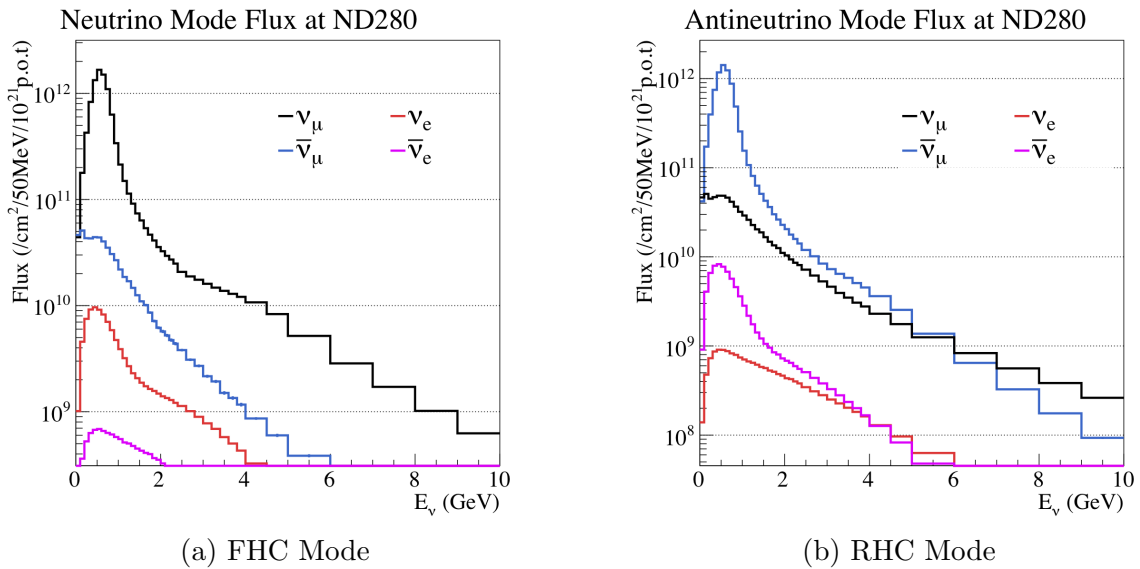
Figure 3.5.: Prediction of ND280 event rate broken down by neutrino species. Figure from [87].

3.1.2. Off Axis Technique

Because neutrinos are produced in a two-body decay, and so some fraction of energy is taken by the second decay product, it is not possible to produce a mono-energetic beam. The oscillation probability depends on the neutrino energy, and so more accurate measurements of oscillation parameters can be made using a neutrino beam with a narrower spread of energies. The neutrino energy is given by [90]:

$$E_\nu = \frac{m_\pi^2 - m_\mu^2}{2(E_\pi - p_\pi \cos\theta_{\pi\nu})}. \tag{3.5}$$

Figure 3.7 shows E_ν as a function of E_π for different values of $\theta_{\pi\nu}$. The E_ν distribution is flatter for larger $\theta_{\pi\nu}$, and so E_ν has a weaker dependence on E_π. At higher angles a narrower range of E_ν can therefore be produced from a wider range of E_π. This effect is highlighted in Figure 3.8, showing the predicted neutrino fluxes on and off-axis. A narrower beam spread is produced at 2.5^o, but with a lower overall rate (the y-axis is normalised). The beam energy and off-axis angle are chosen such that the neutrino energy peaks at ~0.6 GeV, to maximise the probability of oscillation at the far detector by aligning the beam peak with the first oscillation dip. Going off-axis also reduces the wrong sign contamination in the beam.
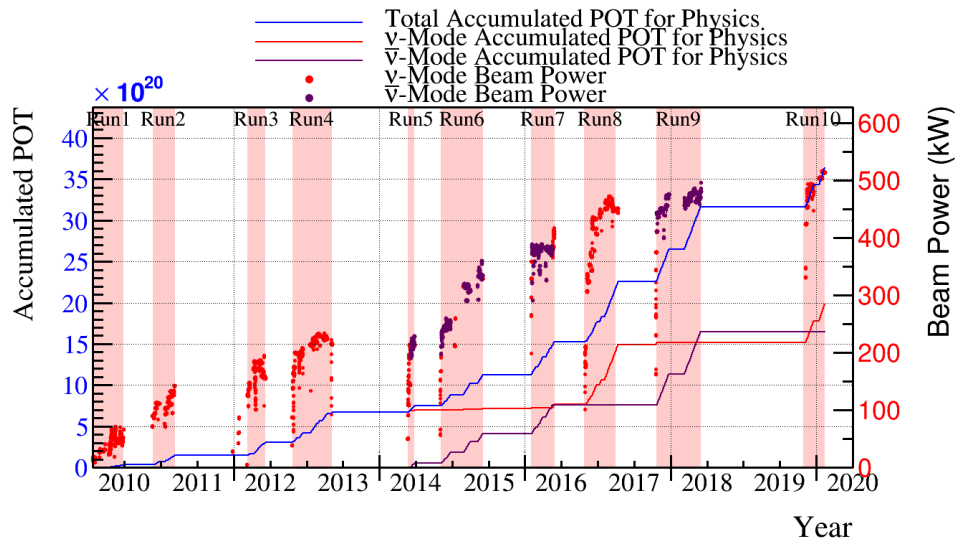
Figure 3.6.: The total accumulated POT and beam power at T2K for runs 2–10. Figure from [89].

3.1.3. The Neutrino Flux Simulation

A data-driven Monte-Carlo (MC) set is produced from beamline simulation to model the neutrino flux. The FLUKA2008 [93] software package was used to simulate 30 GeV proton interactions within the target. Measurements from the NA61/SHINE [94] experiment, which measures interactions of 30 GeV protons in a replica of the T2K target, are used to tune the simulation. The components of the beamline are modelled with the GEANT-4 [95] based JNUBEAM [96] software, and secondary interactions and interactions between particles which exit the target with the surrounding area are simulated with the GCALOR package [97]. Secondary particles are tracked until they decay to neutrinos, fall below the energy threshold to decay to neutrinos, or are absorbed in the beam dump. The predicted neutrino energy distributions for FHC mode are shown in Figure 3.5.

3.2. Near Detectors

The beam is first measured by two near detectors, ND280 and INGRID, 280 m away from the source. At this short distance, the probability of oscillation is negligible and so the unoscillated beam can be measured. The locations of the two detectors within the near detector suite are shown in Figure 3.9.
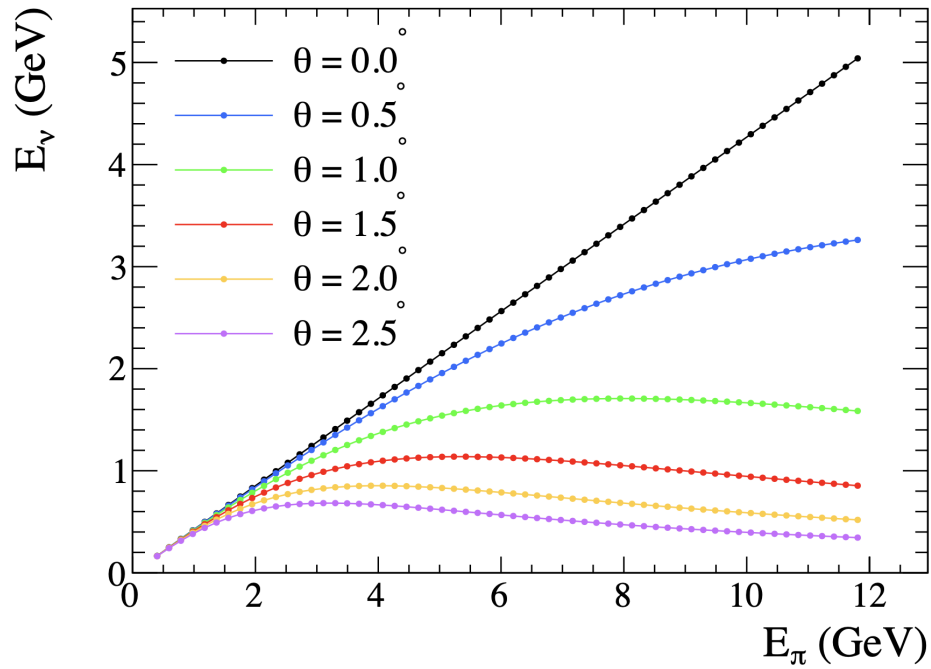
Figure 3.7.: Energy of neutrinos produced in two-body decay as a function of pion energy, for a variety of different off-axis angles. Figure from [91].
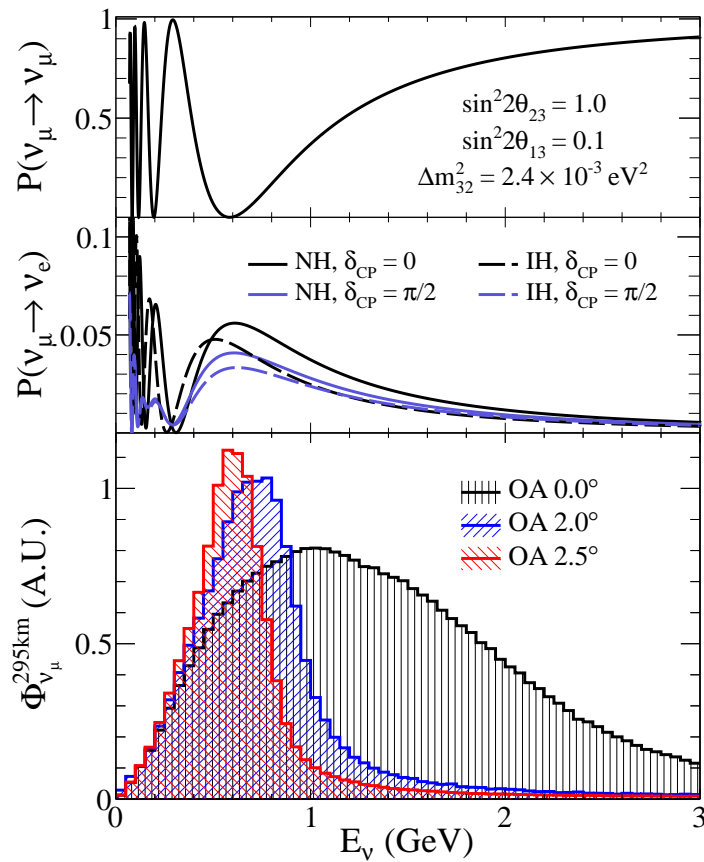
Figure 3.8.: Effect of off-axis angle on the predicted neutrino flux, normalised to arbitrary units, along with the oscillation and survival probabilities of ν_e and ν_μ respectively. Figure from [92].
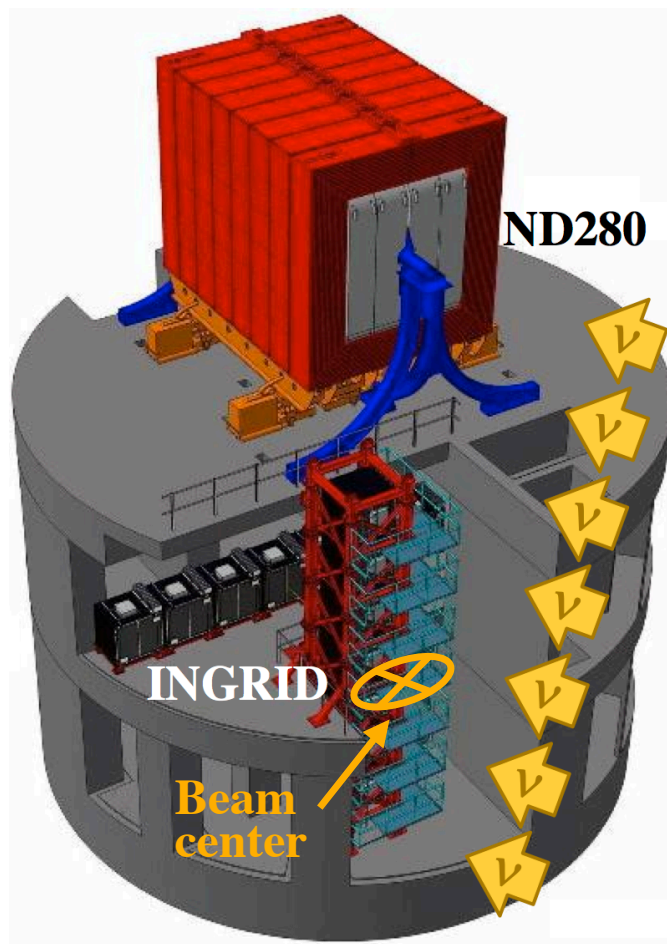
Figure 3.9.: The T2K near detector suite, 280 m from the beam source. Figure from [98].

3.2.1. INGRID

The INGRID detector is located on the beam axis, and is designed to measure the beam profile and direction. From this, the beam angle at ND280 and SK can be determined to within 0.2 mrad, and the beam center to within 5 cm. Neutrino event rates are also measured to within 2%. These measurements are in good agreement with results from MUMON, as shown in Figure 3.10. A precise determination of the beam direction is necessary as a 1 mrad uncertainty corresponds to a 2-3% uncertainty on the beam energy.

INGRID achieves this precision by using 16 identical modules made of interleaved iron and scintillator. These are arranged in a 10×10 m cross shape centered on the primary proton beamline axis, with two modules off-axis outside the cross, as shown in Figure 3.11.

Each module consists of nine iron sheets and 11 tracking scintillator planes, as shown in Fiugre 3.12. Each scintillator plane contains 24 horizontal and 24 vertical bars of plastic scintillator. These bars are threaded with a wavelength shifting (WLS) fibre, which collects photons emitted in the plastic scintillator during energy deposition. The WLS fibres trans-
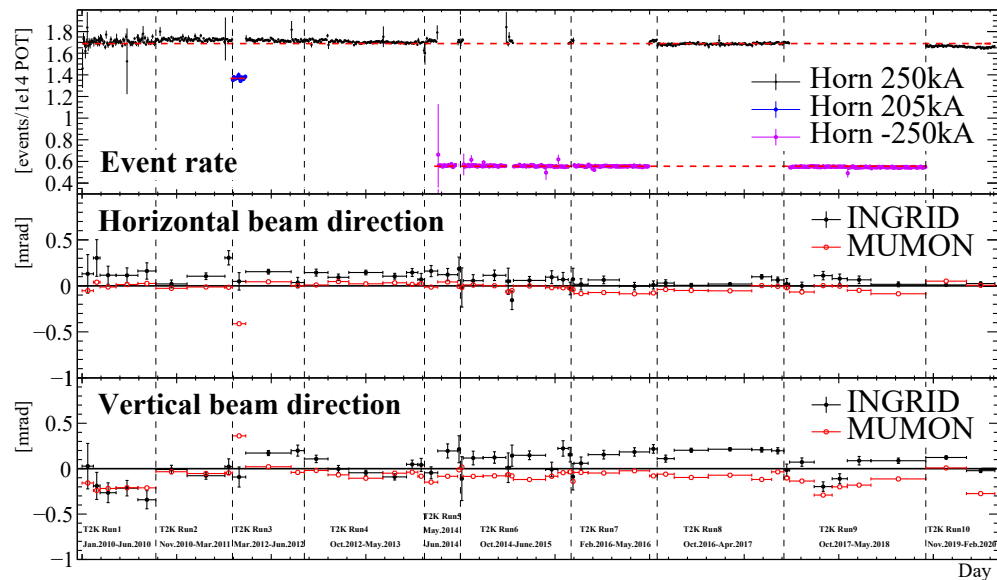
Figure 3.10.: INGRID and MUMON measurements of the beam direction and event rate for runs 1–10. Figure from [99].
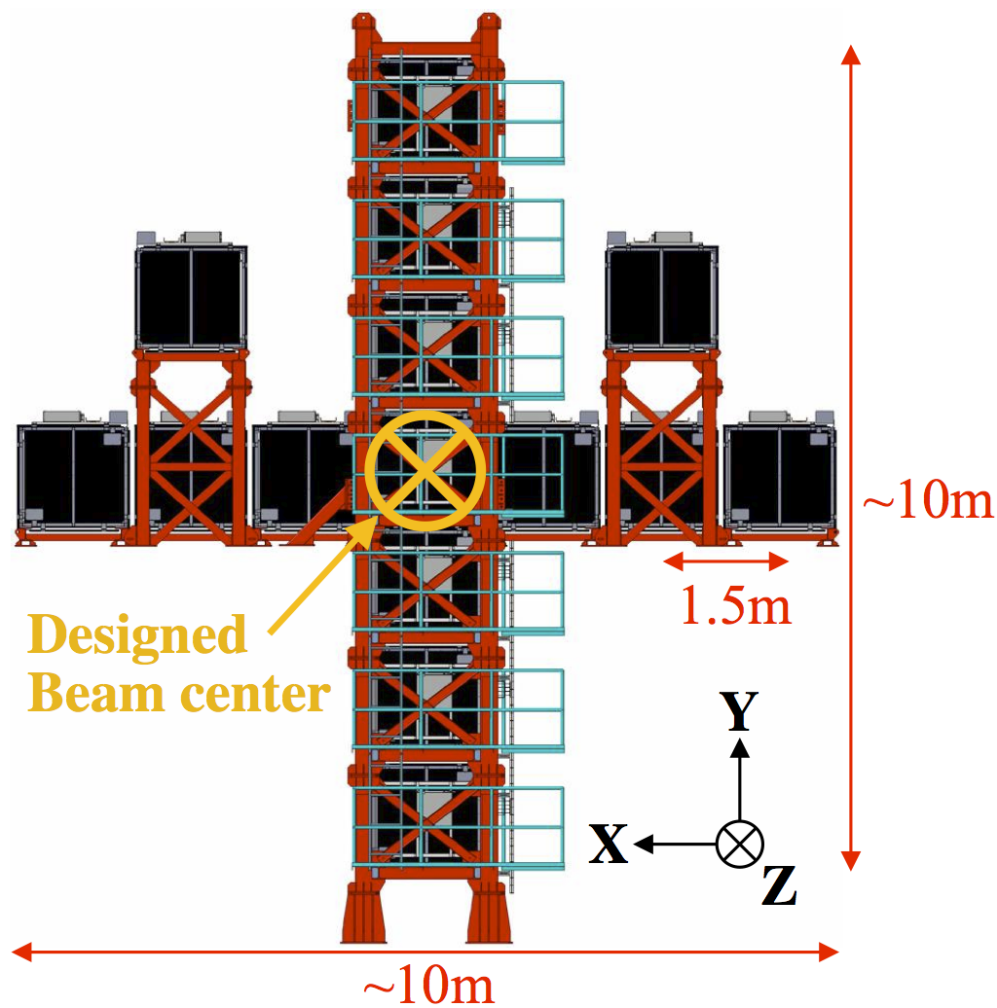
Figure 3.11.: The horizontal, vertical, and off-axis modules of the INGRID detector. Figure from [78].
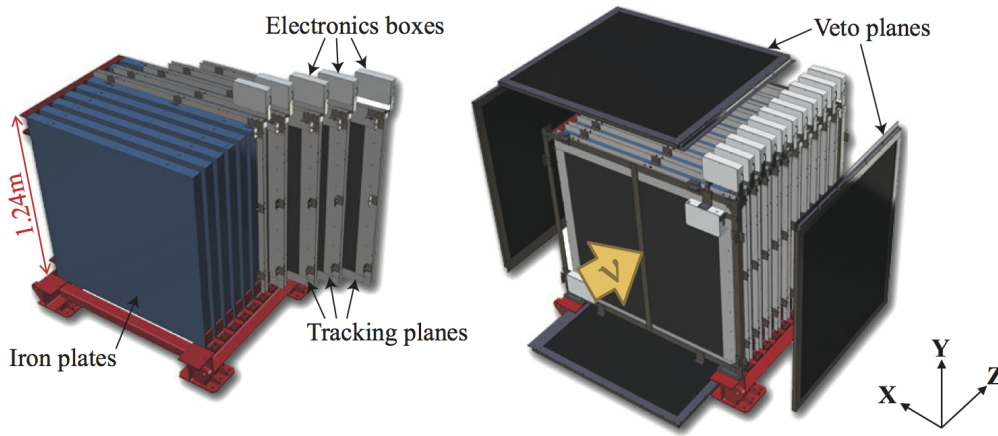
Figure 3.12.: The composition of an INGRID module. Figure from [78].

port photons to multi-pixel photon counters (MPPCs) at the end of each bar, and shift the spectrum of the light to the optimal region for MPPC readout. The MPPCs each contain several hundred avalanche photodiodes with a gain 1×10^6, which convert photons into an electrical signal. This signal is readout by a set of Trip-T front-end electronics boards [78], each of which is connected to 48 MPPCs. The backend electronics are made up of readout merger modules (RMM), which read data from the detector, and clock modules, which send trigger signals and ensure all components of the detector are synchronised.

The layered structure of each module is surrounded by scintillator planes to veto interactions occurring outside the fiducial volume. The total fiducial mass of iron in each module is 7.1 t, sufficient such that at nominal beam intensity, there are enough neutrino events to measure the beam direction on a day-by-day basis.

There is another 17th module, with a slightly different composition to the others. This module, known as the proton module, consists of only scintillator planes, but in smaller bars to give a finer granularity. It is designed to detect muons and protons from CCQE interactions on carbon in the plastic, to improve the MC simulation of the beamline and neutrino interactions. The proton module is located at the center of the cross, and like the other modules, is surrounded by veto planes. Figure 3.13 shows its composition.

3.2.2. ND280

The off-axis near detector, ND280, is designed to detect particles produced in neutrino interactions, to determine the event rate of various interaction modes and measure the unoscillated flux and energy of the beam in the direction of SK. This constrains systematic
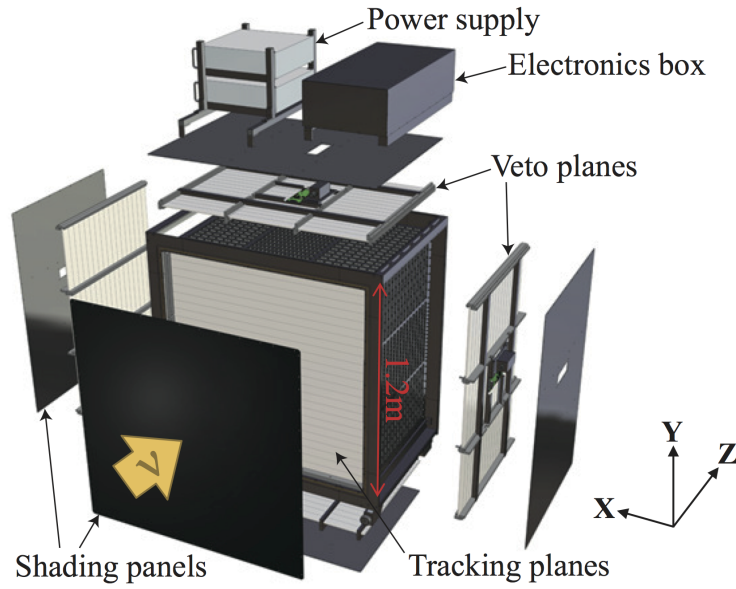
Figure 3.13.: The composition of the proton module in the INGRID detector. Figure from [78].

uncertainties allowing more accurate prediction of the event rates at the far detector. ND280 also measures neutrino interaction cross-sections at the 1 GeV energy scale.

As shown in Figure 3.14, ND280 consists of several sub-detectors. The tracker region contains two time projection chambers (TPCs) between three fine-grained detectors (FGDs). The FGDs provide a target for neutrinos, and track particles close to the interaction vertex. The TPCs identify and measure the momentum of particles, particularly muons, which are produced in the event and leave the FGD the interaction vertex is in. The FGDs and TPCs are described in more detail in Sections 3.2.2.1 and 3.2.2.2.

Upstream of the tracker, the π^0 detector (P0D) detects NC events on water, the same target as in the far detector. This is described in Section 3.2.2.3. Both the tracker region and P0D are surrounded by electromagnetic calorimeters (ECals). The ECals are designed to measure the energy of photons produced in the inner detectors and reconstruct π^0 tracks from the FGDs. They are described in more detail in Section 3.2.2.4.

The whole detector sits within the UA1 magnet which produces a 0.2 T magnetic field. This allows accurate charge and momentum measurements in the TPCs. The magnet yoke is interleaved with scintillator, the side muon range detector (SMRD). This measures high angle muons exiting the detector, cosmic ray muons entering the detector, and interactions in the magnet and surrounding area. The magnet and SMRD are described in Section 3.2.2.5.

The same MPPCs used in INGRID are also used throughout ND280. These were chosen,
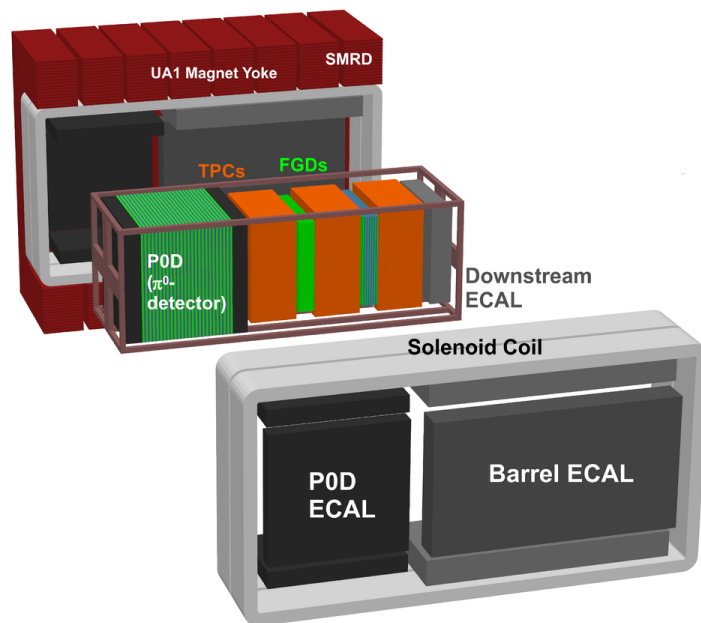
Figure 3.14.: Exploded view of ND280, showing its sub-detectors. Figure from [78].

rather than the more common photomultiplier tubes (PMTs), as they can be used within a magnetic field.

3.2.2.1. The Fine-Grained Detectors

The FGDs provide the primary target for neutrino interactions in ND280. They are designed to be able to measure particles which don't exit themselves and enter a TPC. To achieve this, they are completely active, allowing them to measure vertex activity and short range tracks such as those from recoil protons. Short ranged particles tend to have low momentum and deposit lots of energy per track length, and so the FGDs need to be able to detect charged particles with fine granular resolution, to distinguish between tracks and determine their direction.

The FGDs are also used as a cosmic trigger for stopping pions, which allows the identification and reconstruction of subsequent Michel electrons. They are also used for measurements of the time-of-flight (TOF) of tracks to differentiate between forward-going positive and backward-going negative particles. As well as this, the FGDs have the best timing resolution of all the sub-detectors, and are used as the base for reconstructing tracks which pass through more than one sub-detector.

The two FGDs both contain 1.1 t of target mass, and consist of layers of scintillator bars. Each layer is $186.4 \times 186.4 \times 2.02$ cm, and contains 192 bars in the horizontal direction and

192 bars in the vertical direction. The bars are covered in a reflective coating containing TiO_2, and each are $0.96 \times 0.96 \times 186.4$ cm. A WLS fibre runs down the centre of each bar to an MPPC at the end. The other end is mirrored by a vacuum deposition of aluminium.

The most upstream FGD (FGD1) is composed of 15 scintillator planes. The second FGD (FGD2) has seven scintillator planes separated by layers of hollow corrugated polycarbonate sheets. These are filled with water, providing a water-scintillator hybrid target. The FGDs therefore measure interactions on CH in the scintillator, which is a common target in external neutrino scattering experiments, and H_2O, the far detector target. Nuclear effects cannot be accurately extrapolated between target nuclei, and so it useful to be able measure interaction rates on water at ND280.

The resolution and track length in the FGDs is too low to use dE/dx to identify particles which don't enter the TPCs, so the combination of integrated deposited energy and track length is used. These quantities are compared to values from simulations of different particles, as shown in Figure 3.15, allowing protons to be distinguished from other charged particles. In particular, accurately distinguishing protons and pions which stop in the FGDs is vital for correctly identifying the interaction type. Figure 3.15 shows data from both the neutrino beam, and cosmic trigger. The neutrino beam data contains particles identifiable as protons, muons, and pions, whereas the cosmic trigger data only contains particles identifiable as muons and pions, as would be expected.

3.2.2.2. The Time Projection Chambers

The three TPCs are located either side of each FGD, and provide high resolution tracking for both forward-going and backward-going particles. The majority of the particle identification and momentum measurements at ND280 take place inside them. The multiplicity and direction of tracks can be easily determined, as the TPCs detect events in three dimensions. The momentum of particles can be measured from the bending of tracks in the TPCs due to the magnetic field.

The construction of the three TPCs is identical. They consist of two concentric boxes, as shown in Figure 3.16. The inner box is filled with a drift gas consisting of an $Ar:CF_4:C_4H_{10}$ mixture at 95:3:2. The walls of the inner box form the field cage. The outer box walls are held at ground voltage, and the outer box is filled with CO_2, to provide electrical insulation between the boxes.

The inner box contains the cathode, and the walls parallel to the cathode have a copper pattern, designed to produce a uniform electric field in the module, aligned with the magnetic

(a) Neutrino beam data

(b) Cosmic trigger data
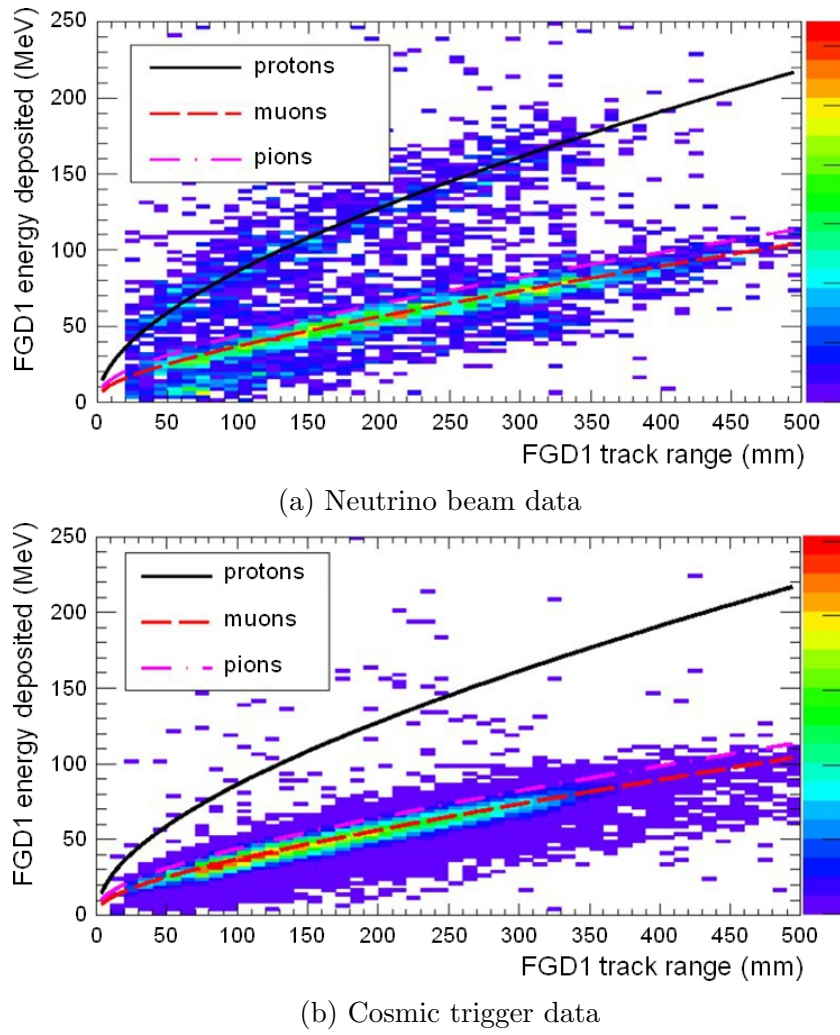
Figure 3.15.: Integrated deposited energy as a function of range for particles stopping in FGD1. The scatter plot shows data while the curves show the MC predictions for protons, muons, and pions. Figures from [100].
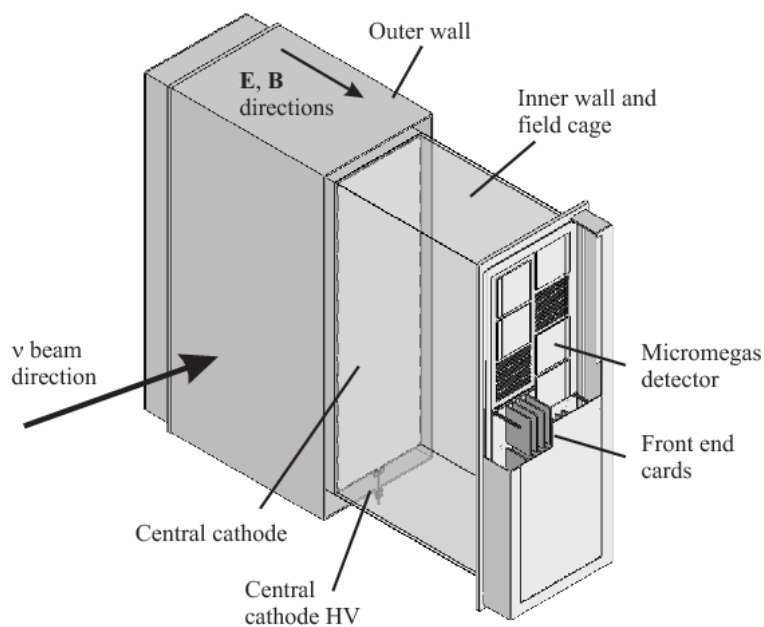
Figure 3.16.: Schematic diagram of a TPC module. Figure from [78].
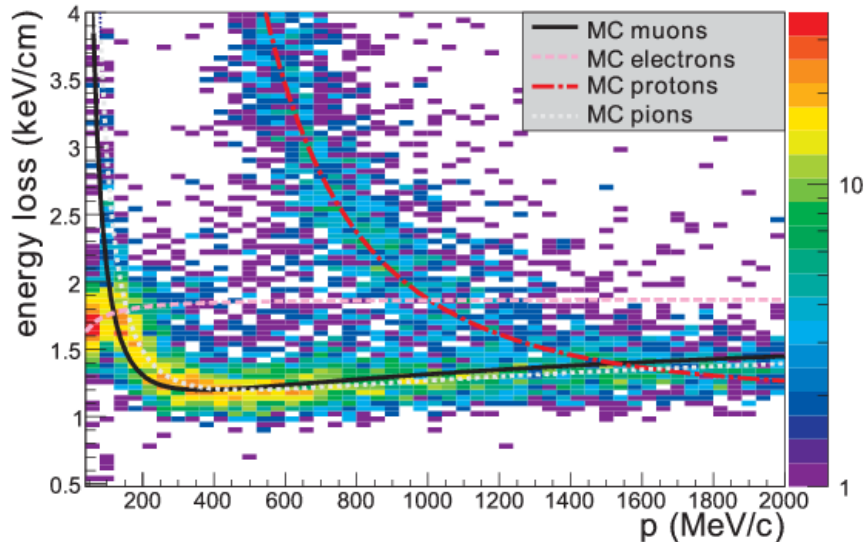
Figure 3.17.: Energy loss as a function of momentum for particles in one TPC. The scatter plot shows data while the curves show the theoretical predictions for protons, electrons, muons, and pions. Figure from [101].

field. The TPC modules are made of non-magnetic materials as not to interfere with the field from the magnet.

When a charged particle passes through a TPC module, ionisation electrons are produced, which drift away from the cathode to readout planes on the walls of the inner box. Here, the electrons are multiplied and detected by micromegas detectors. The time and position of the electron signals at the readout planes give a 3D image of the particle track.

Particle identification is performed in the TPCs using, as in the FGDs, $\mathrm{d}E/\mathrm{d}x$. This is shown in Figure 3.17, as a function of momentum, for different particles. The $\mathrm{d}E/\mathrm{d}x$ resolution is 7.8±0.2%, allowing electrons and muons to be distinguished.

3.2.2.3. The π^0 Detector

The P0D was designed to primarily measure the cross-section of the neutral current interaction $\nu_\mu + N \rightarrow \nu_\mu + N + \pi^0 + X$ on water. This is one of the major backgrounds to ν_e appearance at the far detector, and so it is necessary to measure and constrain it.

The P0D is located upstream of the FGDs and TPCs. It consists of alternating layers of scintillator bars, brass and lead sheets, and water target bags, as shown in Figure 3.18. There are 40 modules, each made up of two triangular scintillator bars. The first layer contains 134 vertical bars, and the second 126 horizontal bars. Each bar contains a WLS fibre which is read out by an MPPC. The horizontal and vertical bars are 234 and 220 cm
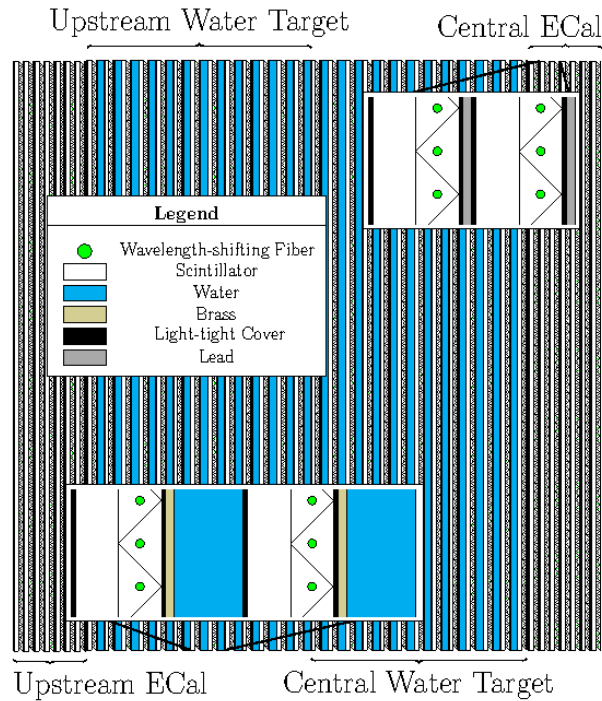
Figure 3.18.: A schematic diagram of the side on view of the P0D. Figure from [78].

long respectively.

The water bags can be emptied, allowing measurements of the cross-section on H_2O by subtraction. The target mass is 16.1 t with water in the bags, and 13.3 t without water. The lead and brass sheets produce electron showers which can be detected, from photons emitted in π^0 decays.

3.2.2.4. The Electromagnetic Calorimeter

The primary purpose of the ECals is to tag and reconstruct π^0s from the FGD, TPC, and P0D, by measuring the energy and direction of photon showers. It is also used to distinguish between pions and muons by shower shape. The ECal's near hermetic coverage of the inner detectors allows full reconstruction of events, and they also provide complimentary particle identification to the TPCs.

There are three sections of the ECal: the downstream ECal (DsECal) located after the last TPC, the barrel ECal surrounding the FGDs and TPCs, and the P0D ECal surrounding the P0D. These are shown in Figure 3.14. The barrel and downstream ECals are tracking calorimeters which reconstruct electromagentic showers, particularly those from high-angle particles from the TPC. The P0D ECal is designed to tag energy escaping from the P0D

and distinguish photons from muons, as the P0D reconstructs showers itself.

Each ECal module consists of alternating layers of scintillator bars and lead sheets. The bars are 40 × 10 mm, and contain a WLS fibre which is read out by an MPPC. The lead sheets are 1.75 mm thick.

The downstream ECal has 34 layers of scintillator, corresponding to 11 electron radiation lengths. Alternate layers are orientated perpendicularly to each other. This gives two 'views', which can be combined to create 3D reconstructed tracks. The downstream ECal is 2300 × 2300 × 500 mm, and its total target mass is 4.80 t.

The six modules of the barrel ECal have 31 layers of scintillator, corresponding to 10 electron radiation lengths. These also alternate orientations similarly to the downstream ECal. The side barrel ECal modules are 4140 × 2500 × 462 mm, and their total target masses are 9.21 t. The top and bottom barrel ECals are 4140 × 1676 × 462 mm, and their total target masses are 6.62 t.

The six modules of the P0D ECal have only six layers of scintillator, each parallel to the beam. These are interspersed with 4 mm thick lead sheets. The side P0D ECal modules are 2898 × 2454 × 155 mm, and their target masses are 2.64 t. The top and bottom P0D ECal modules are 1584 × 2454 × 155 mm, and their target masses are 1.5 t.

3.2.2.5. The UA1 Magnet and Side Muon Range Detector

The TPC, FGD, P0D, and ECal all sit within the UA1 [102] magnet, which provides a 0.2 T magnetic field. This allows the TPCs to measure the momentum of charged particles with a resolution of 10%, and determine their sign, which identifies if the interaction involved a neutrino or anti-neutrino. The magnet consists of water-cooled aluminium coils, which produce the field. The coils are supported by the return yoke, which is made up of 16 C-shaped iron elements. The internal volume of the magnet is 7.0 × 3.5 × 3.6 m, and the external volume is 7.6 × 5.6 × 6.1 m.

The nominal current is 2.7 kA, and this is monitored regularly to accurately calculate the field. This reduces the uncertainty on momentum measurements in the TPC. The uncertainty on the field measurement is 2×10^{-4} T.

Air gaps in the return yoke hold the modules of the SMRD. These are used to track high angle muons and measure their momentum, as well as providing a cosmic trigger. The SMRD consists of 192 horizontal and 248 vertical plastic scintillator modules. The horizontal modules are 9 × 686 × 955 mm, and the vertical modules are 9 × 892 × 955 mm. The scintillator modules contain a WLS fibre which is read out by an MPPC.

3.2.2.6. The Data Acquisition System

The ND280 data acquisition (DAQ) system triggers the readout of information from each of the sub-detectors, as well as the storage of recorded data. The same system is used for both ND280 and INGRID.

There are three trigger requirements during physics runs: the beam trigger, when a beam spill occurs; the Trip-T cosmic trigger, when hits are seen on the opposite sides of the outer detectors (top and bottom SMRD, left and right SMRD, or P0D and downstream ECal) outside the beam window; and the FGD cosmic trigger, when hits are seen in both FGDs outside the beam window. These initiate a fixed time window during which data is recorded. Data is acquired from front-end electronic boards via optical Gigabit links, and event fragments from the sub-detectors are merged and logged.

3.2.3. Near Detector Simulation

The ND280 and INGRID geometries and the paths of final-state particles from neutrino interactions are simulated in GEANT-4. The neutrino interactions themselves are simulated using NEUT [103], an interaction generator written for the Super-Kamiokande and T2K experiments.

A custom software package, ELECSIM, is used to simulate the response of the detector and electronics to energy deposited by particles that have propagated through the detectors. This involves simulating the light emitted from the energy deposition, the transport of that light through the bar and down the WLS fibres, and the subsequent response of the MPPCs. For the TPCs, ELECSIM simulates the electron drift and response of the micromegas detectors.

3.3. Super-Kamiokande

The far detector, Super-Kamiokande, is a water Cherenkov detector located 295 km west of the near detector suite, in the Kamioka mine inside Mt. Ikenoyama. The mine provides 1 km of rock (or 2.7 km equivalent of water), shielding the detector from cosmic ray muons below 1.3 TeV, significantly reducing background rates. The detector is filled with 50 kt of pure water (25 kt fiducial volume), and, like ND280, lies 2.5^o off-axis. Figure 3.19 shows the detector within the mine.

Super-Kamiokande has been searching for proton decay and measuring solar and atmospheric neutrino oscillations since 1996, and has been a far detector for a long baseline accelerator neutrino oscillation experiment since 1999, initially for K2K and now for T2K.
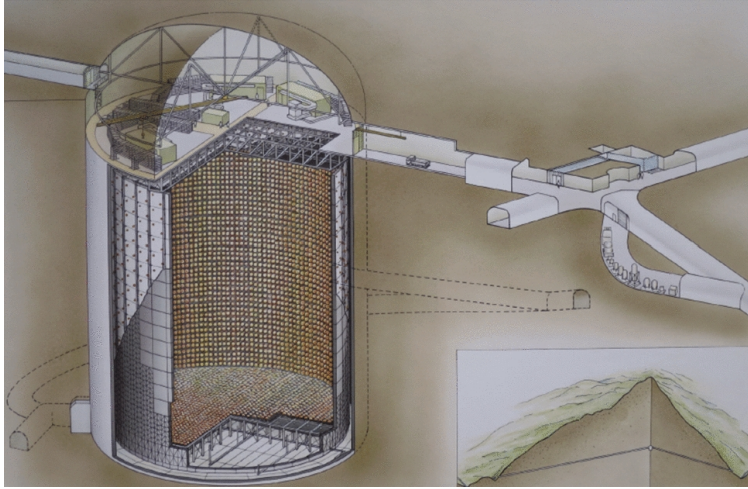
Figure 3.19.: The Super-Kamiokande detector within the Kamioka mine. Figure from [78].

Although it has undergone several updates during this period, the long running operation mean its behaviour is well understood. The atmospheric and cosmic ray muon data provide calibration samples completely separate from the T2K analysis, for which the simulated MC matches data to the percent level.

The detector is divided into the inner (ID) and outer (OD) detectors by a 55 cm cylindrical stainless steel and tyvek framework. The OD is 40.2 m in height with a radius of 35.8 m. It surrounds the ID with 2 m of water, serving as a shield from interactions in the surrounding rock. An active veto for cosmic ray muons is provided by 1885 outward-facing 20 cm PMTs. Despite the fairly sparse PMT coverage, the veto has an efficiency of almost 100%. The beam timing window can be used to identify interactions in the OD from beam neutrinos.

The ID is 36.2 m in height with a radius of 33.8 m. It contains 11,129 inward-facing 50 cm PMTs, each with a combined quantum and collection efficiency of 20%, and a timing resolution of ~2 ns. The PMT coverage of the ID is 40%, which gives enough spatial resolution to sufficiently reconstruct the paths of particles produced in neutrino interactions inside the tank.

When charged particles travel through a medium of refractive index n at a velocity v_p larger than the speed of light in the medium, $v_\gamma = c/n$, Cherenkov radiation is emitted in a cone with opening angle $\theta_c = \arccos\frac{1}{n \cdot v/c}$, along the direction of the particle's path. In water, this is ~42°. Neutrino interactions in the SK tank produce charged particles which, if above an energy threshold, can be detected by the emitted Cherenkov light. Photons from the cone form a ring shape on the walls of the tank, where they are detected by the PMTs. This pattern, along with the hit timing, can be used to reconstruct the interaction vertex

(a) Muon neutrino event.   (b) Electron neutrino event.
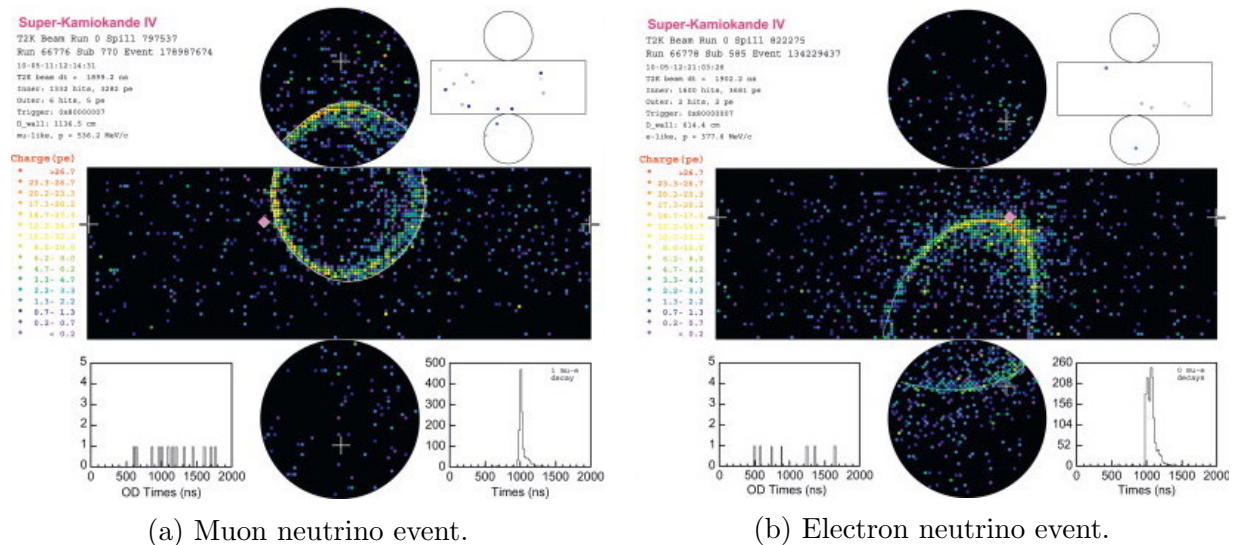
Figure 3.20.: SK ID event display, showing the Cherenkov ring PMT hits for a a) muon, and b) electron neutrino event. Figures from [78].

position, and the charge, momentum and direction of the produced particle.

CCQE interactions are the primary channels used to measure ν_μ disappearance and ν_e appearance at T2K. Discounting final state interactions, these events produce a charged lepton and a proton. The measured momentum and direction of the lepton can be used to reconstruct the neutrino energy, using Equation 2.30, but it is also vital for the lepton flavour to be determined, as this corresponds to the incoming neutrino flavour.

Discrimination between electrons and muons is achieved by separating Cherenkov rings by shape. As electrons are relatively light, they scatter off particles in the water and so travel a more convoluted path to the PMTs. At the T2K beam energies, they will also induce electromagnetic showers. Both of these phenomena cause the Cherenkov light to form a superposition of overlapping rings at the wall of the ID. This results in a fuzzy ring being seen in the PMTs.

Conversely muons, due to their relatively large mass, travel through the water without scattering, and so produce a sharp, clear Cherenkov ring. Examples of electron and muon Cherenkov rings detected at SK are shown in Figure 3.20.

An algorithm [104] is used to distinguish between fuzzy and sharp rings with high efficiency. The probability of a muon neutrino event being misidentified as an electron neutrino event is 0.7%. The number of each different neutrino species events is used to calculate neutrino oscillation parameters in the full T2K oscillation analysis. However, as Super-Kamiokande does not have a magnetic field, it is not able to differentiate between neutrino and anti-

neutrino interactions. The neutrino and anti-neutrino content of the beam is therefore needed to be measured accurately at ND280. In this analysis, only data from ND280 is used, to constrain systematic uncertainties.

3.3.1. Far Detector Simulation

Neutrino interactions in the SK tank are modelled by the NEUT event generator. The simulation of produced particles through the detector is done using the GEANT-3 [105] based SKDETSIM software [106]. As with the beam production, hadronic interactions are simulated with the GCALOR package.

3.4. Summary

This chapter has presented an overview of the T2K experimental setup. A system of three accelerators produce 30 GeV H^+ ions. These are impinged on a graphite target, producing muons, pions, and kaons, which are focused by magnetic horns. Right sign particles are focused onto the beamline, while wrong sign particles are deflected away. The focused mesons then decay to primarily muons and muon neutrinos. A beam dump stops all surviving mesons and muons, leaving a directed beam of muon neutrinos.

The near detectors are located 280 m downstream of the source, to measure the unoscillated neutrino beam. The on-axis INGRID precisely measures the beam profile and direction. It comprises of iron and scintillator modules, arranged in a 10×10 m cross shape centred on the proton beamline axis.

ND280 sits 2.5^o off-axis to measure a narrower neutrino energy peak. It is designed to determine the event rate of various interaction modes, and measure the flux and energy of the beam before oscillation to constrain systematics. The tracker region consists of two FGDs, which provide the neutrino target, interleaved with three TPCs, which measure the momentum of outgoing particles.

Upstream of the tracker region, the P0D detects NC events on water, providing an important constraint on the cross-section of one of the major backgrounds at SK. Both the tracker region and P0D are surrounded by the ECals, which measure the energy of photons produced in the inner detectors.

The whole detector sits within the UA1 magnet, allowing charge and momentum measurements in the TPCs. The SMRD modules sit within the magnet yoke, and measure high angle muons exiting the detector.

The same MPCCs are used throughout INGRID and ND280 to convert photons into an electric signal. The two near detectors also share a common DAQ system, for triggering the readout and storage of data from each sub-detector. Interactions are generated using NEUT, and the detector geometries are simulated in GEANT-4.

Super-Kamiokande measures the flavour composition of the beam after oscillation, and is located 295 km away from the source of the beam, 2.5^o off-axis, and 1 km underground. It consists of a 50 kt tank of pure water, and ∼11,000 PMTs which detect the Cherenkov radiation from leptons produced in neutrino interactions within the tank. Electrons and muons are distinguished by the shape of the Cherenkov rings seen by the PMTs, allowing the flux of electron and muon neutrinos to be measured separately. Neutrino interactions at SK are also generated in NEUT, and the detector geometry is simulated using GEANT-3.

Using the constraint on systematic uncertainties provided by the near detector, oscillation parameters can be extracted from the measurement of the beam at SK.

4. Statistical Treatment

The number of events measured at the far detector is a convolution of the cross-section, flux, detector efficiency, and probability of oscillation. As discussed in Section 2, neutrino interactions are rare and so the cross-section, flux, and detector models all have large systematic uncertainties. The parameters of these models have similar effects as the oscillation parameters being calculated; a change in a single nuisance parameter mimics a change in oscillation parameters in terms of the effect on the kinematic distributions measured.

The aim of the near detector fit is to constrain the cross-section, flux, and detector systematics before oscillation, allowing more precise measurements of the oscillation parameters at SK. However, as there are several hundred nuisance parameters, fitting them all requires careful statistical treatment.

This analysis invokes a Bayesian approach, using the Markov Chain Monte-Carlo (MCMC) method to fit systematic parameters to data. This produces an N-dimensional posterior probability distribution, where N is the number of fit parameters. Post-fit central values and uncertainties are extracted from this distribution by marginalising over all other parameters, one by one. The near detector-only fit in this analysis is used to validate the model and fitting framework, before full joint near and far detector fits can be run.

This chapter describes the statistical treatment in the fit. The general approach to determining parameter values in Bayesian statistics is discussed in Section 4.1. Monte-Carlo methods are introduced in Section 4.2, and the theory behind MCMC is presented in Section 4.3. The methods used to estimate parameter values and assess the model's ability to fit the data are outlined in Section 4.4.

4.1. Bayesian Inference and the T2K Likelihood

In Bayesian statistics, a hypothesis is tested by combining prior information with the likelihood of a dataset. The aim of all Bayesian analyses is to model the probability of both the data and model parameters, to produce a posterior probability distribution $P(\bar{\theta}|D)$, where

4. Statistical Treatment

$\bar{\theta}$ represents the model parameters, and D represents the data. From this, the post-fit parameter values and uncertainties can be extracted using the methods described in Section 4.4.1. The posterior probability distribution is related to the joint probability distribution, $P(D, \bar{\theta}) = P(D|\bar{\theta})\, P(\bar{\theta})$, using Bayes' Theorem:

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

$P(D|\bar{\theta})$ is the probability of measuring the data D given the set of model parameter values $\bar{\theta}$. This is calculated by comparing the number of data events in each bin to the number of events in the Monte Carlo prediction using the given set of model parameters. The Poisson likelihood for a single bin is given by:

$$\mathcal{L}_{Bin} = \frac{\lambda(\bar{\theta})^n e^{-\lambda}}{n!} \Big/ \frac{n(\bar{\theta})^n e^{-n}}{n!} \tag{4.2}$$

where n is the number of observed data events in the bin, and $\lambda(\bar{\theta})$ is the number of predicted MC events for model $\bar{\theta}$ in the bin. The total sample contribution to the log-likelihood is therefore given by:

$$-\log \mathcal{L}_{Sample} = \sum_{Bins} [\lambda(\bar{\theta}) - n + n \log \frac{n}{\lambda(\bar{\theta})}]. \tag{4.3}$$

This is then modified to include the MC statistical uncertainty, which accounts for the fact that there was not an infinite amount of MC generated. An additional penalty is added to the original sample likelihood, and a scaling factor, β, is applied to $\lambda(\bar{\theta})$:

$$-\log \mathcal{L}_{Sample} = \sum_{Bins} [\beta\lambda(\bar{\theta}) - n + n \log \frac{n}{\beta\lambda(\bar{\theta})} + \frac{(\beta - 1)^2}{2\sigma_\beta^2}]. \tag{4.4}$$

Fitting a new parameter, β, for each bin would introduce $\sim$3000 new fit parameters, meaning the total number of parameters would be too high to feasibly fit. Instead, it is assumed β follows a Gaussian distribution for each. By minimising $-\log \mathcal{L}$ with respect to β, β can be calculated analytically for each bin by solving:

$$\beta^2 + (\lambda\sigma_\beta^2 - 1)\beta - n\sigma_\beta^2 = 0, \tag{4.5}$$

where σ_β is the fractional uncertainty on the MC in a given bin. Equations 4.4 and 4.5 are derived in [107].

$P(\bar{\theta})$ contains the prior knowledge of the model parameters, which is driven by previous

and external measurements. This is calculated as either a Gaussian or flat[1] uncertainty for each parameter, along with inter-parameter correlations:

$$-\log \mathcal{L}_{Systematic} = \sum_{Systematics} \frac{1}{2}[(\theta_i - \mu_i)(\mathbf{V})_{ij}^{-1}(\theta_j - \mu_j)], \tag{4.6}$$

where θ_i is the value of parameter i with central value μ_i, and $\mathbf{V}$ is the covariance matrix describing the relation between parameters i and j with $\mathbf{V}_{ij}$. In this analysis, the systematics are grouped by three covariance matrices from the beam flux, cross-section, and ND280 detector models.

The joint probability distribution is therefore given by:

$$P(D|\bar{\theta})P(\bar{\theta}) = \prod \mathcal{L}_{total} = \prod (\mathcal{L}_{Sample} \times \mathcal{L}_{Systematic}), \tag{4.7}$$

and so:

$$-\log \mathcal{L}_{Total} = \sum_{Bins}[\lambda(\bar{\theta}) - n + n\log\frac{n}{\lambda(\bar{\theta})} + \frac{(\beta-1)^2}{2\sigma_\beta^2}] + \sum_{Systematics}\frac{1}{2}[(\theta_i - \mu_i)(\mathbf{V})_{ij}^{-1}(\theta_j - \mu_j)]. \tag{4.8}$$

The integral in Equation 4.1 is often not analytically solvable in practice, and so Monte-Carlo methods are required to sample from the posterior to produce a distribution proportional to the posterior probability distribution up to a normalisation constant.

4.2. Monte Carlo Methods

Monte Carlo simulation can be used to estimate mathematical functions and mimic the operations of complex systems with random sampling and statistical modelling. They provide a solution to the problem of high dimensionality and non-analytically solvable integrals by sampling distributions with a random walk through a given parameter space. Properties of the distribution such as parameter values and integrals can then be approximated from properties of the samples. This is often a much easier process than directly evaluating an integral, and so these methods are used in many fields from climate science, to economics, to computational biology.

The simplest approach for performing an integral is to throw a random point in a region of known volume that encompasses the target region. The reliance on random numbers is

---

[1]Strictly speaking flat uncertainties are actually 'top hat' functions with hard cut-offs. These either represent physical boundaries for the parameter, or are capped at 0 and an arbitrary large number.

where the name 'MC methods' emanates from, referring to the famous casino in Monaco. The fraction of throws within the target region is multiplied by the known volume, to get an approximation of the target volume. The results are not an exact solution, and are dependent on the random throws sufficiently sampling the distribution. From the law of large numbers, the accuracy of this approximation improves with the number of points thrown.

This method is powerful as the full shape of the distribution does not need to be known to perform the calculation, just whether a point is inside or outside the target region. Evaluating a function at a single point is far easier than computing a full integral. However, this simple method can be very inefficient and require a large amount of throws to converge on the solution. If the area is not well chosen, substantial computing time is wasted evaluating points outside the target region which never contribute to the integral. Slow convergence is one of the main drawbacks of MC methods. Several techniques, such as MCMC, have been developed to minimise the unnecessary computation by using a semi-random walk through the parameter space.

4.3. Markov Chain Monte Carlo

A Markov chain is produced by any algorithm that generates a new point, x_i which only depends on the previous generated point, x_{i-1}. The process is truly Markovian if predictions about future steps in the chain can be made by only knowing the current state, just as well as by knowing the whole history of the chain.

If a Markov chain gradually 'forgets' its initial state as more steps are added, it is said to converge. This means the initial state converges to a unique stationary distribution, such that the distribution is independent of the step number. Once convergence has been reached, all ensuing steps are samples from the stationary distribution.

The aim of MCMC methods is to produce a Markov chain with the posterior distribution $P(\bar{\theta}|D)$ as its stationary distribution. The chain steps through N dimensional space, where N is the number of model parameters. A single point in the chain then represents a set of values, one for each parameter, and is defined by a vector $\bar{x}$.

This way, the choosing of points at which to evaluate the distribution is done much more efficiently than just random sampling. The individual steps have a density proportional to the target distribution as the Markov chain performs a semi-random walk through the parameter space, and so less time is wasted sampling areas of low probability. The stationary distribution is only an approximation to the desired posterior, but more closely matches it

as more steps are added.

MCMC methods are dependent on producing a Markov chain that converges. To reach convergence, it is necessary that three conditions are met:

- **Irreducibility:** From any initial state, there is non-zero probability of reaching any other state. This prevents the chain getting stuck in local minima.

- **Aperiodicity:** The chain must not be periodic. This means the chain never gets stuck in a loop between the same states.

- **Recurrence:** All subsequent steps sample from the same stationary distribution once it has been reached. This means once staionarity has been achieved, adding more steps gives a more accurate approximation to the target distribution.

Any Markov chain which satisfies each of these criteria is ergodic, and will eventually converge. The total number of steps needed for convergence however differs from chain to chain. The procedures for testing convergence are described in Section 4.3.3.

Once a Markov chain has been produced which has the posterior distribution $P(\bar{\theta}|D)$ as its stationary distribution, has met the three criteria, and has reached convergence, it can be used to estimate the model parameters $\bar{\theta}$. The main difficulty comes from constructing a chain with the correct stationary distribution. In this analysis, this is achieved using the Metropolis-Hastings Algorithm.

4.3.1. The Metropolis Hastings Algorithm

The Metropolis Hastings algorithm was first developed by N. Metropolis for symmetric proposal distributions in 1953 [108], and was generalised to the asymmetric case by W. K. Hastings in 1970 [109]. It can be used to construct a Markov Chain that satisfies the regularity conditions, and therefore has a stationary distribution. A semi-random walk is directed through the parameter space, such that steps are distributed according to the posterior probability distribution. The algorithm consists of the following method:

1. **Initialisation:** Each parameter is set to its initial value.

2. **Proposal:** A new value is proposed for every parameter, according to the proposal function described in Section 4.3.2.

3. **Acceptance:** The MC is reweighted to the new set of parameters and the test statistic calculated. The acceptance probability, α, is given for the $(n+1)^{th}$ step by:

$$\alpha = min[1, \log \mathcal{L}_{n+1} - \log \mathcal{L}_n]. \tag{4.9}$$

   A random number is thrown from a uniform distribution in the range [0,1]. If this is $\geq \alpha$ the step is accepted. Otherwise the step is rejected, and the parameters are set back to the previous values.

4. **Repeat:** Steps 2 to 3 are repeated N times.

Steps with an improved likelihood are therefore always accepted. Steps to points with a lower posterior probability are less likely to be accepted, but crucially don't have a non-zero acceptance probability to prevent the chain from getting stuck in local minima. In this way, the algorithm builds a distribution of points in the parameter space, with more points in regions of higher posterior probability and fewer points in regions of lower posterior probability. If the parameter space has been sufficiently explored, the density of points in the final chain is therefore proportional to the posterior probability distribution.

4.3.2. Step Proposal

The Metropolis Hastings algorithm ensures that the Markov Chain will always have a stationary distribution regardless of the form of the proposal function. However, it does not ensure that the chain will converge to the stationary distribution quickly. As only steps after convergence are used to sample the posterior probability distribution, results are obtained more efficiently if convergence is reached sooner.

In this analysis the form of the proposal function is a multivariate Gaussian. The central values are the parameter values at the current step, and the widths are the prior uncertainties multiplied by a scaling factor. The value of the scaling factor can be varied for different parameters, and the values used are tuned with respect to the criteria discussed in Section 4.3.3.

Correlations in the prior uncertainties are included in the proposal function. If this were not the case, steps would be likely go into regions which had low prior probability by contravening the correlations. These steps would then be very likely to be rejected, and so convergence would not be reached efficiently. Including the correlations in the proposal function

encourages the algorithm to step into regions of higher posterior probability, increasing the likelihood of steps being accepted.

4.3.3. Chain Diagnostics

The number of steps to run in the Markov Chain is a predetermined value which can be varied. The choice is usually a trade-off of requiring a small number of steps to reduce the computational expense and obtain results in a reasonable time, while still ensuring the chain has converged and sufficiently sampled the parameter space.

The width of the Gaussian proposal functions, known as the step sizes, can be tuned to encourage faster convergence, and so reduce the total number of steps needed. This is again a trade-off; it is important to thoroughly sample all regions, while not stepping out of areas of high posterior probability too often. One index which can be used to test the tuning of the Markov Chain is the acceptance rate of proposed steps. The recommended optimum acceptance rate is 0.23 [110], though this is for chains of lower dimension than those used in this analysis.

The acceptance rate alone, however, is not a good gauge of how well the step sizes are tuned. There are a number of other tools with which convergence is tested after a fit, alongside the acceptance rate. The autocorrelation for each parameter is the correlation between the parameter values at different steps. For points k steps apart in the chain, referred to as a lag of k, the autocorrelation, a, is given by:

$$a = \frac{\Sigma_{i=1}^{N-k}(X_i - \bar{X})(X_{i+k} - \bar{X})}{\Sigma_{i=1}^{N}(X_i - \bar{X})^2}, \tag{4.10}$$

where X_i is the parameter value for the ith step, $\bar{X}$ is the mean parameter value in the chain, and N is the total number of steps. The autocorrelation for a single flux parameter is shown in Figure 4.1 for a number of different step sizes. Increasing the step size causes the autocorrelation to reduce quicker, though this would decrease the acceptance rate. In general, in this analysis the aim was for the autocorrelation of each parameter to reach < 0.2 at a lag of $10,000$.

The trace of the parameter is the plot of its value at each step in the chain. This is used to ensure good mixing has been achieved, and so the parameter space is well explored. The trace of a low energy flux parameter is shown in Figure 4.2 for different step sizes. This shows how decreasing the step size, affects how the chain samples different regions of the posterior distribution.

Finally, batched means are also used to test the chain has converged. These are the mean
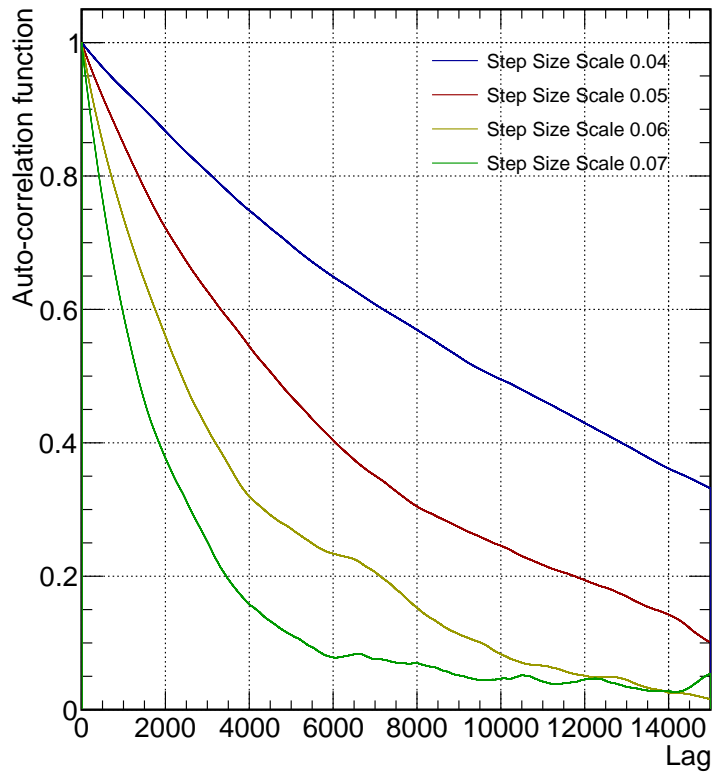
Figure 4.1.: The autocorrelation function for a low energy flux parameter, at different values for the scaling applied to the step size.

of a given parameter over smaller consecutive subsets of steps. Once converged, these should be fairly constant, as shown in Figure 4.3 for a low energy flux parameter. In the second half of the chain, the batched means are within ~2%. If the mean was varying drastically between batches, or if there were regions of consistent bias beyond statistical fluctuation, it would suggest the stationary distribution had not been reached.

In this analysis, the general method used for step size tuning was to individually alter the width of the proposal function for each individual cross-section parameter until the autocorrelations were similar for all parameters. Then a further global scaling was applied to achieve the desired acceptance rate. Step size tuning, however, is not an exact science. When trying to reduce one parameter's autocorrelation, another parameter's could unexpectedly increase due to high dimensional correlations. One could always endeavour to further reduce autocorrelations, but once a reasonable set of tunings was found with good mixing, fairly consistent autocorrelations, and an acceptance rate close to the optimal value, no further time was spent trying to tune further.

Even with a well tuned Markov Chain, it is still desirable to run for as many steps as
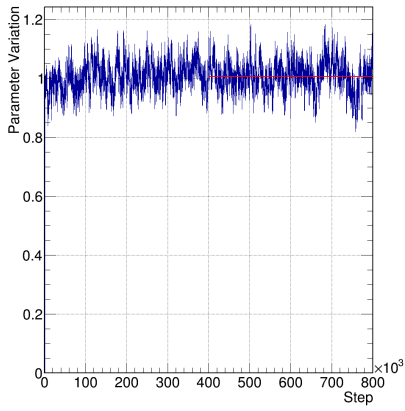
(a) Step Size Scale: 0.03
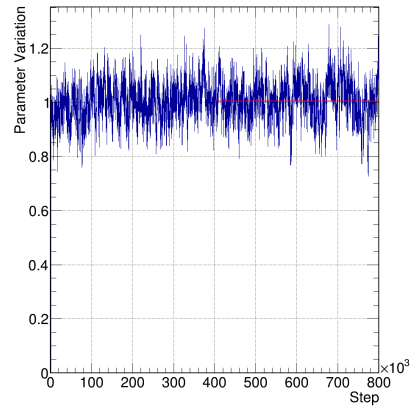
(b) Step Size Scale: 0.04
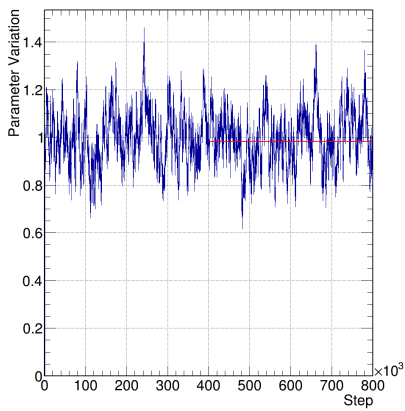
(c) Step Size Scale: 0.075
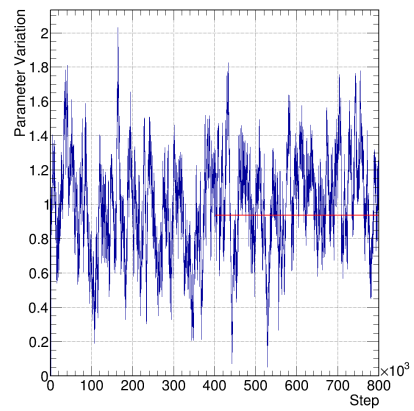
(d) Step Size Scale: 0.10

Figure 4.2.: The traces for a low energy flux parameter for different scalings of the step size. The red lines show the mean for the second half of the chain.
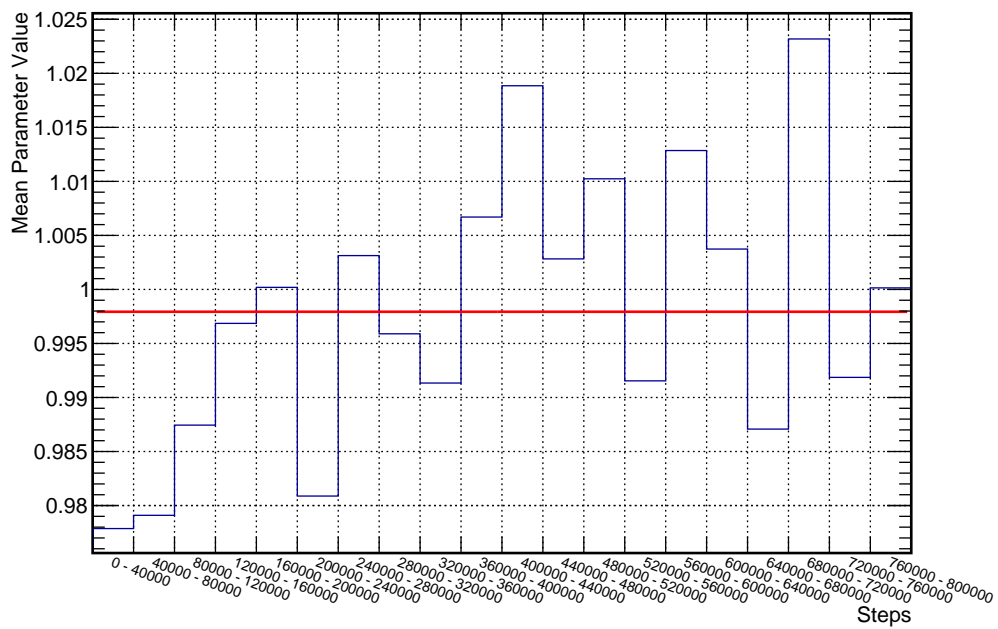
Figure 4.3.: The batched means for a low energy flux at parameter. The red line shows the total mean.
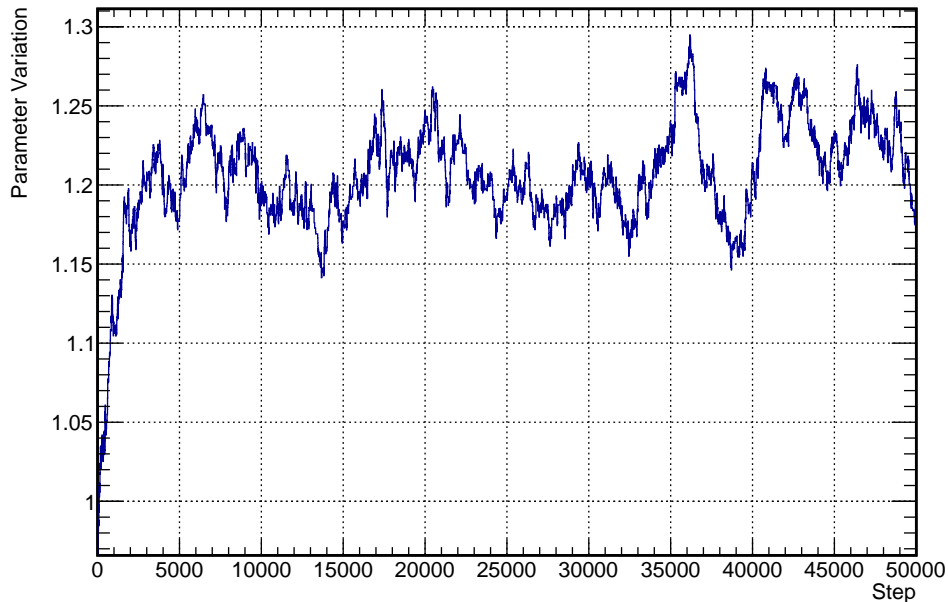
Figure 4.4.: The trace of the first 50,000 steps of high energy flux parameter, showing the initial burn-in phase before reaching the stationary distribution.

possible, to ensure good coverage. In this analysis, six chains of 800,000 steps were run in parallel and then merged. However, not all steps are used in the final results. As the initial values of parameters in the chain are not necessarily in a region of high posterior probability, there is a 'burn-in' period before the chain reaches its stationary distribution.

The batched means and traces for individual parameters can be used to monitor the number of steps in the burn-in. Figure 4.3 shows the means initially being lower than the final converged values, and Figure 4.4 shows the parameter value starting lower than it's final value, but quickly converging after 10,000 steps and then exploring the surrounding region.

The trace of the systematic and sample contributions to the log-likelihood are also a good measure of when the burn-in period has finished. These are shown in Figure 4.5, for six merged chains each with 600,000 steps. The negative log-likelihoods converge once the stationary distribution has been reached after $\sim$20,000 steps.

In general, a large amount of steps from the start of the chain should be rejected, to ensure convergence has been reached for all the steps used. In this analysis, the first 1/4 of steps of all chains are conservatively cut out as burn-in.

(a) Sample contribution

(b) Cross-section penalty contribution

(c) Flux penalty contribution

(d) Detector penalty contribution
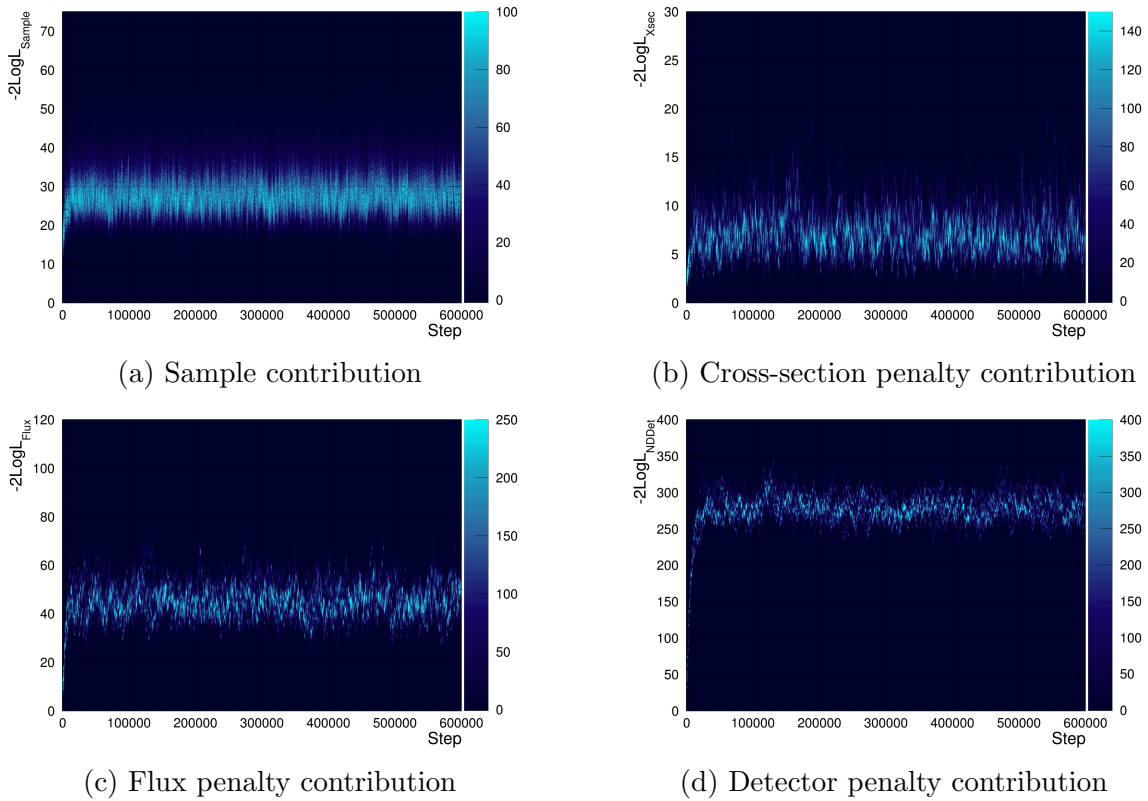
Figure 4.5.: Trace of the different contributions to the LLH for 6 merged chains each of 600,000 steps in total. The LLHs all converge within ∼20,000 steps, though 150,000 are rejected as burn-in to ensure the stationary distribution has been reached.

4.4. Postfit Treatment

In this analysis, the full ∼700 dimensional posterior distribution is the final result which is propagated to the detector. However, interpreting the results for validations before joint fits is not feasible in this high a number of dimensions. Therefore information needs to be removed to be able to intuitively understand the results.

4.4.1. Parameter Value Extraction

Interpreting individual parameter behaviour is achieved by marginalising over all parameters but one, one by one. This is equivalent to integrating the posterior distribution over all parameters but a single parameter of interest. For a parameter θ_i of model $\bar{\theta}$, the marginalised posterior given data, D, is given by:

$$P(\theta|D) = \int P(\bar{\theta}', \theta_i|D) d\bar{\theta}',$$ (4.11)
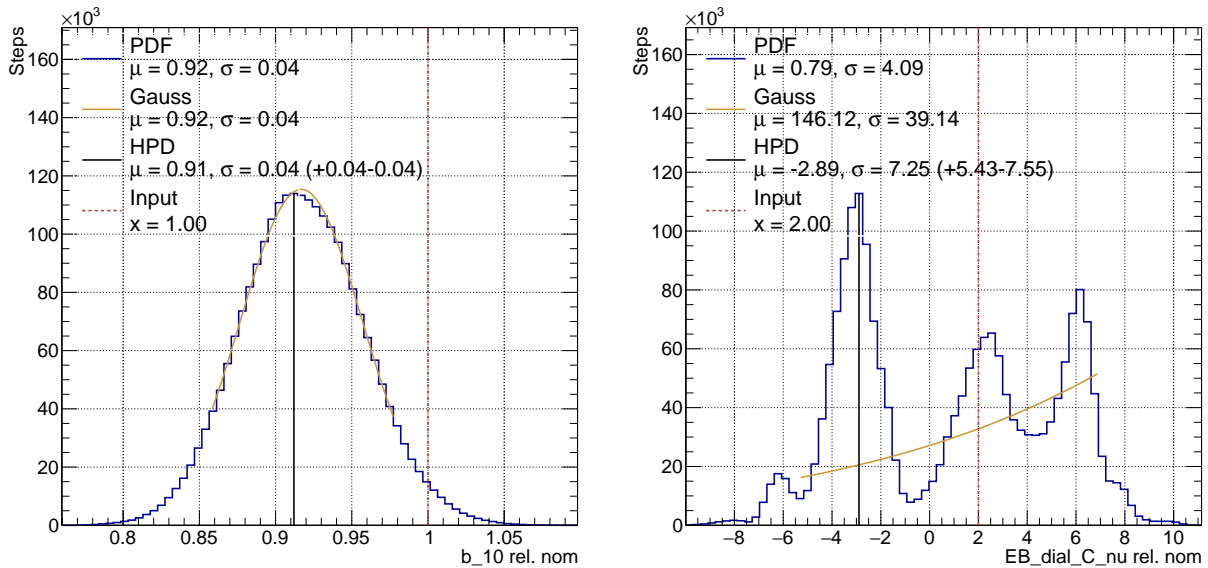
(a) A high energy FHC near detector flux param- (b) The binding energy parameter for neutrino in-
eter.   teractions on carbon.

Figure 4.6.: The 1-dimensional marginalised distribution for two fit parameters, showing the
different methods of parameter extraction. The red lines show the prior central
values, the gold lines show the fitted Gaussian distributions, and the black lines
show the highest posterior density point.

where $\bar{\theta}'$ is the parameter space over all model parameters but θ_i.

Figure 4.6 shows the resulting 1-dimensional projections for two individual fit parameters.
This is equivalent to the parameter value at each step in the Markov Chain after burn-in.
Central values and uncertainties for each parameter are extracted by three different methods.
Firstly, the arithmetic mean and RMS of the histogram are calculated and used as the postfit
parameter value and error. Secondly, the highest posterior density, or mode, of the histogram
is taken as the central value. The number of events in each bin is summed outwards from the
mode to obtain an asymmetric uncertainty. Finally, a Gaussian is fitted to the histogram
and its mean and width are taken as the central value and uncertainty.

For a Gaussian distribution these three methods are equivalent, as shown in Figure 4.6a.
However, non-Gaussian distributions and parameter correlations can lead to non-intuitive
results when marginalised over, moving the region of high probability in the marginal pos-
terior distribution. This is an expected, true effect, which does not indicate any bias in the
fit. The three extracted values are compared to each other, with any differences highlight-
ing non-Gaussian behaviour, as shown in Figure 4.6b. Non-Gaussianity is not necessarily
concerning, but should be understood and considered when interpreting other parameter
results.

(a) Two highly correlated beam parameters. (b) Two less correlated interaction parameters.
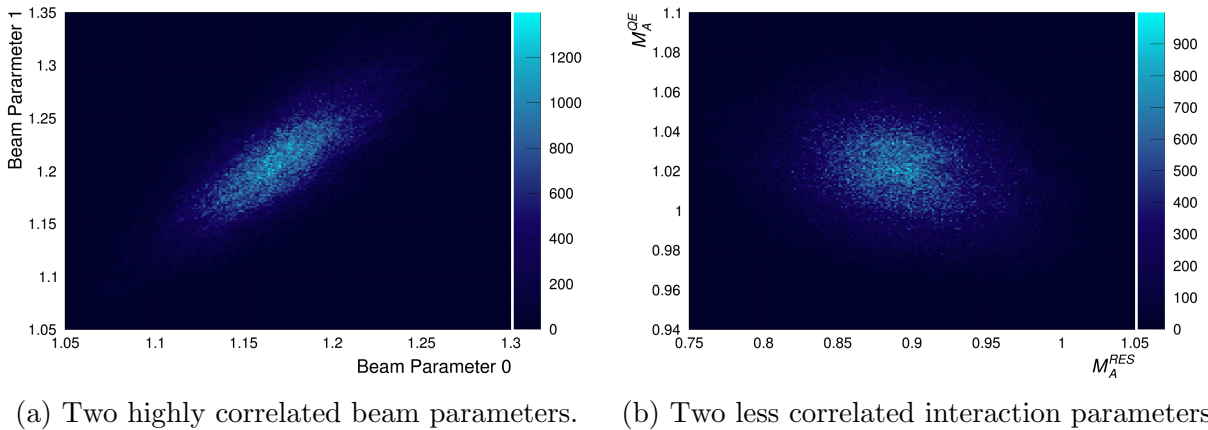
Figure 4.7.: The 2-dimensional marginalised distributions for two pairs of fit parameters.

None of the three extraction methods are incorrect, but none of them contain the whole true result either. Only the full $\sim$700 dimensional distribution contains all the information of the final fit result. As this is what is propagated to SK, differences between the extracted 1D values is not an issue in itself.

The postfit values are also compared to those from the other T2K near detector fitting groups for validation. As the other group finds the single set of parameter values to minimise the test statistic, some differences are expected from marginalisation over non-Gaussian and correlated distributions. However, large discrepancies could indicate differences between the implementation of the fit in the two groups.

4.4.2. Postfit Covariance

A similar process is used to calculate the postfit covariance between parameters. All parameters are marginalised over but two, and this is repeated until each combination of two parameters have been projected onto. Figure 4.7 shows this for two pairs of parameters. This is equivalent to the values of each parameter at each step of the Markov Chain post burn-in. The covariance between the two parameters of interest is calculated from the arithmetic width of the distribution. This is the only method used to extract the covariance, as no shape is assumed for the 2-dimensional posterior.

Although reducing each of these 2-dimensional marginal posteriors to a single number removes information, the covariances are only used for comparisons with the other near detector fitting group and to highlight unexpected strong intra-parameter correlations. The full $\sim$700 dimensional posterior distribution is all that is propagated to the far detector

analysis.

4.4.3. Posterior Predictions

Once the fit has finished, the results are used to produce the posterior predictive distributions with the final fitted parameters. However, as discussed in the previous sections, just marginalising over all but one parameter, one by one, removes significant information from the full posterior. Although this is useful for intuitively interpreting the fit, removing information is not necessary for producing postfit event distributions. The method used in this analysis is as follows:

- Draw 2000 post burn-in steps from the Markov Chain

- For each draw, reweight the MC to the set of parameter values for that step. For each fit bin in each sample, this gives 2000 different number of events

- Fit a Gaussian to the distribution of number of events from all draws in each fit bin

- For each bin in each sample, the mean and width of the fitted Gaussian become the central value and uncertainty for the posterior prediction

In this way, a predictive distribution representative of the draws from the stationary distribution is produced. The comparison of this prediction to the data can be used to see if the model has enough freedom to closely reproduce the data through reweighting. Furthermore, plotting the log-likelihood (LLH) contribution from the posterior prediction to data for each bin in each sample highlights which regions are affecting the LLH calculation the most. It is desired that bins containing the most bins would be having the largest impact on the fit, and when this is not the case it should be fully understood why other regions are having unexpected significance. However, more involved methods are needed to truly test how well the model has been fit to the data.

4.4.4. Goodness of Fit

In this analysis, the distribution of fitted model parameters to data is found. However, this only determines how best to describe the data using the implemented model. If the model does not agree well with the data misleading results can be produced. It is therefore important to be able to assess the goodness of fit for the final result.

As well as comparing the posterior predictive distributions to data, a Bayesian p-value is calculated to see how well the model fits the data. This is done in accordance with the

methods outlined in [111–113]. The 2000 draws used to produce the posterior predictions are used:

- For each fit bin in each sample, the bin contents in the posterior prediction is fluctuated by drawing a random number from a Poisson distribution with a mean equal to the original bin content

- Calculate the LLH between the fluctuated prediction and the prediction

- For each draw:
  - Reweight the MC to the set of parameter values for that step
  - Calculate the LLH between the data and the draw
  - For each fit bin in each sample, the bin contents for the draw is fluctuated by drawing a random number from a Poisson distribution with a mean equal to the original bin content
  - Calculate the LLH between the fluctuated draw and the draw

The LLHs are calculated using Equation 4.8. The Bayesian p-value is then calculated in two different ways:

$$p = \frac{N(-2LLH_{Data,Draw} < -2LLH_{DrawFluc,Draw})}{N_{Total}}, \tag{4.12}$$

and

$$p = \frac{N(-2LLH_{Data,Draw} < -2LLH_{PredFluc,Pred})}{N_{Total}}, \tag{4.13}$$

where $N(-2LLH_{Data,Draw} < -2LLH_{DrawFluc,Draw}$ is the number of draws for which the negative LLH is smaller for the data given the draw than for the fluctuation of the draw given the draw, $N(-2LLH_{Data,Draw} < -2LLH_{PredFluc,Pred})$ is the number of draws for which the negative LLH is smaller for the data given the draw than for the posterior prediction given the fluctuated posterior prediction, and N_{Total} is the total number of draws. The first method gives a measure of how likely we would be to have observed the data we did, or something more extreme, compared to random fluctuations of the model, if the fitted model describes nature. If the selected draws are representative of the posterior distribution, the second method should produce similar p-values to the first.

The p-values can be calculated by plotting the 2-dimensional histograms of $-2LLH_{Data,Draw}$ vs $-2LLH_{DrawFluc,Draw}$ and $-2LLH_{Data,Draw}$ vs $-2LLH_{PredFluc,Pred}$ and

calculating the proportion of steps below the line $y = x$. This is done for each sample individually, and as a total sum for all samples using the LLH contribution from every fit bin.

Neither method of calculating the p-value is individually correct or incorrect, and neither should be interpreted as a binary measure of whether the model can or cannot describe the data. The p-values presented here should therefore not be taken as a single validation of the fit, but used along with final marginalised distribution to interpret the full results. Given that a significant amount of information has to be removed from the full ∼700 dimensional posterior for it to be intuitively understood, using a measure of goodness of fit alongside the marginalised distributions is useful for extracting the full picture of the fit.

Generally a higher p-value is desirable and indicates a better fit to data, but there is no single threshold for which a higher p-value can be determined acceptable for all analyses. There are also deficiencies in this method of calculating the p-value which should be considered when evaluating the goodness of fit. The detector systematics are not varied individually, but instead grouped together and parametrised by their joint effect on event rates in merged fit bins, as described in Section 5.4.3. As these underlying systematics are non-Gaussian, throwing using these merged bins does not necessarily describe the true distribution of these systematics. This could cause the p-value to be lower than would otherwise be measured.

Furthermore, this Bayesian p-value is not the same as a 'traditional' p-value. Here, the p-value answers a very specific question: if the experiment was ran again, how likely is it that data consistent with the post-fit model would be observed? This is, by construction, a stringent test. As SK has much lower statistics than ND280, if this p-value is low it does not necessarily mean that the extrapolation of the near detector result to the far detector is invalid, but the individual p-values for each sample can highlight regions of phase space for which the postfit model is less compatible with the data.

The other near detector fitting group at T2K also produce a p-value, using the method described in [114]. This is a strictly frequentist construction, calculated using throws of the systematics from their priors, giving an indication of how well the prior model can fit the data. As well as this, this p-value uses throws of the underlying detector systematics, rather than the effective bin-by-bin normalisations, giving more accurate variations. The frequentist p-value answers a different question to the Bayesian version: using information on the systematics before the fit is run, how likely is the observed data? The two p-values are therefore not expected to give the same results. For this analysis, the frequentist p-value

is used to determine if an acceptable goodness of fit has been achieved, as it uses the correct detector systematics implementation, while the Bayesian p-value is used to inspect the fit result in more detail and determine for which samples and regions of phase space the fit performs well for.

4.5. Summary

This chapter has presented an overview of the statistical methods used in this analysis. Bayesian statistics are invoked, combining prior information with the likelihood of a sample, to find the set of parameter values which best describe the data.

Markov Chain Monte Carlo is used to fit the systematic uncertainties to data. The Metropolis-Hastings algorithm is used to produce a Markov Chain with a density of points proportional to the posterior probability distribution.

Extensive diagnostics are performed on the Markov Chain to ensure the stationary distribution has been reached. These include step-size tuning and inspecting the autocorrelations, batched means, parameter traces, and burn-in.

Postfit parameter values and uncertainties are extracted from the Markov Chain by three different methods. The first step is to marginalise over all but one parameter, one by one. The arithmetic mean, mode, and mean of a fitted Gaussian to the resulting 1D distribution can then all be used as the parameter's postfit result. The methods can produce different results if the distribution is non-Gaussian, but as the full ~ 700 dimensional distribution is propagated to the far-detector, this is not problematic. The extracted single parameter values are only used for validating the fit and comparing to the other near detector fitting group. Postfit correlation matrices are also produced by marginalising over all but two parameters, for each combination of a pair of parameters.

Posterior predictions are produced by drawing 2000 steps from the Markov Chain and fitting Gaussians to the number of events in each bin, showing the p_μ–$\cos\theta_\mu$ distributions the postfit model produces.

Finally, the goodness of fit is calculated in two separate ways from random draws of the Markov Chain. These are not equivalent to a traditional frequentist p-value, and should not be interpreted as a single binary measure of whether the model can or cannot describe the data. Rather, they are used to determine for which regions of phase space the fit performs well.

5. The Near Detector Fit Setup

5.1. Motivation

The aim of the near detector fit is to constrain systematic uncertainties such that accurate oscillation parameter measurements can be made at SK. The number of neutrinos measured at the far detector, is a convolution of the beam flux, cross-section, detector efficiency, and the probability of oscillation:

$$N_{SK}^{\nu_\beta} = \phi_{SK}^{\nu_\alpha} \cdot \sigma^{\nu_\beta} \cdot \epsilon_{SK}^{\nu_\beta} \cdot P_{PMNS}(\nu_\alpha \to \nu_\beta), \tag{5.1}$$

where $N_{SK}^{\nu_\beta}$ is the number of neutrinos of flavour β, $\phi_{SK}^{\nu_\alpha}$ is the flux of neutrinos of flavour α at SK without oscillation, σ^{ν_β} is the interaction cross-section for neutrinos of flavour β at SK, $\epsilon_{SK}^{\nu_\beta}$ is the detector efficiency of SK for neutrinos of flavour β, and $P_{PMNS}(\nu_\alpha \to \nu_\beta)$ is the probability that a neutrino produced as flavour α oscillates to flavour β before reaching SK.

There are degeneracies between the flux, cross-section, and detector efficiency models, and the PMNS parameters. A change in one of the nuisance parameters mimics the effect of a change in one of the parameters of interest. Therefore it is crucial to constrain the model parameters as much as possible. This is done using near detector data, where the number of neutrinos is measured before oscillation and so the nuisance parameters can be disentangled from the oscillation parameters. The near detector samples also have more data than the SK samples, and using them to constrain nuisance parameters decreases the uncertainty on the predicted event rate at SK from $\sim 15\%$ to $\sim 5\%$. Figure 5.1 shows the SK prediction for the 2018 oscillation analysis with and without the near detector constraint. The uncertainty bands are much narrower when using near detector data in the prediction, showing the significance of the near detector on the full fit. Without the near detector constraint, T2K would not be able to make the world leading oscillation measurements it does.

The parametrised models of the beam, cross-section, and near detector detector efficiencies used in the near detector fit are produced by different groups within T2K. The beam group
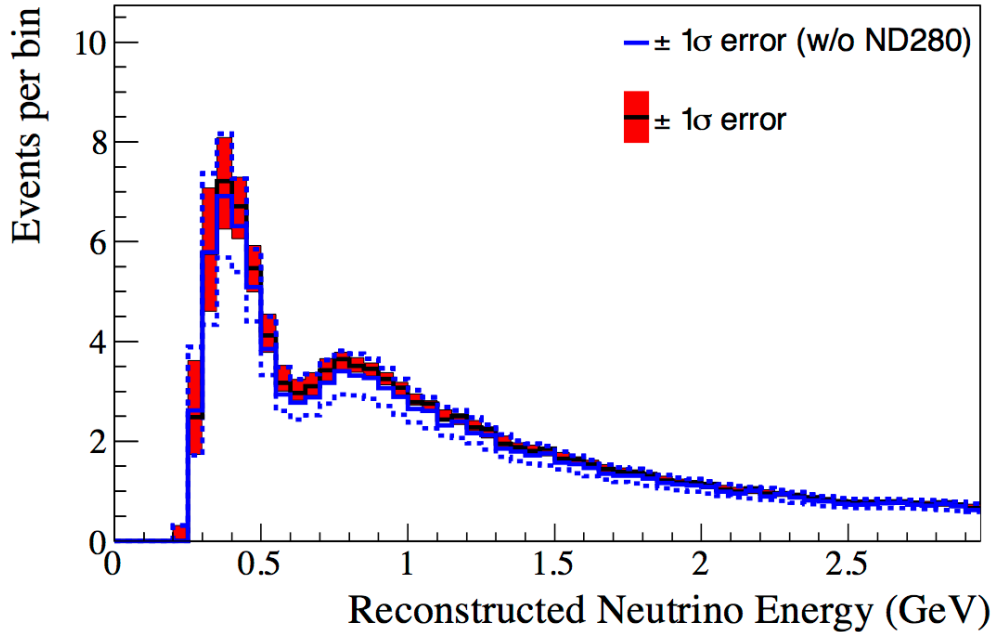
Figure 5.1.: The predicted event rate at SK, with and without the near detector fit constraint from the 2018 oscillation analysis. Using near detector data to reduce systematics narrows the uncertainty on the prediction, allowing more precise oscillation measurements to be made. Figure from [114].

uses data from beamline monitors and the on-axis INGRID to provide the beam flux model, the Systematics, Selections, and Validations (SSV) group provide a model of ND280 detector systematics, and the Neutrino Interactions Working Group (NIWG) provide the interaction model. For the full joint near and far detector fits, the T2K-SK group provides the SK detector model and selections. The parameters of these models affect the log-likelihood in the fit through Equation 4.6.

The MCMC analysis presented in this thesis is performed in the MaCh3 (**Ma**rkov **Ch**ain for a **3** Flavour Oscillation Analysis) fitting framework. MaCh3 is one of two groups that fit near detector data at T2K. The BANFF (**B**eam **A**nd **ND**280 **F**lux Extrapolation Task **F**orce) also perform near detector fits, but use MINUIT [115] to find the global minimum of the likelihood by gradient descent, in contrast to MaCh3's use of MCMC to sample the probability distribution discussed in Section 4.3. The BANFF postfit covariance is propagated to two of the three far detector fitting groups at T2K. This has the disadvantage of assuming that the systematics are Gaussian. MaCh3, by doing joint near and far detector fits, does not assume any shape on the systematics. However, to validate the model and fitting framework, near-detector-only fits are compared with the BANFF.

In this section, the various inputs to the fit are described. The selections and samples used are detailed in Section 5.2, and studies of how they are binned are presented in Section 5.3. The models providing systematic uncertainties are discussed in Section 5.4, and prefit corrections to the MC are described in Section 5.5, before the data used is outlined in Section 5.6.

5.2. Selections

The ND280 selections include both neutrino and anti-neutrino, and FGD1 and FGD2 interactions. Data is then divided by the topology of events, as observed by the detector. The aim of the selections are to group events in such a way that the underlying interaction mode can be identified so that systematics can be applied correctly. They allow us to translate as accurately as possible between what we see in the detector, such as an event with no outgoing pion tracks, to the actual physics of the event, such as a CCQE neutrino interaction. The selection criteria is optimised to select CC-inclusive interactions which originate in FGD1 or FGD2, and contain one reconstructed muon track of negative charge crossing the following TPC.

Events occurring inside FGD1 and FGD2 are separated as they will undergo a different reconstruction procedure, and be affected by different systematics. This is because of the water layers interleaving the plastic scintillators inside of FGD2, which are not present in FGD1. There therefore can be events on oxygen inside FGD2 but not FGD1. This means having separate FGD2 selections can isolate the constraint on systematics affecting events on oxygen, which is important as the far detector is filled with water. As well as this, the geometry of events in the two FGDs is different. Events in FGD1 can leave forward-going tracks which pass through FGD2 and both TPC 2 and 3, whereas events in FGD2 can only leave forward-going tracks in TPC 3.

Events are separated by whether the beam was in FHC or RHC mode as neutrino and anti-neutrino events will be affected by different systematics. For RHC, there are separate selections for wrong sign neutrino events. This is not the case for FHC, as the anti-neutrino cross-section on matter is so much smaller than for neutrinos. As the far detector is not magnetised, it is important to constrain systematics affecting neutrino and anti-neutrino interactions differently at the near detector, and so the separation of these selections is required.

The separation of events by topology allows more accurate identification of the interaction

mode. In this analysis, events are divided into three topological groups: CC 0π, CC 1π, and CC Other. For each FGD there are nine samples, three topologies for each of the three neutrino signs, and so there eighteen near detector samples in total. These samples are binned in the momentum and angle of the final state lepton, as these kinematic variables can be measured with good resolution in ND280. The sample and bin are all the information for a single data event that is used to compare to the MC.

The initial reconstruction of events is performed using an algorithm that fits clusters of hits in each TPC, and then adds hits in the upstream FGD to form a TPC-FGD track. The algorithm is described in more detail in [116]. The specific criteria used to define each selection are detailed in the following section.

5.2.1. FHC ν_μ

The FHC ν_μ selections are designed to initially produce a sample of CC-inclusive interactions which occur in FGD1 or FGD2, and contain one reconstructed muon track of negative charge crossing the following TPC. It is also required that events pass quality cuts and that the highest momentum negative track (HMNT) is identified as a muon. The following cuts define the sample:

- **Data Quality:** The event must occur within bunch time windows of the neutrino beam, and belong to a spill which is entirely flagged as having good global ND280 data quality. Two events occurring within the same spill but different bunches are treated as independent events to avoid pile-up.

- **Total Multiplicity:** There must be at least one reconstructed track crossing the TPC in the event, otherwise there is not enough information to classify the event.

- **Quality and Fiducial Volume:** There must be at least one track reconstructed inside the FGD1 or FGD2 fiducial volume, as this is where the most accurate reconstruction occurs. There must also be at least one track with segments within at least one FGD and at least one TPC. The fiducial volume of the FGDs are defined as:

$$|x| < 874.51 \text{ mm}; |y| < 929.51 \text{ mm}; 136.875 < z < 446.955 \text{ mm}, \qquad (5.2)$$

for FGD1, and:

$$|x| < 874.51 \text{ mm}; |y| < 929.51 \text{ mm}; 1481.45 < z < 1807.05 \text{ mm}, \qquad (5.3)$$

for FGD2. Cuts on x and y also reject events with a vertex within five bars of the edge of the FGD module, and cuts in z reject events in the first module of each FGD. Short tracks with fewer than 19 clusters are also rejected as the reconstruction in the TPC is less reliable.

- **Upstream Background Veto:** Reconstruction failures can cause a muon that started further upstream to produce a track starting in the fiducial volume of one of the FGDs. Events with the second highest momentum track starting at least 150 mm upstream of the muon candidate are rejected to exclude such events. As well as this, events in FGD2 are rejected if a secondary track starts in the fiducial volume of FGD1.

- **Broken Track:** Events can be misreconstructed such that a single muon candidate track originating in the fiducial volume of one of the FGDs is broken into two components, one track only in the FGD and one that starts in the last layers of the FGD and passes through the TPC. The second track is therefore identified as a muon candidate. To cut these events, events with the muon candidate starting within 425 mm of the upstream edge of the FGD are rejected if there is also at least one FGD only track.

- **Muon PID:** Events with the HMNT crossing a TPC, starting inside the fiducial volume of an FGD, and identified as a negative particle, are potential muon candidates. The particle identification (PID) is then determined using the $\mathrm{d}E/\mathrm{d}x$ measured in the TPCs. The energy deposited in the TPC is compared to the amount that would be expected if the particle were a muon, electron, pion or proton, which are shown for positive particles in Figure 3.17. The likelihood for particle type i is calculated by:

$$\mathcal{L}_i = \frac{e^{-Pull_i^2}}{\sum_l e^{Pull_l^2}}, \tag{5.4}$$

where $l = \mu, \pi, e, p$, and the $Pull$ is given by:

$$Pull_i = \frac{\mathrm{d}E/\mathrm{d}x_{measured} - \mathrm{d}E/\mathrm{d}x_{expected,i}}{\sigma(\mathrm{d}E/\mathrm{d}x_{measured} - \mathrm{d}E/\mathrm{d}x_{expected,i})}. \tag{5.5}$$

Electrons are rejecting by requiring:

$$\frac{\mathcal{L}_\mu + \mathcal{L}\pi}{1 - \mathcal{L}_P} > 0.8, \tag{5.6}$$

for tracks with $p < 500$ MeV/c. Protons and pions are cut by requiring:

(a) FGD1 p_μ

(b) FGD1 $\cos\theta_\mu$

(c) FGD2 p_μ
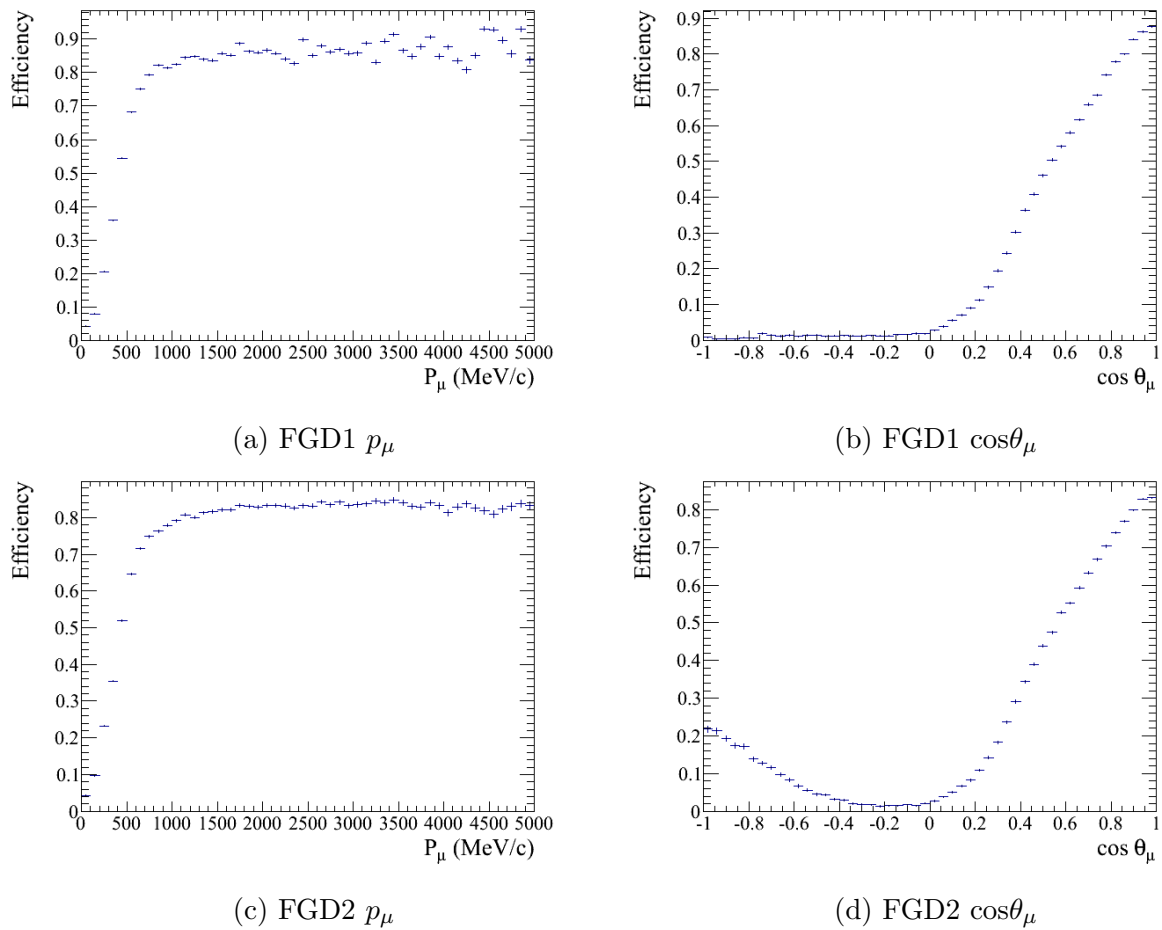
(d) FGD2 $\cos\theta_\mu$

Figure 5.2.: Efficiency of the FGD1 and FGD2 FHC ν_μ CC-inclusive samples, as a function of p_μ and $\cos\theta_\mu$. Figures from [116].

$$\mathcal{L}_\mu > 0.05. \tag{5.7}$$

The efficiency of the FGD1 and FGD2 FHC ν_μ CC-inclusive samples are shown in Figure 5.2. This is defined as the number of selected CC interactions divided by the total number of simulated interactions in the FGD fiducial volume. The efficiencies are higher at larger p_μ and $\cos\theta_\mu$. There is also a small peak at lower $\cos\theta_\mu$ for FGD2.

This CC-inclusive selection is then divided by π multiplicity, which depends on the identification of pions. This is done in the TPCs and FGDs:

- **Secondary Track:** A secondary track separate from the muon candidate must be present in the event. Events with no other track can't have a reconstructed pion, so cannot be treated as CC 1π or Other.

- **Bunch Matching:** The secondary track must be in the same time bunch as the muon candidate track. This cut rejects secondary tracks which are likely not from the same

event.

- **Track Start Matching:** The secondary track must originate from the fiducial volume of the same FGD as the muon candidate track. Secondary tracks starting in a different detector are likely not from the same interaction so are rejected.

- **TPC Matching:** If the secondary and muon candidate tracks start in FGD1, the secondary track must also enter TPC 2. If the secondary and muon candidate tracks start in FGD2, the secondary track must also enter TPC 3. This cut ensures the secondary track is forward-going and long enough to be reconstructed.

- **TPC Quality:** There must be at least 18 clusters in the TPC. This cut ensures the track is large enough to be accurately reconstructed.

- **Pion PID:** The number of charged pions is determined by the number of secondary tracks with PID determined in the TPC corresponding to a pion. For positive tracks, the pion, positron and proton hypotheses are tested. For negative tracks, only the pion and electron hypotheses are tested. The pulls for each hypothesis are calculated and electrons are rejected by requiring:

$$\frac{\mathcal{L}_\mu + \mathcal{L}_\pi}{1 - \mathcal{L}_P} > 0.8, \tag{5.8}$$

for tracks with $p < 500$ MeV/c. Muons and protons are then rejected by requiring:

$$\mathcal{L}_\pi > 0.3. \tag{5.9}$$

The number of neutral pions is determined from the presence of positrons and electrons produced in their decay.

There are two methods by which information from an FGD can be used to identify if a particle with momentum too low or angle too high to enter a TPC is a pion. However, this can only be done for charged pions as electrons and positrons are not reconstructed.

- **FGD Reconstruction:** Secondary tracks in the FGD that don't start in the fiducial volume of the same FGD as the HMNT, or that aren't fully contained within the FGD, are rejected. Tracks that pass this cut and are in the same time bunch as the muon candidate are considered as pion candidates. The deposited energy in the FGD is then used to discriminate charged pions from protons.

    – **Michel Electron:** Lower momentum pions which don't produce a track in an FGD can be identified from the Michel electron produced from the muon produced in the pion decay. To be identified as the delayed signal from a Michel electron, the hit cluster must have at least seven hits in FGD1 or six hits in FGD2, and be outside the beam bunch window[1].

The number of pions identified by the TPC PID, Michel electron tagging, and FGD PID is then used to split the CC-inclusive selection into CC 0π, CC 1π and CC Other:

- **CC 0π:** Contains events with no identified charged pions, electrons or positrons using TPC PID, no Michel electrons or charged pions in either FGD, and one negative muon candidate.

- **CC 1π:** Contains events where the sum of the number of positive pions identified in a TPC and the number of Michel electrons is one. For events with no Michel electrons the sum of positive pions in any TPC or FGD is one. Events with a negative pion, electron, or positron reconstructed in a TPC are rejected. It is also required there is one negative muon candidate.

- **CC Other:** Contains all events in the CC-inclusive selection that do not fall into the CC 0π or CC 1π sample. These are events with one negative muon candidate and either one or more reconstructed negative pions, one or more neutral pions reconstructed as electrons or positrons, or more than one positive pion reconstructed using TPC and FGD information.

5.2.2. RHC $\bar{\nu}_\mu$

The RHC $\bar{\nu}_\mu$ selections are designed to initially produce a sample of CC-inclusive interactions which occur in FGD1 or FGD2, similarly as for the FHC ν_μ samples. However, events must contain one reconstructed muon track of positive charge crossing the following TPC. The selection criteria are similar as those for the FHC ν_μ samples, but have extra cuts to account for the larger wrong sign background due to neutrinos having a larger cross-section on matter than anti-neutrinos:

- **HMT:** The background of ν events producing a π^+ misidentified as a μ^+ are reduced by requiring that the highest momentum positive track (HMPT) is the highest momentum track (HMT) in the event.

---

[1]The muon has a 2.19 μs decay time, causing the delay in the Michel electron signal

- **Upstream Background Veto:** This cut is more stringent for the RHC $\bar{\nu}_\mu$ samples than for FHC ν_μ. Events with tracks entering the fiducial volume of FGD1 from the upstream edge or coming from the P0D or magnet are rejected.

- **Muon PID:** PID is performed using the energy deposited in the TPC and calculating the likelihood for different particle hypotheses, in the same way as for FHC ν_μ samples. However, there are different requirements on the likelihoods for a positive muon candidate to be confirmed. Electrons are rejected by requiring:

$$\frac{\mathcal{L}_\mu + \mathcal{L}_\pi}{1 - \mathcal{L}_P} > 0.9, \tag{5.10}$$

for tracks with $p < 500$ MeV/c. Protons and pions are then rejected by requiring:

$$\mathcal{L}_\mu > 0.1. \tag{5.11}$$

The efficiency of the FGD1 and FGD2 RHC $\bar{\nu}_\mu$ CC-inclusive samples are shown in Figure 5.2. The efficiencies are higher at larger p_μ and $\cos\theta_\mu$.

The pion identification is performed in the same way as for FHC ν_μ, and the CC-inclusive sample is again split into CC 0π, CC 1π and CC Other.

5.2.3. RHC ν_μ

In RHC mode, there are still a significant number of ν events due to the larger ν cross-section, and so a selection of these events is made. The same cuts as for the FHC ν_μ and RHC $\bar{\nu}_\mu$ are applied, but with the following exceptions:

- **HMT:** It is required that the HMNT in the event is the HMT.

- **Upstream Background Veto:** As for the RHC $\bar{\nu}_\mu$ samples, events with tracks entering the fiducial volume of FGD1 from the upstream edge or coming from the P0D or magnet are rejected.

- **Muon PID:** PID is performed in the same way as for the previous samples, using the energy deposited in the TPC and calculating the likelihood for different particle hypotheses. However, there are different requirements on the likelihoods for a negative muon candidate to be confirmed. Electrons are rejected by requiring:

$$\frac{\mathcal{L}_\mu + \mathcal{L}_\pi}{1 - \mathcal{L}_P} > 0.7, \tag{5.12}$$

81

(a) FGD1 p_μ

(b) FGD1 $\cos\theta_\mu$

(c) FGD2 p_μ
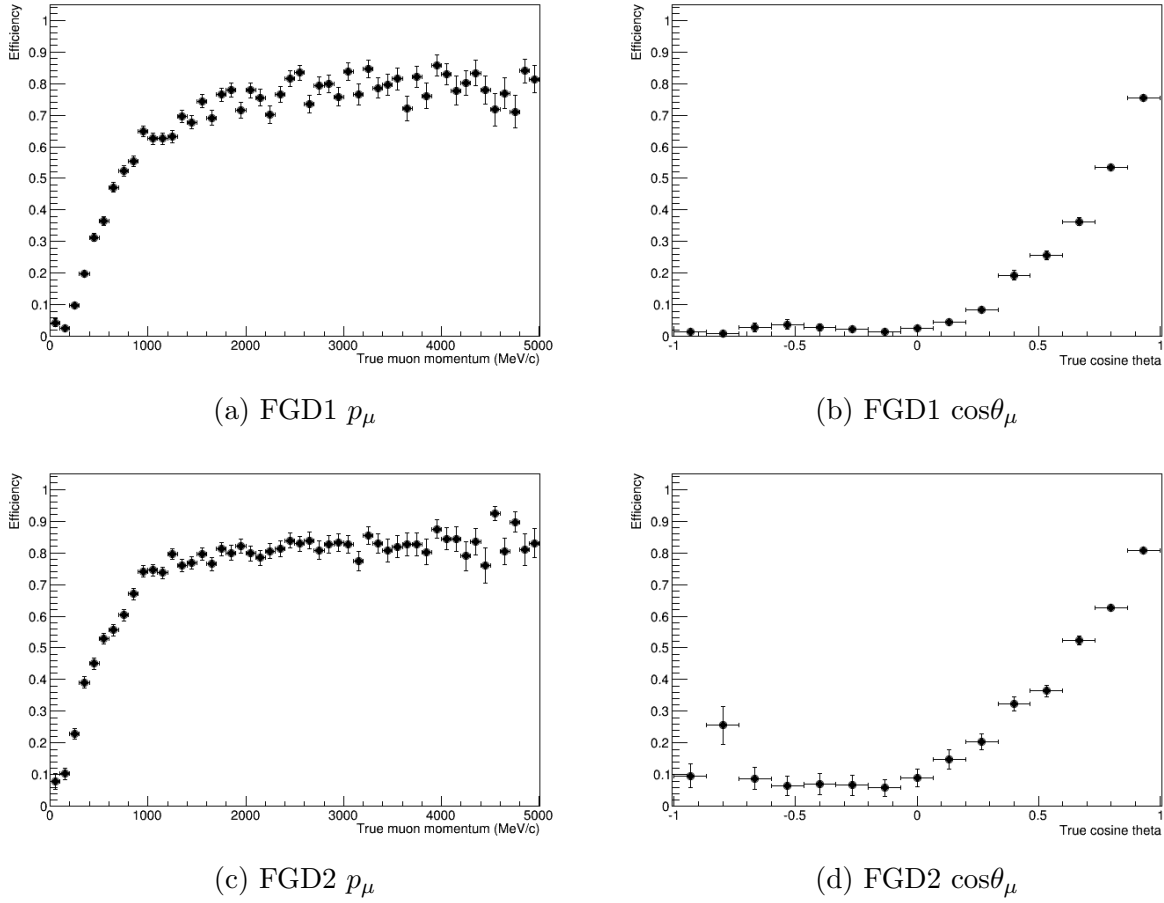
(d) FGD2 $\cos\theta_\mu$

Figure 5.3.: Efficiency of the FGD1 and FGD2 RHC $\bar{\nu}_\mu$ CC-inclusive samples, as a function of p_μ and $\cos\theta_\mu$. Figures from [117].

for tracks with $p < 500$ MeV/c. Protons and pions are then rejected by requiring:

$$\mathcal{L}_\mu > 0.1. \tag{5.13}$$

The efficiency of the FGD1 and FGD2 RHC ν_μ CC-inclusive samples are shown in Figure 5.4. The efficiencies are higher at larger p_μ and $\cos\theta_\mu$.

5.2.4. Updating to RHC Multi π Samples

In previous analyses, the RHC CC-inclusive samples were divided by track, rather than π multiplicity, into CC 1-Track and CC N-Track. Before the 2020 analysis, there were much fewer RHC data events than FHC, and so the RHC samples could not be divided into so many sub-samples. In moving to RHC multi-π, the FHC selection criteria was unchanged, and the RHC selections only changed in the likelihood cuts for rejecting the proton and pion hypotheses in muon identification:

(a) FGD1 p_μ

(b) FGD1 $\cos\theta_\mu$

(c) FGD2 p_μ
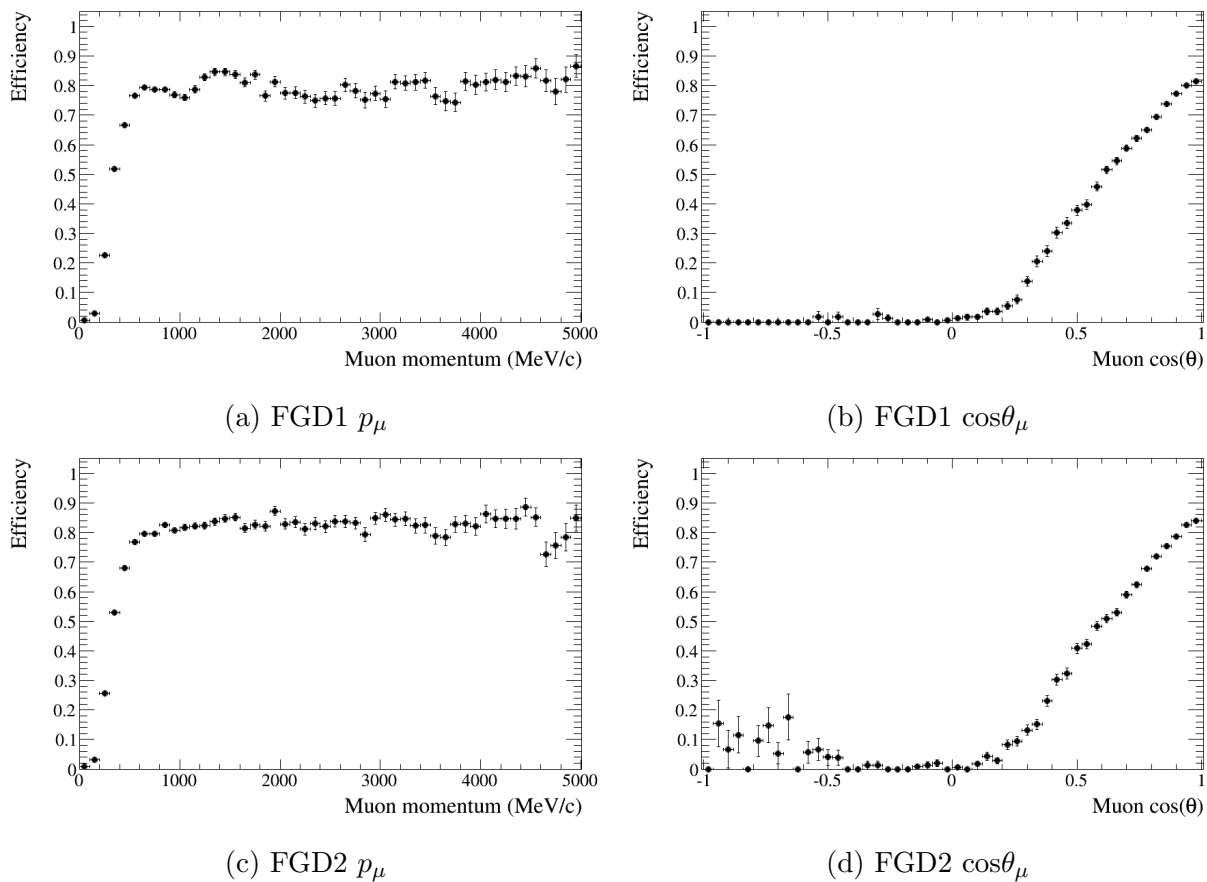
(d) FGD2 $\cos\theta_\mu$

Figure 5.4.: Efficiency of the FGD1 and FGD2 RHC ν_μ CC-inclusive samples, as a function of p_μ and $\cos\theta_\mu$. Figures from [117].

$$0.1 < \mathcal{L}_\mu < 0.7. \tag{5.14}$$

The upper bound was designed to reject low momentum wrong sign muons in RHC events, which could be misreconstructed as positive tracks.

The previous RHC sample splitting proceeded by selecting events with one positive muon and no charged or neutral pions in the CC 1-Track sample, and selecting all other CC-inclusive events in the CC N-Track sample. Validations of the updating of the fitting framework to accommodate the RHC multi-π samples are presented in Appendix J.

5.3. Binning

ND280 events are binned in the reconstructed momentum and angle of the outgoing lepton. The choice of binning is a trade-off of having coarse enough bins to have enough events in each to reduce the statistical error, while being fine enough to have good resolution in the

peak regions. In general, the aim is to have >20 raw MC events in every bin, which is approximately equivalent to having ~1-2 data events. Achieving this using uniform rectangular binning can result in bins outside the region of interest containing the largest amount of events, and contributing the most to the sample log-likelihood. Figure 5.5 shows this effect in the FHC CC 0π and CC 1π sample binnings used for the 2017 oscillation analysis. These plots only show up to 5000 MeV to show the peak regions more clearly. The full binning used here was the same for FGD1 and FGD2, and is as follows:

- **FHC ν_μ CC 0π:**
  $\boldsymbol{p_\mu}$ **(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 θ_μ: -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.

- **FHC ν_μ CC 1π:**
  $\boldsymbol{p_\mu}$ **(MeV/c):** 0, 300, 350, 400, 500, 600, 650, 700, 750, 800, 900, 1000, 1100, 1200, 1500, 2000, 3000, 5000, 30000.
  cos θ_μ: -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.

- **FHC ν_μ CC Other:**
  $\boldsymbol{p_\mu}$ **(MeV/c):** 0, 300, 400, 500, 600, 650, 700, 750, 800, 900, 1000, 1100, 1250, 1500, 1750, 2000, 3000, 5000, 30000.
  cos θ_μ: -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.

- **RHC $\bar{\nu}_\mu$ CC 0π:**
  $\boldsymbol{p_\mu}$ **(MeV/c):** 0, 300, 400, 500, 550, 600, 650, 700, 750, 800, 900, 1000, 1100, 1200, 1500, 2000, 4000, 30000.
  cos θ_μ: -1, 0.6, 0.7, 0.8, 0.85, 0.9, 0.92, 0.93, 0.94, 0.95, 0.96, 0.965, 0.97, 0.975, 0.98, 0.985, 0.99, 0.995, 1.

- **RHC $\bar{\nu}_\mu$ CC 1π:**
  $\boldsymbol{p_\mu}$ **(MeV/c):** 0, 500, 700, 900, 1300, 2500, 30000.
  cos θ_μ: -1, 0.7, 0.8, 0.9, 0.94, 0.96, 0.98, 0.99, 1.

(a) FGD1 FHC ν_μ 0π (b) FGD1 FHC ν_μ 1π

Figure 5.5.: Uniform rectangular binning of the FGD1 CC 0π and FGD1 CC 1π MC samples for T2K runs 2–8.

- **RHC $\bar{\nu}_\mu$ CC Other: p_μ (MeV/c):** 0, 600, 800, 1000, 1250, 1500, 2000, 4000, 30000.

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

- **RHC ν_μ CC 0π:**

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

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

- **RHC ν_μ CC 1π:**

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

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

- **RHC ν_μ CC Other:**

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

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

For example, in the FGD1 CC 0π sample, the bins at low angle, (< 0.6), and low momentum, ($\sim$500 MeV), the bins span a large angle range (-1.0 - 0.6). Having such large bins causes them to contain a large amount of events. However, if they were divided to only span a smaller range of angles, the bins at the same angle but higher momentum would also be divided. As these regions are much more sparsely populated, splitting them further would result in there being insufficient number of events in those bins. The same is true for the high angle bins at higher momentum. These bins cover a large momentum range and are well populated; there are a relatively large amount of events with high momentum forward-going muons. Splitting these bins though would mean the backward-going bins at the same high

momentum would not be large enough and so would be too sparsely populated. This effect becomes more significant with the addition of more data.

5.3.1. Non-Uniform Rectangular Binning Studies

For the 2020 analysis, the MaCh3 near detector framework was updated to be able to use non-uniform rectangular binning[2]. This means the bins can now be any arbitrary shape, including the original uniform rectangular binning which can still be used for validations of the changes to the fitting framework, and cross-group checks with the other near detector fitter which did not move to non-uniform rectangular binning.

The following algorithm was used to define a non-uniform but still rectangular binning for each sample, without having to have the same binning for FGD1 and FGD2:

- The bin edges on the $\cos\theta_\mu$ axis are hard coded and constant, guided by the previous binning, keeping the bins rectangular. The reasons for this are discussed later in this section.

- For each $\cos\theta_\mu$ row, scroll across from 30 GeV down 0 GeV in 100 MeV steps.

- Once 50 unscaled MC events are reached, start a new bin.

- If the last bin in a row (the lowest momentum bin) has <20 unscaled MC events, merge with the previous bin. Scroll through this merged bin in 5 MeV steps and split once half the events in the bin are reached.

The algorithm was tuned using data and MC from runs 2–6. The aim of this process was to produce as uniform a distribution of events across the bins as possible. This was not always possible, as regions of high density would require bins so small that they go below the resolution of the detector. The hard coded $\cos\theta$ bin edges and momentum step sizes of 100 MeV were driven by this minimum bin size limit. The resolutions are calculated by plotting the reconstructed vs true kinematic variables, as shown in Figure 5.6, and taking the RMS at different slices of the 2D Gaussian. These RMSs, shown in Figures 5.7, give a gauge of the detector resolution, and so the minimum bin width for each variable in different regions.

For momentum, the RMS is fairly constant at approximately 100 MeV, for a momentum >1000 MeV. It then reduces linearly between 1000 and 400 MeV, before levelling off at approximately 60 MeV below 300 MeV. Similarly, for the angle, the RMS is constant at approximately 0.08 below 0.96. It then reduces linearly up to a $\cos\theta_\mu$ of 1.0. However, given

---

[2]Mechanically, this meant using ROOT [118] TH2Poly objects instead of TH2Ds.

(a) Final state lepton momentum.

(b) Final state lepton angle.
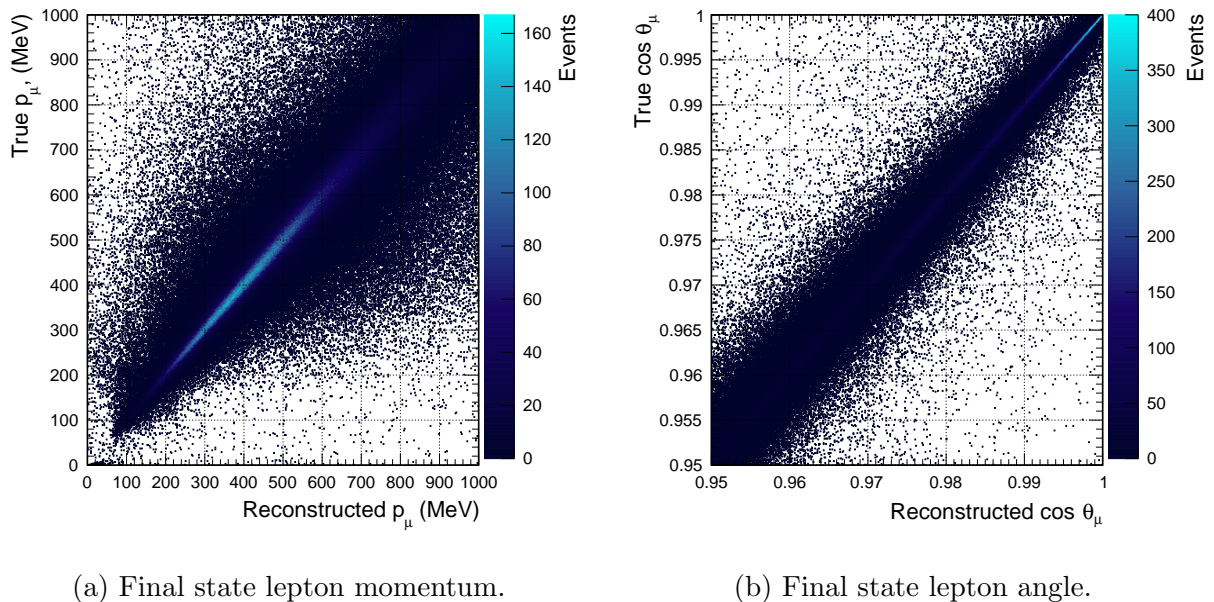
Figure 5.6.: True vs reconstructed lepton kinematic variables of CC-inclusive MC events from T2K runs 2–8.

there would be a large uncertainty on the gradient for each variable's RMS, rather than varying the minimum bin size to trace out the change in RMS as closely as possible, it was safer to have a constant minimum size for all regions. This was chosen to be 0.01 $\cos\theta_\mu$ in angle, and 100 MeV in momentum to be sure they are above the resolution in all regions. Furthermore, as the $\cos\theta_\mu$ systematics are better controlled than for p_μ, and for aiding the simplicity of optimising the binning algorithm, the bins were kept constant in $\cos\theta_\mu$. Although these bins are conservatively large in regions with the most data, studies in reducing the bin sizes showed diminishing returns in improvements to sensitivity.

The distribution of events binned by the scheme produced using the algorithm for the FGD1 CC 0π and CC 1π samples are shown in Figure 5.8. Full templates of the binning for each sample are shown in Appendix B. Particularly for the FHC ν_μ and RHC $\bar{\nu}_\mu$ CC 0π samples, the bins containing the largest amount of events are in the peak regions at high angle and $\sim$500 MeV. However, what is of more importance is the significant reduction in ranges of the z axis scales compared to the uniform binning. For example, the non-uniform binning for the RHC $\bar{\nu}_\mu$ CC 0π samples may not look like the most aesthetic representation of distributions, but the bin with the most events contains less than double the amount of the bin with the least. For the uniform binning, the bin with the most events contains $\sim$250 times more than the amount in the bin with the least events. Reducing this range prevents

(a) Final state lepton momentum.

(b) Final state lepton angle.
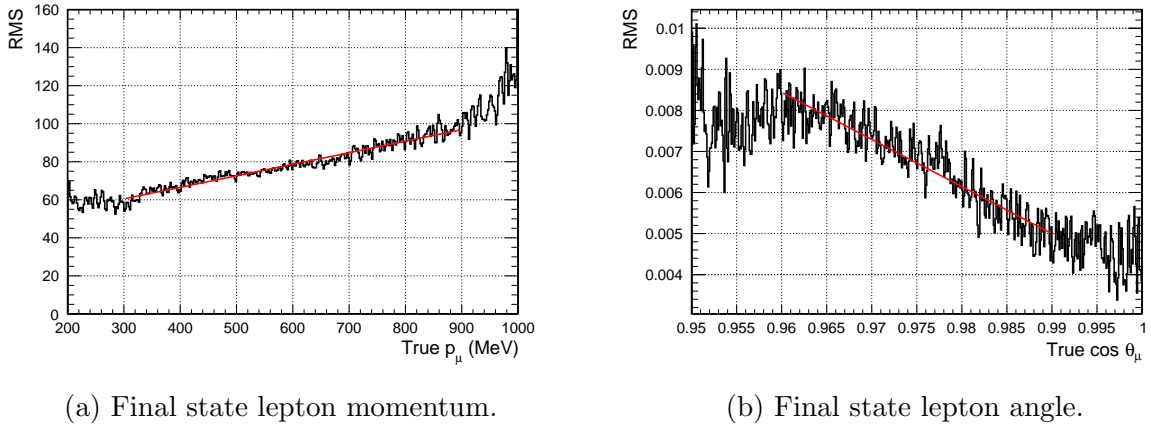
Figure 5.7.: The RMS of the true vs reconstructed lepton kinematic variables for CC-inclusive MC events from T2K runs 2–8, at different values of the true variables.

(a) FGD1 FHC ν_μ 0π

(b) FGD1 FHC ν_μ 1π
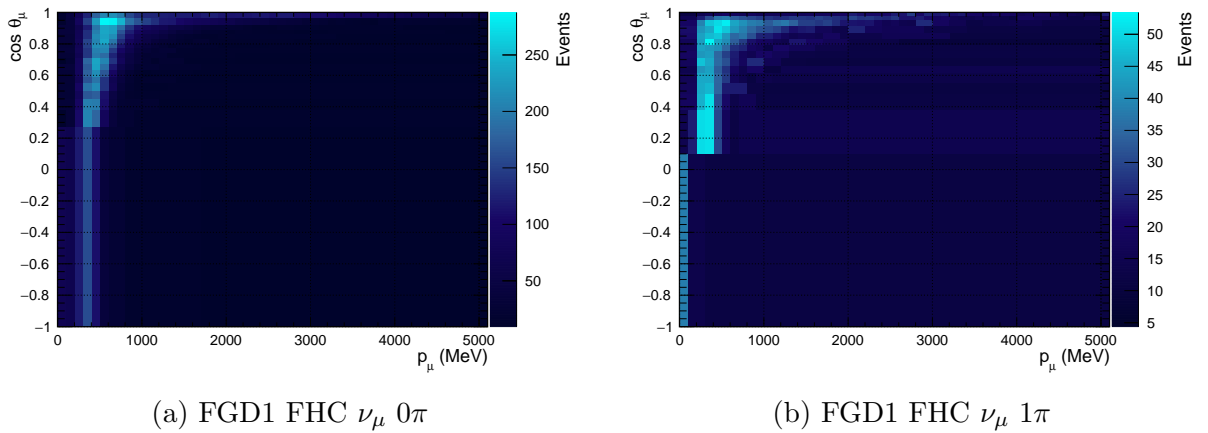
Figure 5.8.: Non-uniform rectangular binning of the FGD1 CC 0π and FGD1 CC 1π MC samples for T2K runs 2–8.

bins outside the peak contributing more to the LLH than the region of interest.

Using non-uniform binning significantly improves the sensitivity of the fit to changes in parameter values. This is evident in the log-likelihood scans shown in Figure 5.9 for two selected interaction and flux parameters. A single parameter is set to different values while all others are kept at nominal, and the sample contribution to the LLH between the reweighted MC and nominal MC is calculated for each parameter value. The LLH scan process, and the parameters themselves, are described in more detail in Section 6.2 and Section 5.4 respectively.

The narrower likelihood distributions show that by using the non-uniform binning, moving a parameter further away from the nominal value will be less favoured than by using uniform binning. The ratio panel shows that the improvement is fairly constant across the range of the scan.

(a) 2p2h Shape ^{16}O

(b) Pion FSI Single Charge Exchange

(c) ND RHC ν_μ 2500-30000 MeV Flux

(d) ND RHC $\bar{\nu}_e$ 1500-2500 MeV Flux
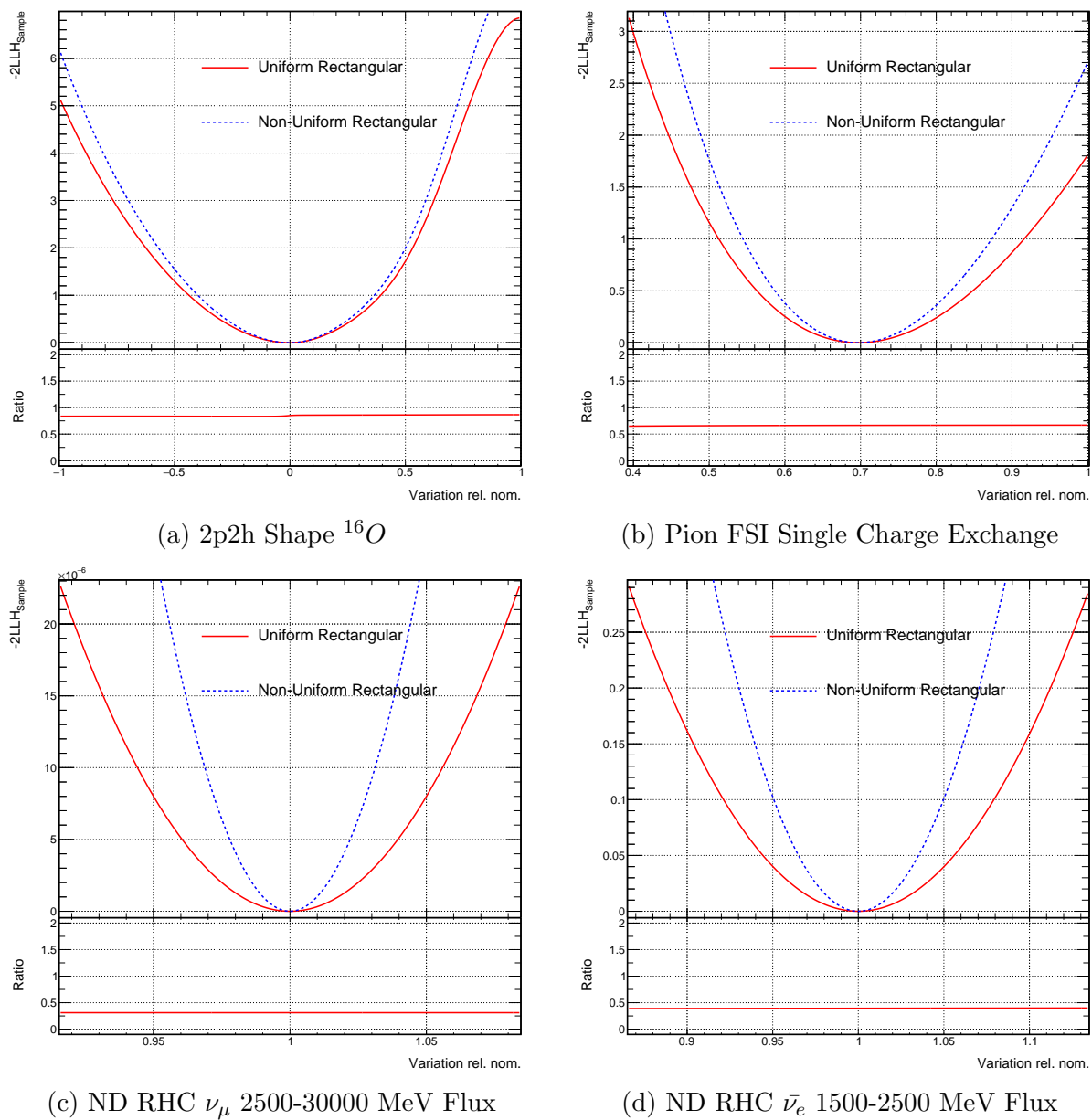
Figure 5.9.: Comparison of LLH scans using uniform and non-uniform rectangular fit binning, for two selected interaction and beam parameters.

Further studies could be undergone to develop a non-rectangular fit binning. For instance, given the underlying distribution of events, using bins of constant Q^2 in the p_μ–$\cos\theta_\mu$ space may allow a better representation of events. However, given that the bin sizes are already close to the resolution of the detector, and that several iterations of non-uniform rectangular binning showed diminishing returns on improvements in sensitivity, this was not investigated for this analysis.

5.4. Systematics

The purpose of the near detector fit is to constrain systematic uncertainties so that accurate oscillation measurements can be made at the far detector. It is therefore vital that these systematics are understood.

In the near detector analysis, there are three sources of systematic uncertainties:

- **Interaction:** Systematics from uncertainties on neutrino interaction cross-sections.

- **Flux:** Systematics from uncertainties on the neutrino beam flux.

- **Detector:** Systematics from uncertainties on the ND280 detector response and reconstruction.

The models of these three groups of systematics are parametrised, and each enter the fit through a covariance matrix. As a parameter is pulled from its prior value, a penalty is added to the likelihood, as shown in Equation 4.6. For systematics that are known to be constrained from external data, the likelihood penalty takes the form of a Gaussian. If there is no, or conflicting, data about a systematic, the likelihood penalty is just a constant. This is referred to as a 'flat' prior uncertainty.

5.4.1. Interaction

The MC prediction is produced using the NEUT 5.4.0 generator [103], as discussed in Section 3.2.3. The model for each interaction mode is therefore based on the models implemented in NEUT, with several modifications to tune to external data and additional theoretical calculations.

The uncertainties of these models are applied to the interaction modes they correspond to on an event by event basis. They are each parametrised as either shape or normalisation uncertainties. For normalisation parameters, the weight applied to the event is just the value of the parameter. However, for shape parameters, the weight applied at different values of the parameter depends on the kinematic variables of the event. Splines[3] are used to translate from a change in parameter to a change in weight. These are produced by evaluating the change in weight for an event at evenly spaced values of the parameter, and interpolating between them.

The parametrisation for each interaction mode proceeds using the following models:

---

[3]The ROOT TSpline3 class is used here.

- **CC Quasi-Elastic:**

  The nominal MC is generated using a Spectral Function (SF) nuclear model from [119]. There is one splined parameter, M_A^{QE}, for the axial mass in the dipole form factor. It's prior is informed by bubble chamber data [120].

  There are also eight normalisation parameters applied, for different bins in Q^2. The lowest five have width 0.05 GeV^2, from 0.00 GeV^2 up to 0.25 GeV^2. The highest three span the ranges from 0.25-0.50 GeV^2, 0.50-1.00 GeV^2, and > 1.00 GeV^2. Their prior central values and uncertainties are tuned by MINERvA [121] data.

  The neutrino energy of CCQE events is calculated using Equation 2.30, which is highly dependent on having an accurate value of the binding energy of the target nuclei. Four parameters are used to fit the binding energies (for target ^{12}C/^{16}O and $\nu/\bar{\nu}$). These are neither shape nor normalisation parameters, and are described in more detail in Section 5.4.1.1.

- **2-particle-2-hole:**

  2p2h interactions are generated in NEUT using the Nieves model [122]. Two shape parameters apply to 2p2h interactions, one for events on ^{12}C and one for events on ^{16}O. At one extreme the distribution of energy and momentum transfer is entirely non-pionless-Δ-decay-like in accordance with the Nieves model, and at the other extreme tit is entirely pionless-Δ-decay-like, in accordance with the Martini model [123]. These two shape parameters have a 30% correlation.

  There are additional shape parameters, to account for differences in the energy dependence of 2p2h interactions. At one extreme the shape is entirely consistent with the Nieves model, and at the other it is consistent with the Martini model. There are four of these parameters in total: $\nu/\bar{\nu}$ and high/low E_ν. However, studies showed these could not be constrained using near detector data alone, so are fixed in near detector only fits. For the joint near and far detector fits in the oscillation analysis, the parameters are free in the fit.

  Three normalisations are also applied, one for ν events, one for $\bar{\nu}$, and one for ^{12}C $\to$^{16}O scaling. The latter is applied to events on ^{16}O multiplicatively with the other normalisations.

- **CC Resonance:**

  The Rein-Sehgal model [124] is used to describe resonant π production in NEUT. Splines are used to parametrise the resonance axial mass, M_A^{Res}, the normalisation of

the axial form factor, C_5^A, and the size of the $I = 1/2$ non-resonant background to $I = 3/2$ resonant π production. The prior central values and uncertainties are tuned using bubble chamber data [125]. There is a prior anti-correlation of 83% between M_A^{Res} and C_5^A, 1% between M_A^{Res} and $I = 1/2$, and 31% between C_5^A and $I = 1/2$. In the joint near and far detector fits, the $I = 1/2$ parameter is split into two, one for anti-neutrino events with $p_\pi < 200$ MeV, and one for all other $I = 1/2$ non-resonant background events. In the near detector only fit, the low p_π parameter is not fit, and the other $I = 1/2$ parameter applies to all $I = 1/2$ non-resonant background events.

- **CC Coherent Scattering:**

  The Rein-Seghal model [126] is used to describe coherent scattering events. However, measurements at MINERνA show a 30% difference in cross-section from this model. Two normalisation parameters are fit, one for CC events on ^{12}C and one for CC events on ^{16}O, each with a 30% prior uncertainty to account for this discrepancy. These are 100% correlated.

- **CC Deep Inelastic Scattering and Multi-π:**

  The CC DIS and multi-π cross-section is calculated from 'Structure Functions' of the nucleus, which themselves are constructed using 'Parton Distribution Functions' (PDFs). The PDFs describe the probability of finding a quark with a given fraction of the nucleon momentum inside the nucleon. In NEUT, these are constructed using the GRV98 [127] model with corrections from Bodek and Yang [128]. Two shape parameters are applied, one for DIS and one for multi-π events, to account for uncertainty in the reliability of these corrections. One extreme corresponds to fully applying the corrections, and the other corresponds to not applying them at all.

  Another shape parameter is applied to multi-π events, to account for differences in the π multiplicity models in different generators. If the π multiplicity model changes, this directly alters the multi-π cross-section (as it is required multi-π events contain ≥ 2 πs). At one extreme, the multi-π cross-section is entirely reweighted to the AGKY model [129] which has a smooth transition between low and high W^2 parametrisations, and at the other extreme it is entirely the nominal custom model in NEUT which has a hard cut-off between high and low W^2 parametrisations.

  Two normalisation parameters are applied to CC DIS and multi-π events, one for ν and one for $\bar{\nu}$ interactions. This is to account for a difference between the high energy CC-inclusive cross-sections in NEUT and the PDG world average. The prior uncertainty

is 3.5% for ν and 6.5% for $\bar{\nu}$.

- **CC Miscellaneous:**

  A normalisation parameter with 100% uncertainty is applied to CC $1K$, 1η, and 1γ events.

- **CC Inclusive:**

  Two normalisations are applied to CC interactions with $0.4 < E_\nu < 0.6$ GeV, one for ν and one for $\bar{\nu}$ events. This is to account for the fact that the relative effect of the Coulomb corrections, described in Section 5.5, is smaller at higher momentum. The uncertainties on these systematics are 2% and 1% for ν and $\bar{\nu}$ respectively, and they are 100% correlated.

- **Neutral Current:**

  As for CC Coherent scattering, NC scattering events receive a normalisation with 30% uncertainty. NC 1γ interactions are modelled with the Rein-Sehgal CC Res model. The cross-section is half the value calculated with more recent models [130], and so the prior weight applied is 2.0. As there is no external data to constrain the cross-section, the prior uncertainty is 100% of the nominal weight.

  NC π production is simulated using the same Rein-Sehgal model as for CC π production, and the same parameters are applied: C_5^A, M_A^{Res} and $I = 1/2$.

  NC DIS, multi-π, 1η, and $1K$ are grouped together and receive the same normalisation parameter (NC Other), with a 30% uncertainty. In the full joint analysis, separate NC Other parameters are applied to near and far detector events. These parameters are uncorrelated.

- **Electron (Anti-)Neutrino:**

  As all the other systematic uncertainties are determined for ν_μ or $\bar{\nu}_\mu$ interactions, a normalisation is applied to all ν_e and $\bar{\nu}_e$ events. This is to account for any unmodelled effects which affect $\nu_e/\bar{\nu}_e$ events but not $\nu_\mu/\bar{\nu}_\mu$. As these processes may be different for neutrinos and anti-neutrinos, there are two normalisations, one for ν_e and one for $\bar{\nu}_e$. The prior uncertainty is 2.8% and the two parameters are 50% anti-correlated, as calculated in [131].

- **Final State Interactions:**

  The propagation of πs produced in neutrino interactions is simulated as a cascade implementation of the model described in [132]. There are five shape parameters
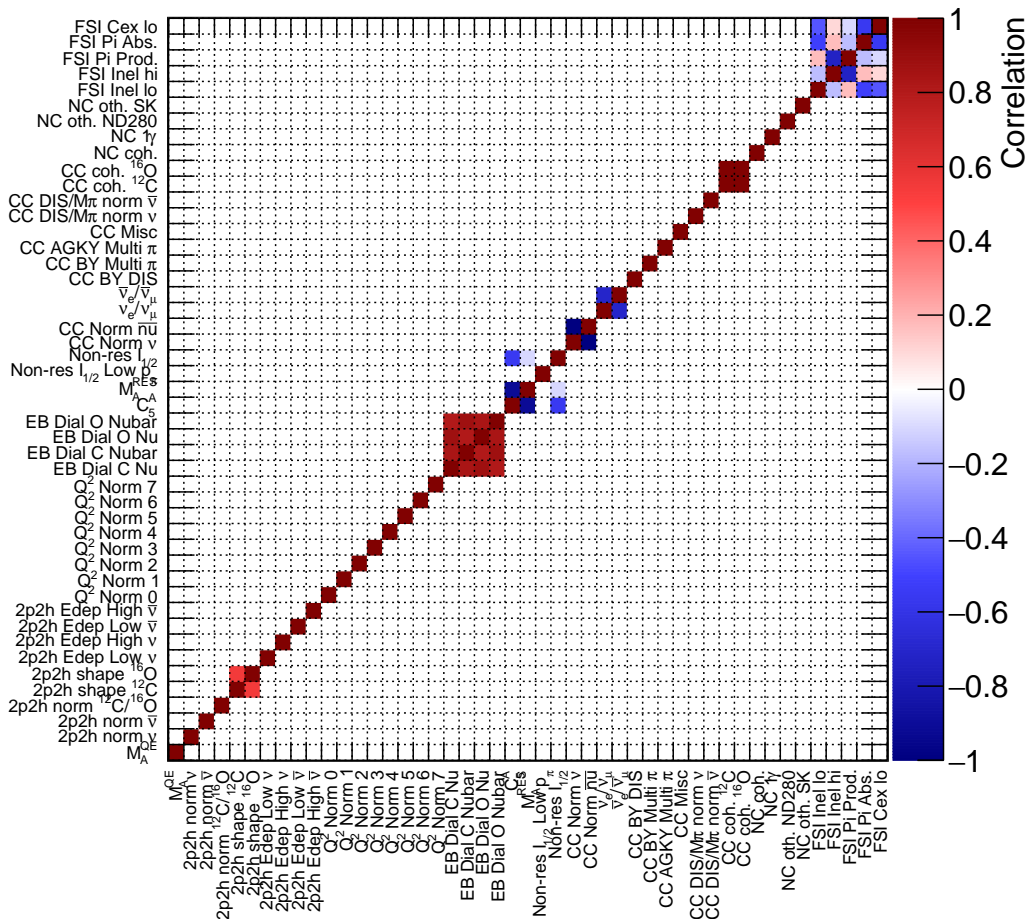
Figure 5.10.: The cross-section correlation matrix.

representing the probability of different interactions at each step in the cascade. The five types of final state interaction parametrised are quasi-elastic scattering (at low and high energy), π production, π absorption, and charge exchange. The prior central values and uncertainties are tuned to π-nucleon scattering data [133].

When nucleon final state interactions produce πs, the πs are propagated with the above systematics. However, nucleon final state interactions are not accounted for in the MC.

A summary of the interaction parameters is shown in Table 5.1. The full prefit cross-section correlation matrix is shown in Figure 5.10, showing the central values, uncertainties, and correlations of the parameters described in this section.

5.4.1.1. Binding Energy

Energy is required to remove a nucleon from a nucleus, to overcome the nucleus force holding the nucleus together. This binding energy can be measured in electron scattering experi-

Parameter	Events	Prior (Nominal) Central Value	Prior Uncertainty	Prior Shape	Type
M_A^{QE}	CCQE	1.03 (1.21) GeV	0.06 GeV	Gaus	Shape
2p2h Norm ν	2p2h, ν	1.0	1.0	Flat	Norm.
2p2h Norm $\bar{\nu}$	2p2h, $\bar{\nu}$	1.0	1.0	Flat	Norm.
2p2h Norm ^{12}C to ^{16}O	2p2h on ^{16}O	1.0	0.2	Gaus	Norm.
2p2h Shape ^{12}C	2p2h on ^{12}C	1.0	3.0	Gaus	Shape
2p2h Shape ^{16}O	2p2h on^{16}O	1.0	3.0	Gaus	Shape
2p2h E dep low E ν	-	-	-	-	-
2p2h E dep high E ν	-	-	-	-	-
2p2h E dep low E $\bar{\nu}$	-	-	-	-	-
2p2h E dep high E $\bar{\nu}$	-	-	-	-	-
Q^2 **0**	CCQE; $0.0 < Q^2 < 0.05$ GeV^2	0.495 (1.0)	1.0	Flat	Norm.
Q^2 **1**	CCQE; $0.05 < Q^2 < 0.10$ GeV^2	0.695 (1.0)	1.0	Flat	Norm.
Q^2 **2**	CCQE; $0.1 < Q^2 < 0.15$ GeV^2	0.78 (1.0)	1.0	Flat	Norm.
Q^2 **3**	CCQE; $0.15 < Q^2 < 0.2$ GeV^2	0.89 (1.0)	1.0	Flat	Norm.
Q^2 **4**	CCQE; $0.2 < Q^2 < 0.25$ GeV^2	0.93 (1.0)	1.0	Flat	Norm.
Q^2 **5**	CCQE; $0.25 < Q^2 < 0.5$ GeV^2	1.0	0.11	Gaus	Norm.
Q^2 **6**	CCQE; $0.5 < Q^2 < 1.0$ GeV^2	1.0	0.18	Gaus	Norm.
Q^2 **7**	CCQE; $Q^2 < 1.0$ GeV^2	1.0	0.40	Gaus	Norm.
$E_b\nu$ **C**	CCQE on ^{12}C; $E_\nu < 4$ GeV; ν	27 (25) MeV	6 MeV	Gaus	p_μ Shift
$E_b\bar{\nu}$ **C**	CCQE on ^{12}C; $E_\nu < 4$ GeV; $\bar{\nu}$	25 MeV	6 MeV	Gaus	p_μ Shift
$E_b\nu$ **O**	CCQE on ^{16}O; $E_\nu < 4$ GeV; ν	31 (27) MeV	6 MeV	Gaus	p_μ Shift
$E_b\bar{\nu}$ **O**	CCQE on ^{16}O; $E_\nu < 4$ GeV; $\bar{\nu}$	27 MeV	6 MeV	Gaus	p_μ Shift
M_A^{RES}	CC Res, NC π^0, NC $\pi^\pm$	1.07 (0.95) GeV	0.15 GeV	Gaus	Shape
C_5^A	CC Res, NC π^0, NC $\pi^\pm$	0.96 (1.01)	0.15	Gaus	Shape
$I_{1/2}$ **non-res**	CC Res, NC π^0, NC $\pi^\pm$	0.96 (1.30)	0.40	Gaus	Shape
$I_{1/2}$ **non-res low p_π**	-	-	-	-	-
CC norm ν	CC; $0.4 < E_\nu < 0.6$ GeV; ν	1.0	0.02	Gaus	Norm.
CC norm $\bar{\nu}$	CC; $0.4 < E_\nu < 0.6$ GeV; $\bar{\nu}$	1.0	0.01	Gaus	Norm.
CC ν_e/ν_μ	CC; ν_e	1.0	0.028	Gaus	Norm.
CC $\bar{\nu}_e/\bar{\nu}_\mu$	CC; $\bar{\nu}_e$	1.0	0.028	Gaus	Norm.
CC BY DIS	CC DIS; $W < 4.0$ GeV	1.0	1.0	Gaus	Shape
CC BY multi-π	CC multi-π; $1.6 < W < 2.0$ GeV	1.0	1.0	Gaus	Shape
CC AGKY mult	CC multi-π; $1.6 < W < 2.0$ GeV	1.0	1.0	Gaus	Shape
CC misc	CC1γ, CC1K, CC1η	1.0	1.0	Gaus	Norm.
CC DIS, multi-π norm ν	CC DIS, CC multi-π, ν	1.0	0.035	Gaus	Norm.
CC DIS, multi-π norm $\bar{\nu}$	CC DIS, CC multi-π, $\bar{\nu}$	1.0	0.065	Gaus	Norm.
CC coh^{12}C	CC Coherent on ^{12}C	1.0	0.3	Gaus	Norm.
CC coh ^{16}O	CC Coherent on ^{16}O	1.0	0.3	Gaus	Norm.
NC coh	NC Coherent	1.0	0.3	Gaus	Norm.
NC 1γ	NC 1γ	1.0	1.0	Gaus	Norm.
NC other near	NC DIS, multi-π, 1K, 1η	1.0	0.3	Gaus	Norm.
NC other far	-	-	-	-	-
π **FSI QE**	CC Res, NC π^0, NC $\pi^\pm$	1.069	0.313	Gaus	Shape
π **FSI QE high E**	CC Res, NC π^0, NC $\pi^\pm$	1.824	0.859	Gaus	Shape
π **FSI Abs**	CC Res, NC π^0, NC $\pi^\pm$	1.002	1.102	Gaus	Shape
π **FSI Prod**	CC Res, NC π^0, NC $\pi^\pm$	1.404	0.432	Gaus	Shape
π **FSI CX**	CC Res, NC π^0, NC $\pi^\pm$	0.697	0.305	Gaus	Shape

Table 5.1.: The interaction parameters used in this analysis.
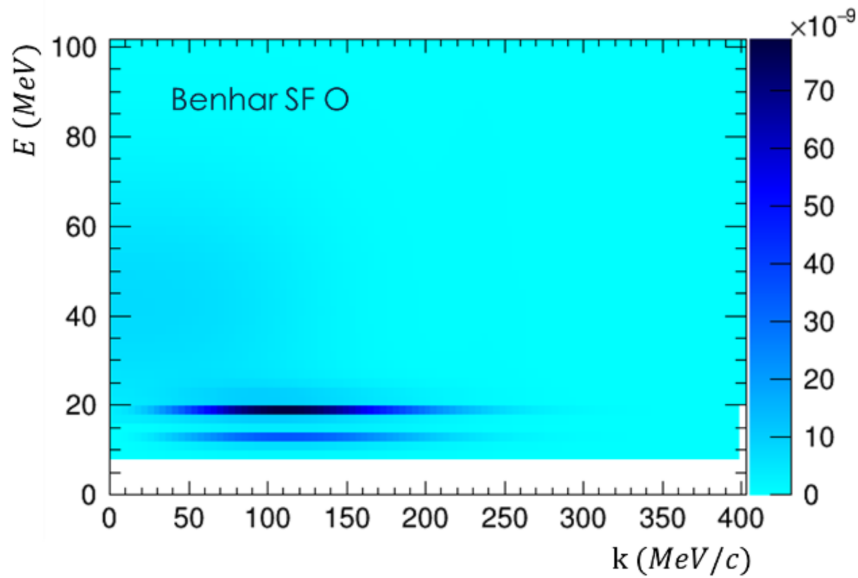
Figure 5.11.: Removal energy ('*E*') at different values of the initial nucleon momentum ('k') for the ground state in the SF model. Figure from [120].

ments, but is not precisely known.

The uncertainty on the binding energy was the largest individual systematic in the previous oscillation analysis. It contributed 3.7% to the uncertainty on the predicted relative number of electron neutrino and electron anti-neutrino candidates at SK (with no decay electrons) [134]. Studies of simulated datasets also showed that the uncertainty had a significant impact on the credible intervals of the measured oscillation parameters [135]. This indicated that the old cross-section model parametrisation could not account for changes to the binding energy, instead misattributing them to the flux and 2p2h model parameters. To reduce its impact, a new treatment of the uncertainty was implemented for this analysis.

In the SF nuclear model, there isn't a single nuclear binding energy systematic uncertainty. Instead, external data is used to inform distributions of the nuclear binding energy and initial nucleon momentum, which contain peaks corresponding to the nuclear shell structure. The peaks are shown for different initial nucleon momentum for the ground state in Figure 5.11.

Ideally, there would be systematic parameters for the height, width, and position of each of the peaks for each target nucleus. However, this was not feasible to implement on the time scale of this analysis. Instead, a global uncertainty of 6 MeV is applied to the removal energy in the SF nuclear model.

These binding energy systematics are implemented as neither shape nor normalisation parameters in the T2K oscillation analysis. Instead, the offset to the removal energy is propagated to a change in the p_μ–$\cos\theta_\mu$ distributions by directly shifting p_μ. It has been

shown that changes in the removal energy do not cause significant changes in the distribution of θ_μ [120].

The SF model used was constrained using data from ee', p experiments, which only applies to initial state protons, and therefore anti-neutrino CCQE interactions. The spectral function of initial state neutrons in neutrino CCQE interactions can not be constrained in the same way. The offset between the SF and Relativistic Mean-Field model predictions [136] for neutrons is $\sim$4 MeV for oxygen and $\sim$2 MeV for carbon.

There are therefore four binding energy parameters applied to CCQE events: for ^{12}C/^{16}O and $\nu/\bar{\nu}$. The $\bar{\nu}$ parameters have a prior central value of 25 MeV for ^{12}C, and 27 MeV for ^{16}O. The central value of the ν parameters are offset from their $\bar{\nu}$ counterparts by 2 and 4 MeV for ^{12}C and ^{16}O respectively. The parameters are correlated with each other as follows:

- $E_b\nu$ **O:** 70% with $E_b\bar{\nu}$ O, 77.77% with $E_b\nu$ C, 65.27% with $E_b\bar{\nu}$ C

- $E_b\bar{\nu}$ **O:** 70% with $E_b\nu$ O, 65.27% with $E_b\nu$ C, 77.77% with $E_b\bar{\nu}$ C

- $E_b\nu$ **C:** 77.77% with $E_b\nu$ O, 65.27% with $E_b\bar{\nu}$ O, 70% with $E_b\bar{\nu}$ C

- $E_b\bar{\nu}$ **O:** 65.27% with $E_b\nu$ C, 77.77% with $E_b\bar{\nu}$ O, 70% with $E_b\nu$ C

These values are derived in [120].

The effect of the $E_b\nu$ C parameter on the FGD1 FHC CC0π sample is shown in Figure 5.12. Here, the parameter is set to $\pm 1\sigma$ while all other parameters are kept at nominal, and the ratio to the nominal MC is taken. As the binding energy parameter is increased, the final state lepton momentum is decreased. This is seen in the increase in number of events at lower momentum in Figure 5.12a. The opposite effect is seen for decreasing the binding energy parameter in Figure 5.12b.

To validate the implementation of this new type of parameter, Asimov fits [137] (where the nominal MC is fitted to itself) with binding energy included were run. The Asimov fitting process is described in more detail in Section 6.4. The 1D posterior distributions for each of the 4 four parameters are shown in Figure 5.13, with a fitted Gaussian distribution. These validations were performed using the uniform binning described in Appendix B. As expected, the postfit values are close to the nominal values, and the distributions are approximately Gaussian.

However, for fits to data the posteriors contain many discontinuities as shown in Figure 5.14. This is not seen for other parameters. When reweighting, either by spline or normalisation, the change in the log-likelihood is continuous. As the parameter varies, the

(a) $+1\sigma$

(b) -1σ
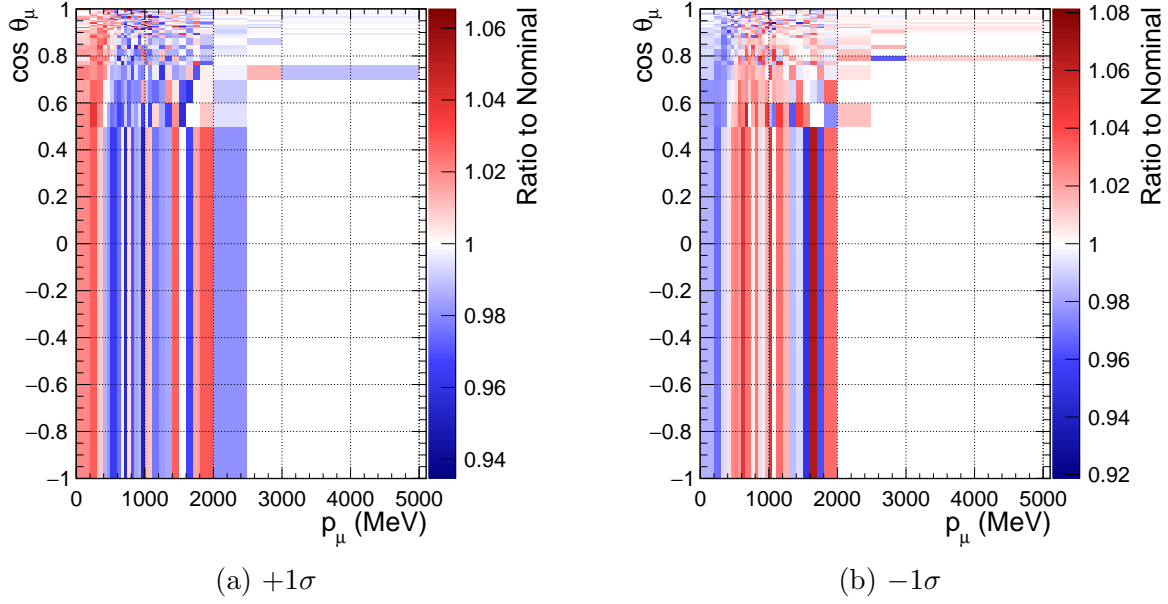
Figure 5.12.: Ratio of the FGD1 FHC CC 0π sample with $E_b\nu$ C parameter set to $\pm 1\sigma$ to the nominal MC.

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
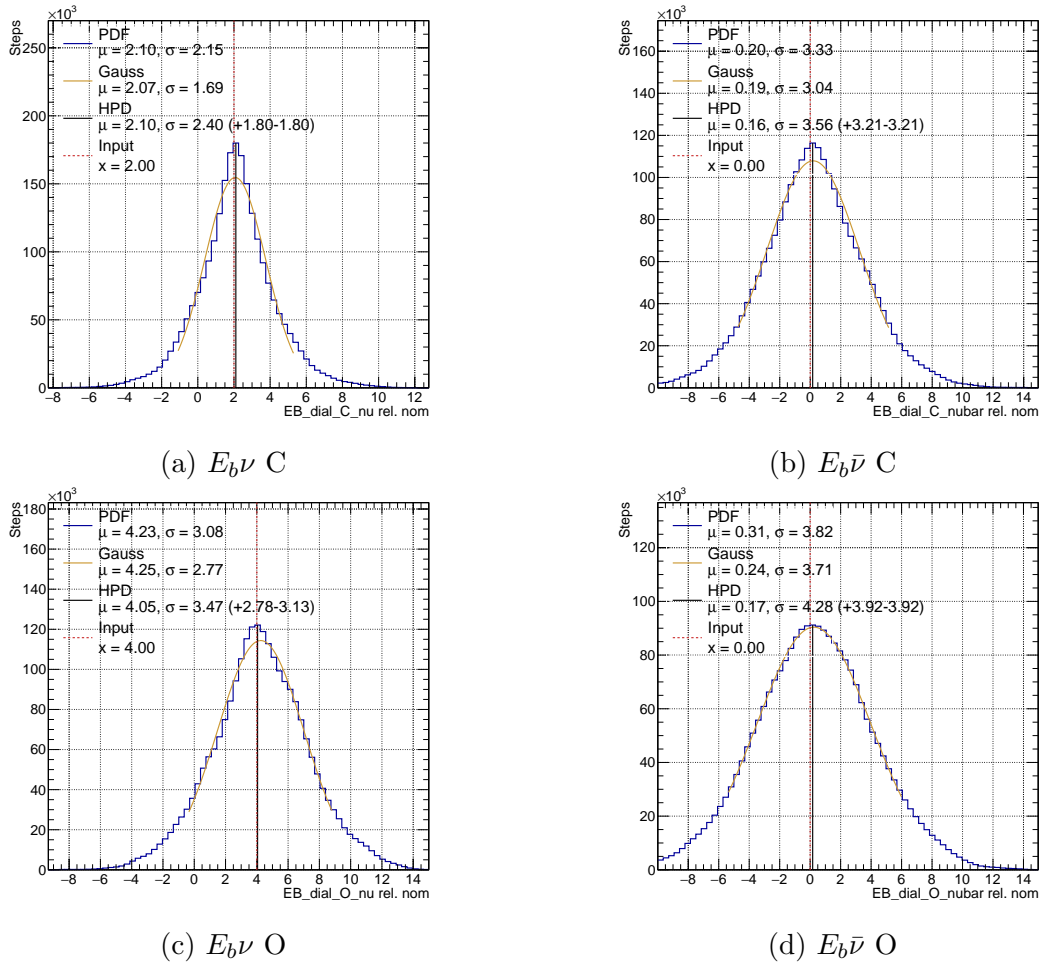
Figure 5.13.: Posterior distributions for the binding energy parameters from an Asimov fit.

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
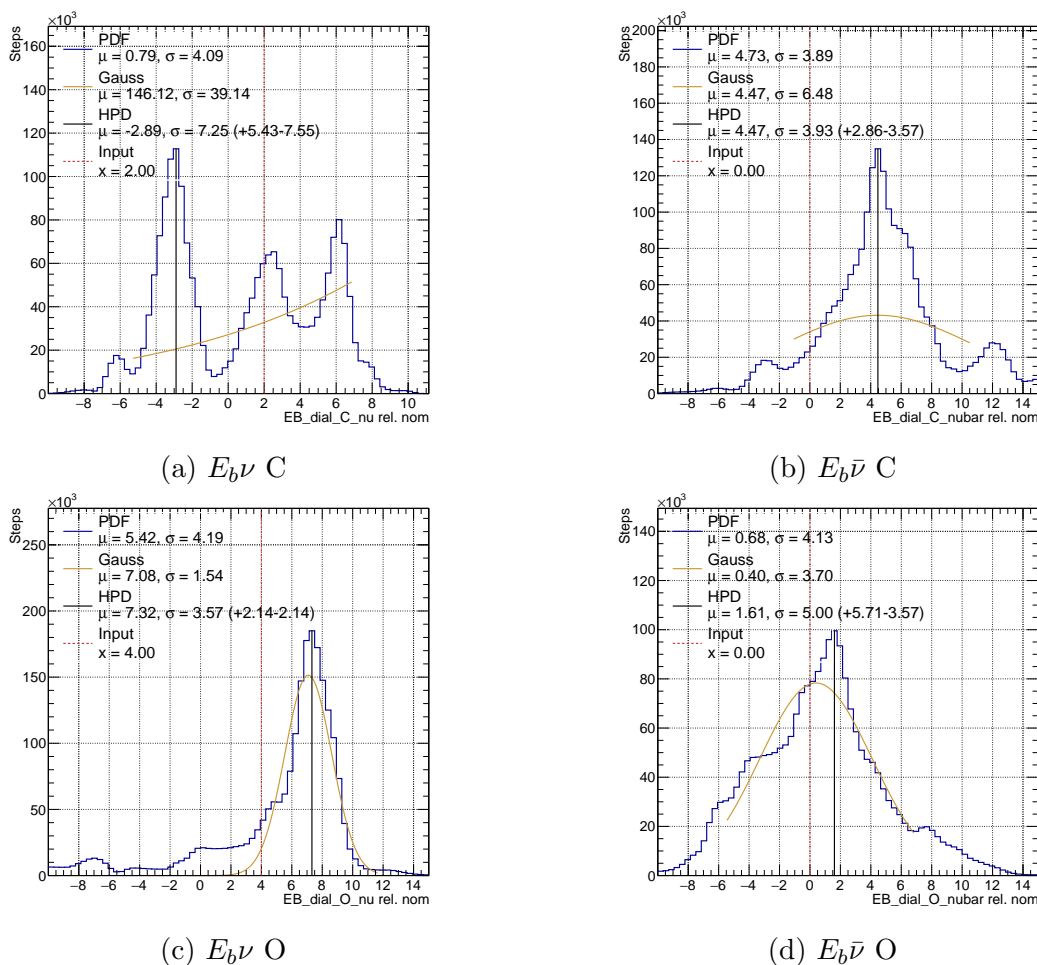
Figure 5.14.: Posterior distributions for the binding energy parameters from a data fit.

change in the penalty contribution, and the change in the number of events in each bin, and therefore the sample contribution, are both smooth. However, for this direct shift in lepton momentum, the sample contribution to the likelihood only changes if an event crosses a bin boundary. There are therefore threshold values of each parameter where several events cross boundaries and cause discontinuous changes in the log-likelihood.

To show this is what causes the discontinuities, a fluctuated version of the nominal MC was produced. This was done by setting the number of events in each bin to be a random number from a Poisson distribution, with a mean equal to the nominal bin content. This ensured the number of events in each bin was an integer. The fluctuated MC was fitted to the nominal MC, and the posterior distributions for the E_b parameters are shown in Figure 5.15. The fact that these are non-Gaussian and discontinuous, despite them being continuous and Gaussian in the regular Asimov fit, shows that the effect is caused by the discrete shifts to integer events.

Furthermore, Figure 5.15 also shows the E_b distributions for a fluctuated Asimov fit using

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
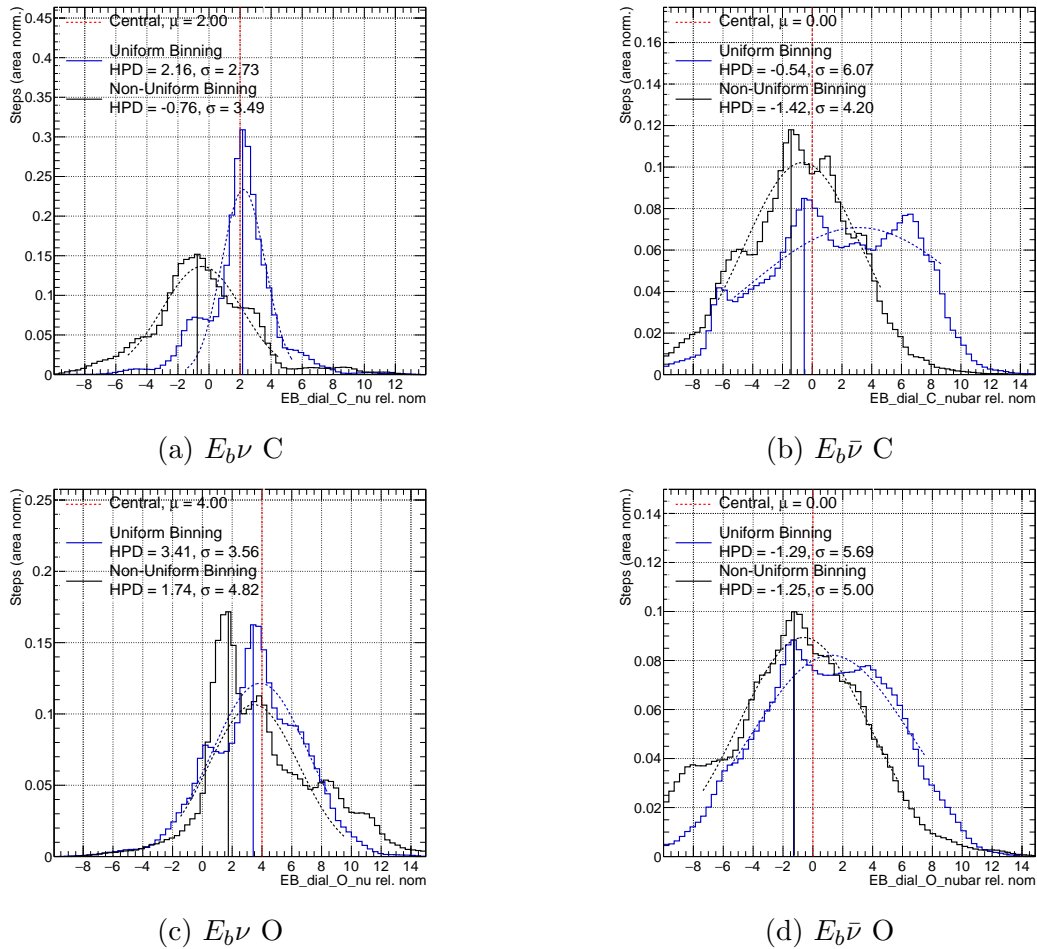
Figure 5.15.: Posterior distributions for the binding energy parameters from fits to fluctuated Asimov data using uniform and non-uniform rectangular fit binnings.

the non-rectangular binning described in Section 5.3.1. The peaks in the distributions, caused by the shifting of discrete events across bin boundaries, move positions between the two fits. This shows the location of the peaks in E_b is very dependent on the binning. This is much more of a significant effect for E_b, as the parametrisation directly shifts events rather than reweighting them. Changes in the likelihood therefore only occur when the events cross bin boundaries, and so the parameter is sensitive to where those boundaries are. Different binnings can therefore produce different fit results.

As the BANFF fit uses a gradient descent, the discontinuities in the binding energy parameter likelihoods can prevent the fit from converging. To avoid this, rather than directly shift the outgoing lepton momentum of events the BANFF use an effective reweighting. This aims to smoothly replicate the shifting of events by increasing and decreasing the events in a continuous manner. Template distributions were produced for each sample for a number of knot points for each parameter. Bin-by-bin splines were then produced by interpolating

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
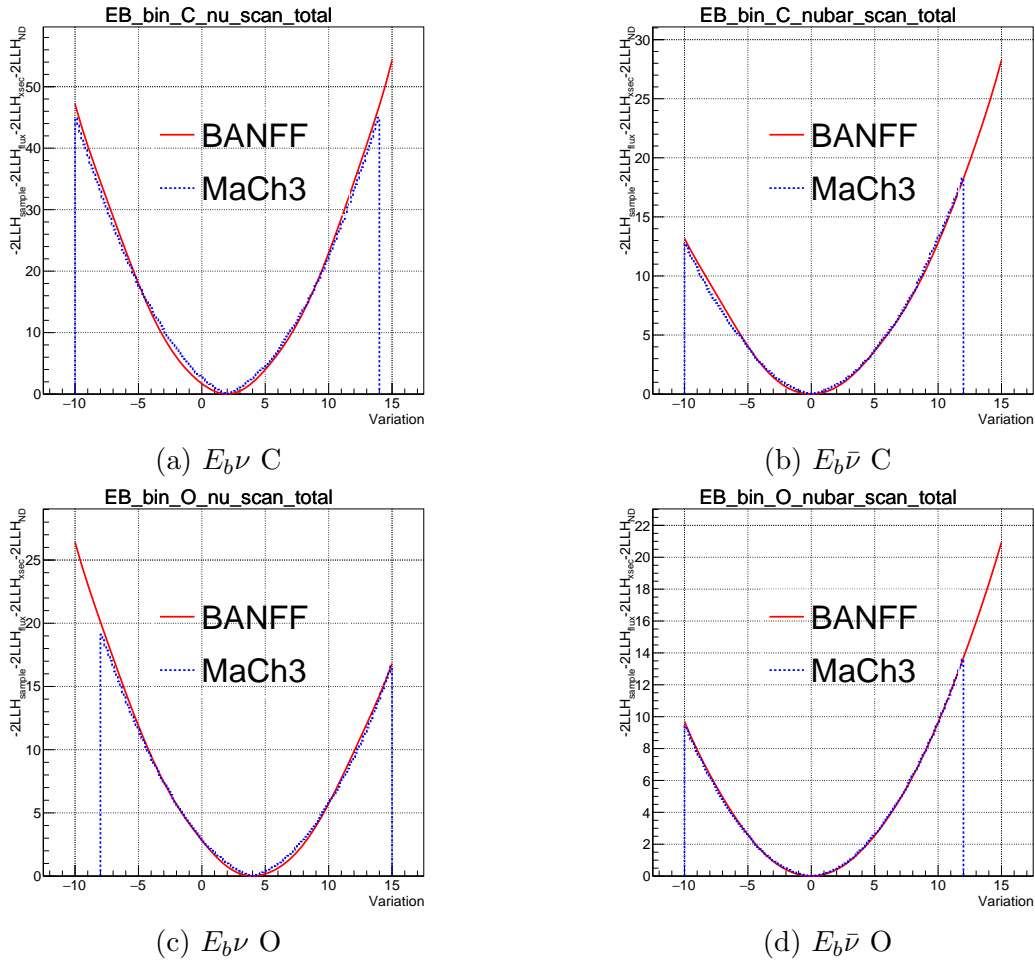
Figure 5.16.: Comparison of log-likelihood scans for the two near detector analysis. The good, but not perfect, agreement shows that the effective bin-by-bin reweighting the BANFF use will closely, but not perfectly, replicate the effect of the event-by-event direct shifts used in MaCh3.

between the number of events in each bin at different knot values. This process is described in more detail in [138].

Figure 5.16 shows log-likelihood scans for the binding energy parameters in both the MaCh3 and BANFF frameworks. For the BANFF fit, these are smoother, but they closely trace the MaCh3 distributions. This is the intended behaviour of the splines: to smooth over discontinuities while retaining the overall shape. However, the fact that there is not perfect agreement shows that the spline process cannot completely represent the direct momentum shifts. The implementation of the binding energy uncertainty used in MaCh3 is therefore a more accurate method.

The treatment of the binding energy is one of the main differences between the two near detector fitters, and one of the main advantages of this analysis over the BANFF.

In previous analyses, the binding energy systematic was applied as a shape parameter.

The change in removal energy was propagated to a change in the p_μ–$\cos\theta_\mu$ distributions by reweighting events. However, this method broke down at extreme values of the kinematic variables as changes in the removal energy could change the allowed phase space in unphysical ways. With this implementation, the removal energy was one of the dominant uncertainties in the T2K oscillation analysis. With the direct lepton momentum shift, the removal energy is now a sub-dominant systematic uncertainty, as shown in Section 6.8.3. The treatment of the binding energy is therefore also one of the main improvements made of this analysis compared to previous T2K near detector fits.

This is the first time a systematic of this type has been implemented in a T2K analysis. Now that the framework is in place to accommodate such parameters, more systematics could directly shift event kinematics in the future, as discussed in more detail in Section 5.4.3.3. This work could therefore pave the way for improvements to subsequent analyses.

In summary, the new treatment of the binding energy gives a more accurate implementation of the systematic, significantly reduces the uncertainty, and shows that fitting direct shifts to events is a viable method of applying new parameters in the future.

5.4.2. Flux

The flux systematics are determined using the simulation described in Section 3.1.3. New data is used to regularly update and improve the modelling. This comes from external experiments such as NA61/SHINE [94], the T2K beam monitors, and the on-axis near detector INGRID. This analysis is the first to use flux systematics developed from simulation tuned to data from a full T2K target replica, and not just a thinner replica version.

There are six sources of flux uncertainty:

- Alignment of the proton beam with the target.

- Number of protons on target.

- Interactions of protons and produced hadrons with the target.

- Alignment of the target with the focusing horns.

- The horn current and produced magnetic field.

- Modelling of materials in the target and decay volume.

The fractional sizes of the different sources of ND280 flux systematic are shown in Figure 5.17, for different neutrino signs. The hadron production contribution dominates, and the

total uncertainty is ∼10% around the beam peak energy at 600 MeV. The proton beam alignment becomes more significant around 1 GeV. The black dotted line shows the total flux uncertainty using the previous version of the model. This also used the full replica target tuning; the improvement comes from using the latest FLUKA [93] release. Similar results are seen for the SK flux uncertainties, shown in Figure 5.18.

The flux uncertainty model is parametrised as 100 true neutrino energy bin normalisations, split by neutrino species, horn current mode, and detector. The binnings are as follows:

- **ND280 + SK, FHC ν_μ + 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 $\bar{\nu}_\mu$ + RHC ν_μ:**

  E_ν^{true}: 0, 0.7, 1, 1.5, 2.5, 30

- **ND280 + SK, FHC ν_e + RHC $\bar{\nu}_e$:**

  E_ν^{true}: 0, 0.5, 0.7, 0.8, 1.5, 2.5, 4, 30

- **ND280 + SK, FHC $\bar{\nu}_e$ + RHC ν_e:**

  E_ν^{true}: 0, 2.5, 30

In total there are 100 flux parameters, 50 for ND280 and 50 for SK. The SK flux systematics are used in the near detector fit because of their high correlations with their ND280 counterparts. All flux systematics have a prior central value of 1.0, and Gaussian prior uncertainty with width equal to the standard deviation in the fractional covariance matrix, shown in Figure 5.19 for the ND280 FHC ν_μ parameters. There are larger prefit correlations for parameters with similar energies, particularly at higher energies. The full flux covariance matrix for all samples is shown in Appendix C.

5.4.3. Detector

The detector systematics are outlined in this section. Firstly, the physical underlying uncertainties are described in Section 5.4.3.1. The application of these systematics, by studying their impact on adjacent fit bins, is discussed in Section 5.4.3.2, along with a brief investigation into the validity of the assumptions used in this method. The algorithm used to determine the detector binning is defined in Section 5.4.3.3, and the results of fits using different detector binnings are then compared in Section 5.4.3.4, before the final covariance matrices are presented in Section 5.4.3.5. Finally, potential improvements to the treatment of detector systematics for future analyses are discussed in Section 5.4.3.6.
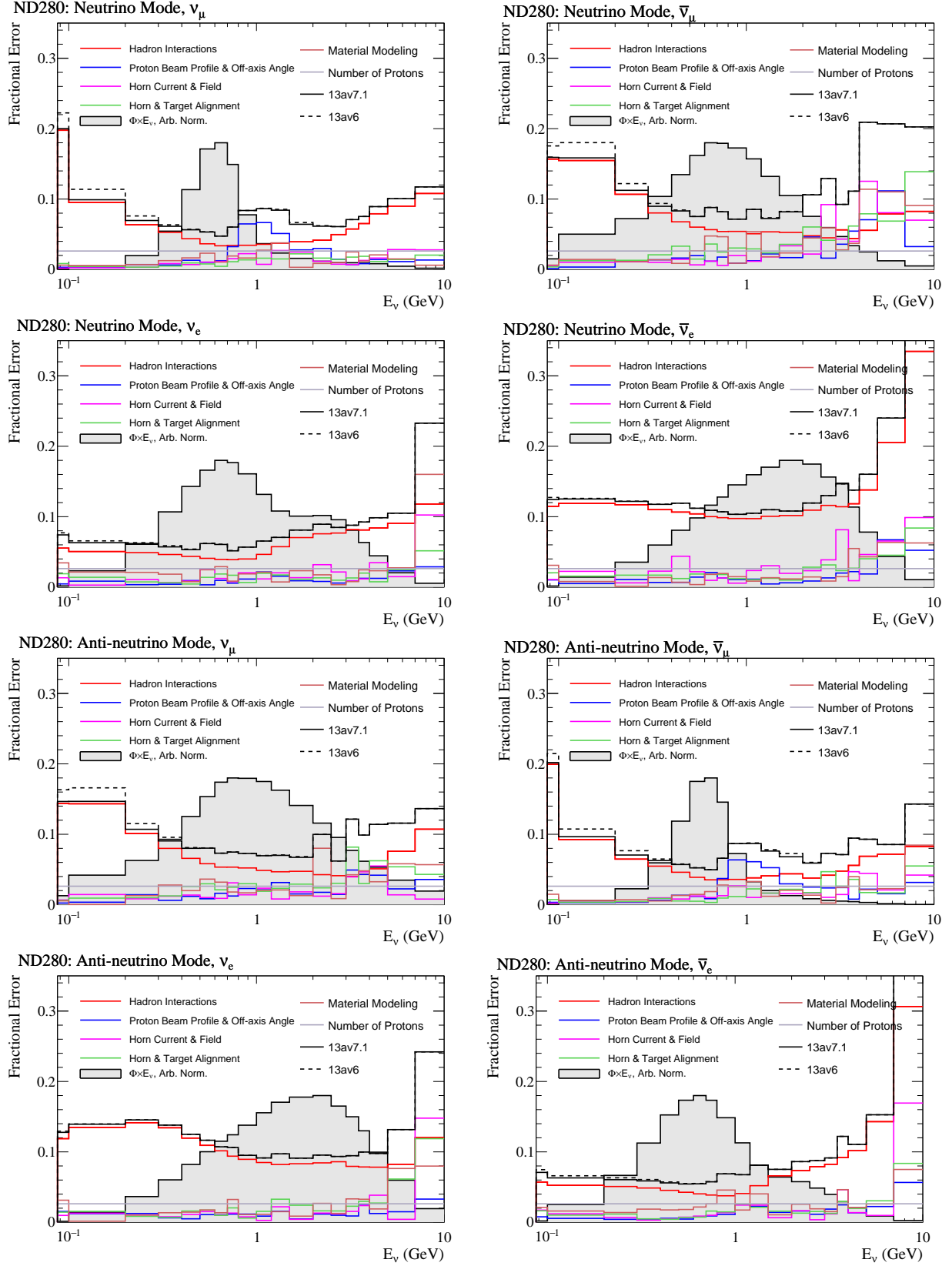
Figure 5.17.: Relative sizes of the sources of uncertainties in the ND280 flux parameters. "13av7.1" (black solid line) is the version used in this analysis, and is compared with the previous version, "13v6" (dotted line). Figures from [139].
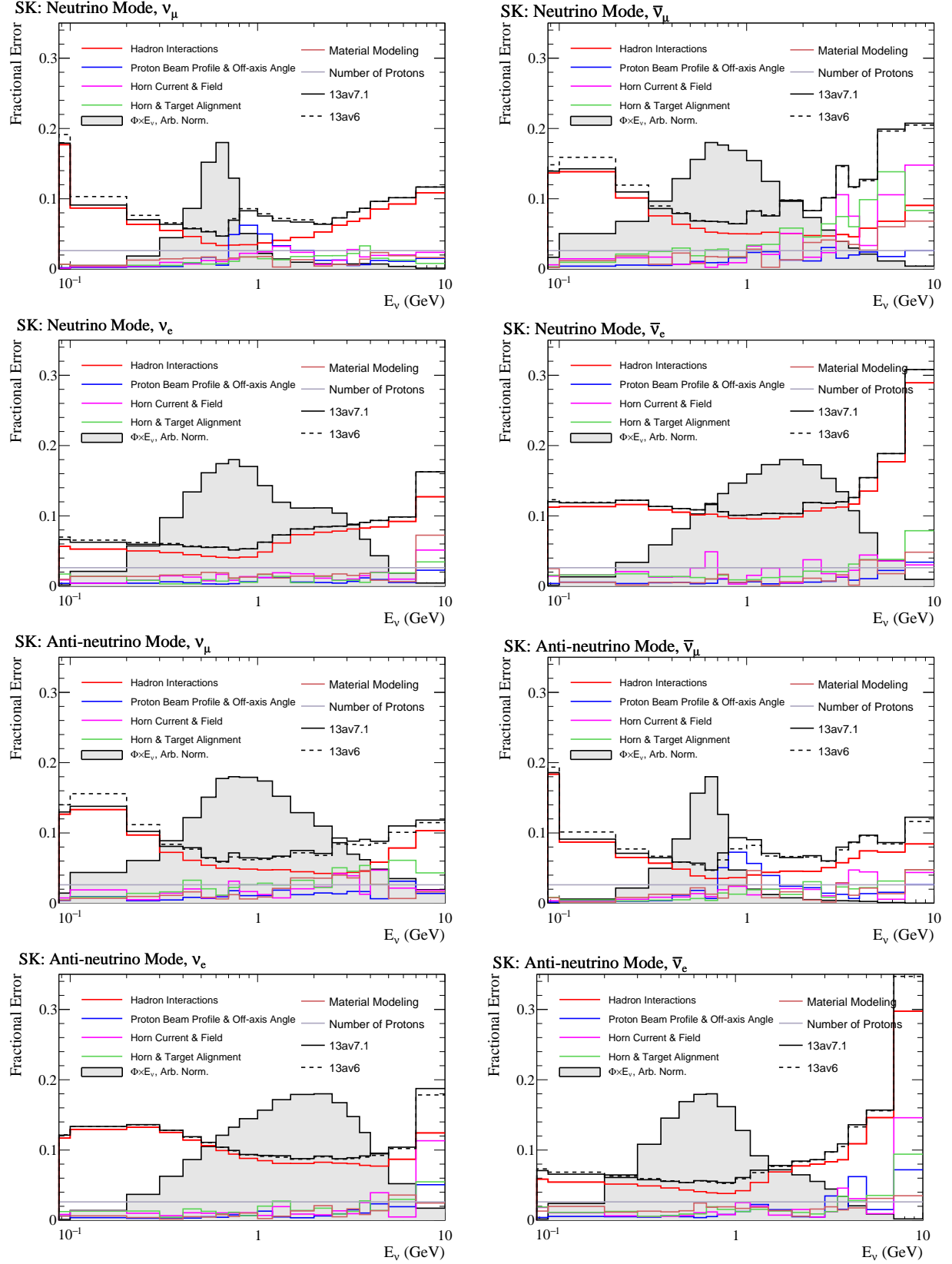
Figure 5.18.: Relative sizes of the sources of uncertainties in the SK flux parameters. "13av7.1" (black solid line) is the version used in this analysis, and is compared with the previous version, "13v6" (dotted line). Figures from [139].
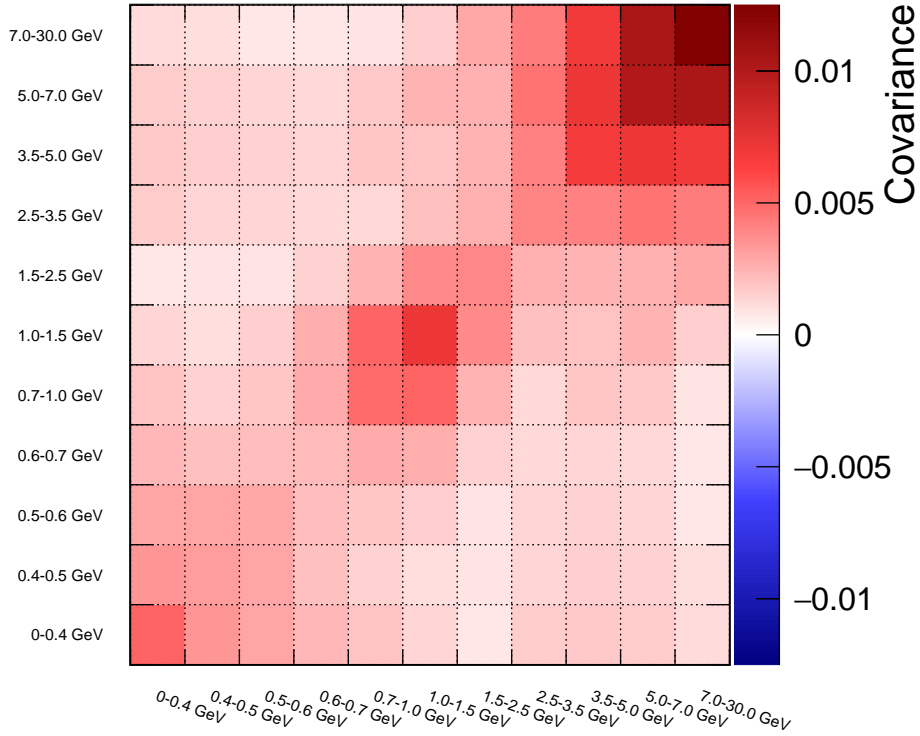
Figure 5.19.: The flux covariance matrix for the ND280 FHC ν_μ parameters.

5.4.3.1. Underlying Systematics

The ND280 detector systematics are modelled using bin normalisation parameters. The underlying systematics are initially varied to see the change to the number of events in each p_μ–$\cos\theta_\mu$ bin. These systematics are divided into 17 groups:

- **TPC Field Distortions:** The magnetic field applied in ND280 is not perfectly uniform, and so there are field distortions in the TPCs. These are measured by calibration lasers with the magnet both on and off.

- **TPC Momentum Scale:** The uncertainty in the magnetic field causes an uncertainty in the measured momentum within the TPC. Four hall probes within ND280 provide scaling factors applied to the field strength on the MC.

- **TPC Momentum Resolution:** There is a discrepancy between the MC and data momentum resolutions in the TPCs, which is not well understood. To account for this, MC TPC tracks are smeared so that the resolution matches that of data.

- **TPC PID:** The pulls between the measured and predicted energy loss is used to

identify particles in the TPC, as discussed in Section 3.2.2.2. The difference between the mean MC and data pulls is applied as a systematic uncertainty on the TPC PID.

- **TPC Cluster Efficiency:** There are differences between the reconstruction of clusters of hits in the data and MC. An uncertainty is applied, correlated between horizontal and vertical clusters to account for this. It is calculated using control samples of both beam and cosmic trigger events.

- **TPC Tracking Efficiency:** The efficiency of the TPC reconstruction algorithm successfully merging hits into tracks is measured for control samples of both beam and cosmic trigger events. The difference between the efficiencies obtained for data and MC is applied as a systematic.

- **TPC Charge ID Efficiency:** There are two sources of uncertainty in the TPC charge identification: the efficiency of the initial TPC charge being correct, and the probability of the TPC charge sign being reversed in the overall charge identification. The TPC Charge ID uncertainty is applied as the probability of the overall charge being different from the TPC reconstructed charge.

- **TPC-FGD Matching Efficiency:** The efficiency of matching FGD and TPC tracks was calculated using a control sample of events with a high angle with respect to the neutrino beam, that passed through at least two TPCs. The difference in the efficiency found for data and MC is applied as a systematic.

- **FGD PID:** The pulls between the measured and predicted energy loss is used to identify particles in the FGD, as discussed in Section 3.2.2.1. The difference between the mean MC and data pulls is applied as a systematic uncertainty on the FGD PID.

- **FGD Time of Flight:** For tracks passing through FGD1 and FGD2, the hit times in each FGD are used to determine the direction of the track. The uncertainty on the time of flight can therefore affect which FGD the event was reconstructed as having occurred in. Analysis, rather than control, samples were used to measure the time of flight for data and MC events. All reconstructed time of flights are smeared with the discrepancy as an uncertainty.

- **FGD Hybrid Tracking Efficiency:** The efficiency of reconstructing FGD-only tracks in the presence of FGD-TPC matched tracks is calculated for a sample set of GEANT-4 generated stopping protons and pions in a control sample of events with either one reconstructed track entering the TPC, or two tracks which both enter the

TPC. The difference between the efficiencies for data and MC is applied as a systematic uncertainty.

- **Michel Electron Efficiency:** The efficiency of detecting Michel electrons depends on the probability of the electron producing enough hits in the FGD to pass the selection cut, and the purity of the cut itself. The efficiency was measured for a control sample of cosmic trigger events, and was defined as the probability to detect a Michel electron that was expected from the presence of a stopped muon in the FGD.

- **Out of Fiducial Volume (OOFV) Background:** Events outside the fiducial volume can be misreconstructed as being being inside the fiducial volume. These could be events that occurred in the first two layers of FGD1, the first layer of FGD2, or in one of the other sub-detectors. The background rate was calculated for beam trigger events, and the discrepancy between the measurement for data and MC is applied as a systematic uncertainty.

- **Sand Muon Background:** Interactions from beam neutrinos can occur in the sand outside the near detector pit but look similar to events in the FGDs, forming a background rate. The rate from a dedicated simulation is compared to data to calculate the associated uncertainty.

- **Pile-Up:** Out-of-fiducial volume events being coincident with in-fiducial volume CC-inclusive events in the FGDs can lead to CC-inclusive events being rejected by the external veto cut described in Section 5.2. The difference in number of events per bunch in the data and MC is applied as systematic uncertainty to account for this effect.

- **Pion Secondary Interactions (SI):** Pions produced by neutrino interactions at ND280 can interact within the detector. This causes pion detection inefficiencies. An uncertainty is applied to account for this effect, calculated from the difference between pion SI cross-sections measured in data and MC.

- **FGD Mass:** The uncertainty on the FGD masses affects the number of target nuclei, and so can change the total event rate. Differences between the measured and simulated FGD masses are applied as a systematic uncertainty.

Each of these systematics is applied in one of three different ways:

Systematic Source	Type	Prior Shape
B Field Distortions	Observable	Flat
TPC Momentum Scale	Observable	Gaussian
TPC Momentum Resolution	Observable	Gaussian
TPC PID	Observable	Gaussian
TPC Cluster Efficiency	Efficiency-Like	Gaussian
TPC Tracking Efficiency	Efficiency-Like	Gaussian
TPC Charge ID Efficiency	Efficiency-Like	Gaussian
TPC-FGD Matching Efficiency	Efficiency-Like	Gaussian
FGD PID	Observable	Gaussian
FGD ToF	Observable	Gaussian
FGD Hybrid Tracking Efficiency	Efficiency-Like	Gaussian
Michel Electron Efficiency	Efficiency-Like	Gaussian
OOFV Background	Normalisation	Gaussian
Sand Muon Background	Normalisation	Gaussian
Pile-Up	Normalisation	Gaussian
Pion Secondary Interactions	Normalisation	Gaussian
FGD Mass	Normalisation	Gaussian

Table 5.2.: ND280 detector systematics, and their propagation type and prior uncertainty shape.

- **Observable Variable Systematic:** These are smearings that are applied to reconstructed variables. The selection algorithm is then rerun, and so smeared events can change their topology, selection, and which track the lepton candidate is.

- **Efficiency-Like Systematics:** These are uncertainties on detection and reconstruction efficiencies, which are applied as weights to the event after selection.

- **Normalisation Systematics:** These are overall normalisation changes applied directly to events to scale rates up or down.

The type of each of the 17 systematics is shown in Table 5.2, along with whether they have a Gaussian or flat prior uncertainty.

5.4.3.2. Applying the Detector Systematics

In theory, the systematics described above could be applied on an event by event basis and fitted individually. However, this is not computationally feasible on the timescales
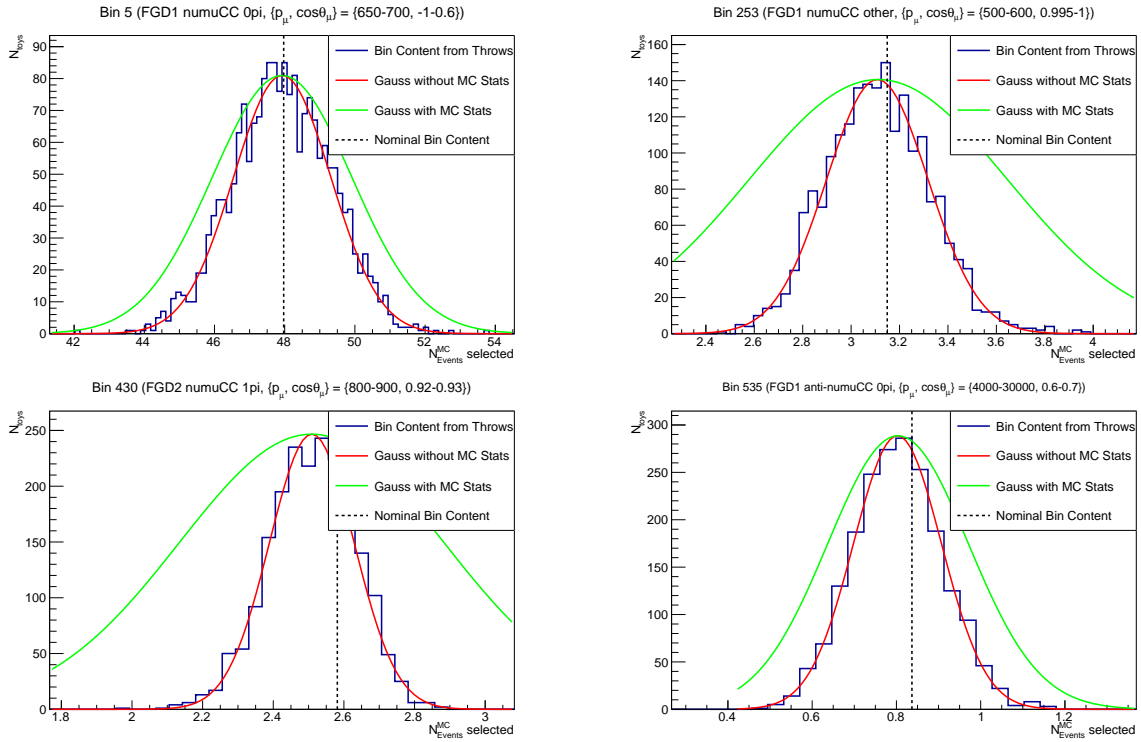
Figure 5.20.: Distribution of number of events in selected Gaussian distributed bins after 2000 throws of all detector systematics. The red and green lines show Gaussians fitted with and without the MC statistical uncertainty included, and the dotted black line shows the nominal number of events.

required for the oscillation analysis. Instead, each systematic is varied 2000 times and the MC reweighted. This gives a distribution of 2000 number of events for each p_μ–$\cos\theta_\mu$ bin. The mean and width of a Gaussian fitted to each of these distributions becomes the prior central value and uncertainty for a normalisation parameter applying to that bin. However, the number of fit bins is large, and so to reduce the number of fit parameters, adjacent bins with similar responses to the systematic variations are merged, so the detector binning is coarser than the fit binning. Studies of different detector binnings are presented in Section 5.4.3.3.

This process assumes the shape of the underlying systematics are Gaussian. This is the case for the majority of bins, four of which are shown in Figure 5.20. In previous analyses, the MC statistics uncertainty was included in the ND280 detector covariance instead of as an extra term in the log-likelihood calculation, and so a Gaussian fitted to the distribution of number of events in each bin with and without the MC stats uncertainty are shown.

However, several bins exhibit non-Gaussian behaviour, as shown in Figure 5.21.

To investigate which of the underlying detector systematics is causing the non-Gaussianity, each was switched off one-by-one and the variations repeated. It was found that when all
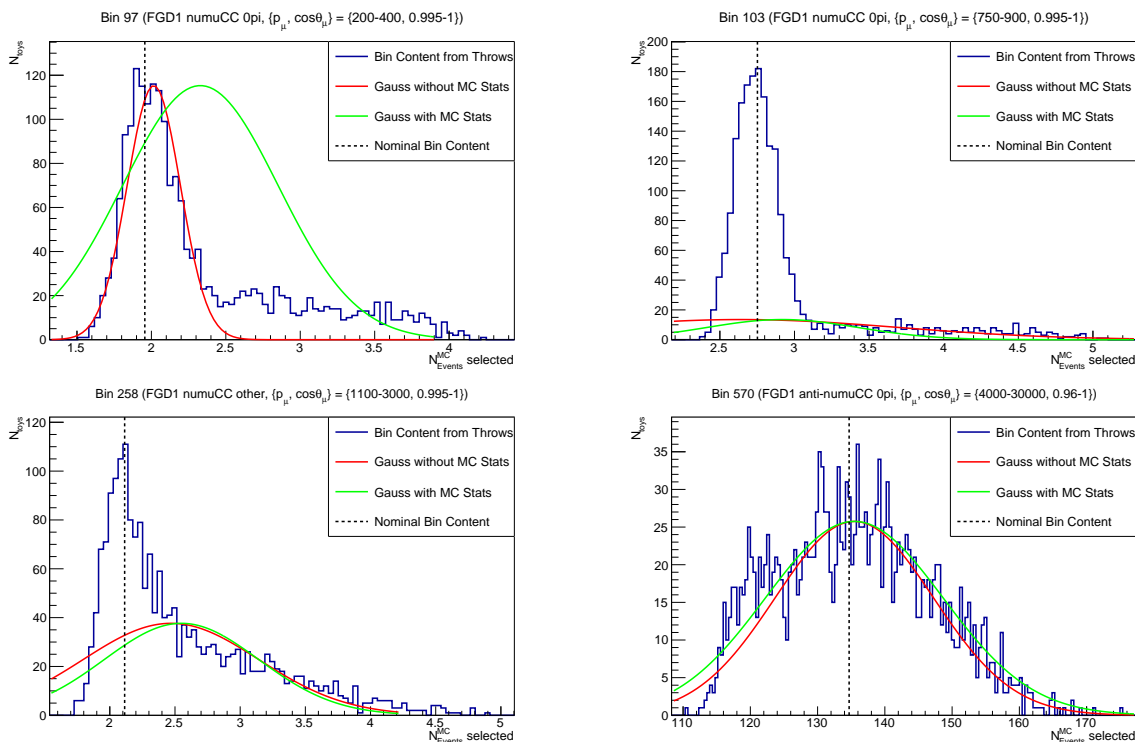
Figure 5.21.: Distribution of number of events in selected non-Gaussian distributed bins after 2000 throws of all detector systematics. The red and green lines show Gaussians fitted with and without the MC statistical uncertainty included, and the dotted black line shows the nominal number of events.

systematics but the Pion SI are left on, the distributions in the misbehaving bins are more Gaussian. These are shown in Figure 5.22. This suggests that the pion SI is the only one of the detector systematics with a significantly non-Gaussian distribution. As the effect only manifests in a small number of bins, it is not too concerning for this analysis.

5.4.3.3. Detector Binning Algorithm

The choice of detector binning is a trade-off of having as close to the fit binning as possible to obtain more accurate results, without introducing too many parameters to feasibly fit. To investigate the effect of merging detector bins, detector covariances with different binnings were produced. The binnings were produced by requiring different criteria to merge adjacent bins. These criteria were based on the number of events in a bin before applying the detector systematics, the change in the number of events by applying the systematics, and the difference between the covariance of bins. These correspond to merging bins with few events, bins where the effect of the systematics are small, and bins with similar response to the systematics respectively. The exact criteria for merging bins used were:
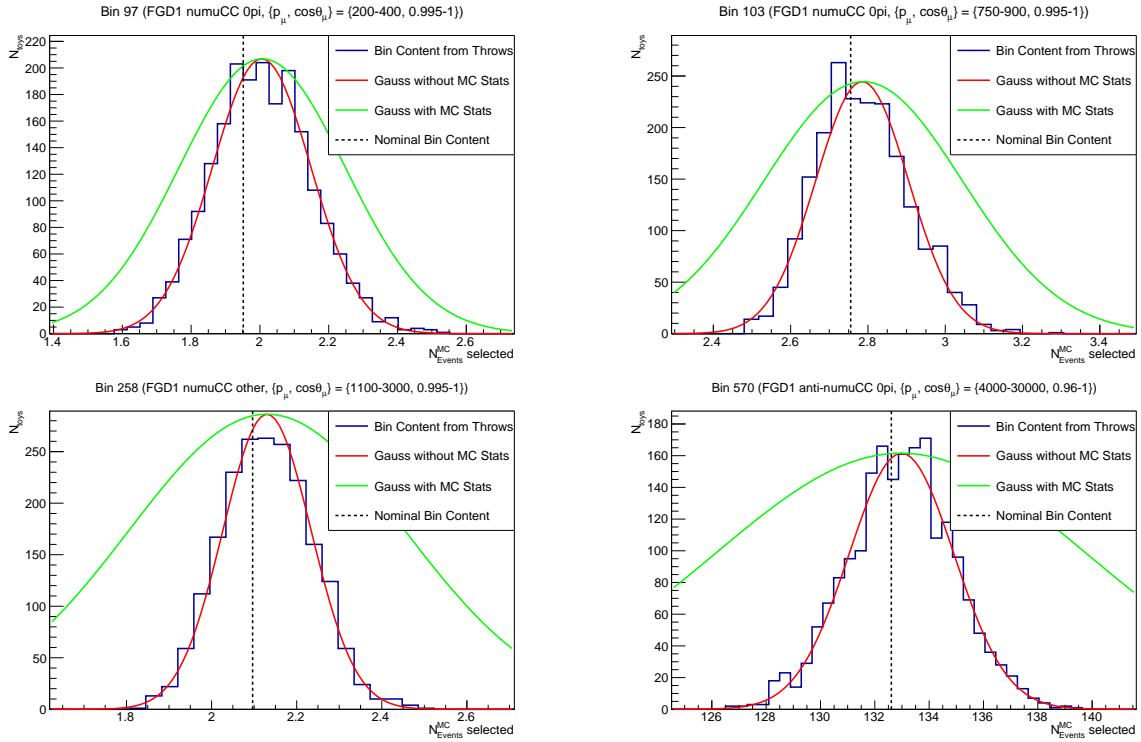
Figure 5.22.: Distribution of number of events in selected bins after 2000 throws of all detector systematics but the pion SI. The red and green lines show Gaussians fitted with and without the MC statistical uncertainty included, and the dotted black line shows the nominal number of events.

- < 1 event in a bin, or < 1 event change from applying the systematics, or $< 10\%$ difference in covariance between bins. This produced 179 merged bins.

- < 1 event in a bin, or < 1 event change from applying the systematics, or $< 5\%$ difference in covariance between bins. This produced 574 merged bins.

- < 1 event in a bin, or < 0.5 event change from applying the systematics, or $< 5\%$ difference in covariance between bins. This produced 1347 merged bins.

A covariance using the fit binning as the detector binning was also produced. As in these studies the fit binning used was the uniform-rectangular set defined in Appendix B, this corresponded to 4238 bins.

5.4.3.4. Detector Binning Fits

Fits were run using each of the four detector covariances, using an intermediate cross-section model including parts of the 2017 analysis parametrisation described in [140], and parts of the 2020 analysis parametrisation described in Section 5.4.1. To avoid tuning the detector binning on data, fake data was produced by setting the cross-section parameters to their

best fit values from the 2017 analysis and reweighting the runs 2–6 nominal MC. The MC was then fitted to this fake data.

The result of the fits are shown in Figure 5.23 for the ND280 FHC flux parameters, Figure 5.24 for the cross-section parameters, and Appendix A for all parameters. Although there are several differences between the postfit parameter values, there is no consistent trend of having more bins being closer to the fit binning result, and they are all consistent within uncertainties. Although the 179 fit bin fit is often the most different from the 4238 (fit binning) detector bin fit.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$
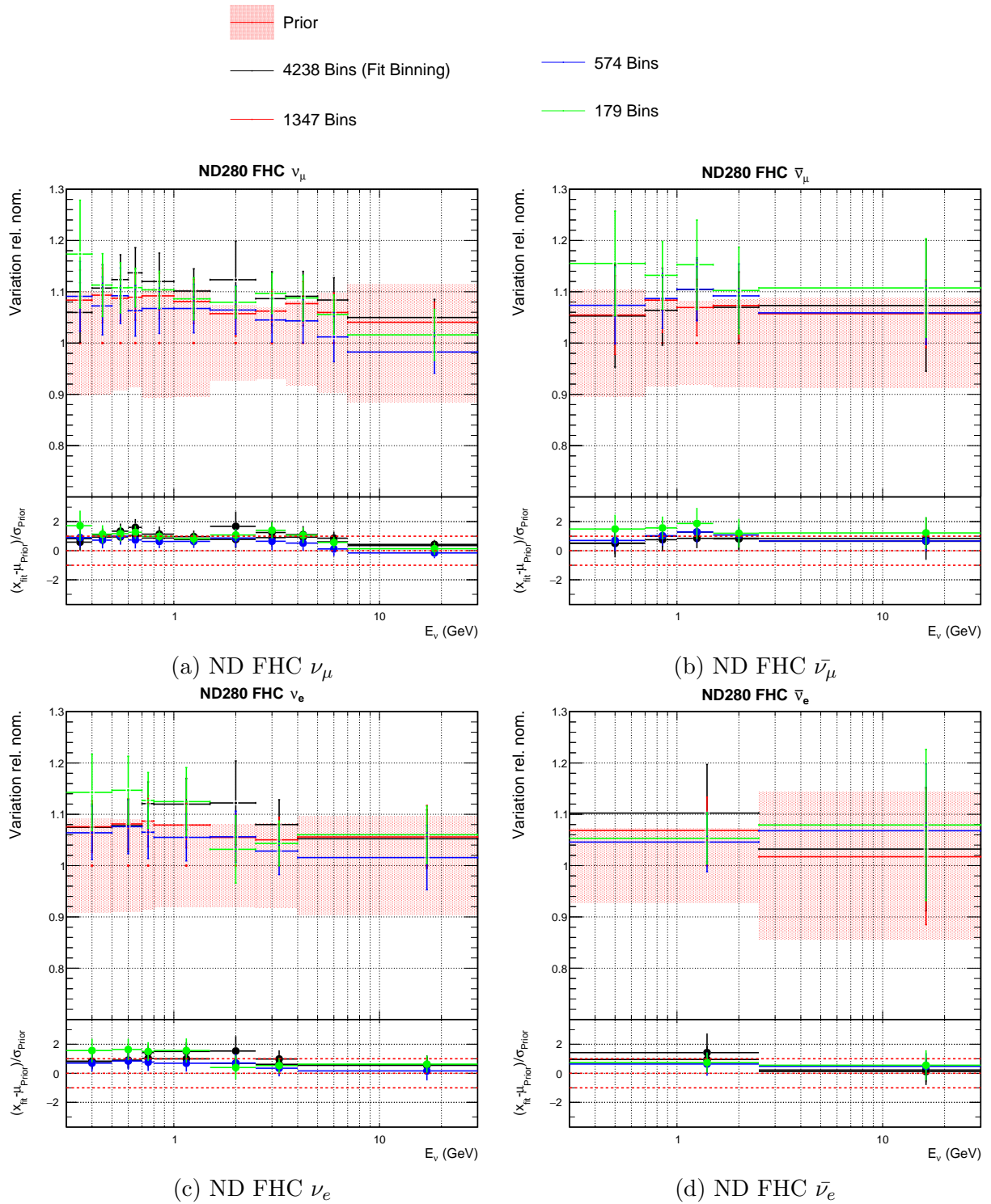
Figure 5.23.: ND280 FHC flux parameters for fake data fits using different detector binnings.

(a) CC 0π

(b) CC 1π, ν_e, CC DIS, and CC coh.

(c) NC

(d) π FSI and E_b
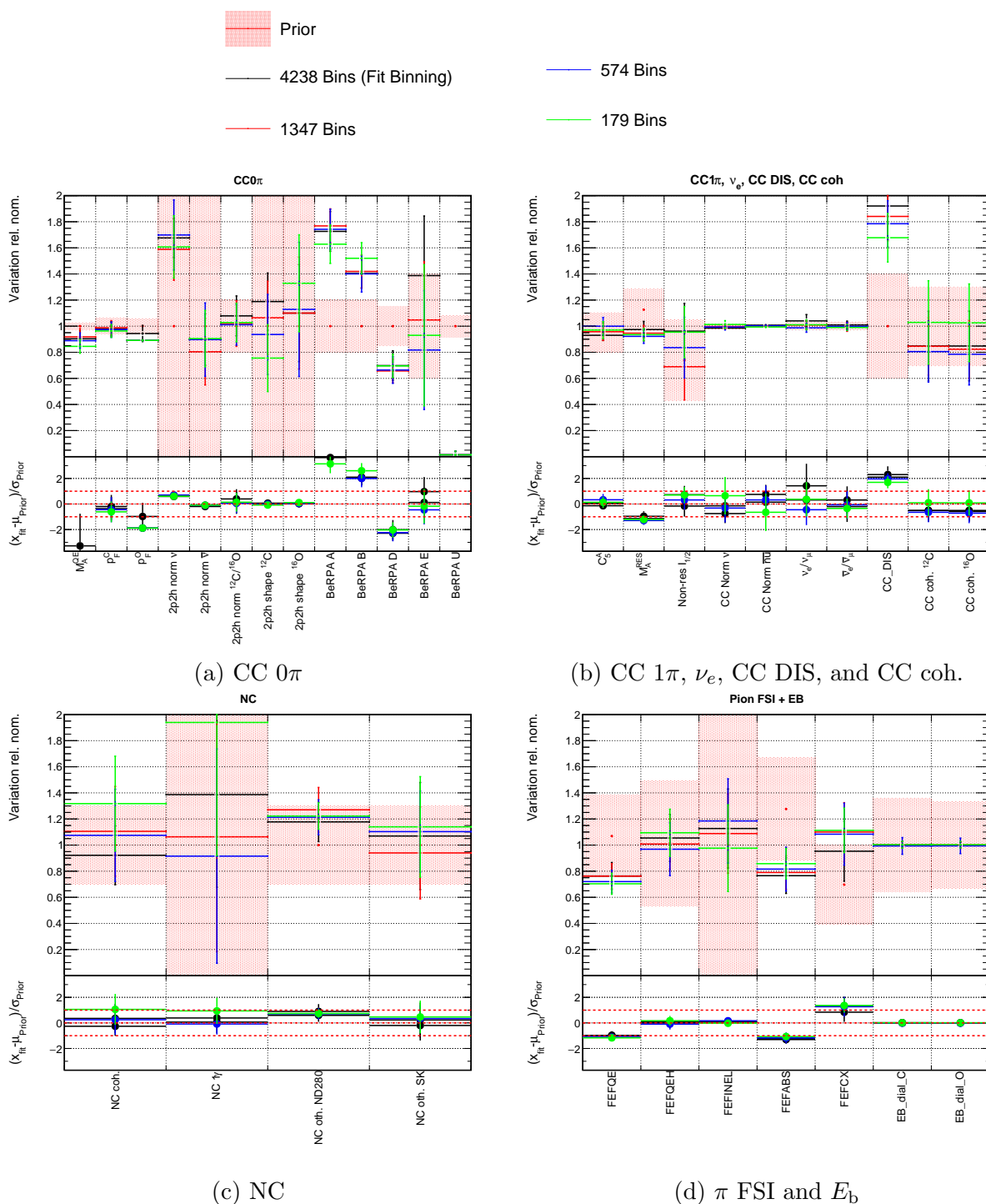
Figure 5.24.: Interaction parameters for fake data fits using different detector binnings.

Furthermore, when the postfit chains are used to produce posterior predictive distributions at the far detector, there is very little difference seen, as shown in Figure 5.25. This suggests that there are several regions of minima in the $\sim$700 dimensional parameter space which correspond to the same result at SK.

(a) 1R$_\mu$

(b) RHC 1R$_\mu$

(c) 1R$_e$

(d) RHC 1R$_e$

(e) 1R$_e$ 1d.e.
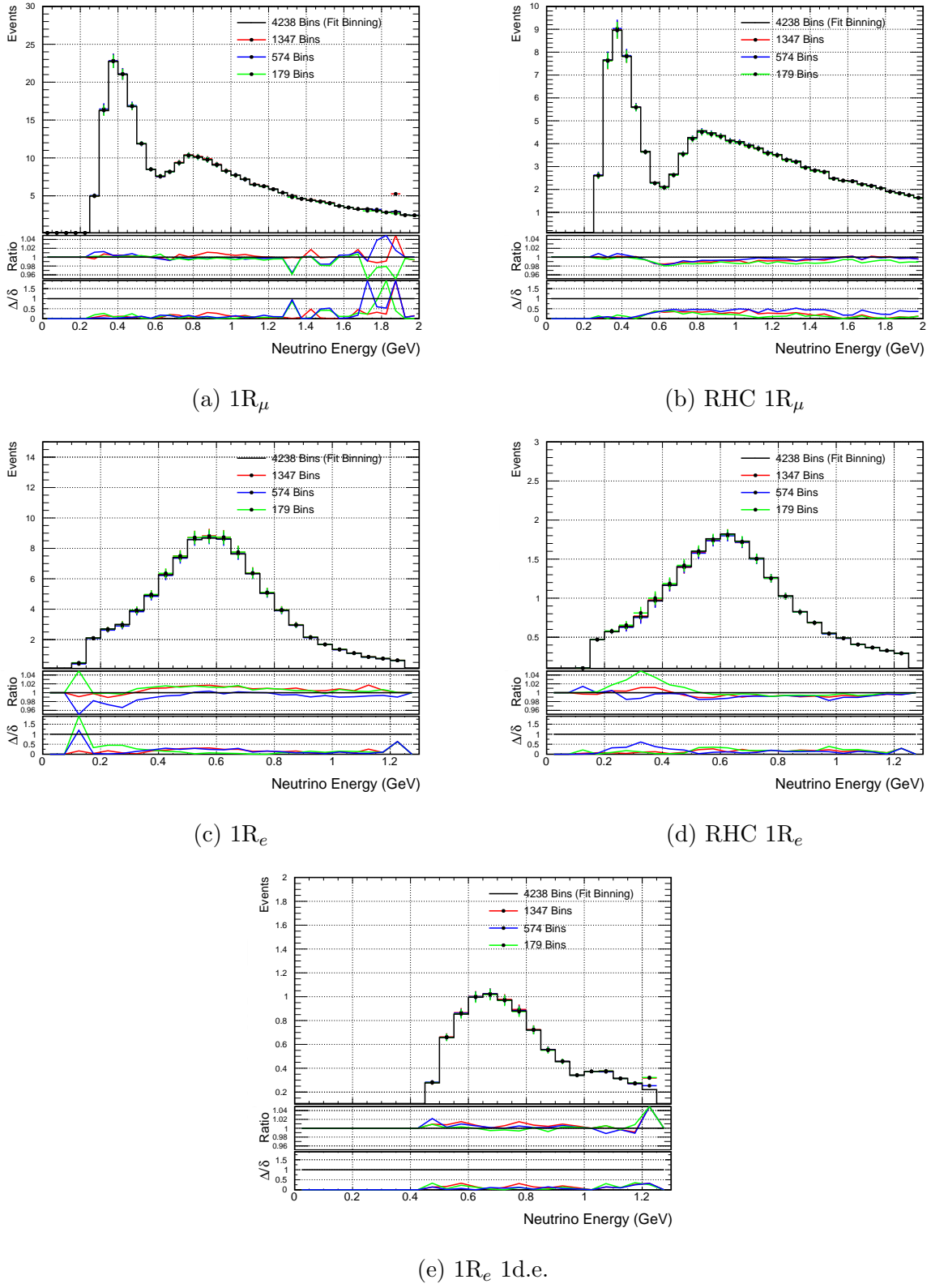
Figure 5.25.: SK posterior predictive distributions from near detector fits using different binnings for the detector covariance.

5.4.3.5. Detector Systematics Covariance Matrices

As the far detector prediction was robust to the different detector bin mergings, and previous analyses have shown that having ~ 600 detector parameters does not cause any issues with fit convergence[4], the 574 bin detector covariance was chosen as the optimum bin merging to use in this analysis.

The binning is as follows:

- **FHC ν_μ CC 0π:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0., 300., 1000., 1250., 1500., 2000., 3000., 5000., 30000.

  $\cos\theta_\mu$: -1.0, 0.6, 0.8, 0.85, 0.9, 0.92, 0.98, 0.99, 1.0

- **FHC ν_μ CC 1π:**

  $\boldsymbol{p}_\mu$ **(MeV/c):** 0., 300., 400., 700., 800., 1000., 1500., 2000., 5000., 30000.

  $\cos\theta_\mu$: -1.0, 0.6, 0.8, 0.9, 0.92, 0.94, 0.96, 0.98, 0.99, 1.0

- **FHC ν_μ CC Other:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0., 300., 400., 700., 800., 900., 1250., 2000., 3000., 5000., 30000.

  $\cos\theta_\mu$: -1.0, 0.6, 0.8, 0.85, 0.9, 0.92, 0.96, 0.98, 0.99, 1.0

- **RHC $\bar{\nu}_\mu$ CC 0π:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0., 300., 2000., 4000., 30000.

  $\cos\theta_\mu$: -1., 0.6, 0.8, 0.9, 0.96, 1.

- **RHC $\bar{\nu}_\mu$ CC 1π:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0., 500., 30000.

  $\cos\theta_\mu$: -1, 0.7, 1.

- **RHC $\bar{\nu}_\mu$ CC Other:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0., 600., 800., 30000.

  $\cos\theta_\mu$: -1., 0.7, 0.95, 0.97, 1.

- **RHC ν_μ CC 0π:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0., 300., 1500., 30000.

  $\cos\theta_\mu$: -1., 0.7, 1.

- **RHC ν_μ CC 1π:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0., 600., 800., 30000.

  $\cos\theta_\mu$: -1, 0.7, 1.

---

[4]For the 2017 analysis, 556 detector bins were used.

- **RHC ν_μ CC Other:**

  p_μ (MeV/c): 0., 600., 30000.

  $\cos\theta_\mu$: -1., 0.7, 1.

In this analysis, another detector covariance was also used for separate fits ran in parallel with the 574 detector bin fits. These fits used the full non-uniform fit binning as the detector binning. As the MC statistics uncertainty was taken out of the detector covariance for this analysis, reducing the terms in the diagonal of the covariance, the number of fit parameters can be reduced using principal component analysis, as discussed in Section 5.4.3.6. This could then allow the full fit binning detector covariance to be used in joint near and far detector fits, which was not previously possible as it would introduce too many fit parameters. This was not done for this analysis, but the full fit detector covariances are used in parallel near detector fits (which can accommodate more fit parameters) to study the potential impact.

The 574 binned detector covariance is shown for the FGD1 CC 0π sample in Figure 5.26. For each momentum range, the bins correspond to the angle ranges, in increasing order, defined above for the 574 merged bins. Bins with similar angles are strongly correlated with each other, and this correlation increases with angle, particularly for bins which apply to the same momentum range. There is a slight anti-correlation between the highest angle bins, and some of the lower angle bins. Bins with similar momentum ranges are also more strongly correlated to each other than bins corresponding to higher or lower momentum.

The two full detector covariances used are shown in Appendix C, showing similar patterns for all samples.

5.4.3.6. Future Detector Systematics Improvements

The treatment of the detector systematics is the weakest point of this analysis. It is currently not computationally feasible to include each uncertainty directly in the fit, but it is also not possible to include ~4000 detector parameters by using the same fit binning for the detector bins. Therefore a compromise is needed, and this is achieved with the bin merging process. Although the results in Section 5.4.3.4 show that the analysis is fairly robust to different binnings, the studies in Section 5.4.3.3 show that the underlying assumption that the shapes of all the systematics are Gaussian is not strictly true.

In future analyses, this could be overcome by directly applying each underlying systematic individually, rather than by using the overall bin normalisations. The implementation of the binding energy parameter, described in Section 5.4.1.1, shows that fitting systematics which directly shift the kinematic variables of events is possible within the MaCh3 framework.

117
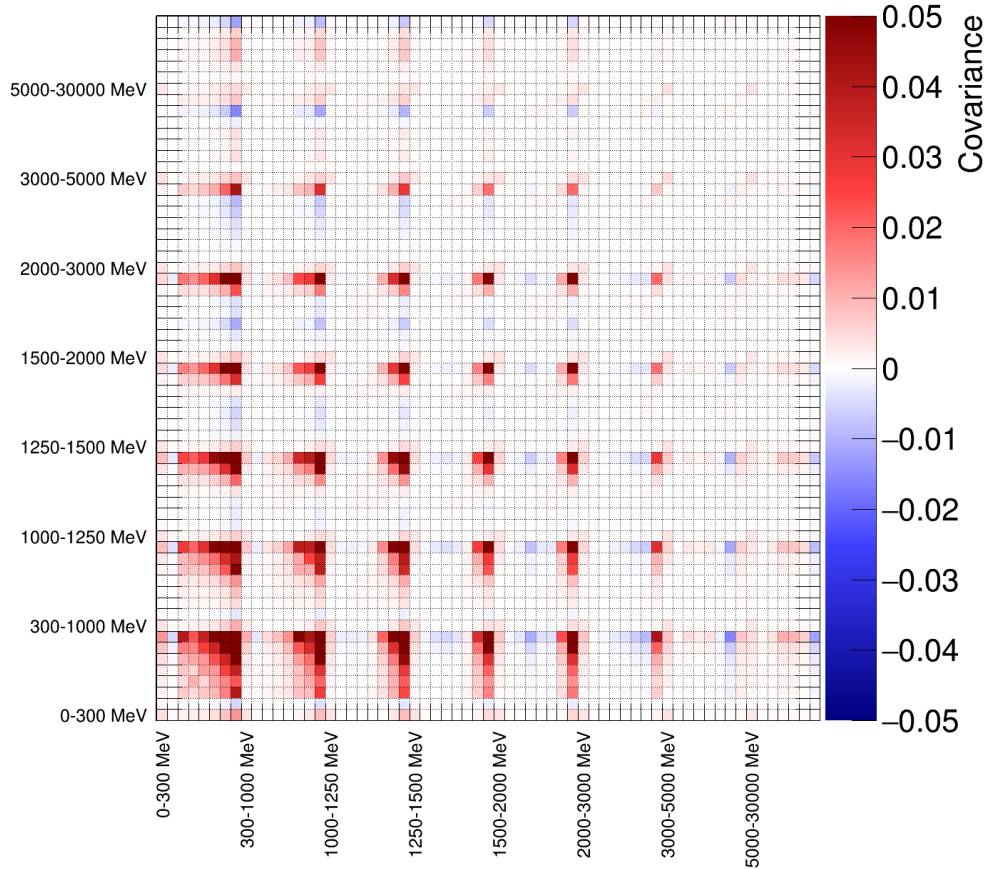
Figure 5.26.: The ND280 detector covariance matrix with 574 merged bins for the FGD1 CC 0π sample, produced using runs 2–9 MC. Bins are ordered in increasing angle ranges for each momentum range. The full matrix for all samples is shown in Appendix C.

This technique could be used particularly for the detector uncertainties that are applied as Observable Variable systematics. These already smear reconstructed variables to produce the normalisation parameters which are then fitted, so it would be natural to directly modify the appropriate variable in the fit. This would allow each detector systematic to be fit without assuming the underlying shapes are all Gaussian.

On a shorter time-scale, the treatment of the detector systematics can be improved by using Principal Component Analysis (PCA) to only fit a certain amount of parameters without having to merge detector bins. PCA produces N fit parameters (where N is the full amount of fit bins), but in principle component basis where each parameter is orthogonal. The parameters can then be cut by the size of their corresponding eigenvalue, to reduce the total number to a more feasible amount.

With this method, the bin merging is avoided, allowing the systematics to more accurately

apply to the correct events, while the number of fit parameters is kept to a manageable number. This is feasible now that the MC statistics uncertainty is no longer included in the detector covariance, reducing the diagonal term in the matrix. Studies have already begun on how many principle components can be cut without an excessive impact on the fit result [138], though this work did not converge in time for this analysis.

5.4.3.7. Detector Systematics Summary

In summary, there are 17 underlying detector systematics which are currently varied 2000 times and applied to the MC, to produce bin-by-bin normalisations. This process assumes the shape of each parameter is Gaussian, which is not strictly true, particularly for the Pion SI uncertainty, but this only affects a small number of bins. The bins are merged to reduce the number of fit parameters, which reduces the accuracy with which the systematics are applied to the correct events, but this is shown to not have a significant impact on the fit results. In future analyses, the bin merging can be avoiding by using PCA to reduce the number of parameters without reducing the number of detector bins. Ultimately, fitting each uncertainty individually, and allowing them to directly shift event kinematics in a similar way to the binding energy parameter will improve the treatment of the detector systematics.

5.5. Prefit Corrections and Scalings

A number of corrections and reweightings are applied to the MC prediction to account for well understood discrepancies between MC and data. These 'one-time' weights are applied before the fit, and are not varied in it.

- **POT:** A scaling is applied to every MC event to account for the fact that more MC POT is produced for each run than data POT. The weight applied to each event depends on the run, and is the ratio of the total data POT with good data quality flag to MC POT in the run.

- **Pile Up:** The detector efficiency in data is lower than in MC because of the coincidence between neutrino events and interactions involving sand muons. This needs to be applied as MC neutrino interactions within the magnet are simulated separately from sand muons. The correction applied depends on the run, and which FGD the event occurred in.

- **Detector:** Precise control studies are used to inform corrections due to well known

Target	p_{μ^+} Shift (MeV)	p_{μ^-} Shift (MeV)
C	+2.6	-3.6
O	+3.3	-4.3

Table 5.3.: Momentum shifts applied to final state leptons in CC events.

hardware and reconstruction deficiencies in the TPC dE/dx and PID. These are applied on an event by event basis, and are discussed in more detail in [116].

- **Flux:**   A correction is made to the nominal neutrino flux for tunings to updated replica target data. The value of the weight applied depends on the run and neutrino energy.

- **Coulomb:**   When the lepton produced in a CCQE neutrino interaction leaves the nucleus, it is either electrostatically attracted to, or repulsed by, the nucleus, depending on its charge. This increases or decreases the momentum of the lepton. To model this, a shift in momentum is applied to CC events. The value of the shift is tuned to electron scattering data [141], and depends on the target nucleus and sign of the lepton, as shown in Table 5.3.

All of these one-time corrections, along with the systematics that vary in the fit (apart from the Coulomb shifts and E_b parameters), are applied to the raw MC as multiplicative weights on an event-by-event basis. This produces the nominal MC prediction which is ultimately fitted to the data.

5.6. Data

T2K has been taking data in distinct run periods since 2010, in both FHC and RHC beam mode. The beam power has been steadily increasing over this time, as shown in Figure 3.6.

MC is produced for each run period individually, so that run-by-run effects such as the beam and detector configurations, and tunings to in-situ beam measurements, can be accounted for.

Table 5.4 shows the amount of MC POT produced in each run, along with the amount of data taken with good data quality flag (defined as POT collected with all sub-detectors

Run	Beam Mode	Data POT ($\times 10^{19}$)	MC POT ($\times 10^{19}$)
2a	**FHC**	3.59	167.99
2w	**FHC**	4.34	120.38
3	**FHC**	15.81	307.77
4a	**FHC**	17.83	361.23
4w	**FHC**	16.43	361.22
5	**RHC**	4.35	221.10
6	**RHC**	34.09	346.99
7	**RHC**	24.38	333.00
8a	**FHC**	41.50	361.10
8w	**FHC**	15.81	254.23
9	**RHC**	20.54	245.61
Total	**FHC**	115.31	1933.89
	RHC	83.36	1146.69
	FHC+RHC	198.67	3080.59

Table 5.4.: Collected and generated POT for the run periods used in this analysis.

of ND280, and the data acquisition system, online). Over runs 2–9, the overall efficiency of good data taking was ∼69%.

Significantly more MC is produced than data, with the weights and corrections described in the previous sections bringing the MC prediction comparable to data.

Runs with an 'a' or 'w' suffix refer to whether the P0D was filled with air (a) or water (w) during the run. These are separated, despite being part of the same global run, as the MC production requires different geometries to be used for these two different configurations. Runs without a suffix had a consistent P0D filling for the entire run. This was water for runs 5, 7, and 9, and air for runs 3 and 6.

The most recent data taking period, run 10, finished in February 2020, but had not been through the full processing in time for this analysis. Run 1 was not included as not all of the sub-detectors of ND280 were online.

5.7. Summary

This chapter has presented an overview of the inputs to the fit and the fitting framework itself. Prior to the near detector fit, the uncertainty on the predicted event rate at SK is ∼ 15%. Fitting the systematics at the near detector reduces this to ∼ 5%, without which,

5. The Near Detector Fit Setup

T2K would not be able to make the precise oscillation measurements it does.

There are two near detector analyses at T2K. The analysis presented in this thesis uses MCMC methods to sample the probability distribution, while the other uses a gradient descent method to find the global minimum.

ND280 data is divided into 18 samples depending on the topology of the event, the beam mode, and which FGD the interaction took place in. These are designed to separate different interaction modes so that systematic uncertainties can be applied accurately. This analysis is the first to use the same topological samples for FHC and RHC data.

These selections are binned in p_μ–$\cos\theta_\mu$. In previous analyses, the use of uniform rectangular binning caused the bins with the highest number of events to be outside the peak region. To avoid this, non-uniform rectangular bins are used in this analysis. An algorithm was written and tuned to produce small enough bins in the peak region to have good resolution, while still having sufficient events in the bins in sparse regions to reduce the statistical uncertainty. The use of non-uniform binning is shown to improve the constraint on systematics significantly.

There are three groups of systematics applied in the fit: cross-section, flux, and detector. The cross-section systematics are provided by the Neutrino Interactions Working Group (NIWG). One of the main improvements to the cross-section model since the last analysis is in the treatment of the binding energy. A new type of parameter has been implemented into the fit, which directly shifts the lepton momentum of an event to represent the uncertainty in the binding energy. Many studies and validations of this parameter have been performed as part of this thesis.

As the binding energy systematic produces discontinuities in the likelihood distribution, it cannot be fitted as direct momentum shifts in the other near detector fitting framework, and so bin by bin splined response functions are used instead. The non-uniform binning is also only used in this analysis. These are the two main advantages of the analysis presented in this thesis over the other near detector fitter.

The flux systematics are provided by the beam group in the form of a covariance matrix. These consist of 100 different normalisations, binned by neutrino energy, species, beam mode, and detector.

The detector systematics are provided by the ND280 Selection, Systematics, and Validations group. These are then varied 2000 times, and the effect on the number of events in each bin used to form normalisation parameters. To reduce the total number of fit parameters, bins with similar responses are merged. The fit is shown to be fairly robust to different bin

mergings. However, in future analyses, using PCA to cut parameters without reducing the number of bins will allow the systematics to be applied more accurately.

The process of throwing the detector systematics to produce normalisations assumes the shape of all underlying systematics is Gaussian. This is shown to not be strictly true. In future analyses, applying each underlying systematic individually will avoid this assumption. The studies of the binding energy parameter show that uncertainties can be fitted as direct shifts to event kinematics, and this could be a natural way of applying some of the detector systematics in the future.

6. 2020 Near Detector Fit Results

The results of the analysis are presented in this chapter, starting with the nominal MC prediction and fit validations in Sections 6.1, 6.2, 6.3, and 6.4. The data fit is presented in 6.5, and the posterior predictive distributions and p-values are shown in 6.6. Results from fits with different binnings are then compared in Section 6.7. Finally, the impact on the sensitivity to oscillation parameters are shown in Section 6.8.

6.1. Nominal MC

The data, unweighted MC, and nominal MC event rates for each sample are shown in Table 6.1. The CC 0π samples are consistently underestimated in the MC prediction by $\sim$15–20%, the CC 1π samples are overestimated by $\sim$5–10% for FHC ν and RHC ν, and underestimated by $\sim$5% for RHC $\bar{\nu}$, and the CC Other samples are underestimated by $\sim$20–30%. The differences to data are consistent across the FGDs to within $\sim$5%. Overall, the MC prediction is 15% lower than the observed data.

The number of unweighted events for each interaction mode are shown in Table 6.2. CCQE is the most common mode, making up $\sim$50% of all events.

The 2D nominal MC distributions for each sample are shown in Figure 6.1. The non-uniform-rectangular binning defined in Appendix B is used to bin the samples for the main results.

The projection of these distributions onto the p_μ axis are shown in Figures 6.2, 6.3, and 6.4, along with the interaction mode breakdown and data.

The ratio of data to MC for fluctuates for the CC 0π and CC Other samples, but is consistently >1. It is slightly increased at the peak momentum for FHC and RHC ν, and decreased at the peak for RHC $\bar{\nu}$. The ratio for the CC 1π samples is more flat in momentum, but shows a small fluctuation <1 at low momentum for FHC ν, and >1 for RHC ν and $\bar{\nu}$. The behaviour is similar across the FGDs.

The FHC ν and RHC $\bar{\nu}$ CC 0π samples are dominated by the target interaction modes CCQE and 2p2h. However, for RHC ν, there is a large contamination of CC 1π events. The

Sample	Raw MC	Nominal MC	Data	Data/MC
FGD1 FHC ν **CC 0π**	524093	27951.1	33443	1.20
FGD1 FHC ν **CC 1π**	127176	8358.97	7713	0.92
FGD1 FHC ν **CC Other**	99730	7031.47	8026	1.14
FGD2 FHC ν **CC 0π**	521757	27556.2	33156	1.20
FGD2 FHC ν **CC 1π**	103305	6723.98	6281	0.93
FGD2 FHC ν **CC Other**	94164	6454.68	7700	1.19
FGD1 RHC $\bar{\nu}$ **CC 0π**	115456	7270.56	8388	1.15
FGD1 RHC $\bar{\nu}$ **CC 1π**	9272	694.32	698	1.01
FGD1 RHC $\bar{\nu}$ **CC Other**	16790	1286.78	1472	1.14
FGD2 RHC $\bar{\nu}$ **CC 0π**	112390	7036.71	8334	1.18
FGD2 RHC $\bar{\nu}$ **CC 1π**	8533	624.76	650	1.04
FGD2 RHC $\bar{\nu}$ **CC Other**	15616	1176.62	1335	1.18
FGD1 RHC ν **CC 0π**	41789	3035.85	3594	1.13
FGD1 RHC ν **CC 1π**	14304	1159.02	1111	0.96
FGD1 RHC ν **CC Other**	12733	1073.16	1344	1.25
FGD2 RHC ν **CC 0π**	41554	3013.01	3433	1.14
FGD2 RHC ν **CC 1π**	11472	930.64	926	1.00
FGD2 RHC ν **CC Other**	11954	1000.03	1245	1.24
Total	1882090	112378	128849	1.15

Table 6.1.: MC and data event rates for the ND280 samples.

Interaction	Number of Events
CCQE	827104
2p2h	134298
CC 1π	462170
CC coherent	14065
CC multi-π	174069
CC DIS	185284
CC miscellaneous	26643
NC 1π^0	3476
NC 1$\pi^\pm$	15218
NC coherent	271
NC 1γ	11
NC Other	61334
Total	1903943

Table 6.2.: MC event rates broken down by interaction mode.

FHC ν and RHC $\bar{\nu}$ CC 1π samples are dominated by the target interaction modes CC 1π, CC coherent, and CC multi-π, but for RHC ν, the 1π sample has a significant number of CC DIS events. The CC Other samples are populated mainly by the target interaction modes CC DIS, CC multi-π, and CC miscellaneous, but with a significant number CC 1π and CC coherent events for FHC ν and RHC $\bar{\nu}$.

The projections onto the $\cos\theta_\mu$ axis are shown in Appendix D, along with the data and

(a) FGD1 FHC ν_μ 0π (b) FGD1 FHC ν_μ 1π (c) FGD1 FHC ν_μ Other

(d) FGD2 FHC ν_μ 0π (e) FGD2 FHC ν_μ 1π (f) FGD2 FHC ν_μ Other

(g) FGD1 RHC $\bar{\nu}_\mu$ 0π (h) FGD1 RHC $\bar{\nu}_\mu$ 1π (i) FGD1 RHC $\bar{\nu}_\mu$ Other

(j) FGD2 RHC $\bar{\nu}_\mu$ 0π (k) FGD2 RHC $\bar{\nu}_\mu$ 1π (l) FGD2 RHC $\bar{\nu}_\mu$ Other

(m) FGD1 RHC ν_μ 0π (n) FGD1 RHC ν_μ 1π (o) FGD1 RHC ν_μ Other

(p) FGD2 RHC ν_μ 0π (q) FGD2 RHC ν_μ 1π (r) FGD2 RHC ν_μ Other
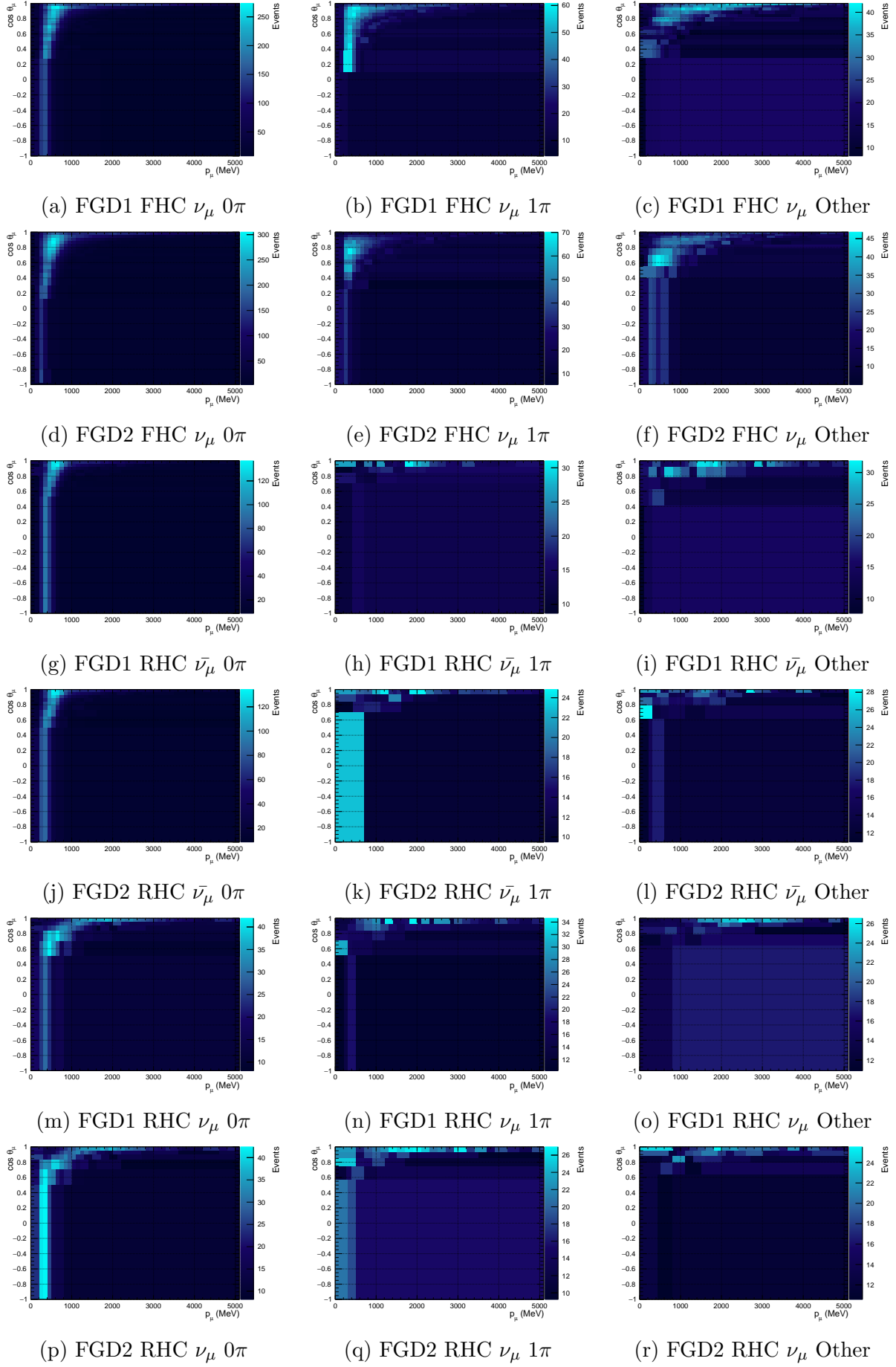
Figure 6.1.: p_μ–$\cos\theta_\mu$ distributions for the nominal MC with non-uniform rectangular binning.
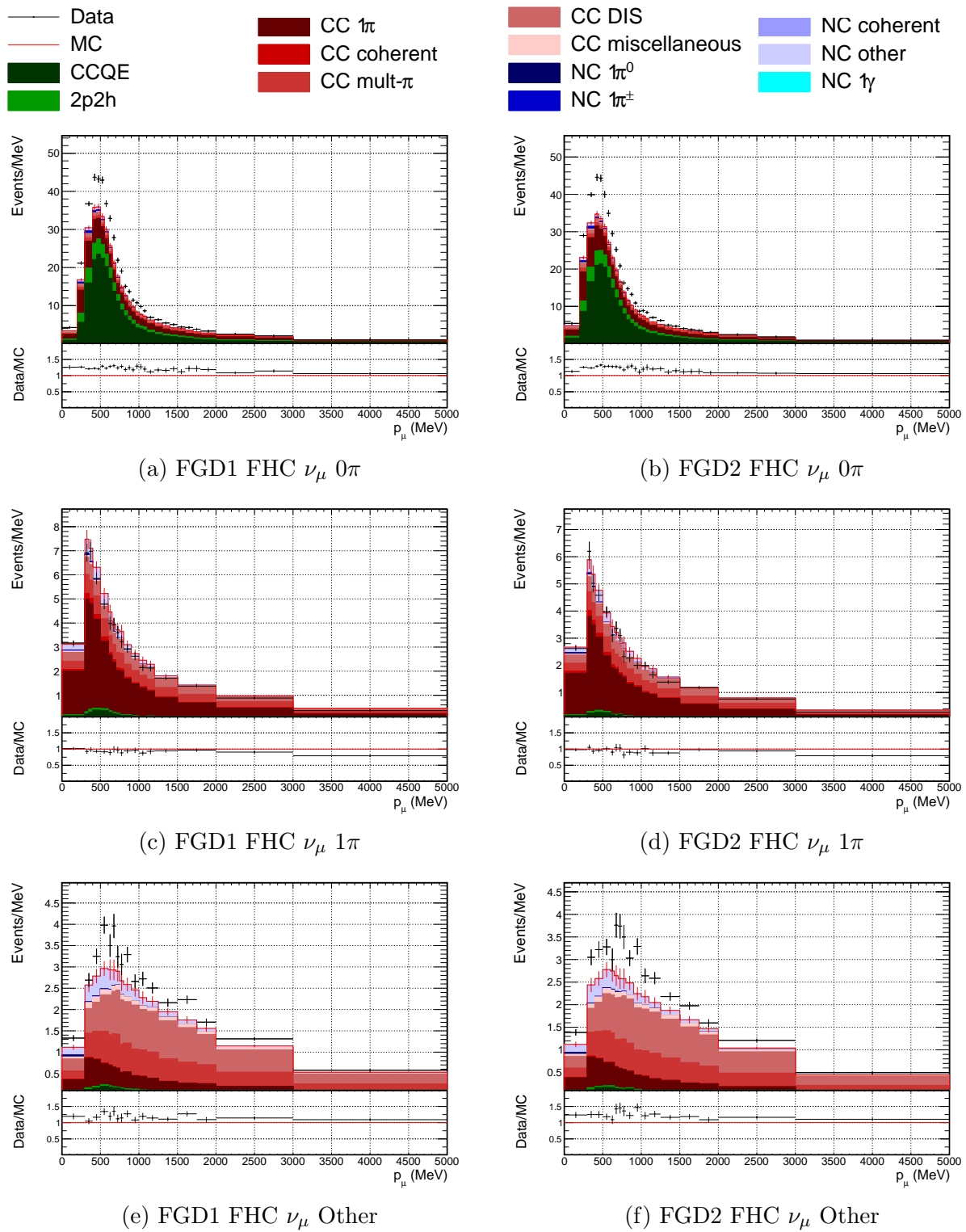
Figure 6.2.: p_μ projections of data and nominal MC broken down by interaction mode for FHC selections.

interaction mode breakdown.

The ratio of data to MC for CC 0π and CC Other samples again fluctuates, but always remains > 1. For the CC 1π samples, the ratio is more flat, but at high angle oscillates between the MC over and underestimating the data. The behaviour is consistent across
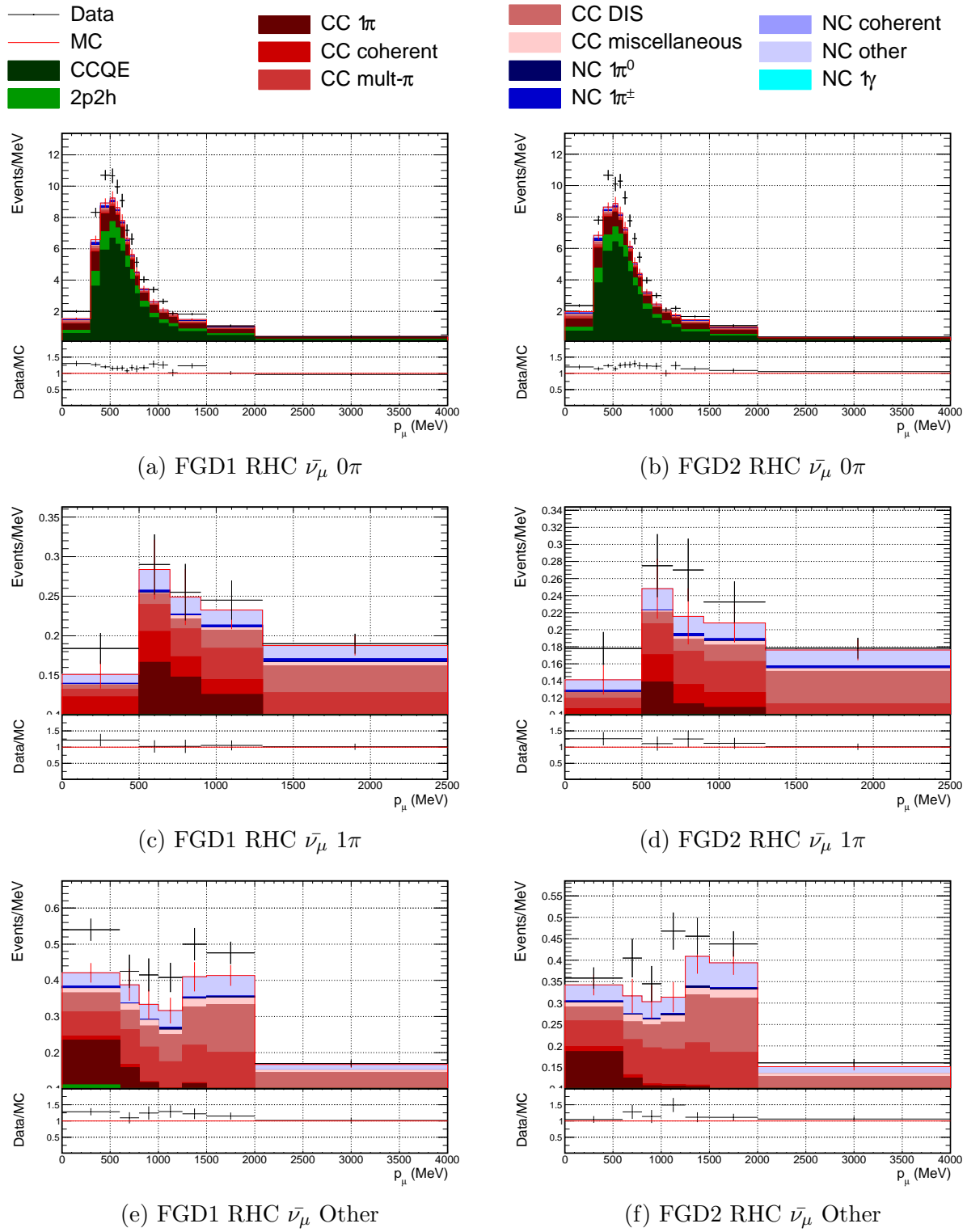
(a) FGD1 RHC $\bar{\nu}_\mu$ 0π

(b) FGD2 RHC $\bar{\nu}_\mu$ 0π

(c) FGD1 RHC $\bar{\nu}_\mu$ 1π

(d) FGD2 RHC $\bar{\nu}_\mu$ 1π

(e) FGD1 RHC $\bar{\nu}_\mu$ Other

(f) FGD2 RHC $\bar{\nu}_\mu$ Other

Figure 6.3.: p_μ projections of data and nominal MC broken down by interaction mode for RHC $\bar{\nu}_\mu$ selections.

FGD1 and FGD2, showing that the strengths and weaknesses of the modelling are similar for both FGDs.
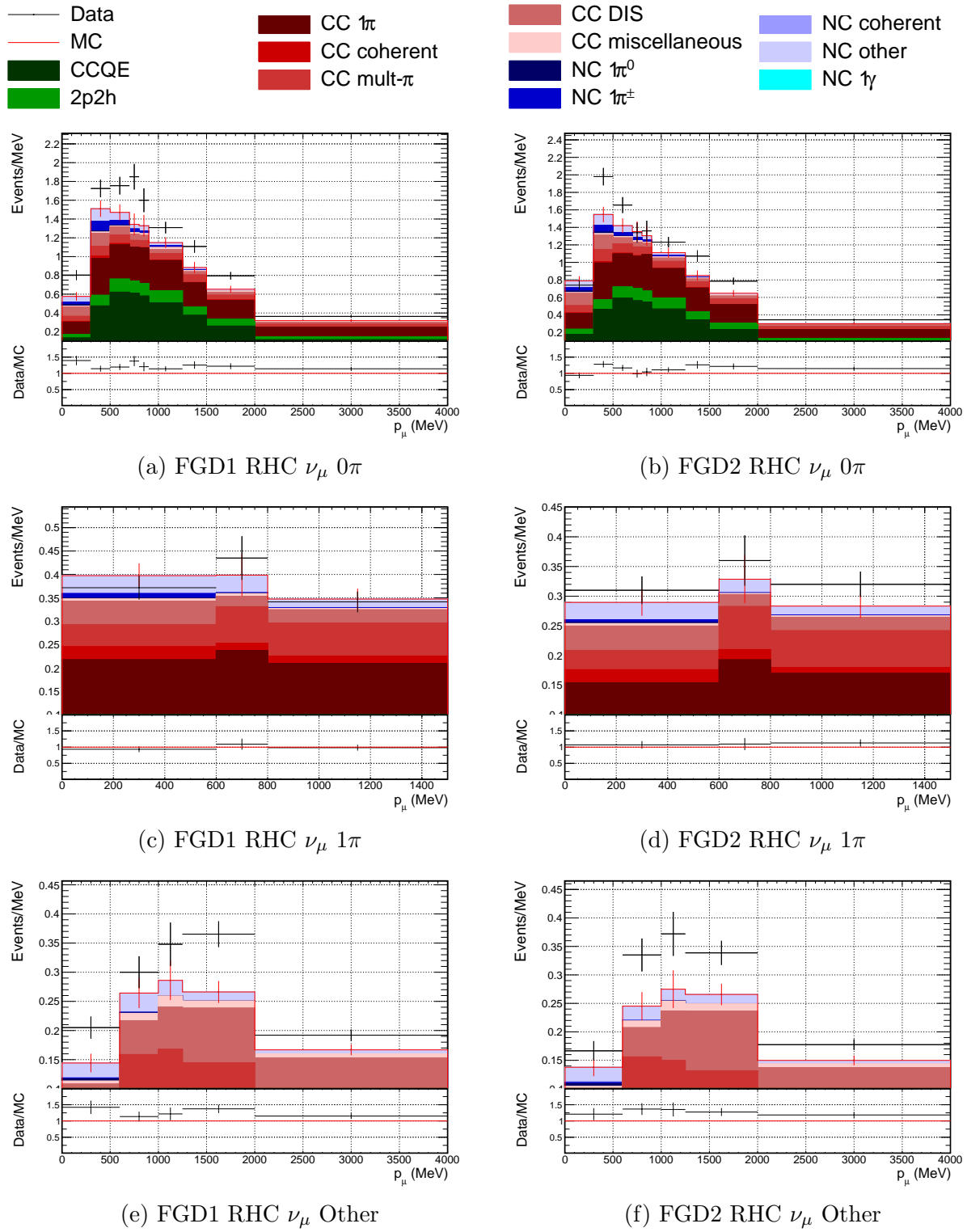
Figure 6.4.: p_μ projections of data and nominal MC broken down by interaction mode for RHC ν_μ selections.

6.2. Log-likelihood Scans

As described in Section 4.4.1, the marginalisation effects from extracting correlated and non-Gaussian parameters from the full posterior distribution can cause the fit to appear biased.

A full Asimov fit alone, described in Section 6.4, is therefore not a good method of validating the framework.

Log-likelihood scans are also run as part of the validations. The nominal MC is set as the data, and each systematic parameter is varied one at a time to 150 equally spaced points from -1σ to +1σ. At each step, the MC is reweighted and the total likelihood from all contributions calculated. Only the diagonal terms of the covariance matrices are used for the penalty contribution, as otherwise varying one parameter alone could invoke significant penalties from correlations. The scans are therefore not a fully accurate measure of the sensitivity of the fit to constrain each systematic, but a useful validation of the framework.

After each scan, the parameter is reset, and the next parameter in question varied. The likelihood response is expected to be fairly Gaussian for each parameter, as the prior uncertainty is either Gaussian or flat, and most parameters are expected to have a symmetric effect on the number of events in individual bins. The minimum should be at the prior central value of the parameter, and the log-likelihood here should be 0, as at this point the reweighted MC is identical to the nominal MC. No variation of a single parameter should be able to produce a set of distributions more similar to the nominal MC than itself.

The log-likelihood scans for four selected interaction parameters are shown in Figure 6.5. As expected, the test statistic minimises to 0 at the prior central value of each parameter. The penalty contribution to the log-likelihood dominates for the CC normalisation parameter, due to the prior uncertainty being so small. Conversely, the 2p2h ^{12}C to ^{16}O normalisation parameter has a weaker prior and therefore a larger contribution from the sample likelihood. The likelihood for the $0.05 < Q^2 < 0.10 \, \mathrm{GeV}^2$ normalisation parameter is entirely dominated by the sample contribution, as the prior is flat. The CC DIS and multi-π $\bar{\nu}$ normalisation parameter has more balanced contributions from both the sample and penalty likelihoods.

The log-likelihood scans for four selected flux parameters are shown in Figure 6.6. The test statistic again minimises to 0 at the prior central value of each parameter, as expected. These parameters all have tight prior uncertainties, and so the penalty terms dominate the likelihoods. For the SK flux parameters, there is no sample contribution to the likelihood. This is expected as the SK flux parameters should have no effect on the ND280 samples (apart from through the correlations with ND280 flux parameters, which are not included in these scans).

The log-likelihood scans for four selected ND280 detector parameters are shown in Figure 6.7. As expected, the test statistics all minimise to 0 at the prior central value of each parameter. The prior dominates for all regions of p_μ–$\cos\theta_\mu$ in each sample. For the higher
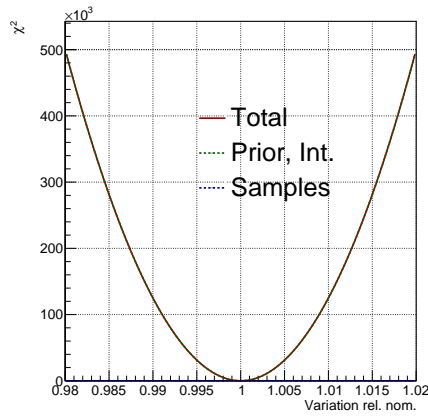
(a) CC normalisation ν
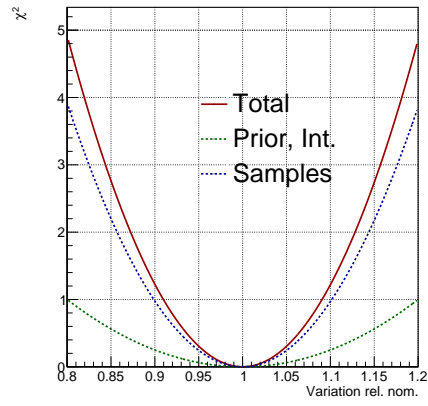
(b) 2p2h ^{12}C to ^{16}O normalisation
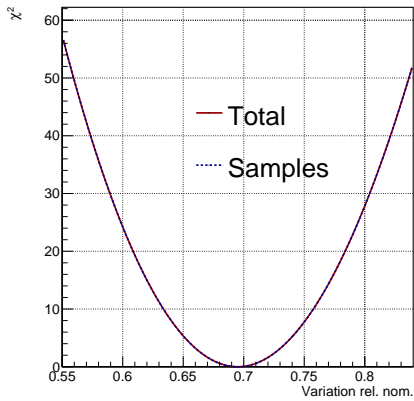
(c) $0.05 < Q^2 < 0.10$ normalisation
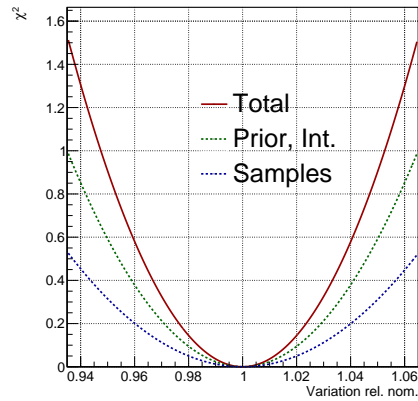
(d) CC DIS and multi-π normalisation

Figure 6.5.: Log-likelihood scans for selected interaction parameters. CC norm. ν has a tight prior uncertainty that dominates the likelihood, whereas the low Q^2 normalisations have a flat prior so the sample is the only contribution. The 2p2h ^{12}C to ^{16}O and CC DIS and multi-π normalisations have significant contributions from both the sample and prior uncertainty.

statistic regions (eg. FGD1 FHC ν_μ CC 0π: 300–1000 MeV, 0.92–0.98), the overall constraint is larger than for the lower statistic regions, (eg. FGD1 FHC ν_μ CC 1π: 5000–30000 MeV, -1.0–0.6).

The sample, prior, and total log-likelihood distributions were compared with the other near detector fitting group, and good agreement was found for all parameters.

6.3. Parameter Variations

As a further validation of the fitting framework and models, the parameters are again each set to $\pm 1\sigma$, one by one, while all others are held at nominal. Instead of the change in likelihood, here the effect on the event distributions in p_μ–$\cos\theta_\mu$ is inspected.

One varied interaction parameter for each sample is shown in Figure 6.8. The combinations

(a) ND280 FHC ν_μ 1–1.5 GeV

(b) ND280 FHC ν_e 0.5–0.7 GeV

(c) ND280 RHC ν_e 2.5–30 GeV

(d) SK RHC $\bar{\nu}_\mu$ 0.5–0.6 GeV
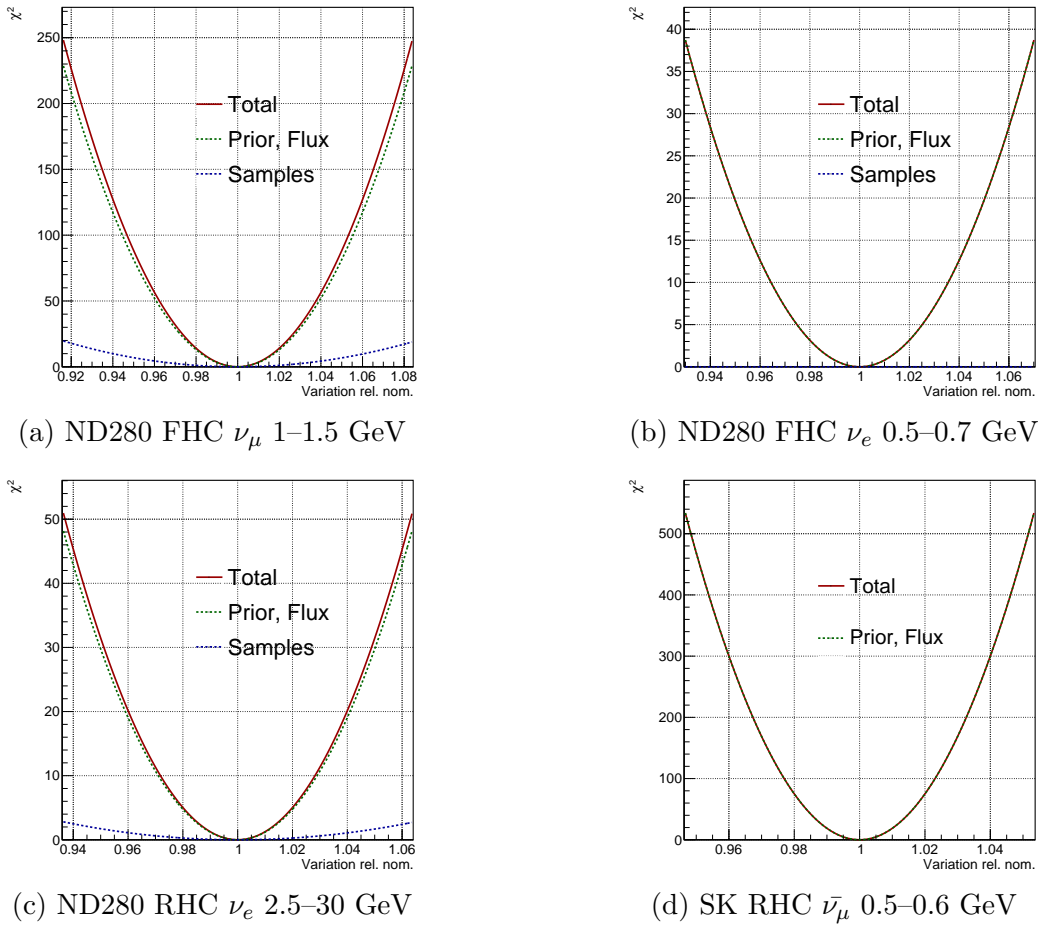
Figure 6.6.: Log-likelihood scans for selected flux parameters.

of parameter and sample were selected such that the parameter controls interactions targeted by the sample. The parameter is therefore expected to have a significant and well-understood impact on the shown sample. The selection of parameter and sample was also made such that no parameters are shown more than once. The parameter in question is set to $+1\sigma$ above its nominal value, and the ratio of the reweighted MC to the nominal MC is taken.

The 2p2h ν normalisation, Q^2 normalisations, and M_A^{QE} parameters all have a Q^2 dependence in the response of event distributions when set to $+1\sigma$. M_A^{QE} and $Q^2 > 1.0$ GeV^2 have a larger effect at high Q^2, while the $0.00 < Q^2 < 0.05$ GeV^2 controls the lower Q^2 region, as would be expected.

The π FSI, and 2p2h shape ^{12}C parameters reduce the number of events in the shown samples, despite being set higher than nominal. This is because they are spline parameters not normalisation parameters, and so the weight applied can be lower for a higher parameter value. The AGKY mult-π, C_5^A, M_A^{RES}, BY DIS, BY mult-π, and $I_{1/2}$ uncertainties are all shape parameters which cause an increase in events in the samples shown when set to $+1\sigma$.

133

(a) FGD1 FHC ν_μ CC 0π: 300–1000 MeV, 0.92–0.98

(b) FGD1 FHC ν_μ CC 1π: 5000–30000 MeV, -1.0–0.6

(c) FGD2 RHC $\bar{\nu}_\mu$ CC Other: 800–30000 MeV, 0.97–1.0

(d) FGD1 RHC ν_μ CC Other: 600–30000 MeV, -1.0–0.7
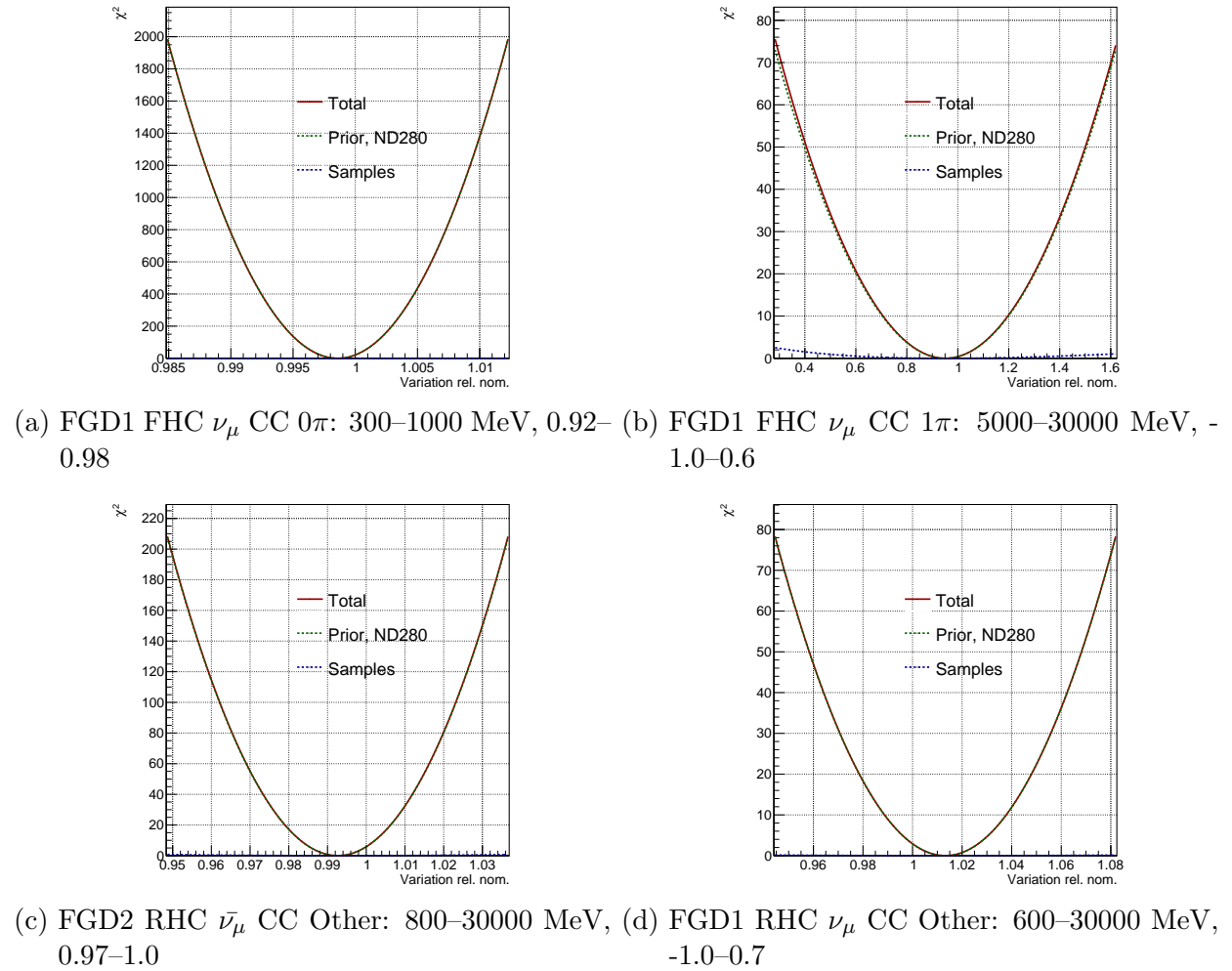
Figure 6.7.: Log-likelihood scans for selected ND280 detector parameters.

The $E_{\rm b}$ ^{16}O $\bar{\nu}$ parameter causes an increase in events at low momentum, and decrease at higher momentum, as the events are directly shifted and not just reweighted.

The DIS normalisations, CC coh. ^{16}O, and CC misc. normalisations all increase events at high angle when set to $+1\sigma$. The DIS normalisations effect higher momentum events, as would be expected as DIS interactions tend to involve higher energies. The CC misc. parameter has a large impact despite only affecting a small number of events, because of its large uncertainty. When it is set to $+1\sigma$ it is therefore significantly higher than at nominal.

All the variations are causing changes to the event distributions in the regions each parameter would be expected to. The total number of events in each sample at each variation was compared with the other near detector fitting group to verify that each parameter is behaving in the same way in each framework. Good agreement was found for all parameters, for all samples. For these cross-group validations, the uniform binnning defined in Appendix B was used.

(a) FGD1 FHC ν_μ 0π: 2p2h Norm. ν += 1σ (b) FGD1 FHC ν_μ 1π: π FSI QE High E += 1σ (c) FGD1 FHC ν_μ Other: CC AGKY Mult. += 1σ

(d) FGD2 FHC ν_μ 0π: $Q^2 >$ 1.0 GeV^2 += 1σ (e) FGD2 FHC ν_μ 1π: CA5 += 1σ (f) FGD2 FHC ν_μ Other: DIS/mult π Norm. ν += 1σ

(g) FGD1 RHC $\bar{\nu}_\mu$ 0π: MAQE += 1σ (h) FGD1 RHC $\bar{\nu}_\mu$ 1π: M_A^{RES} += 1σ (i) FGD1 RHC $\bar{\nu}_\mu$ Other: BY DIS += 1σ

(j) FGD2 RHC $\bar{\nu}_\mu$ 0π: E_b ^{16}O $\bar{\nu}$ += 1σ (k) FGD2 RHC $\bar{\nu}_\mu$ 1π: $I_{1/2}$ += 1σ (l) FGD2 RHC $\bar{\nu}_\mu$ Other: DIS/mult–π Norm. $\bar{\nu}$ += 1σ

(m) FGD1 RHC ν_μ 0π: 2p2h Shape ^{12}C += 1σ (n) FGD1 RHC ν_μ 1π: CC Coh ^{12}C += 1σ (o) FGD1 RHC ν_μ Other: CC Misc. += 1σ

(p) FGD2 RHC ν_μ 0π: 0.00 < Q^2 < 0.05 GeV^2 += 1σ (q) FGD2 RHC ν_μ 1π: π FSI Abs += 1σ (r) FGD2 RHC ν_μ Other: BY mult-π += 1σ

Figure 6.8.: Ratio of each sample to nominal with one parameter set to +1σ. The selected parameters shown all affect interactions which the sample they are shown for target.

135

6.4. Asimov Fit

The Asimov dataset[1] is defined as the MC prediction with all systematic parameters set to their nominal values [142]. For an Asimov fit, the nominal MC prediction is set to be the 'data', and is then fitted to itself. This is completely unphysical, as there can be a non-integer number of 'data' events, but means there are no statistical fluctuations in the dataset, and the expected result of the fit is known. Therefore any deviations from the expected result indicate problems with the fitter. The results can also be used to obtain the maximum sensitivity of the fit. The constraint on each parameter shows the reduction in systematic uncertainties that would be achieved if the models perfectly described the true data. This represents the maximum possible constraint, as the sample and parameter likelihoods are each maximised for the same set of parameter values.

The results of the Asimov fit for the ND280 FHC flux parameters are shown in Figure 6.9. As described in Section 4, the fit does not find a single best-fit set of parameters, but single parameter values are extracted from the posterior distribution by marginalising over all but one parameter, one by one. Marginalisation effects, whereby marginalising over non-Gaussian parameters shifts the highest posterior density for a given parameter, cause the postfit parameter values to not exactly equal the nominal inputs, but the discrepancies are small. The flux parameters with the largest discrepancies are those that apply to rarer events at ND280, such as for high energies, and so are constrained mostly by the prior uncertainty only.

As the ND280 samples all target ν_μ or $\bar{\nu}_\mu$, the ν_e and $\bar{\nu}_e$ parameters are mainly constrained only by the prior, so the pre and post Asimov fit uncertainties are very similar. As there is no wrong-sign FHC sample, there is also little constraint beyond the prior for FHC $\bar{\nu}_\mu$.

The constraint on the SK flux parameters comes entirely from the correlation with the ND280 ones, and so the postfit values have the same behaviour, as shown in Appendix F, along with the ND280 RHC flux parameters.

The results for the interaction parameters are shown in Figure 6.10. All parameters stay close to their nominal values, as would be expected. There are several parameters which show small deviations, but these are again parameters which are not well constrained, such as NC 1γ, as there are few events of the relevant interaction mode at ND280. The 2p2h energy dependence and low π momentum $I_{1/2}$ parameters are not fitted at the near detector, so aren't shown here.

---

[1]Named after the Isaac Asimov short story, *Franchise*, in which an individual is chosen as the sole voter as their views represent those of the whole population.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$
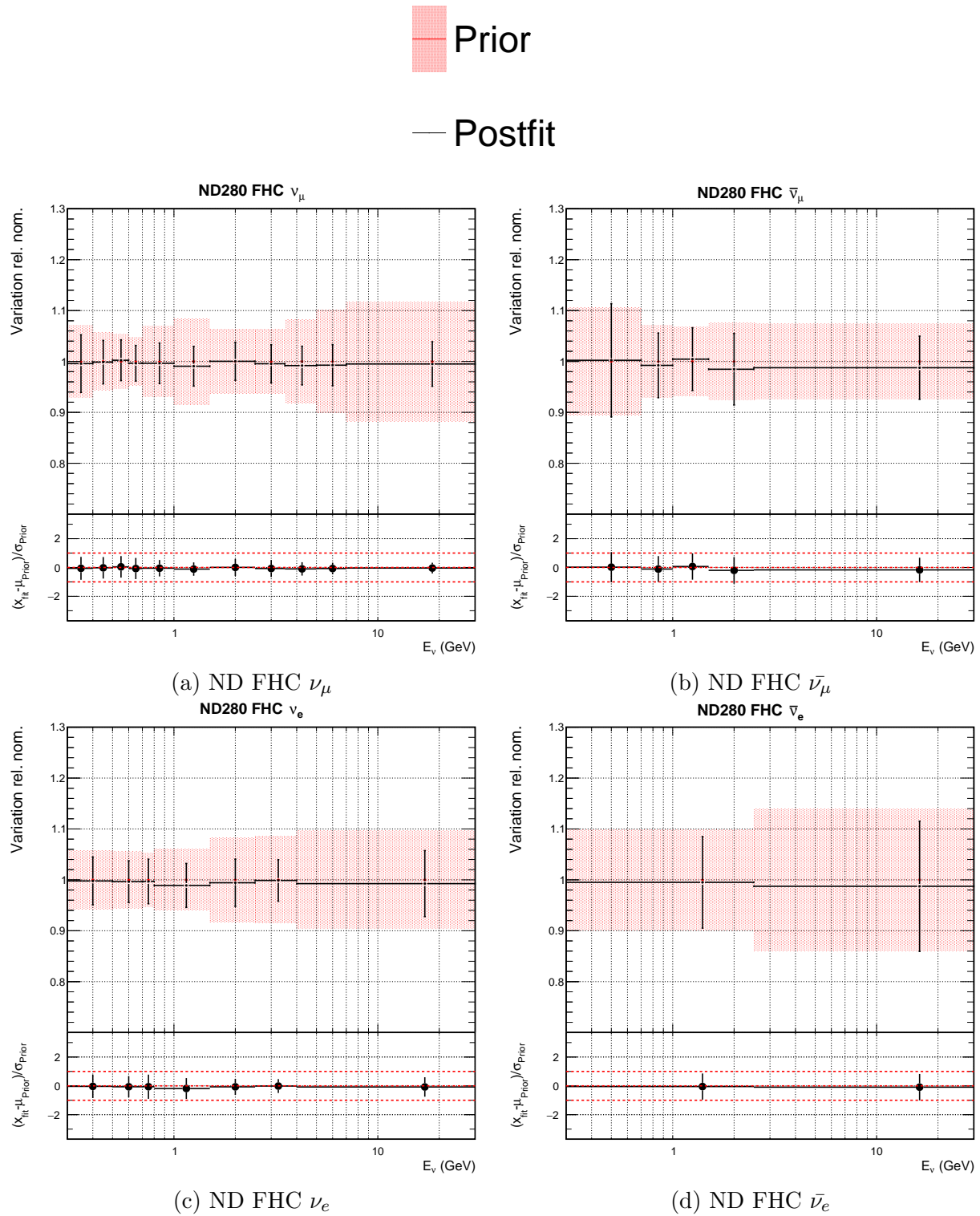
Figure 6.9.: ND280 FHC flux parameters for the Asimov fit.

The uncertainty on the majority of parameters has been reduced by the fit. The parameters which are not constrained are either due to there being a very small number of (or 0) events affected by them in ND280 samples, such as NC Other SK (which only applies to SK events), or there being a very strong prior uncertainty, such as the CC ν and $\bar{\nu}$, ν_e/ν_μ, and multi-π and DIS normalisations.

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
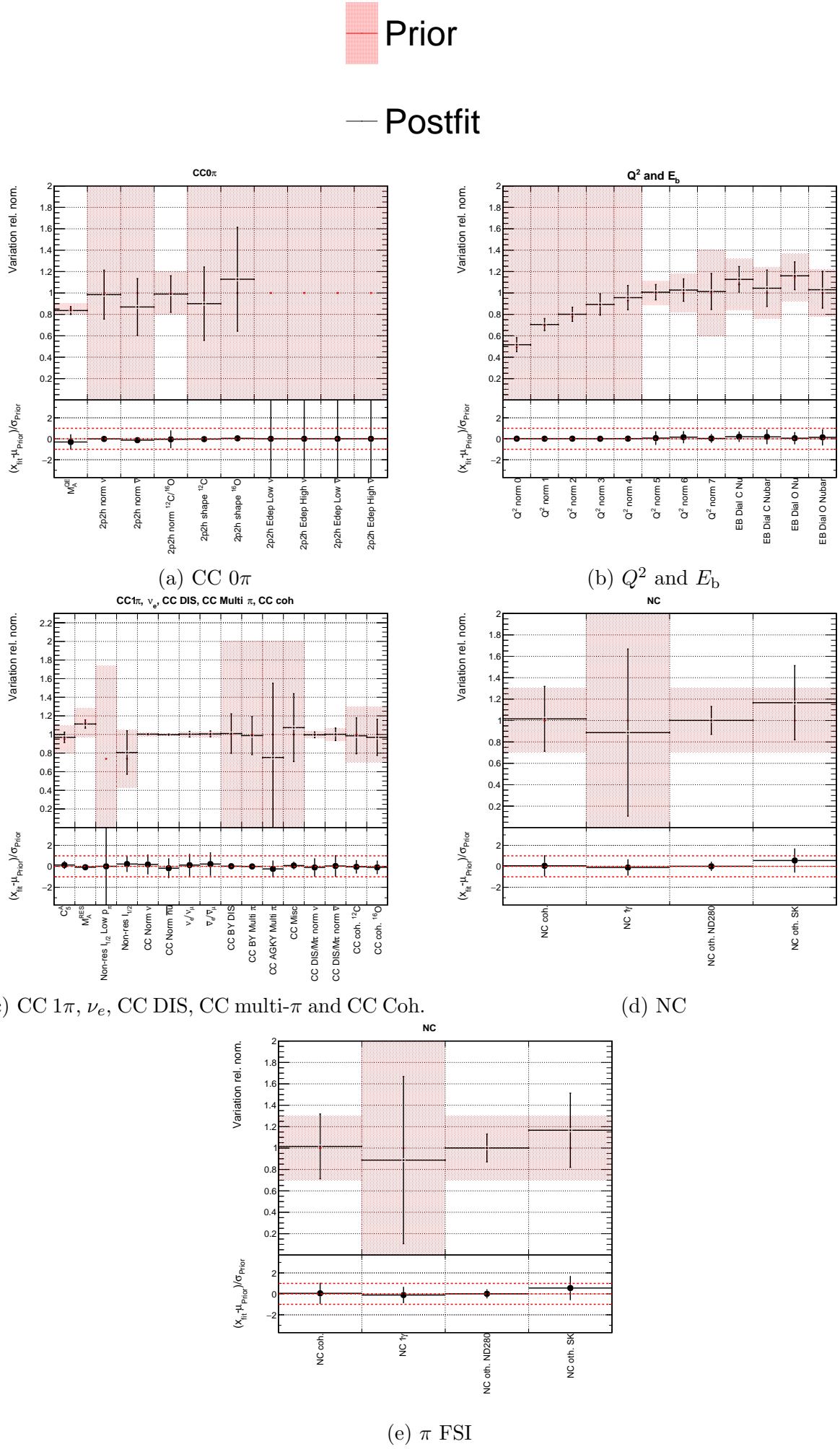
Figure 6.10.: Interaction parameters for the Asimov fit.
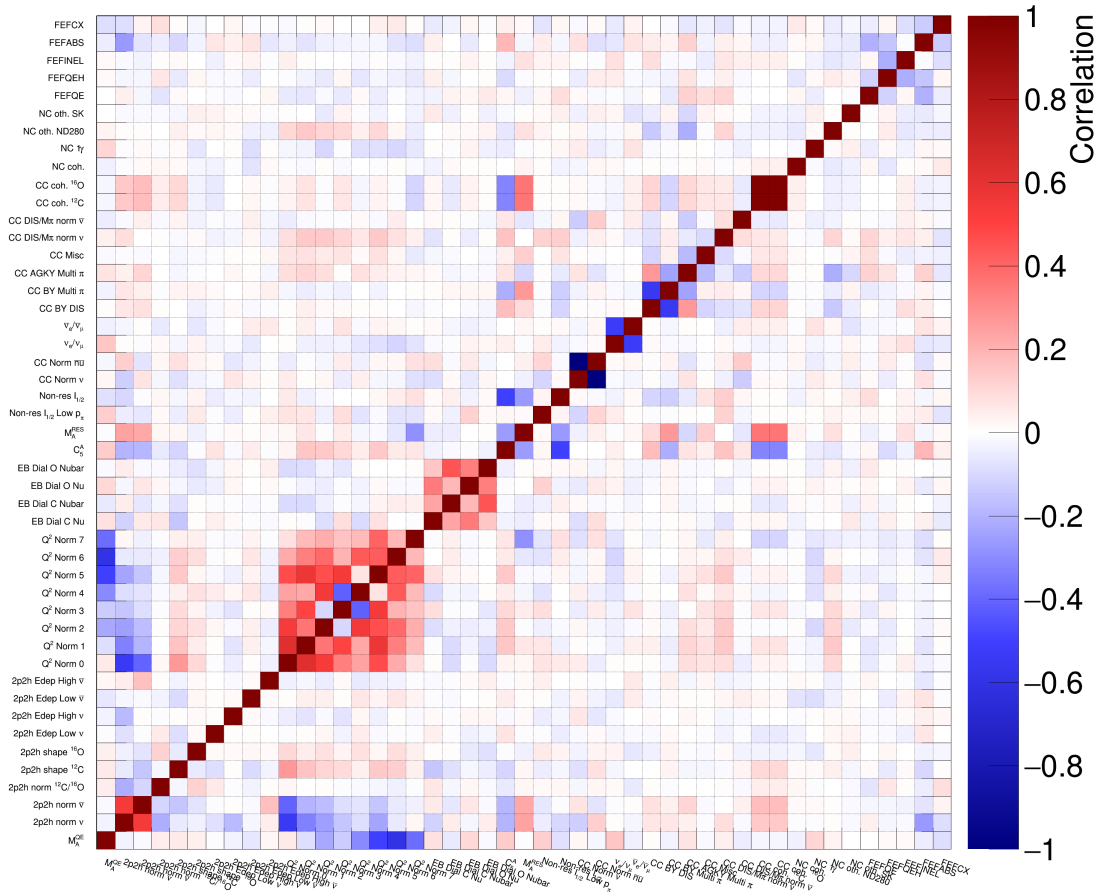
Figure 6.11.: Asimov postfit correlation matrix for interaction parameters.

The postfit correlation matrix for the interaction parameters is shown in Figure 6.11, and both the flux and interaction parameters together are shown in Appendix G. There are strong correlations between uncertainties which control interactions with similar topologies, such as the M_A^{QE}, 2p2h, and Q^2 parameters. These all affect different regions of Q^2, but also all correlate with the flux parameters, causing them to correlate with other. The CC coherent parameters correlate with each other, and the CC 1π parameters. There are strong internal correlations for the E_b, π FSI and single π production parameters. There are also slight correlations between CCQE and CC 1π parameters, due to the contamination of CC 1π events in the CC 0π samples. The flux parameters have strong internal correlations from their priors, and are anti-correlated with many interaction parameters, particularly normalisations.

6.5. Data Fit

6.5.1. Prior Predictions

Prior predictions are produced using a similar method to the posterior predictions described in Section 4.4.3. However, instead of using draws from the Markov Chain, correlated throws of the fit parameters are made. For the parameters with Gaussian priors, the throws are from a Gaussian with the same central value and width as the prior. For the parameters with flat priors, the throws are from a uniform distribution between physical bounds. For each of 2000 throws, the nominal MC is reweighted to the thrown parameter values. Each bin in each sample therefore has 2000 different number of events, from which the central value and uncertainty is used to build the prediction in the same way as for the posterior predictions. This method has the advantage of incorporating the prior uncertainties when inspecting how well the nominal model fits the data, which just looking at the nominal MC does not do. The prior prediction therefore gives a better gauge of how significant the discrepancies between the unfitted model and the data are. The nominal MC and prior prediction are not expected to be identical as they are constructed differently; the latter builds predictions from the prior and then averages whereas the former just takes the prior at the central value, but large discrepancies would be surprising.

The prior prediction can also be used to compare to the posterior predictive distributions, to show how the fit has changed not just the shape of the predictions, but also the uncertainties. The constraining of systematics by the fit is expected to reduce the uncertainties on the prediction.

The p_μ projections of the prior predictions for the FHC samples are shown in Figure 6.12. The rest of the samples, and $\cos\theta_\mu$ distributions are shown in Apprendix E.

As expected from the comparisons of the nominal MC to data, the prior predictive distributions underestimate the data significantly, particularly in the peak region around $p_\mu \sim 600$ MeV for the CC 0π and CC Other samples. For the CC 1π samples, there are regions of significant overestimation. The levels of discrepancy are not concerning though, if the prior model perfectly described the data with small uncertainties the fit would not be needed. The prior predictions are compared to the posterior predictions in Table 6.3.
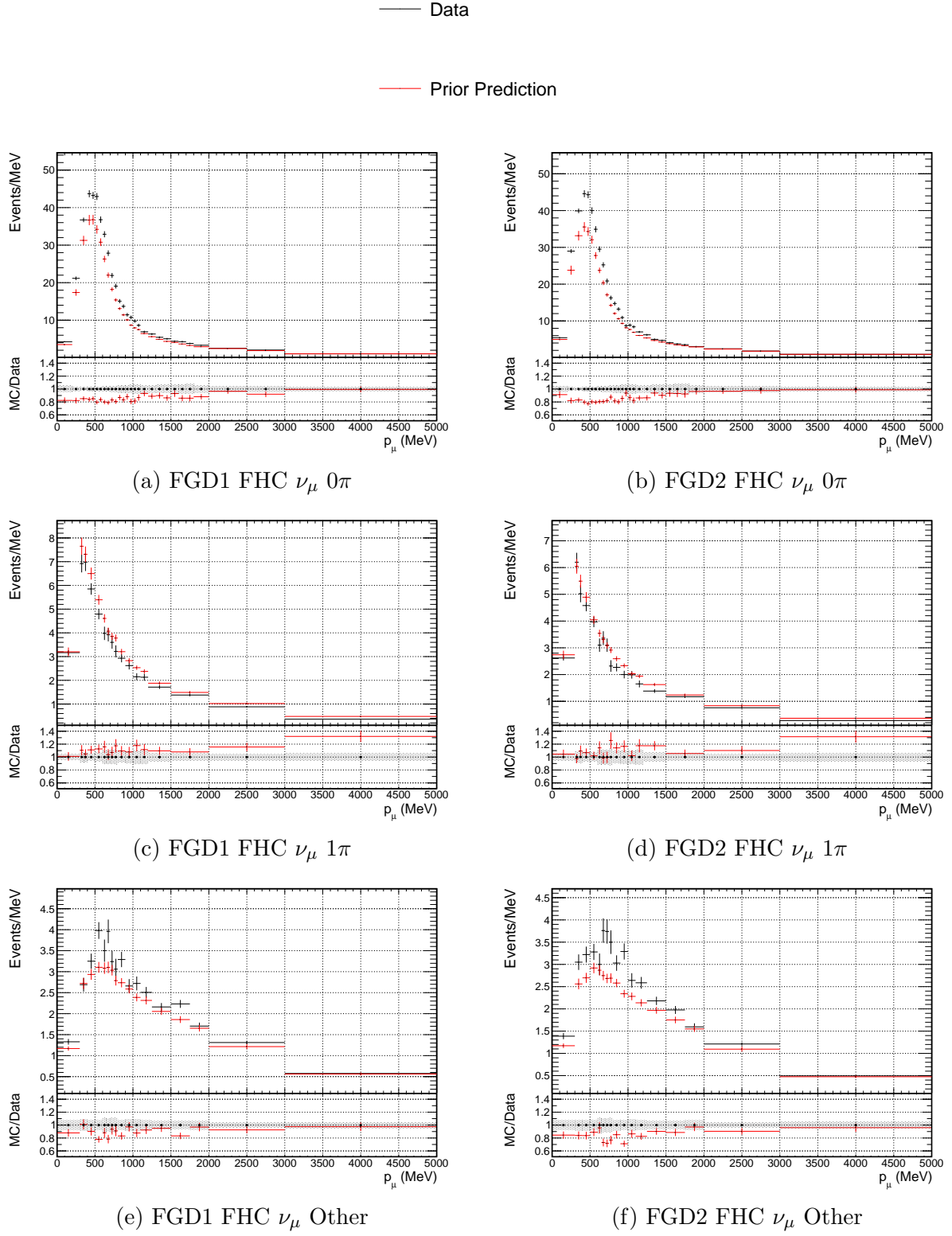
(a) FGD1 FHC ν_μ 0π

(b) FGD2 FHC ν_μ 0π

(c) FGD1 FHC ν_μ 1π

(d) FGD2 FHC ν_μ 1π

(e) FGD1 FHC ν_μ Other

(f) FGD2 FHC ν_μ Other

Figure 6.12.: p_μ projections of the prior predictive distributions and data for FHC ν_μ selections.

6.5.2. Fit Results

The MC was then fitted to the real data. The postfit ND280 FHC flux parameter values are shown in Figure 6.13. The full fit results are shown in Appendix F. For FHC ν_μ and ν_e, there is a pull of $\sim$10% below 1 GeV. The pull decreases as the energy decreases, and falls below nominal at higher energies. A similarly high pull is seen for the FHC $\bar{\nu}_\mu$ and RHC ν_μ parameters, and this is fairly constant in energy.

For FHC $\bar{\nu}_e$ and RHC ν_e, the pull is $\sim$8% for the high energy parameter, but the low energy parameter is slightly closer to nominal. The RHC $\bar{\nu}_\mu$ and $\bar{\nu}_e$ parameters are also pulled significantly upwards, to $\sim$5–10% decreasing with energy.

Similar behaviour is seen for the ND280 and SK parameters, as would be expected due to their prefit correlations.

Although many of the flux parameters are pulled significantly away from their prior central values, and beyond the prefit $\pm 1\sigma$ range, these results do not represent a strong bias in the fit. As the flux parameters are so strongly correlated, a pull in one translates to many of them moving in similar ways. The flux penalty contribution to the log-likelihood at each step in the Markov Chain is shown in Figure 6.14. The stationary distribution is at -LLH$\approx$50, which for 100 flux parameters corresponds to $\sim$1 unit of χ^2 per degree of freedom.

As seen in the Asimov fits, there is little constraint beyond the prior uncertainties for the ν_e, $\bar{\nu}_e$ and FHC $\bar{\nu}_\mu$ flux parameters.

The interaction parameters are shown in Figure 6.15. The M_A^{QE} parameter is pulled above its prior central value to much be closer to the nominal generated value (corresponding to 1.2 GeV^2). The 2p2h normalisations are all consistent with the nominal value within the postfit uncertainty. 2p2h shape ^{12}C is the only 2p2h parameter pulled significantly away from nominal, to $\sim$1.7, favouring the Martini model. This is not consistent with the 2p2h shape ^{16}O parameter, which is much closer to nominal.

The Q^2 normalisations all sit slightly above their prior central values, favouring a smaller suppression, and the shape of the increase in parameter value with increasing Q^2 is similar to the priors. The $0.25 < Q^2 < 0.50$ GeV^2 parameter is the only Q^2 normalisation pulled significantly away from the prior.

The 1D distributions for the $E_{\rm b}$ parameters are shown in Figure 6.16. Although the distributions are non-Gaussian, making it difficult to extract a single central value, the arithmetic means are all within 1σ of the prior.

The M_A^{RES} parameter is pulled down -2σ, of its prior uncertainty, while the other single

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$
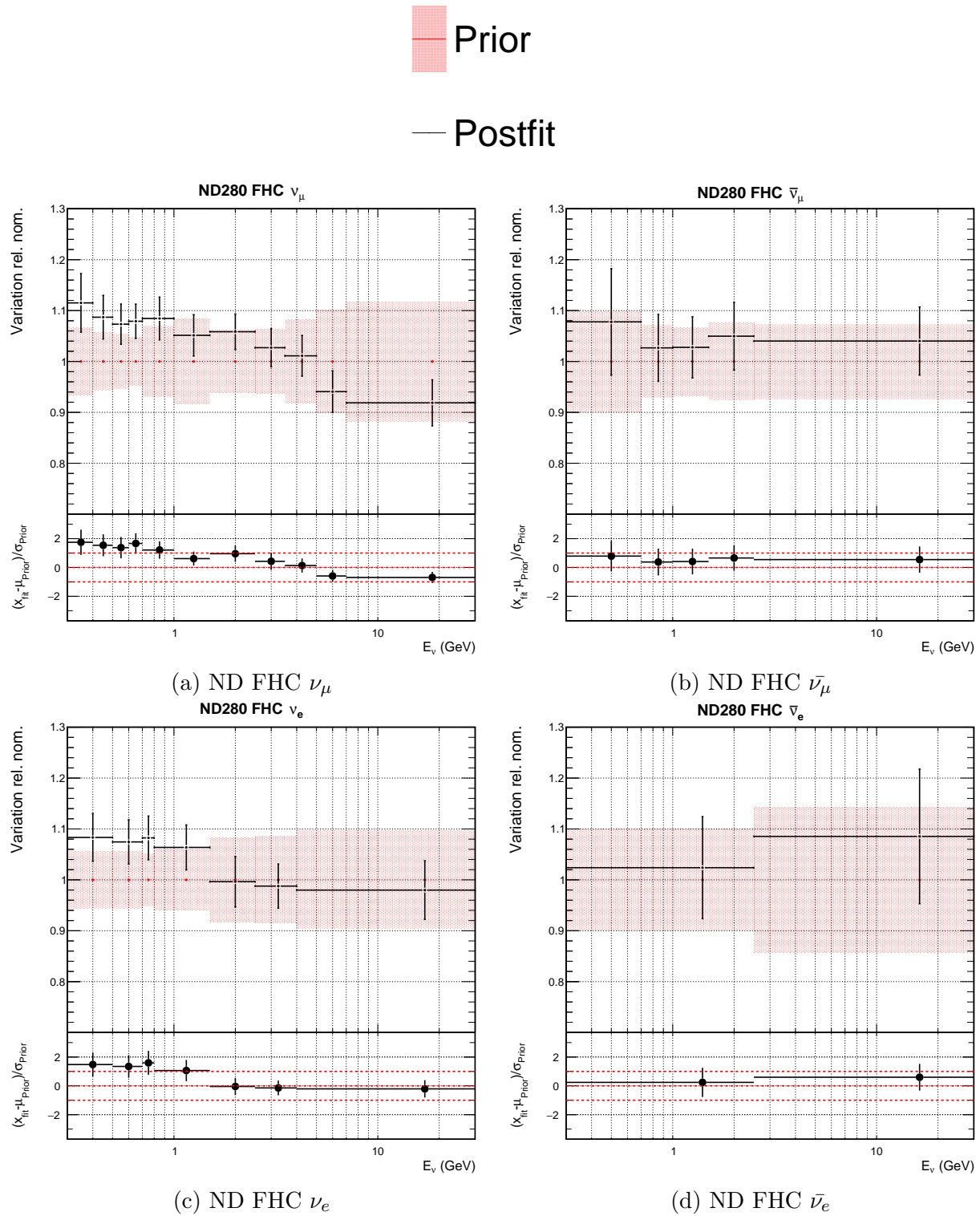
Figure 6.13.: ND280 FHC flux parameters for the data fit.

π production parameters are all consistent with their nominal values. This could suggest M_A^{RES} is soaking up deficiencies in the single π production model.

There is tension between the BY corrections, with the BY DIS parameter being pulled to the edge of its 100% prior uncertainty, and the BY mult-π parameter staying at its nominal value. The CC misc. parameter is also pushed high, but has a large prior uncertainty. The
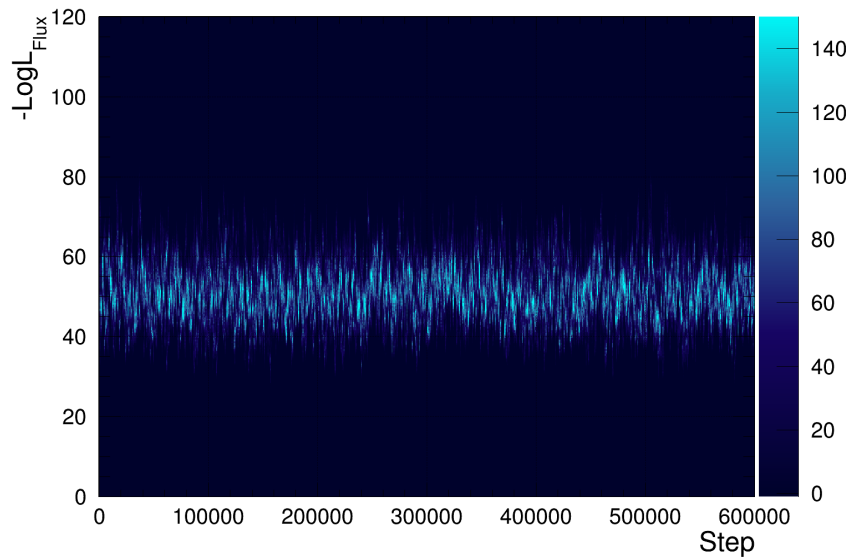
Figure 6.14.: Flux penalty contribution to the log-likelihood at each step in the data fit.

other parameters targeting events in the CC Other samples are very consistent with their prior central values. The CC coherent parameters are pulled down by $\sim 1\sigma$.

The CC coherent, NC, and π FSI parameters are all within 1σ of their nominal values, apart from NC Other ND280. This covers a number of different interaction types, and so a more sophisticated treatment may be needed for future analyses.

The post-data-fit correlation matrix for the interaction parameters is shown in Figure 6.17, and both the flux and interaction parameters together are shown in Appendix G. The overall trends are similar to what was seen for the Asimov fit in Figure 6.11. The fluxes are strongly internally correlated, and anti-correlated with interaction normalisations.

M_A^{QE} correlates with the lowest Q^2 normalisation, which decreases as Q^2 increases, becoming a strong anti-correlation for the higher Q^2 parameters. This is expected as M_A^{QE} affects higher Q^2 events, with the low Q^2 anti-correlation likely due to the mutual correlation with the flux parameters. M_A^{QE} now also correlates highly with the E_{b} parameters more strongly than in the Asimov fit.

The strength of the E_{b} correlations and anti-correlations has increased since the Asimov fit. E_{b} is correlated with the low energy, and anti-correlated with the high energy flux parameters. This is because as the E_{b} parameters increase, the number of low lepton momentum events increases as events shifts to lower momentum. This can be compensated by low energy flux parameters decreasing, as lower energy neutrino events are likely to produce lower momentum leptons, and so the anti-correlations arise. The opposite is true for higher

(a) CC 0π

(b) Q^2 and E_{b}

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC coh.

(d) NC

(e) π FSI
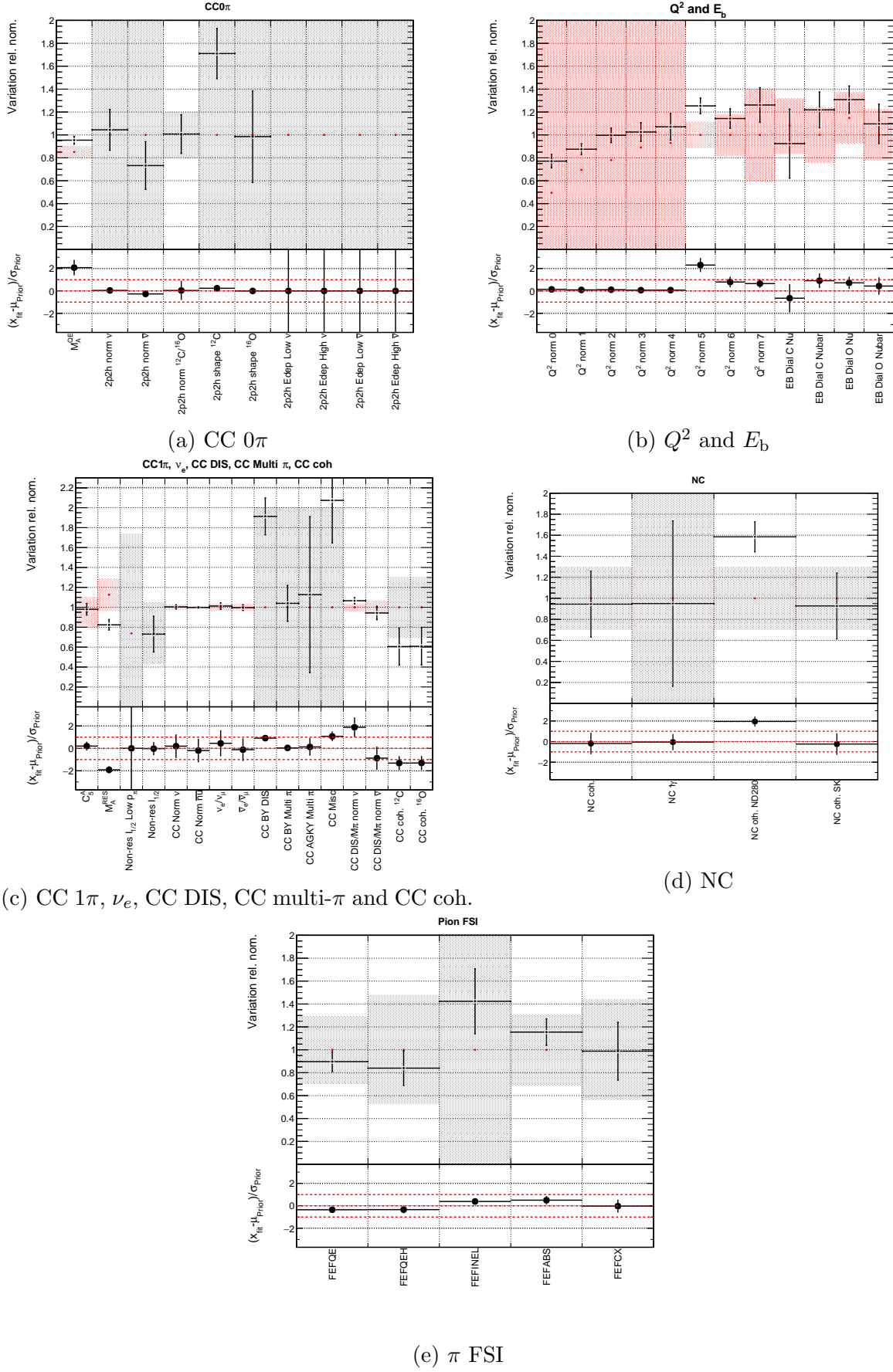
Figure 6.15.: Interaction parameters for the data fit.

(a) $E_\mathrm{b}\nu$ C

(b) $E_\mathrm{b}\bar{\nu}$ C

(c) $E_\mathrm{b}\nu$ O

(d) $E_\mathrm{b}\bar{\nu}$ O
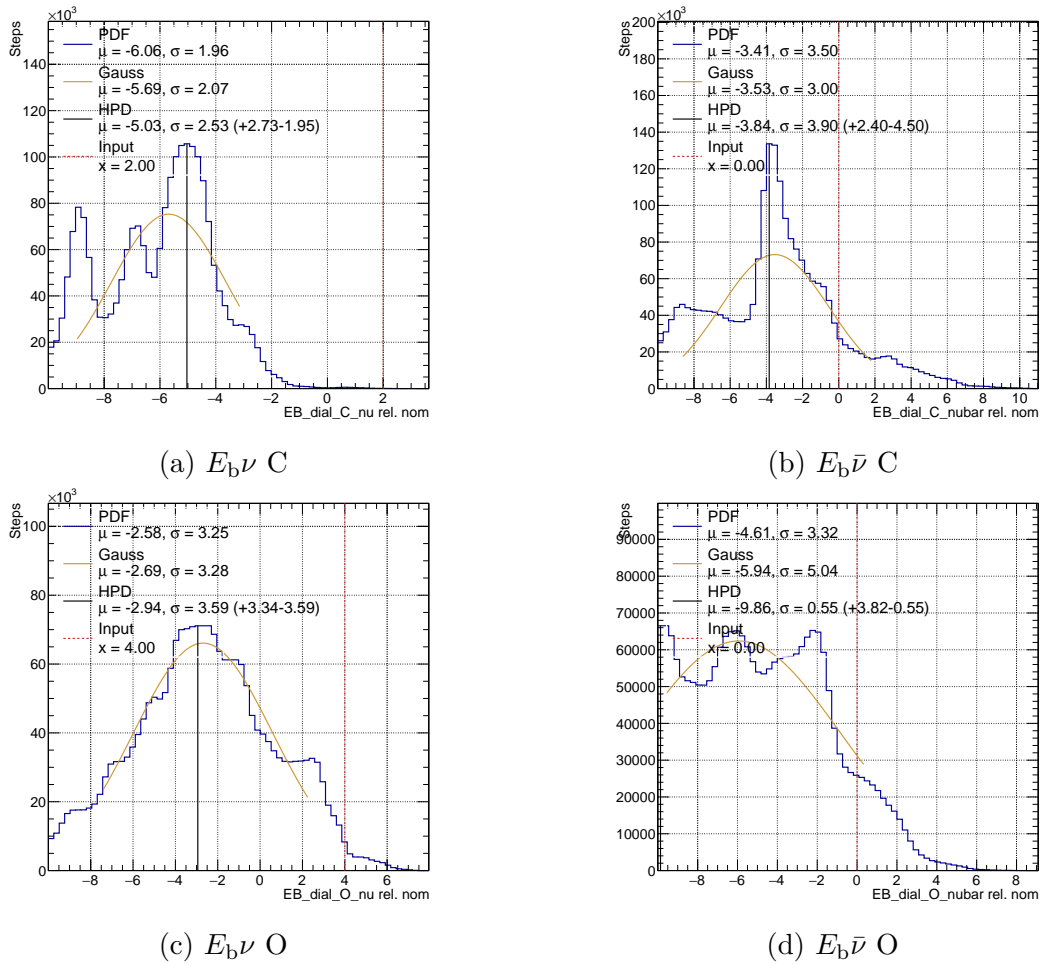
Figure 6.16.: Posterior distributions for the binding energy parameters from the data fit.

energies, causing positive correlations.

There are strong correlations between the 2p2h and CC 1π parameters. This is likely due to final state interactions in which a π is absorbed, causing CC 1π events to be detected as CC 0π. The 2p2h shape parameters are more anti-correlated than for the Asimov fit. 2p2h shape ^{12}C is still correlated with the Q^2 normalisations, with increasing strength at lower Q^2, but this has decreased for 2p2h shape ^{16}O. Both 2p2h shape parameters are less anti-correlated with the fluxes than they were in the Asimov fit.

The M_A^{RES} and C_A^5 parameters have a stronger anti-correlation with each other than in the data fit, and both are now even more strongly correlated and anti-correlated with the CC coh. parameters.

The Q^2 normalisations, DIS and multi-π, and π FSI parameters all have internal correlations and anti-correlations, with a slight increase in strength compared to the Asimov fit.
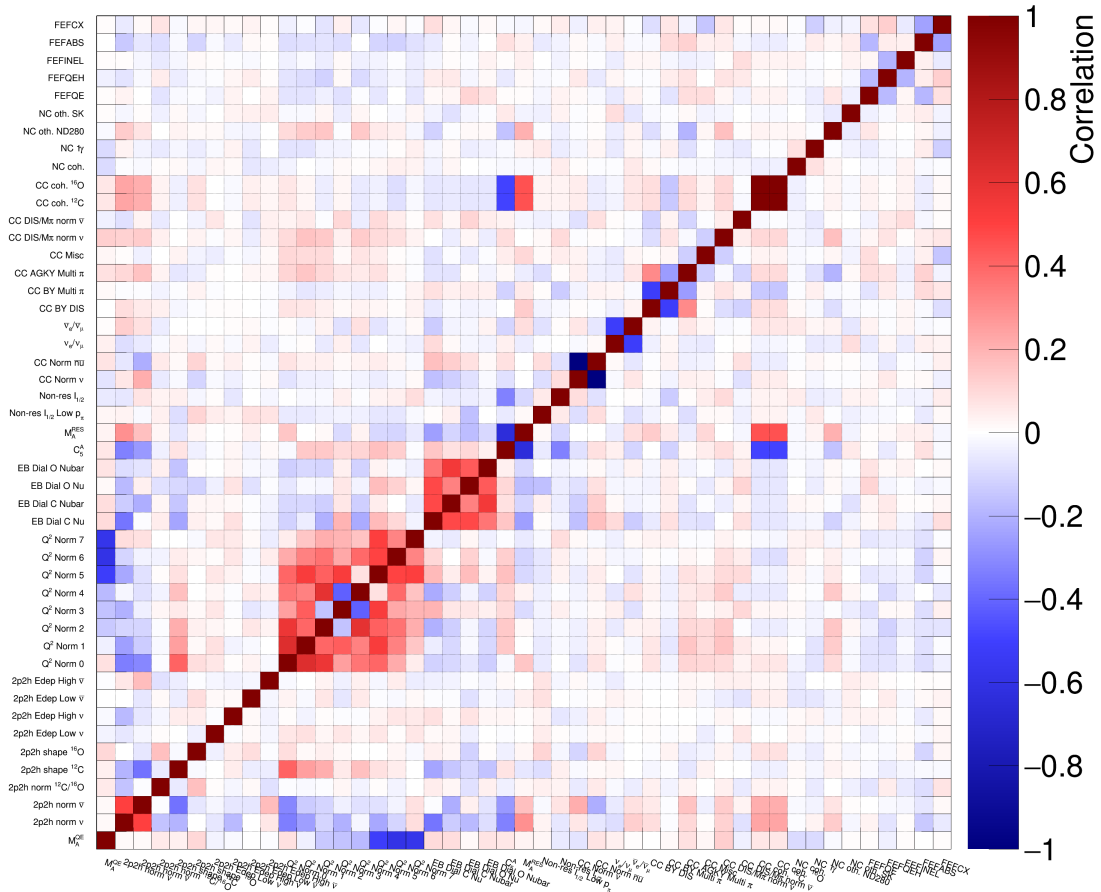
Figure 6.17.: Data postfit correlation matrix for interaction parameters.

6.6. Posterior Predictions

Posterior predictions are produced using the method described in Section 4.4.3. The p_μ projections of the posterior predictions for each sample are shown in Figures 6.18–6.20, along with the prior predictions and data. The $\cos\theta_\mu$ projections are shown in Appendix H. There is significant improvement in the agreement with data for the posterior predictions compared to the prior predictions. Particularly in the momentum peak around $p_\mu \sim 600$ MeV, the prediction is closer to data. There are still regions of underestimation in the CC 0π and CC Other samples, and overestimation in CC 1π samples, but these are less strong than for the prior prediction. The error band is also reduced significantly for the posterior compared to the prior for all samples, showing how the constraint on systematic uncertainties from the fit has reduced the uncertainty in the prediction.

This is confirmed by the reduction in -2LLH for the posterior prediction compared to for the prior prediction, shown in Table 6.3. The uncertainties on the total event rate for all

147

samples is also reduced significantly. The fractional errors in Table 6.4 show the overall ND280 event rate uncertainty has been reduced from 9.32% to 0.29% by the fit.

Sample	Data	Prior			Posterior			-2LLH$_{Sample}$ Prior	-2LLH$_{Sample}$ Posterior
FGD1 FHC ν CC 0π	33443	28912.3	±	3049.9	33383.9	±	161.6	1699.87	430.09
FGD1 FHC ν CC 1π	7713	8691.5	±	1013.2	7914.6	±	67.3	436.38	318.79
FGD1 FHC ν CC Other	8026	7343.3	±	1004.0	7933.5	±	71.4	519.01	292.16
FGD2 FHC ν CC 0π	33156	28461.0	±	2998.9	33151.9	±	166.2	1801.15	463.30
FGD2 FHC ν CC 1π	6281	6965.6	±	791.5	6418.2	±	57.0	411.05	312.24
FGD2 FHC ν CC Other	7700	6740.4	±	893.0	7301.6	±	69.0	541.39	376.69
FGD1 RHC $\bar{\nu}$ CC 0π	8388	7665.0	±	872.3	8443.4	±	70.7	506.12	229.49
FGD1 RHC $\bar{\nu}$ CC 1π	698	736.0	±	94.2	679.0	±	14.5	64.84	46.53
FGD1 RHC $\bar{\nu}$ CC Other	1472	1360.3	±	179.1	1468.9	±	23.6	116.38	94.88
FGD2 RHC $\bar{\nu}$ CC 0π	8334	7393.5	±	816.7	8204.3	±	68.2	522.58	206.31
FGD2 RHC $\bar{\nu}$ CC 1π	650	660.2	±	84.1	638.4	±	12.2	54.42	58.74
FGD2 RHC $\bar{\nu}$ CC Other	1335	1251.9	±	164.9	1378.0	±	19.8	120.95	84.06
FGD1 RHC ν CC 0π	3594	3175.4	±	333.8	3575.8	±	39.1	193.31	135.04
FGD1 RHC ν CC 1π	1111	1216.9	±	144.7	1151.8	±	14.6	65.29	54.46
FGD1 RHC ν CC Other	1344	1131.0	±	153.9	1291.7	±	17.3	95.29	87.45
FGD2 RHC ν CC 0π	3433	3151.3	±	329.4	3522.0	±	37.7	152.49	153.84
FGD2 RHC ν CC 1π	926	977.1	±	116.8	916.7	±	11.3	57.27	61.94
FGD2 RHC ν CC Other	1245	1058.5	±	147.0	1190.8	±	15.0	78.36	69.43
Total	128849	117237.9	±	10925.7	128562.2	±	378.2	7436.15	3475.44

Table 6.3.: Prior and posterior predictive event rates and log-likelihood to data.

Sample	$\delta N/N(\%)$ Prior	$\delta N/N(\%)$ Posterior
FGD1 FHC ν CC 0π	10.55	0.48
FGD1 FHC ν CC 1π	10.45	0.85
FGD1 FHC ν CC Other	13.67	0.90
FGD2 FHC ν CC 0π	10.54	0.50
FGD2 FHC ν CC 1π	11.36	0.88
FGD2 FHC ν CC Other	13.25	0.94
FGD1 RHC $\bar{\nu}$ CC 0π	11.38	0.84
FGD1 RHC $\bar{\nu}$ CC 1π	12.80	2.14
FGD1 RHC $\bar{\nu}$ CC Other	13.17	1.61
FGD2 RHC $\bar{\nu}$ CC 0π	11.05	0.83
FGD2 RHC $\bar{\nu}$ CC 1π	12.74	1.91
FGD2 RHC $\bar{\nu}$ CC Other	13.17	1.44
FGD1 RHC ν CC 0π	10.51	1.09
FGD1 RHC ν CC 1π	11.89	1.27
FGD1 RHC ν CC Other	13.61	1.34
FGD2 RHC ν CC 0π	10.45	1.07
FGD2 RHC ν CC 1π	11.95	1.23
FGD2 RHC ν CC Other	13.89	1.26
Total	9.32	0.29

Table 6.4.: Fractional uncertainties on the prior and posterior predictive event rates.
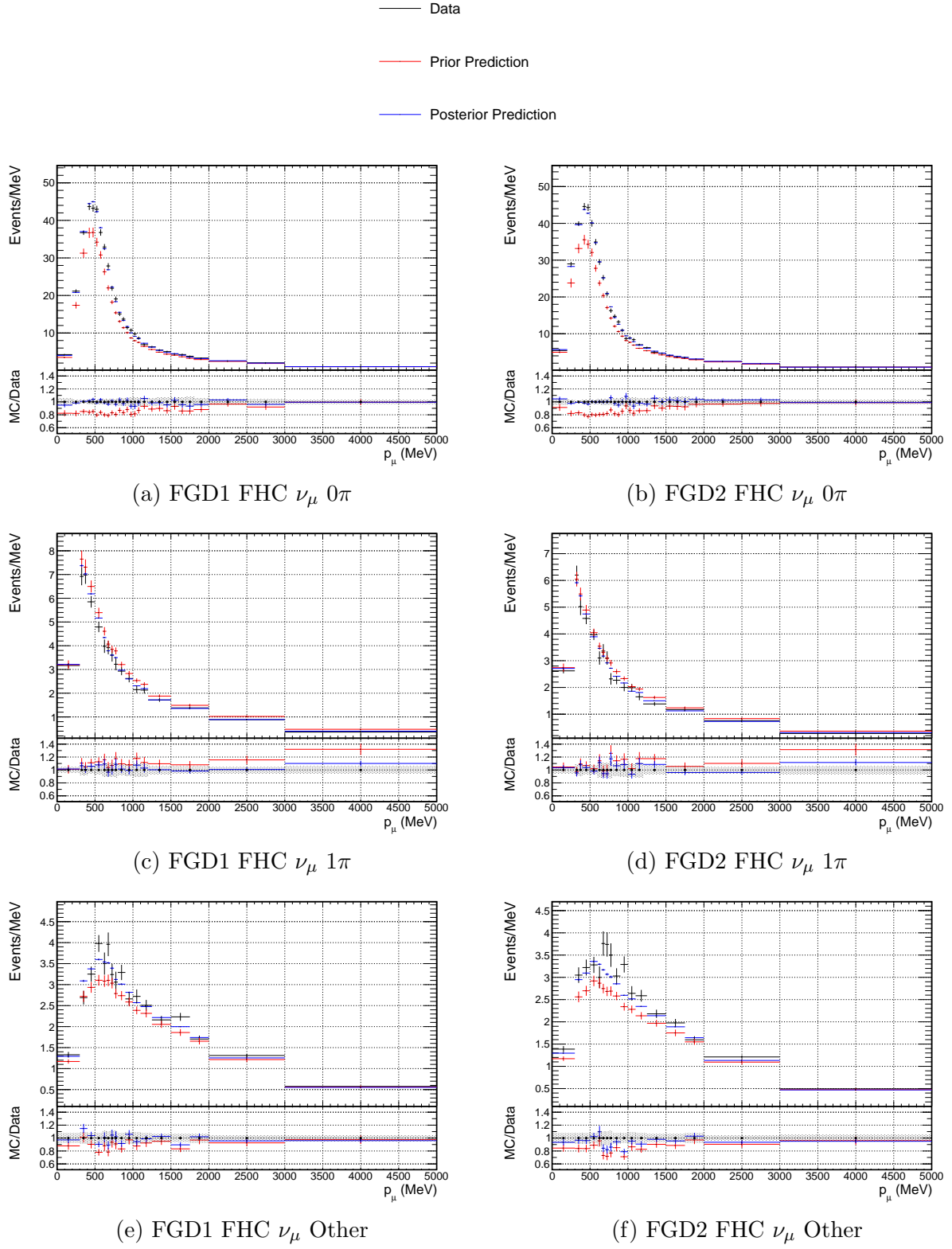
Figure 6.18.: p_μ projections of the prior and posterior predictive distributions and data for FHC ν_μ selections.

(a) FGD1 RHC $\bar{\nu}_\mu$ 0π

(b) FGD2 RHC $\bar{\nu}_\mu$ 0π

(c) FGD1 RHC $\bar{\nu}_\mu$ 1π

(d) FGD2 RHC $\bar{\nu}_\mu$ 1π

(e) FGD1 RHC $\bar{\nu}_\mu$ Other

(f) FGD2 RHC $\bar{\nu}_\mu$ Other

Figure 6.19.: p_μ projections of the prior and posterior predictive distributions and data for RHC $\bar{\nu}_\mu$ selections.

(a) FGD1 RHC ν_μ 0π

(b) FGD2 RHC ν_μ 0π

(c) FGD1 RHC ν_μ 1π

(d) FGD2 RHC ν_μ 1π

(e) FGD1 RHC ν_μ Other

(f) FGD2 RHC ν_μ Other
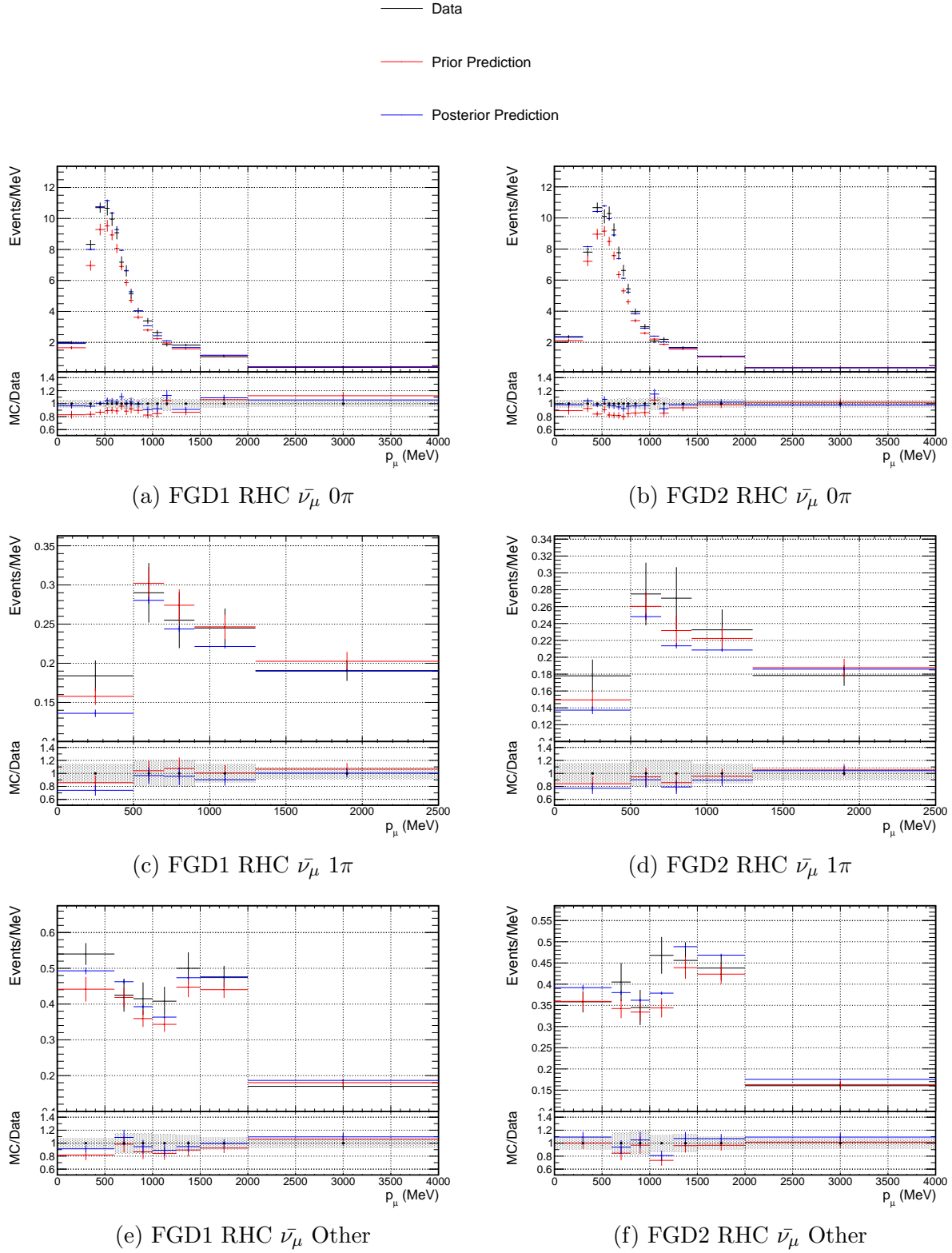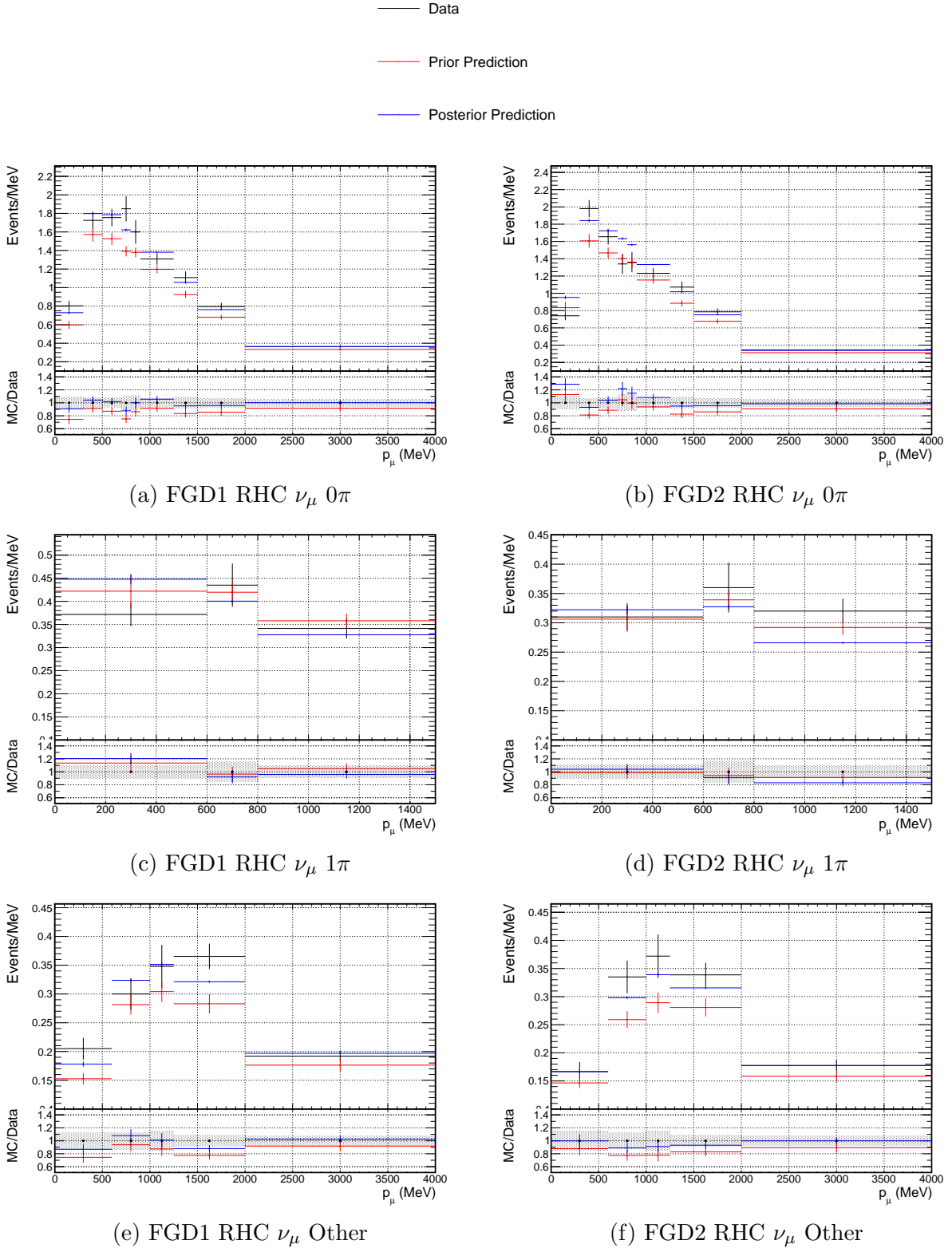
Figure 6.20.: p_μ projections of the prior and posterior predictive distributions and data for RHC ν_μ selections.

6.6.1. Posterior Predictive p-values

Posterior predictive p-values were calculated using the methods described in Section 4.4.4. As previously discussed, this Bayesian p-value measures how likely it would be for data described by the postfit model to be observed if the experiment was repeated with the same statistics. It is therefore a more stringent test than the traditional frequentist p-value.

Discouragingly, the total p-value is 0.00 for both Bayesian methods, as shown in Figure 6.21. The $y = x$ line is shown in red, below which steps contribute to the p-value. This is likely caused by throwing the detector systematics using the merged bins, despite the underlying systematics having non-Gaussian shape, as shown in Section 5.4.3. Looking at the contributions from each sample, the individual p-values are low for many samples, as shown in Table 6.5.

All but one of the CC 0π samples have a p-value $< 5\%$, and for FGD2 FHC ν and FGD1 RHC $\bar{\nu}$ the p-values $= 0.0\%$. The CC 1π samples are generally higher, but are $< 5\%$ for both FHC ν, and the FGD2 RHC $\bar{\nu}$ samples. For CC Other, only the FGD1 FHC ν, FGD2 RHC $\bar{\nu}$ and FGD2 RHC ν samples have p-values $> 0.05\%$. The p-values are not consistent across FGD1 and 2, suggesting differences in how well the two are modelled. The CC Other samples had the lowest p-value of the FHC selections in the previous oscillation analysis [114], and this is now also the case for the RHC CC Other selections. The overall p-value was also 0.00 in the previous analysis.

As the total p-values are constructed from the sum of all the sample likelihoods, a high likelihood for one sample can dominate the overall p-value. A low p-value for a single sample can therefore drive the total p-value to be 0.0%. The total p-value should therefore not be interpreted as an average across all samples, and so the final value of 0.0% does not mean that the model is entirely unsuitable to fit to all the data, just that there is at least one region of selections or kinematic phase space which is not well described by the posterior.

The two methods of calculating the p-value are consistent. Although they are not all identical, they follow the same trends and are within 5% of each other for every sample. This suggests the 2000 drawn steps describe the posterior distribution sufficiently.

To investigate the cause of the low p-values, the LLH contributions from each bin were inspected for each sample, as shown in Figure 6.22. As would be expected there are more regions with higher contribution in the samples with low p-values. For example, comparing the FGD1 and FGD2 FHC ν CC Other samples, there is a very large contributions from an individual bin in FGD2, which has a significantly lower p-value, which doesn't appear for
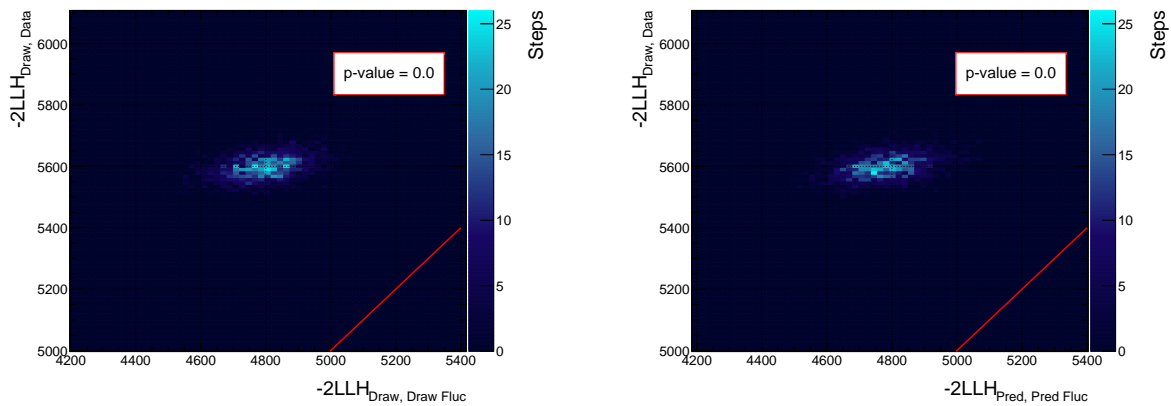
Figure 6.21.: Posterior predictive p-values from the data fit. The fraction of steps below the line $y = x$, shown in red, is the p-value

FGD1.

However, it is difficult to identify any definitive trends which could explain all the low goodness of fits. Comparing the CC 0π samples, the FGD1 and FGD2 RHC ν samples have similarly uniform distributions, despite the p-value being much higher for FGD1. The FHC ν and RHC $\bar{\nu}$ samples do all several high contributing bins though, in line with their lower p-values.

For CC 1π, the distributions for the low p-value FHC ν samples both have large contributions from individual bins, but in different regions of p_μ–$\cos\theta_\mu$ space. The RHC $\bar{\nu}$ samples have similarly sporadic high contributing bins, despite the p-value for FGD1 being much higher than for FGD2. The RHC ν log-likelihood distributions are more uniform, but there are a few higher contributing bins in FGD2, despite the high p-value.

Comparing the CC Other samples, there is a single bin with a large contribution for FGD2 RHC $\bar{\nu}$ which does not appear for FGD1, despite FGD2 having the higher p-value. For RHC ν, the p-value is much higher for FGD2, but the distributions are similarly uniform.

Overall, the high log-likelihoods aren't coming from a consistent region in p_μ–$\cos\theta_\mu$. Although there are individual bins with higher contributions, when considering so many fit bins this is to be expected, and they are not all grouped together. The low p-values are therefore not just driven by a single set of kinematic variables being badly modelled or reconstructed.

As discussed in Section 4.4.4, there is no definitive gauge for what does or does not constitute an acceptable value for the goodness of fit. Furthermore, as the Bayesian p-value measures how likely it would be that a repeat of the experiment would observe data consistent with the postfit model, and as such is a more stringent test than a traditional

| | *p*-value | |
Sample	Fluctuation of Draw	Fluctuation of Prediction
FGD1 FHC ν CC 0π	0.005	0.004
FGD1 FHC ν CC 1π	0.042	0.042
FGD1 FHC ν CC Other	0.334	0.288
FGD2 FHC ν CC 0π	0.000	0.000
FGD2 FHC ν CC 1π	0.008	0.010
FGD2 FHC ν CC Other	0.000	0.001
FGD1 RHC $\bar{\nu}$ CC 0π	0.000	0.001
FGD1 RHC $\bar{\nu}$ CC 1π	0.268	0.236
FGD1 RHC $\bar{\nu}$ CC Other	0.001	0.004
FGD2 RHC $\bar{\nu}$ CC 0π	0.010	0.005
FGD2 RHC $\bar{\nu}$ CC 1π	0.010	0.003
FGD2 RHC $\bar{\nu}$ CC Other	0.050	0.051
FGD1 RHC ν CC 0π	0.246	0.213
FGD1 RHC ν CC 1π	0.500	0.516
FGD1 RHC ν CC Other	0.008	0.009
FGD2 RHC ν CC 0π	0.037	0.035
FGD2 RHC ν CC 1π	0.095	0.043
FGD2 RHC ν CC Other	0.097	0.077
Total	0.000	0.000

Table 6.5.: Posterior predictive *p*-values for each sample.

p-value, combined with the use of merged detector bins, the low *p*-values presented here do not necessarily mean the fit results are invalid.

The frequentist *p*-value, calculated in the BANFF framework, was significantly higher, at 0.74 [138]. This result does suggest that the data is consistent with the prior model, and has increased since the last oscillation analysis [114]. As previously discussed, this frequentist *p*-value is the more traditional interpretation of the goodness of fit, indicating the compatibility of the model to fit the data, and the construction uses a more rigorous treatment of the detector systematics. Therefore, as the frequentist *p*-value was high, and the low Bayesian *p*-value was driven by the CC Other and RHC CC 0π samples rather than being consistently 0.00, it can be concluded that an acceptable level of goodness of fit has been achieved.

6.6.2. Propagating to SK

To see the effect of the near detector fit on the full oscillation analysis, posterior predictive distributions at SK were produced. The same process is used as for the near detector pos-

(a) FGD1 FHC ν_μ 0π

(b) FGD1 FHC ν_μ 1π

(c) FGD1 FHC ν_μ Other

(d) FGD2 FHC ν_μ 0π

(e) FGD2 FHC ν_μ 1π

(f) FGD2 FHC ν_μ Other

(g) FGD1 RHC $\bar{\nu}_\mu$ 0π

(h) FGD1 RHC $\bar{\nu}_\mu$ 1π

(i) FGD1 RHC $\bar{\nu}_\mu$ Other

(j) FGD2 RHC $\bar{\nu}_\mu$ 0π

(k) FGD2 RHC $\bar{\nu}_\mu$ 1π

(l) FGD2 RHC $\bar{\nu}_\mu$ Other

(m) FGD1 RHC ν_μ 0π

(n) FGD1 RHC ν_μ 1π

(o) FGD1 RHC ν_μ Other

(p) FGD2 RHC ν_μ 0π

(q) FGD2 RHC ν_μ 1π

(r) FGD2 RHC ν_μ Other
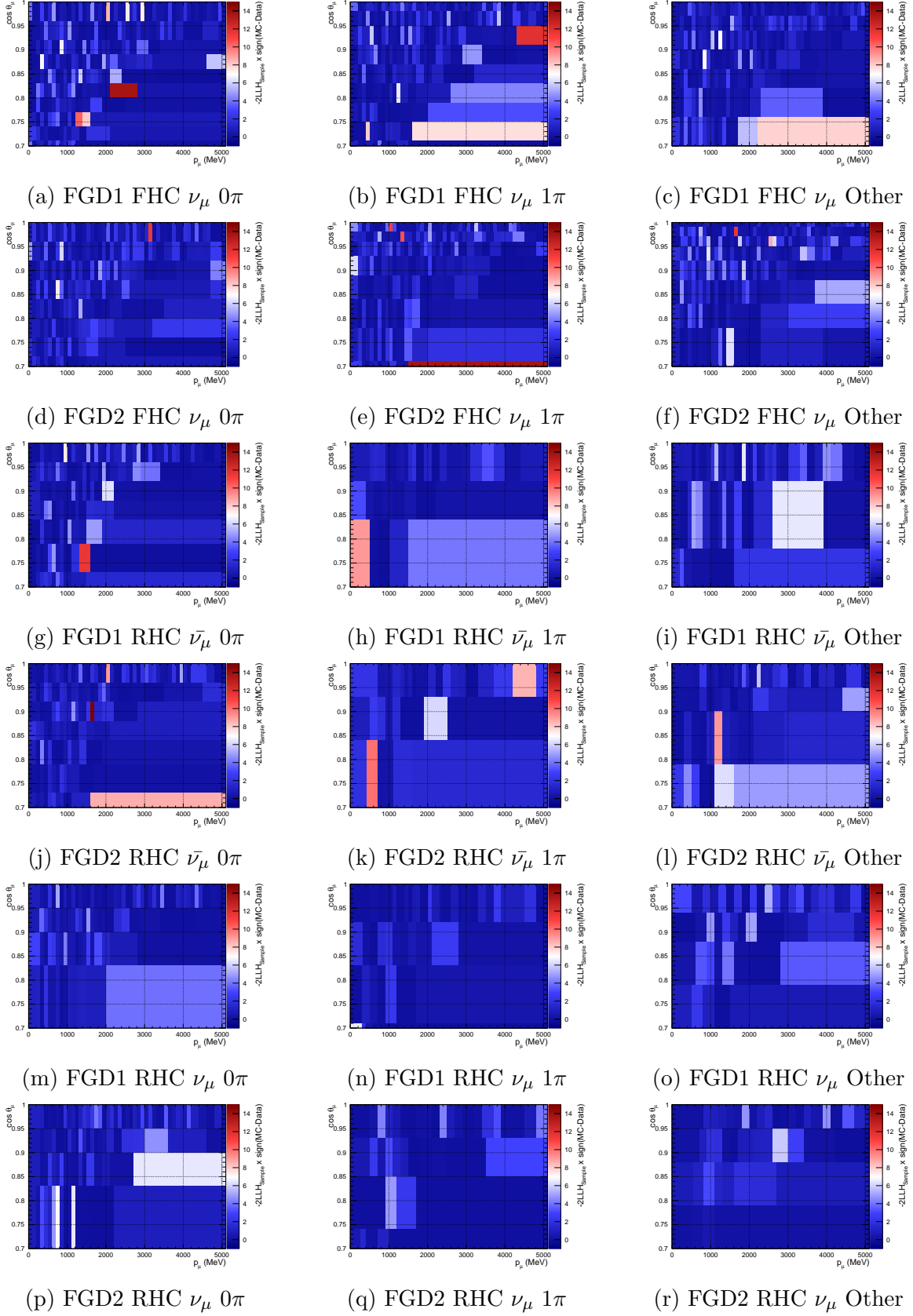
Figure 6.22.: Contributions to the sample log-likelihood from each fit bin.

terior predictions, using 2000 draws from the near detector Markov Chain, but reweighting the nominal MC for SK rather than ND280. Only the 50 SK flux parameters, and 36 interaction parameters which apply to SK events (not parameters for ^{12}C or ND280 only) are propagated. Nominal values for the SK detector uncertainties, described in [143], and the oscillation parameters in Table 6.6, are used to produce these predictions.

The prior predictive distributions were also produced, using the same method as for ND280, but throwing the SK flux, SK detector, and interaction uncertainties, before reweighting the SK MC.

The SK MC used corresponds to the same run periods as for ND280, but for a higher POT as the detector was online for a larger proportion of these runs. The total data POT is 1.966×10^{21} for FHC and 1.635×10^{21} for RHC.

The total prior and posterior event rates for each of the SK samples are shown in Table 6.7. The total number of events has been increased beyond the prior uncertainty for all samples except 1Re 1d.e. This is expected given the nominal ND280 MC underestimated the data. The uncertainties for all samples have been reduced significantly, as shown in Table 6.8, and the overall event rate uncertainty has been reduced from 11.84% to 2.31% by the near detector fit.

Parameter	Value
$\sin^2\theta_{12}$	0.307
$\sin^2\theta_{23}$	0.528
$\sin^2\theta_{13}$	0.0218
Δm^2_{12}	7.53×10^{-5} eV^2
Δm^2_{23}	2.509×10^{-5} eV^2
δ_{CP}	-1.601

Table 6.6.: Oscillation parameter values used to produce the SK posterior predictions.

	Event Rates			
Sample	**Prior**		**Posterior**	
1R$_\mu$	286.62 ± 38.24		347.67 ± 8.05	
RHC 1R$_\mu$	120.82 ± 14.61		134.81 ± 3.59	
1R$_e$	73.52 ± 9.96		95.49 ± 3.69	
RHC 1R$_e$	14.13 ± 1.76		16.17 ± 0.65	
1R$_e$ 1d.e.	10.48 ± 2.34		8.89 ± 0.47	
Total	505.56 ± 59.88		603.03 ± 13.96	

Table 6.7.: Prior and posterior predictive SK event rates.

Sample	$\delta N/N(\%)$	
	Prior	Posterior
1R$_\mu$	13.33	2.32
RHC 1R$_\mu$	12.09	2.66
1R$_e$	13.55	3.86
RHC 1R$_e$	12.46	4.02
1R$_e$ 1d.e.	22.33	5.29
Total	11.84	2.31

Table 6.8.: Fractional uncertainties on the prior and posterior predictive SK event rates.

The prior and posterior distributions are shown in Figure 6.23. There is significant enhancement in the the 1R$_\mu$ samples at all energies for both FHC and RHC. For RHC, the posterior prediction in the oscillation dip at $E_{\text{rec}} \sim 0.6$ GeV agrees with the prior. The decrease in depth of the dip in FHC will directly impact the measurement of Δm_{32}^2, but will not affect $\sin^2\theta_{23}$ as the location of the dip in E_{rec} is unmoved.

The 1R$_e$ samples are also consistent for FHC and RHC, both showing an enhancement of ~ 1.5 at the peak energy, around $E_{\text{rec}} \sim 0.6$ GeV. Below the peak the posterior predictions are within the uncertainty of the priors, whereas above the peak the enhancement is slightly above the uncertainty for both samples. The posterior prediction is consistently within the prior uncertainty for the 1Re 1d.e. sample, consistent with the nominal MC prediction for the ND280 CC 1π samples being closer to data than for CC 0π and CC Other.

The error band on the prediction has been reduced for all samples across all energies, particularly in the dip in the 1R$_\mu$ samples and peak in 1R$_e$ samples. This shows how the near detector fit has significantly constrained the SK prediction.

6.7. Comparing Different Fit and Detector Binnings

Non-uniform fit binning was used for the fit in Section 6.5 to improve the sensitivity of the analysis. To study the impact this has had, an additional fit was run using the uniform binning defined in Appendix B. As well as this, the detector binning was updated to match the non-uniform fit binning, instead of using merged-uniform rectangular bins. This has the advantage of allowing the detector systematics to more accurately apply to the events they should, but introduces a vast amount of additional fit parameters.

In total, three fits were run, including the main fit, to see the full effect of the binning

(a) 1R$_\mu$

(b) RHC 1R$_\mu$

(c) 1R$_e$

(d) RHC 1R$_e$

(e) 1R$_e$ 1d.e.
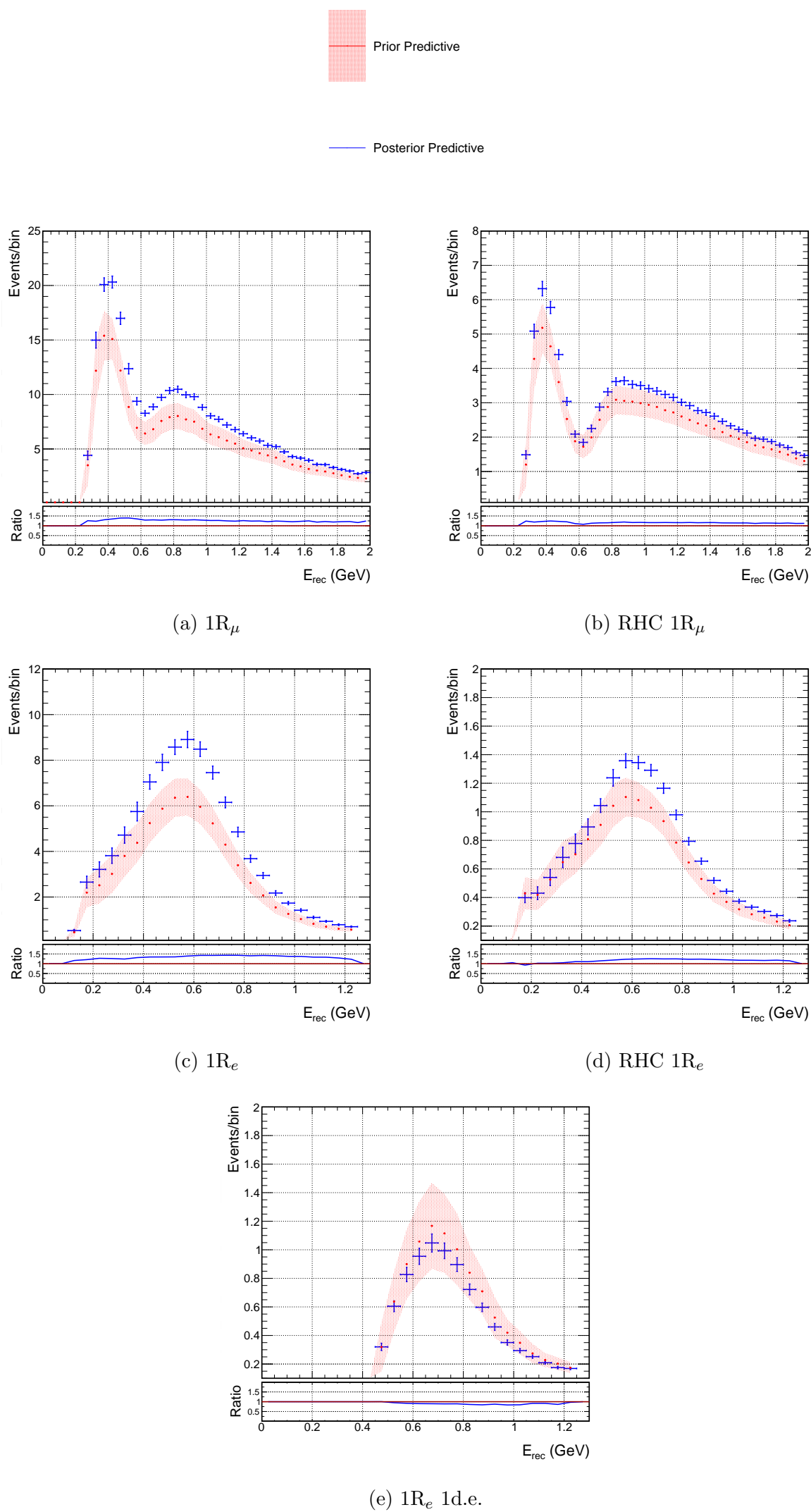
Figure 6.23.: Prior and posterior predictive SK distributions.

changes:

- **Fit Binning:**   Uniform rectangular bins

  Detector Binning: Merged-uniform rectangular bins

- **Fit Binning:**   Non-uniform rectangular bins

  Detector Binning: Merged-uniform rectangular bins

  (This is the fit presented in Section 6.5)

- **Fit Binning:**   Non-uniform rectangular bins

  Detector Binning: Non-uniform rectangular fit bins

All other inputs, and the fitting framework itself, are identical for the three fits.

6.7.1. Nominal MC

The nominal MC p_μ–$\cos\theta_\mu$ distributions, binned uniformly, are shown in Figure 6.24 for the FGD1 and FGD2 FHC CC 0π samples. Comparing to the non-uniform binning in Figure 6.1, there is a less uniform distribution of events, with the bins with the largest number of events being outside the peak. The overall range for each sample is drastically larger compared to the non-uniform binning.

6.7.2. Asimov Fits

The nominal MC was initially fitted to itself, using both the merged-uniform and non-uniform detector binnings. This served as a test of the potential sensitivity, as well as validating the changes to framework to accommodate non-uniform fit and detector binning.

The results for the ND280 FHC flux parameters are shown in Figure 6.25, and the full results are in Appendix F. As expected, the postfit parameter values are very close to the prior central values. Marginalisation effects cause some small deviations from the exact prior, but these are no larger than would be expected from what was seen in previous analyses. The postfit uncertainties are approximately the same in all the fits, indicating similar levels of sensitivity.

The interaction parameters are shown in Figure 6.26. The postfit values are again close to the nominals for all parameters, with small differences between the fits from marginalisation effects. The sizes of the uncertainties are similar for each of the fits.

(a) FGD1 FHC ν_μ 0π

(b) FGD2 FHC ν_μ 0π
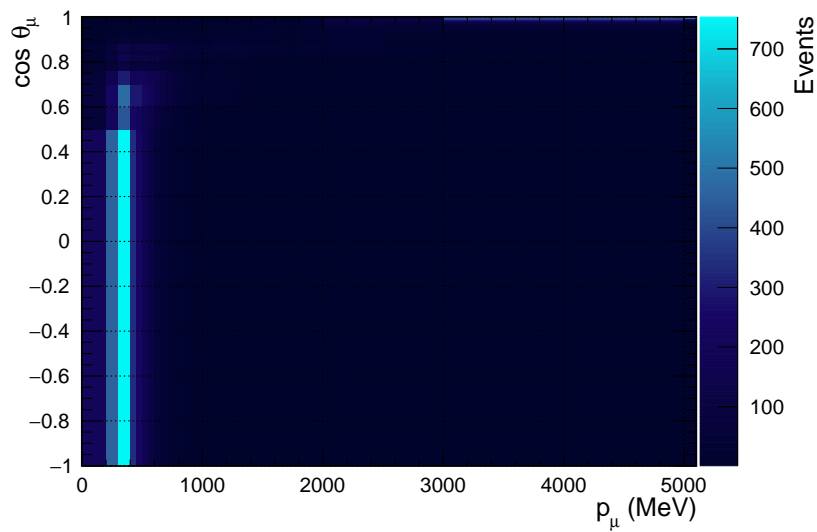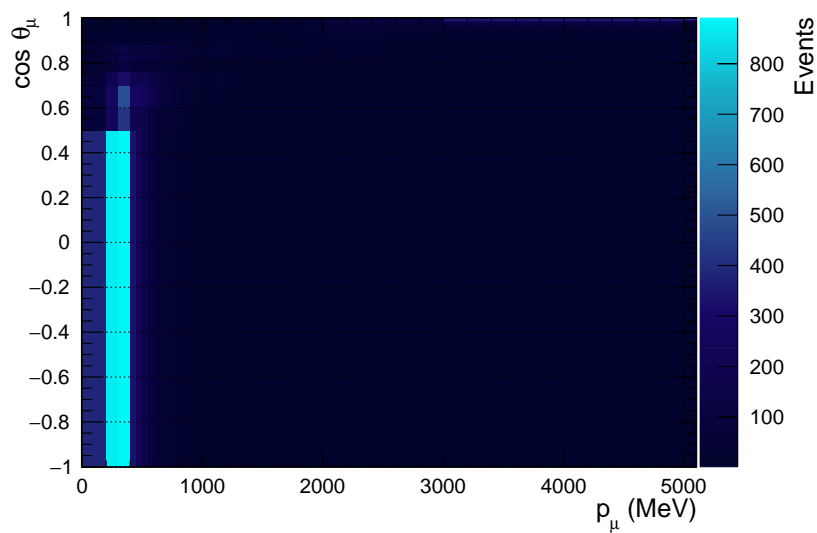
Figure 6.24.: Nominal MC distributions for the FGD1 and FGD2 CC 0π samples with uniform binning. All samples are shown in Appendix D.

6.7.3. Data Fits

The MC prediction was then fitted to the data using each binning, and compared to the previous fit described in Section 6.5. The ND280 FHC flux and interaction parameter results for all three fits are shown in Figures 6.27, and 6.28. The full results are shown in Appendix F.

The flux parameters are mostly compatible between the three fits. Similar high pulls are seen for each, particularly at low energy. The shape of the pulls in energy change slightly, with the two non-uniform fit binning fits lying closer to each other than for the uniform-

(a) ND FHC ν_μ (b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e (d) ND FHC $\bar{\nu}_e$
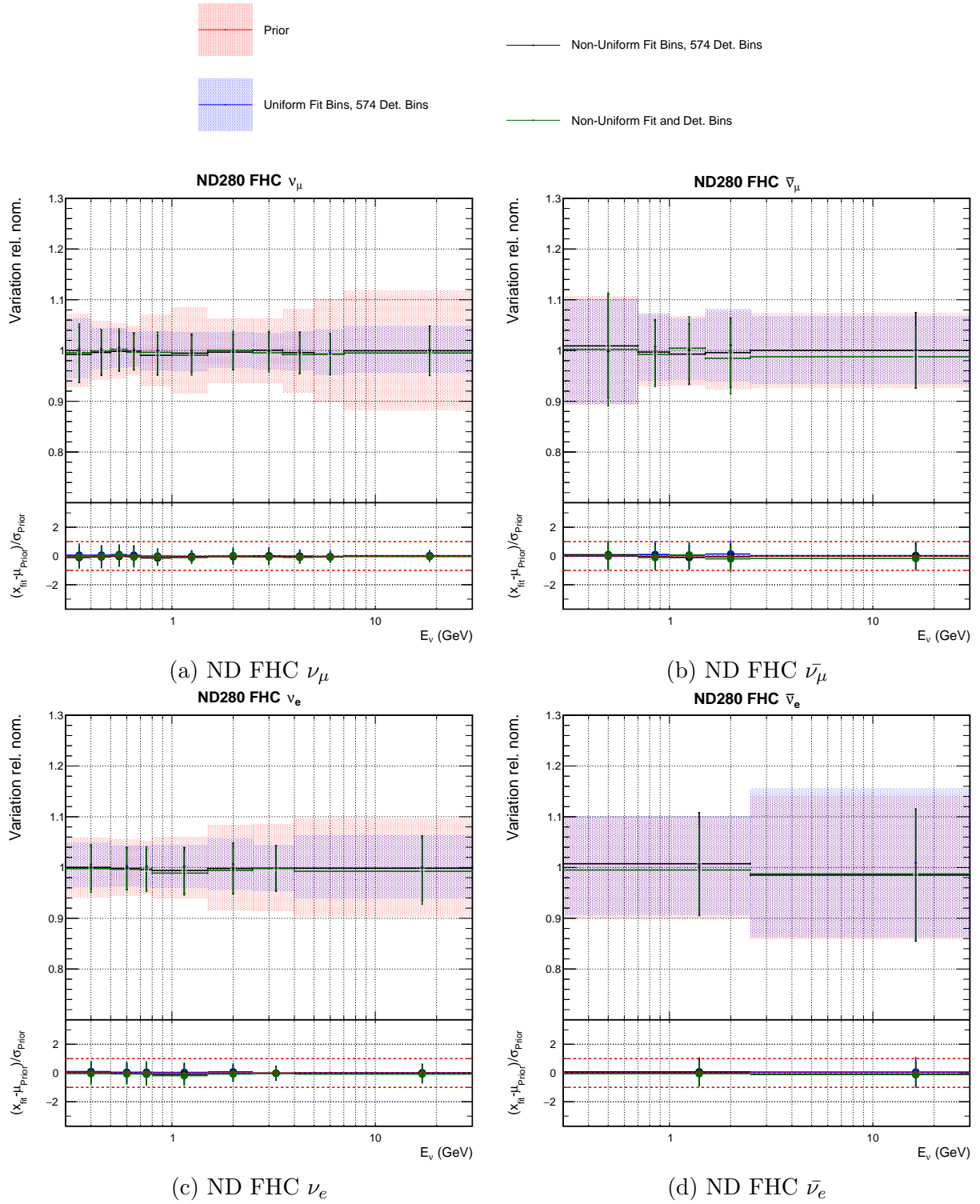
Figure 6.25.: Comparison of ND280 FHC flux parameters for the Asimov fits with different fit and detector binnings.

binning. Only at very high energy for FHC ν_μ are the postfit values outside 1σ of each other, but this region has fairly low statistics and has a very small impact on the SK prediction.

The interaction parameters show more differences. M_A^{QE} is very consistent between the fits, but the 2p2h normalisations are pushed higher for the two non-uniform binning fits.

(a) CC 0π

(b) Q^2 and $E_{\rm b}$

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
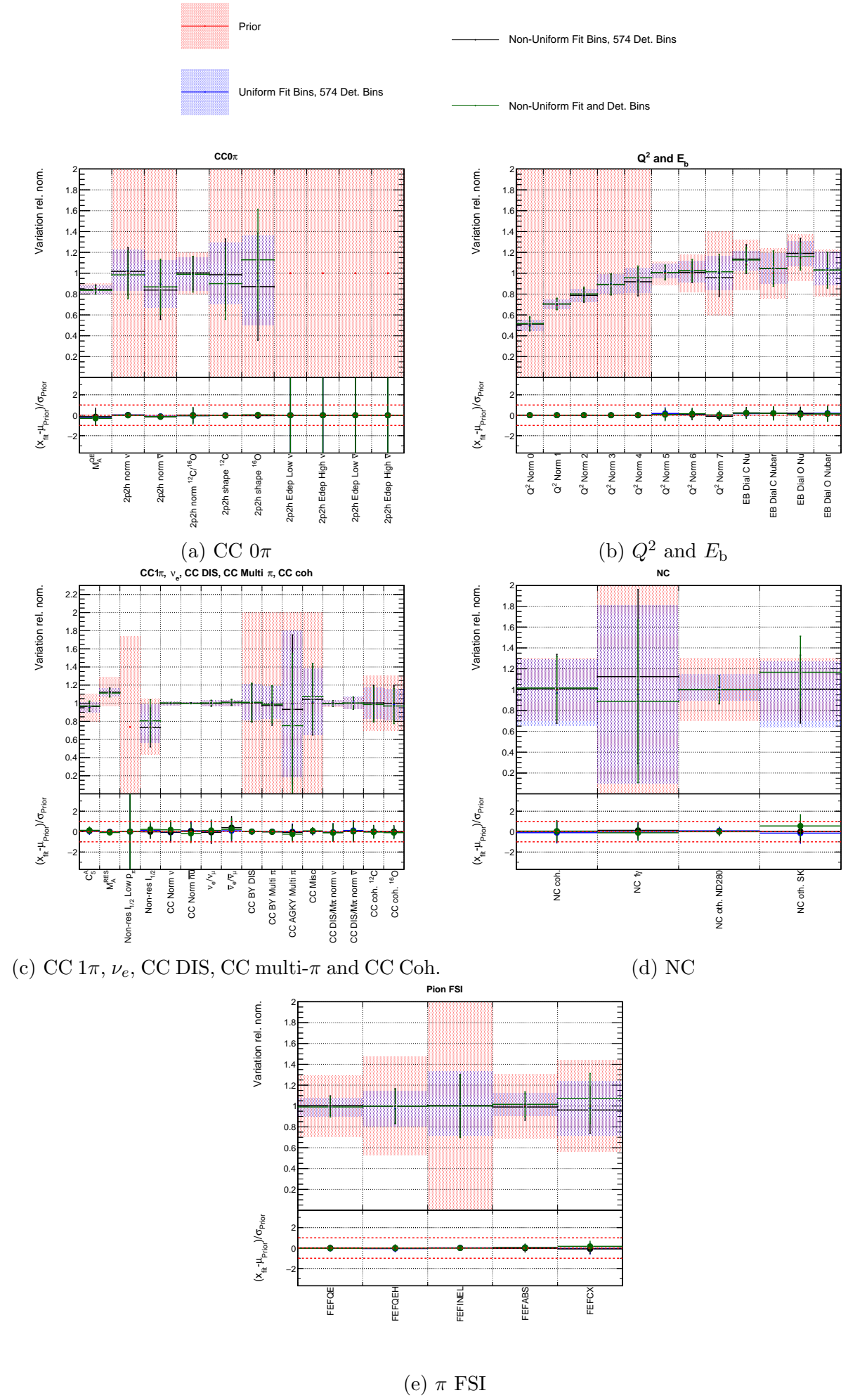
Figure 6.26.: Comparison of interaction parameters for the Asimov fits with different fit and detector binnings.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$
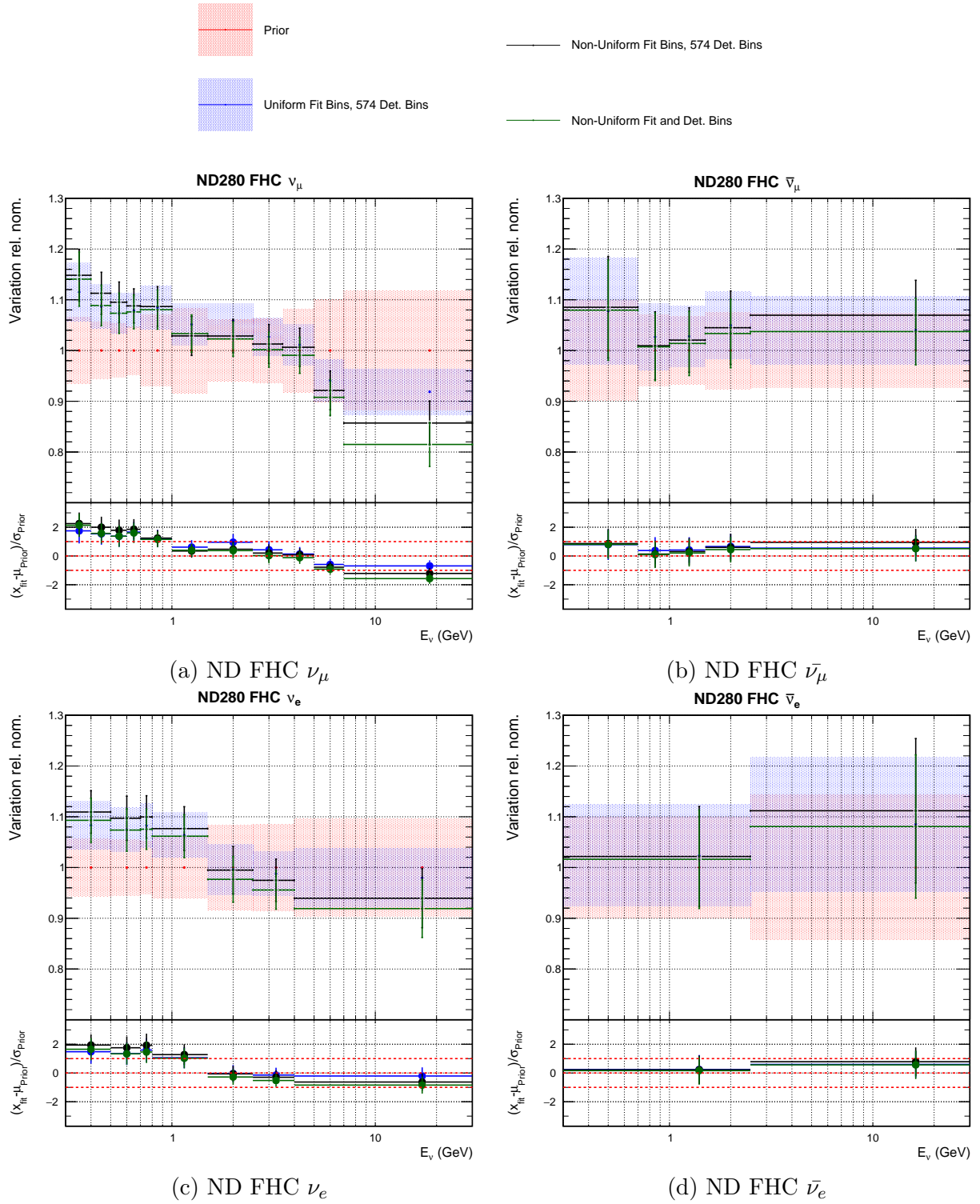
Figure 6.27.: Comparison of ND280 FHC flux parameters for the data fits with different fit and detector binnings.

The 2p2h shape parameters show opposite behaviour, with 2p2h shape C being lower for the two non-uniform binning fits, while 2p2h shape O is higher.

The Q^2 normalisations are fairly consistent. At lower Q^2, the non-uniform binned fits are slightly closer to nominal, but the postfit values are within 1σ of each other for all three fits.

(a) CC 0π

(b) Q^2 and $E_{\rm b}$

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
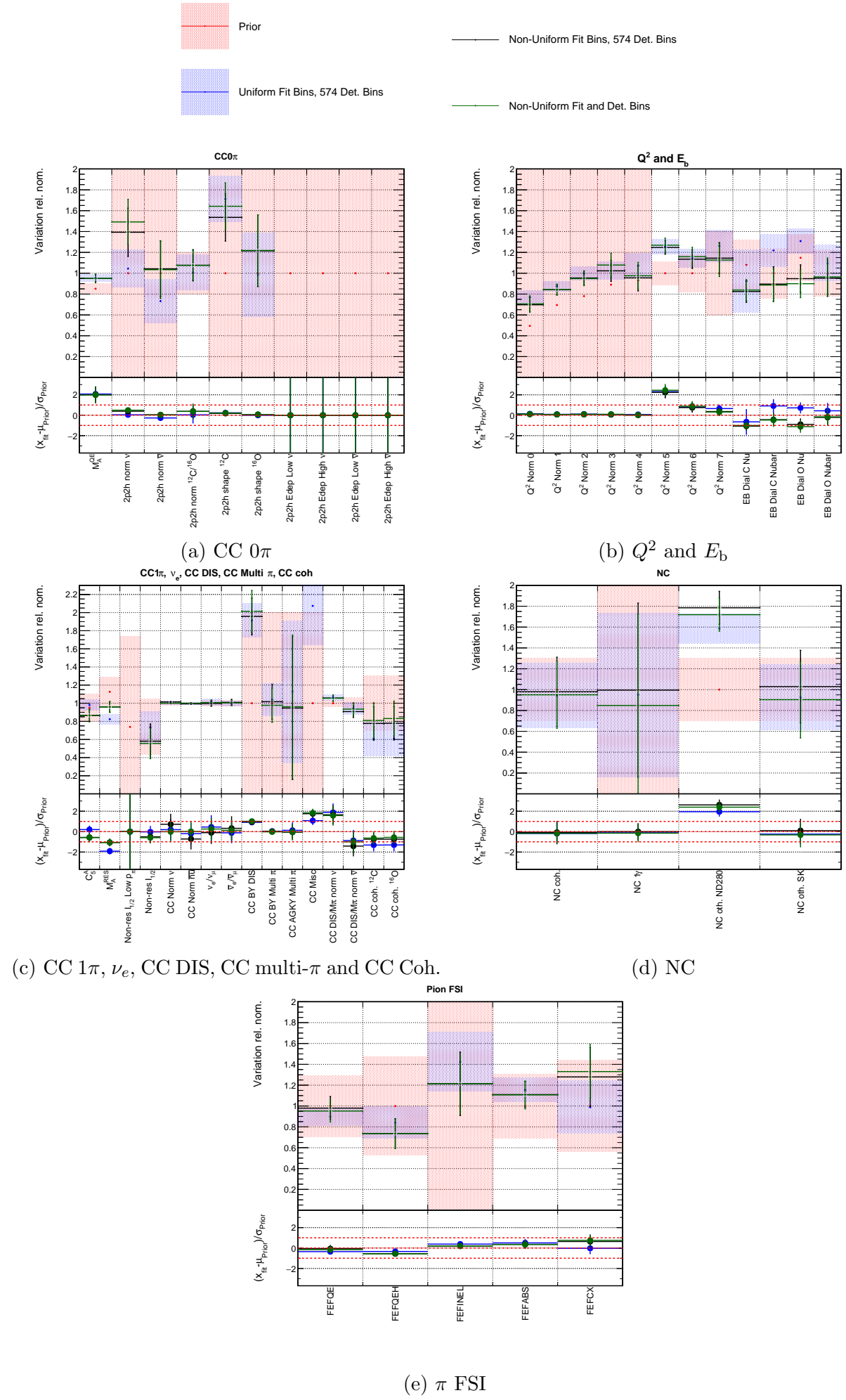
Figure 6.28.: Comparison of interaction parameters for the data fits with different fit and detector binnings.

(a) $E_\mathrm{b}\nu$ C

(b) $E_\mathrm{b}\bar{\nu}$ C

(c) $E_\mathrm{b}\nu$ O

(d) $E_\mathrm{b}\bar{\nu}$ O
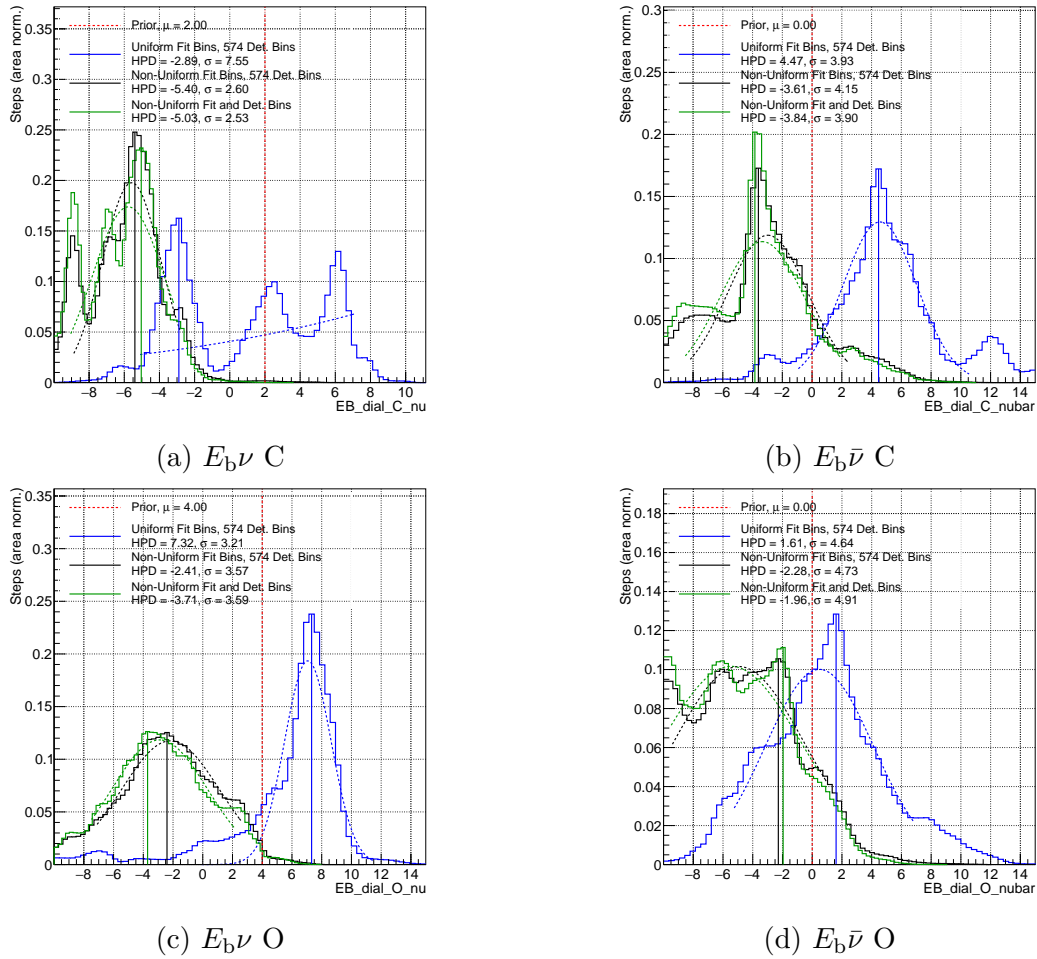
Figure 6.29.: Comparing the posterior distributions for the binding energy parameters from the data fits with different binnings.

The 1D postfit distributions for the E_b parameters are shown in Figure 6.29. The underlying distributions are extremely similar for the two non-uniform fit binning fits, but are very different to the uniform binning fit. This shows that the change in fit binning causes significant changes in the distribution. Comparing the two fits using the 574 merged detector bins (shown in blue and black in Figure 6.29), the systematics applied are identical and all that is changing is the fit binning, but the distributions are completely different, with a different number of peaks, in different positions, and with different relative sizes.

Of the single π parameters, C_A^5 and $I_{1/2}$ are lower for the non-uniform fits, but M_A^{RES} is higher. These differences are slightly larger than the postfit uncertainties. The DIS and multi-π parameters are largely consistent, and the CC coh. parameters are closer to nominal for the non-uniform binning fits.

NC 1γ is the only NC parameter to have any significant differences, but this is not surprising as there are low statistics for this interaction. The π FSI parameters have some

differences, but all are within the postfit uncertainties. The quasi-elastic, inelastic, and absorption parameters are slightly closer to nominal for the non-uniform fits, but the quasi-elastic higher energy and charge exchange parameters are pushed slightly further away.

Like the flux parameters, there is an overall trend of the postfit values for the two non-uniform fit binnings being very similar. This suggests that the criteria used for the merging of fit bins to produce the detector binning was appropriate, as the merging has not had a large impact on the final results.

6.7.4. Posterior Predictions

The posterior predictive event rates for the three fits are shown in Table 6.9. For each sample, the event rates and uncertainties are similar for each fit. The fractional uncertainties are shown in Table 6.10, which are also very similar. The uncertainty on the event rate is slightly smaller for having non-uniform fit and detector binnings for several samples, but not by very much. There is a large reduction in -2LLH for the non-uniform fit binning, but this is mostly due to the reduced number of bins.

The posterior predictive p-values for all three fits are shown in Table 6.11. Here, only the values calculated using the fluctuated draws are shown. These are still low for several samples, and the overall p-value for all fits is still 0.000.

However, several samples have a larger p-value for the non-uniform binning fits. The FGD1 FHC ν CC Other sample is 0.000 for the uniform binning, but is 0.334 with non-uniform fit binning, and 0.530 for non-uniform fit and detector binning.

The CC 0π samples are largely unchanged by the different binnings. FGD1 RHC ν, is the only 0π sample with a significant increase for using non-uniform binnings, but the uniform binning p-value is still $>5\%$.

The CC 1π samples see more differences, particularly for the non-uniform fit and detector binnings fit. The RHC ν sample p-values are 0.011 and 0.019 for the uniform binning, but increase to 0.716 and 0.286 for the non-uniform fit and detector binnings.

The CC Other samples are significantly higher for the non-uniform binning for FGD1 FHC ν and FGD2 RHC $\bar{\nu}$, but are slightly lower for FGD1 RHC ν.

In general the largest differences in p-values occur for samples which already had a 'good' value, but there a several which are $\sim0\%$ for the uniform fits, but $>5\%$ for the non-uniform fits. For the majority of samples, the non-uniform fit and detector binning produces a larger p-value than the non-unform fit and uniform detector binning, despite the fit results being very similar.

Sample	Posterior Predictive Event Rates			-2LLH		
	Uniform Fit and Det. Bins	Non-Uniform Fit, Uniform Det. Bins	Non-Uniform Fit, and Det. Bins	Uniform Fit and Det. Bins	Non-Uniform Fit, Uniform Det. Bins	Non-Uniform Fit, and Det. Bins
FGD1 FHC ν CC 0π	33375.4 ± 159.6	33383.9 ± 161.6	33385.1 ± 163.0	877.22	430.09	425.81
FGD1 FHC ν CC 1π	7951.9 ± 64.2	7914.6 ± 67.3	7919.1 ± 67.4	293.07	318.79	308.41
FGD1 FHC ν CC Other	7917.0 ± 71.4	7933.5 ± 71.4	7926.4 ± 70.8	430.51	292.16	286.97
FGD2 FHC ν CC 0π	33125.4 ± 157.6	33151.9 ± 166.2	33159.3 ± 164.6	869.49	463.30	455.58
FGD2 FHC ν CC 1π	6434.1 ± 56.8	6418.2 ± 57.0	6447.1 ± 57.4	315.77	312.24	309.32
FGD2 FHC ν CC Other	7281.8 ± 64.8	7301.6 ± 69.0	7300.8 ± 62.9	404.94	376.69	370.88
FGD1 RHC $\bar{\nu}$ CC 0π	8431.2 ± 69.5	8443.4 ± 70.7	8446.1 ± 69.0	377.92	229.49	219.71
FGD1 RHC $\bar{\nu}$ CC 1π	677.1 ± 13.4	679.0 ± 14.5	680.2 ± 10.8	58.99	46.53	45.72
FGD1 RHC $\bar{\nu}$ CC Other	1468.4 ± 23.0	1468.9 ± 23.6	1472.3 ± 21.5	93.92	94.88	90.71
FGD2 RHC $\bar{\nu}$ CC 0π	8179.5 ± 66.3	8204.3 ± 68.2	8199.9 ± 66.9	381.58	206.31	196.29
FGD2 RHC $\bar{\nu}$ CC 1π	633.8 ± 11.3	638.4 ± 12.2	642.5 ± 10.8	55.89	58.74	50.98
FGD2 RHC $\bar{\nu}$ CC Other	1371.7 ± 19.6	1378.0 ± 19.8	1371.6 ± 18.6	119.59	84.06	79.86
FGD1 RHC ν CC 0π	3576.0 ± 37.6	3575.8 ± 39.1	3570.8 ± 37.6	133.15	135.04	134.34
FGD1 RHC ν CC 1π	1156.2 ± 14.7	1151.8 ± 14.6	1150.5 ± 14.7	60.10	54.46	50.82
FGD1 RHC ν CC Other	1285.2 ± 18.3	1291.7 ± 17.3	1291.8 ± 18.4	60.51	87.45	81.08
FGD2 RHC ν CC 0π	3531.9 ± 36.6	3522.0 ± 37.7	3512.4 ± 37.1	136.10	153.84	151.90
FGD2 RHC ν CC 1π	923.0 ± 11.1	916.7 ± 11.3	912.2 ± 11.0	56.28	61.94	56.98
FGD2 RHC ν CC Other	1190.7 ± 15.9	1190.8 ± 15.0	1194.8 ± 15.5	53.13	69.43	66.24
Total	128510.3 ± 367.5	128562.2 ± 378.2	128580.1 ± 371.1	4778.16	3475.44	3381.60

Table 6.9.: Posterior predictive event rates and log-likelihood to data for the non-uniform fit binning fits.

Sample	$\delta N/N(\%)$		
	Uniform Fit and Det. Bins	**Non-Uniform Fit, Uniform Det. Bins**	**Non-Uniform Fit and Det. Bins**
FGD1 FHC ν CC 0π	0.49	0.48	0.49
FGD1 FHC ν CC 1π	0.81	0.85	0.85
FGD1 FHC ν CC Other	0.90	0.90	0.89
FGD2 FHC ν CC 0π	0.48	0.50	0.50
FGD2 FHC ν CC 1π	0.88	0.88	0.89
FGD2 FHC ν CC Other	0.89	0.94	0.86
FGD1 RHC $\bar{\nu}$ CC 0π	0.82	0.84	0.82
FGD1 RHC $\bar{\nu}$ CC 1π	1.98	2.14	1.59
FGD1 RHC $\bar{\nu}$ CC Other	1.57	1.61	1.46
FGD2 RHC $\bar{\nu}$ CC 0π	0.81	0.83	0.82
FGD2 RHC $\bar{\nu}$ CC 1π	1.78	1.91	1.68
FGD2 RHC $\bar{\nu}$ CC Other	1.43	1.44	1.36
FGD1 RHC ν CC 0π	1.05	1.09	1.05
FGD1 RHC ν CC 1π	1.27	1.27	1.28
FGD1 RHC ν CC Other	1.42	1.34	1.42
FGD2 RHC ν CC 0π	1.04	1.07	1.06
FGD2 RHC ν CC 1π	1.20	1.23	1.21
FGD2 RHC ν CC Other	1.34	1.26	1.30
Total	0.29	0.29	0.29

Table 6.10.: Fractional uncertainties on the posterior predictive event rates for fits with non-uniform binnings.

The SK posterior predictive event rates are shown in Table 6.12. The 1R$_e$ 1d.e. sample is the only sample for which the non-uniform fit and detector binning prediction is not higher than the non-uniform fit and uniform detector binning prediction. However, the differences are smaller than the uncertainty for all samples. The event rates for the two non-uniform binning fits are within uncertainty of the event rates for the uniform binning fit for every sample. The uncertainties on the SK posterior predictive event rates are shown in Table 6.13. The uncertainty on the overall rate, and most sample rates, is slightly smaller for the uniform binning fit.

The posterior predictive distributions are shown in Figure 6.30. There is a general trend of the non-uniform fit binning fits producing very similar predictions at SK to each other, with more prominent differences to the prediction from the uniform fit binning, as was seen in the postfit parameter values.

For the FHC 1R$_\mu$ sample, the height of the peak is slightly higher and is at a slightly lower energy for the uniform fit binning. The dip is at the same energy in all three fits, but is lower and wider for the uniform fit binning. At energies above the dip, the three predictions are very consistent, with the non-uniform fit and detector binning prediction being a very small amount higher. The differences in the RHC 1R$_\mu$ sample are very similar as for FHC 1R$_\mu$, but the height and location of the peaks are identical in all three fits.

Sample	*p*-value (Fluctuation of Draw)		
	Uniform Fit and Det. Bins	Non-Uniform Fit, Uniform Det. Bins	Non-Uniform Fit and Det. Bins
FGD1 FHC ν CC 0π	0.023	0.005	0.023
FGD1 FHC ν CC 1π	0.106	0.042	0.118
FGD1 FHC ν CC Other	0.000	0.334	0.530
FGD2 FHC ν CC 0π	0.041	0.000	0.000
FGD2 FHC ν CC 1π	0.026	0.008	0.004
FGD2 FHC ν CC Other	0.002	0.000	0.005
FGD1 RHC $\bar{\nu}$ CC 0π	0.000	0.000	0.004
FGD1 RHC $\bar{\nu}$ CC 1π	0.073	0.268	0.617
FGD1 RHC $\bar{\nu}$ CC Other	0.056	0.001	0.020
FGD2 RHC $\bar{\nu}$ CC 0π	0.001	0.010	0.035
FGD2 RHC $\bar{\nu}$ CC 1π	0.187	0.010	0.139
FGD2 RHC $\bar{\nu}$ CC Other	0.001	0.050	0.085
FGD1 RHC ν CC 0π	0.079	0.246	0.518
FGD1 RHC ν CC 1π	0.011	0.500	0.716
FGD1 RHC ν CC Other	0.203	0.008	0.079
FGD2 RHC ν CC 0π	0.055	0.037	0.010
FGD2 RHC ν CC 1π	0.019	0.095	0.286
FGD2 RHC ν CC Other	0.338	0.097	0.436
Total	0.000	0.000	0.000

Table 6.11.: Posterior predictive *p*-values for each sample, from the non-uniform fit binning fits.

Sample	Event Rates								
	Uniform Fit and Det. Bins			Non-Uniform Fit, Uniform Det. Bins			Non-Uniform Fit and Det. Bins		
1R$_\mu$	347.46	$\pm$	6.92	347.67	$\pm$	8.05	351.25	$\pm$	7.13
RHC 1R$_\mu$	134.94	$\pm$	3.00	134.81	$\pm$	3.59	136.57	$\pm$	3.12
1R$_e$	94.56	$\pm$	3.16	95.49	$\pm$	3.69	96.63	$\pm$	3.18
RHC 1R$_e$	15.96	$\pm$	0.55	16.17	$\pm$	0.65	16.33	$\pm$	0.55
1R$_e$ 1d.e.	9.04	$\pm$	0.42	8.89	$\pm$	0.47	8.62	$\pm$	0.42
Total	601.96	$\pm$	11.60	603.03	$\pm$	13.96	609.41	$\pm$	11.76

Table 6.12.: Prior and posterior predictive SK event rates.

Sample	$\delta N/N(\%)$		
	Uniform Fit and Det. Bins	Non-Uniform Fit, Uniform Det. Bins	Non-Uniform Fit and Det. Bins
1R$_\mu$	1.99	2.32	2.03
RHC 1R$_\mu$	2.22	2.66	2.28
1R$_e$	3.34	3.86	3.30
RHC 1R$_e$	3.45	4.02	3.37
1R$_e$ 1d.e.	4.65	5.29	4.87
Total	1.93	2.31	1.93

Table 6.13.: Fractional uncertainties on the prior and posterior predictive SK event rates.

The predictions for the FHC 1R$_e$ sample has a higher peak for the uniform fit binning fit. Above and below the peak energy, the non-uniform fit binning predictions are slightly higher. This is a larger effect at lower energies.

For the RHC 1R$_e$ sample, the three predictions are very consistent. For the uniform fit binning, the event rate is slightly smaller at low energies than for the other fits, but the difference is very small.

The predictions for the 1Re 1d.e. sample are also compatible across all energies. However, the peak is lower for the non-uniform fit and detector binning. This is the only region in any of the samples where the two non-uniform fit binning predictions aren't closer to each other than the uniform binning. Here, the non-uniform fit and merged-uniform detector binning event rate is slightly lower than the for the uniform binning, but this difference is smaller than the difference to the non-uniform fit and detector binning.

Overall, the three fit predictions are very compatible, with similar shapes and differences to the prior prediction. All the differences between the three posterior predictions are within the predicted uncertainties.

As the aim of the non-uniform binning was to improve the sensitivity, it is perhaps disappointing that the uncertainties on the event rates at SK are not smaller for the non-uniform binning fits. However, looking at just the event rate does not take into account the changes in shape and so does not describe the the full impact on the sensitivity to oscillation measurements. The shape changes in the predicted event distributions cause the $\sin^2\theta_{23}$ contours to be tighter, particularly for the normal hierarchy, indicating an improvement in constraint, as discussed in Section 6.8.2.

6.8. Oscillation Parameter Sensitivity

To see the full effect of all the updates to the near detector fit on the oscillation analysis, joint Asimov fits were run. These fit the ND280 and SK MC to the ND280 data and SK nominal MC.

Like for the SK posterior predictions, the oscillation parameters are set to the values in Table 6.6, and the SK detector parameters are set to their nominal values. The priors on $\sin^2\theta_{23}$, Δm^2_{32}, δ_{CP}, and $\sin^2\theta_{13}$ are flat for these fits. The solar parameters, $\sin^2\theta_{12}$ and Δm^2_{12}, which T2K has little sensitivity to, have Gaussian priors from the PDG [72].

The ND280 fit binning used is the non-uniform binning defined in Appendix B. However, the detector binning is the 574 merged-uniform bins, as using the fit binning as detector
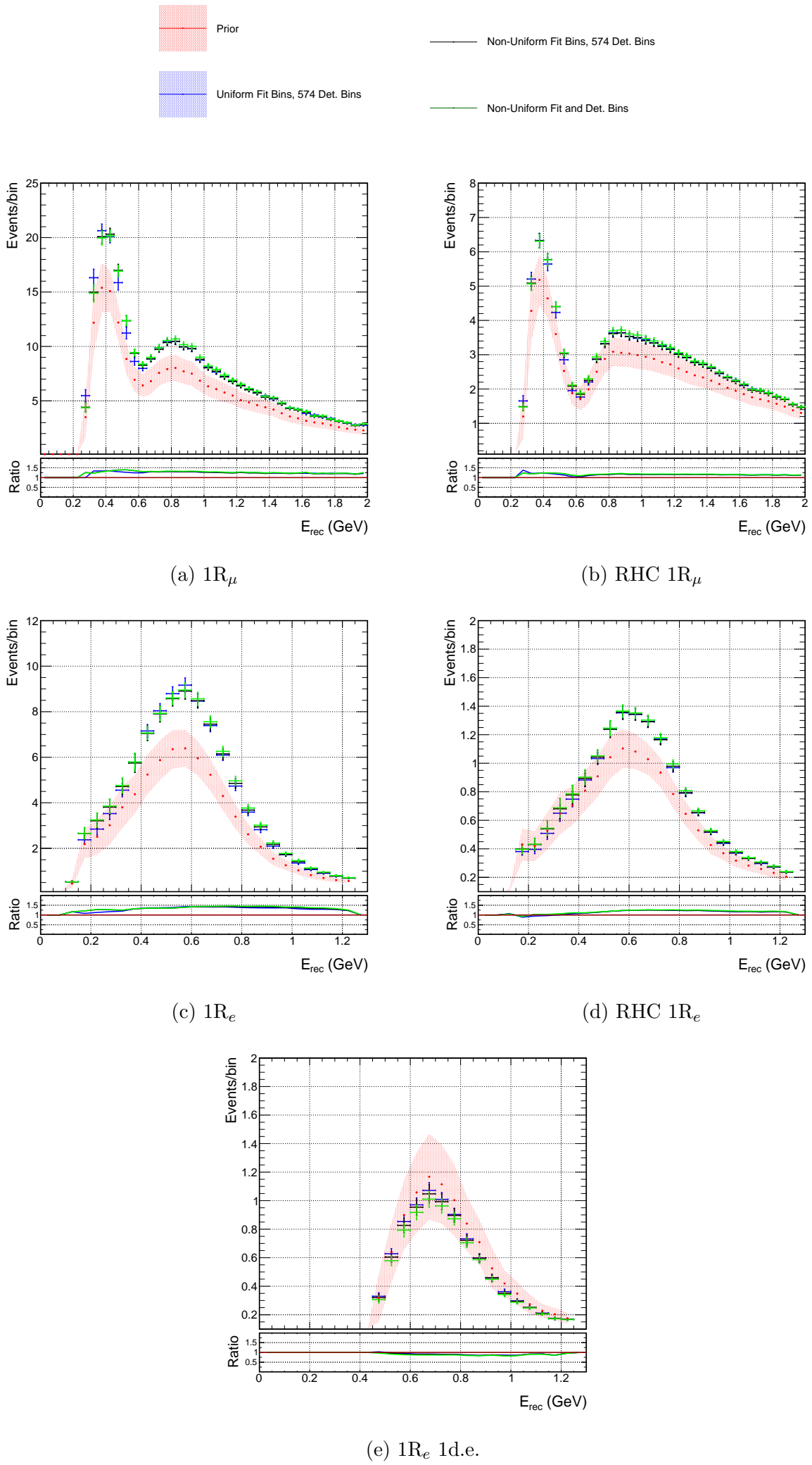
(a) 1R$_\mu$

(b) RHC 1R$_\mu$

(c) 1R$_e$

(d) RHC 1R$_e$

(e) 1R$_e$ 1d.e.

Figure 6.30.: Posterior predictive SK distributions with non-uniform binnings.

binning introduces too many parameters for the full joint fit to converge in a feasible amount of time. As the joint fits have an increased number of samples and systematics, longer Markov Chains are needed to reach the stationary distribution. The joint fit presented here consists of ∼21 million steps.

The interaction and flux systematics are compared to the near detector only fit in Figures 6.31 and 6.32, with the full results being shown in Appendix F. The postfit central values are very similar and the uncertainties are not reduced further, showing that almost the entire constraint on the systematic parameters is coming from the near detector samples. Interestingly, many parameters have a slightly larger postfit uncertainty for the joint fit than for the near detector only fit. This is likely due to correlations with SK only parameters causing marginalisation effects.

6.8.1. Comparison to the 2018 Oscillation Analysis

The joint Asimov fit results were compared to those from the 2018 oscillation analysis, with the improvement in sensitivity coming from the increase in data and updates to the input models, fitting framework, and systematic treatments described in this thesis. Part of the improvement will also be due to the increased SK data and updated SK detector covariance.

Figure 6.33 shows the posterior probability distribution for the disappearance parameters, Δm_{23}^2 and $\sin^2\theta_{23}$, with all other parameters marginalised over. Figure 6.33a shows steps for the normal hierarchy, with $\Delta m_{23}^2 < 0$, and Figure 6.33b shows steps for the inverted hierarchy, with $\Delta m_{23}^2 > 0$. The solid dashed lines show the 90% and 68% credible intervals respectively. The best fit points are shown by the triangles, for both analyses. Where the contours are jagged and not smooth, this is due to more steps being needed to fully fill out the posterior distribution.

There is significant improvement on the potential constraint for both parameters, with the 90% interval for this analysis now lying almost entirely inside the 68% interval for the previous analysis, for both hierarchies.

The comparison of the appearance parameters, $\delta_{\rm CP}$ and $\sin^2\theta_{13}$, are shown in Figure 6.34. The constraint on $\sin^2\theta_{13}$ has improved significantly, but not by as much as the disappearance parameters. The $\delta_{\rm CP}$ constraint is very similar for the two analyses. This is because the appearance parameters are dominated by the statistical uncertainty, and so the considerable updates to the systematic treatments has a smaller effect. The one dimensional posterior probability distributions for $\delta_{\rm CP}$ are shown in Figure 6.35. There is a slightly narrower peak for the 2020 analysis, but this a very small difference.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

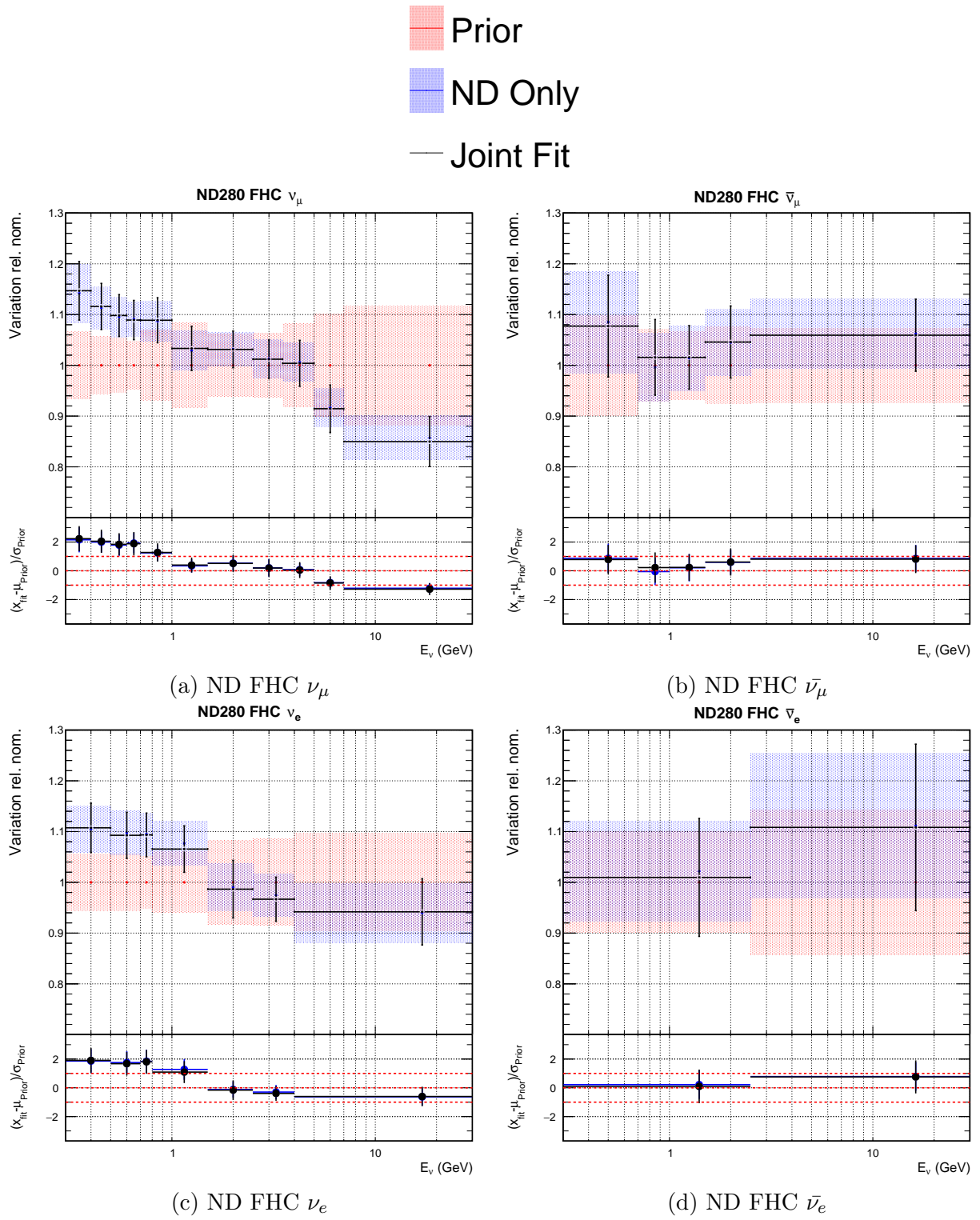(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

Figure 6.31.: ND280 FHC flux parameters for the joint and near detector only fits.

To show exactly where the improvement in sensitivity is coming from, one of the far detector fitting groups produced contours adding in updates in a piecewise manner [144]. These contours include a prior [72] on $\sin^2\theta_{13}$ using data from reactor experiments, which also had an update for this analysis. The results in Figures 6.33–6.35 only use T2K data, and so do not agree exactly with these contours. Furthermore, these were produced using one of

(a) CC 0π

(b) Q^2 and E_b

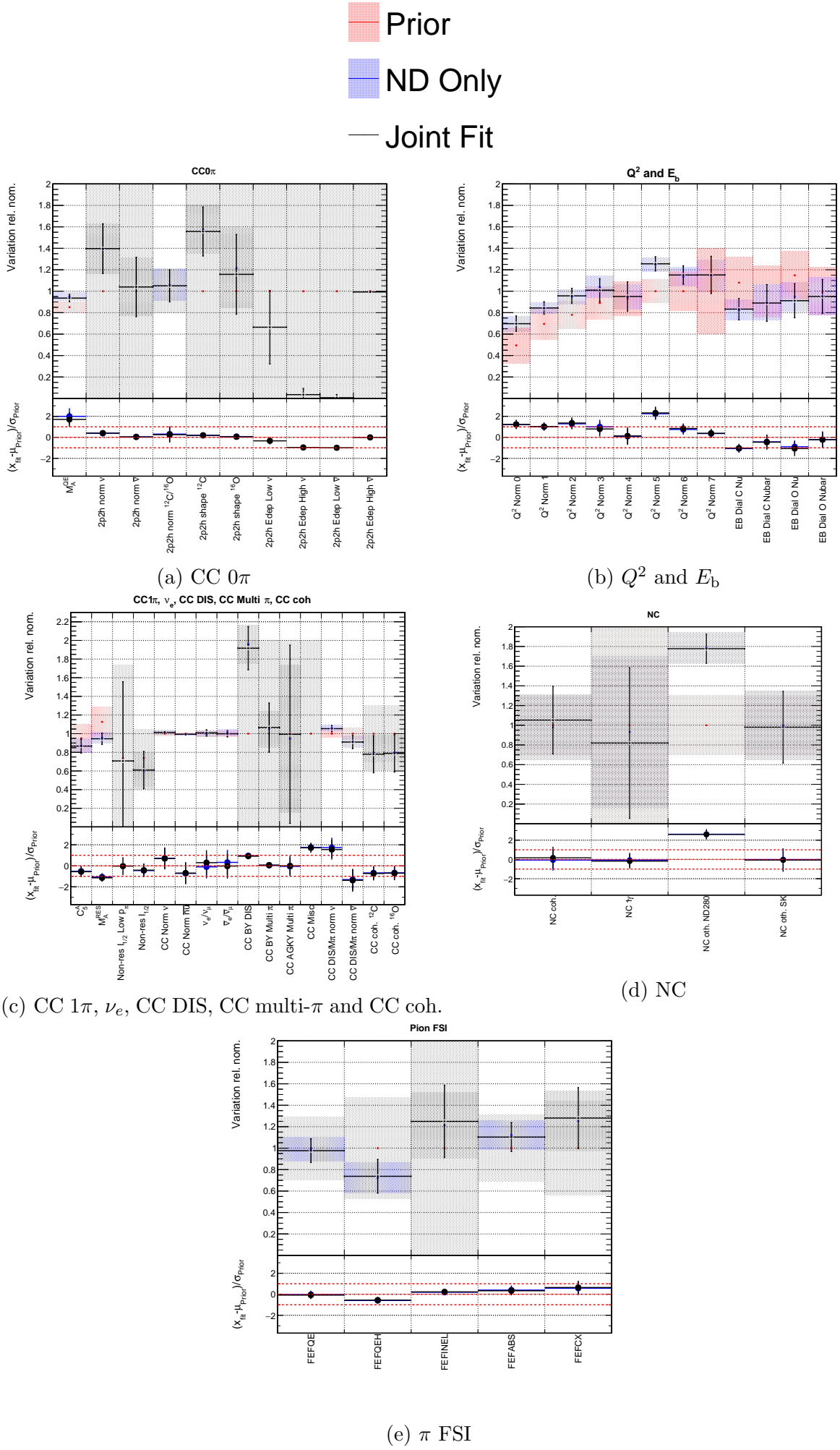(c) CC 1π, ν_e, CC DIS, CC multi-π and CC coh.
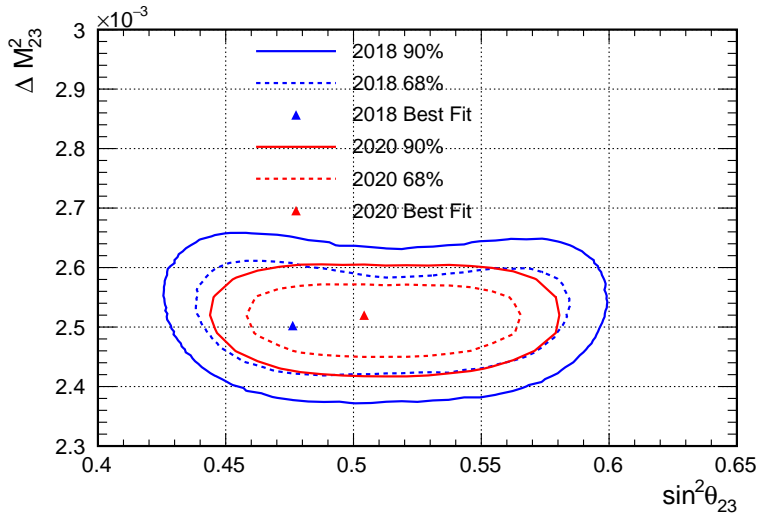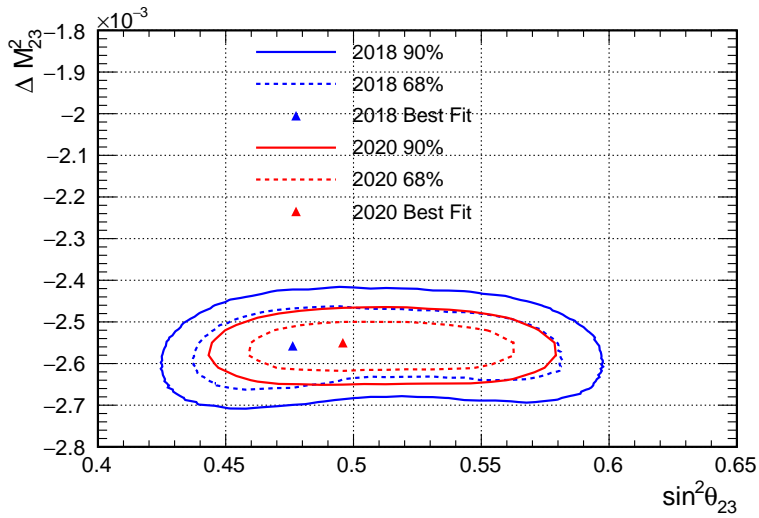
(d) NC

(e) π FSI

Figure 6.32.: Interaction parameters for the joint and near detector only fits.
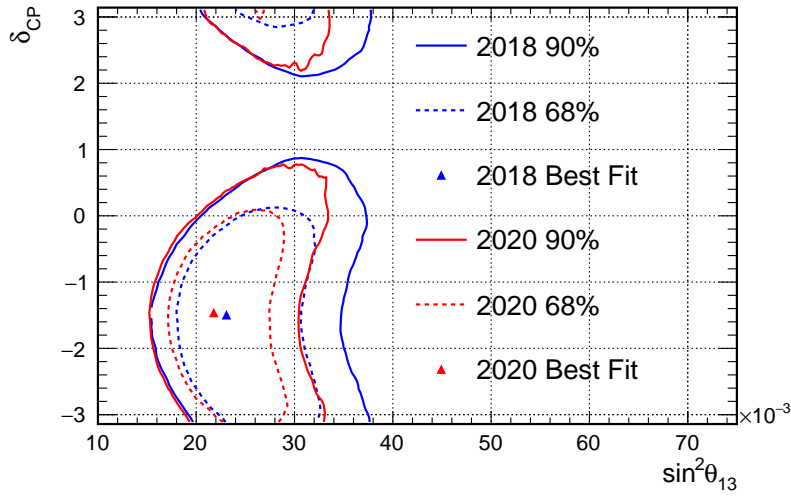
(a) Normal Hierarchy

(b) Inverted Hierarchy

Figure 6.33.: Comparison of the disappearance parameter credible intervals from joint Asimov fits for the 2018 and 2020 oscillation analyses. The solid (dashed) lines show the 90% (68%) credible intervals.

the frequentist far detector fitting groups which takes the BANFF near detector fit result as an input. There are therefore expected to be differences to the Markov Chain result because of differences in the implementation. Marginalisation effects, the different near detector fit binning, and the different E_b treatments will all cause discrepancies in the results. However, the relative size of changes from the different updates will be the same for each analysis.

The disappearance parameters in Figure 6.36 show a large increase in constraint for fixing the binding energy parameter, demonstrating how big an impact the old treatment of the uncertainty had on the oscillation analysis. There are then smaller improvements for moving
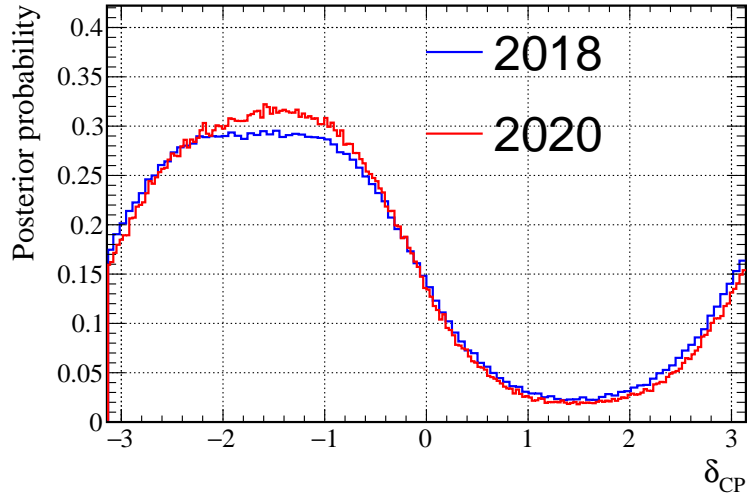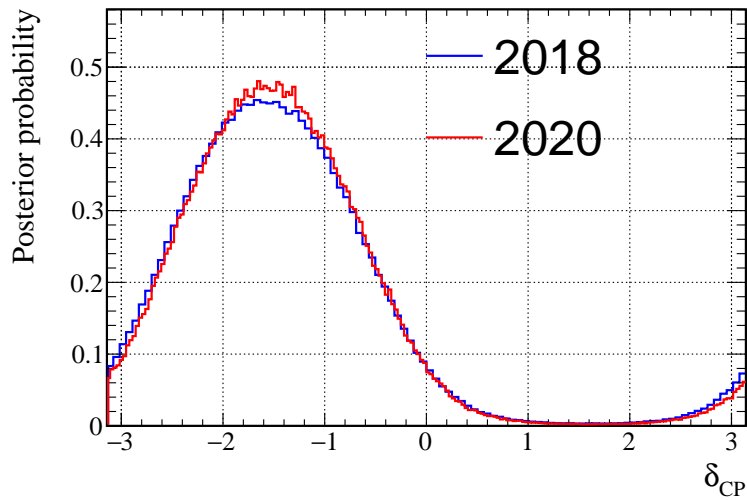
(a) Normal Hierarchy
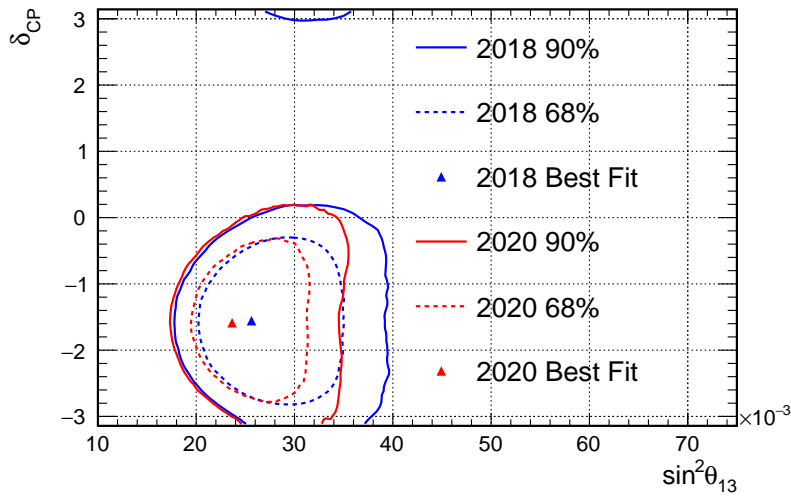
(b) Inverted Hierarchy

Figure 6.34.: Comparison of the appearance parameter credible intervals from joint Asimov fits for the 2018 and 2020 oscillation analyses. The solid (dashed) lines show the 90% (68%) credible intervals.

to the updated interaction model (including the new binding energy parameter), and the new near detector fit results. The update to the reactor constraint, which sets a tighter prior on $\sin^2\theta_{13}$, does not have a significant effect on the disappearance parameters. Finally, adding the new SK data improves the sensitivity, but the change is not as large as for fixing the binding energy. There are similar relative improvements from each update for the normal and inverted hierarchies.

For the appearance parameters in Figure 6.37, the update to the reactor constraint has the most significant impact. The contours get tighter in $\sin^2\theta_{13}$, but slightly wider in δ_{CP}.
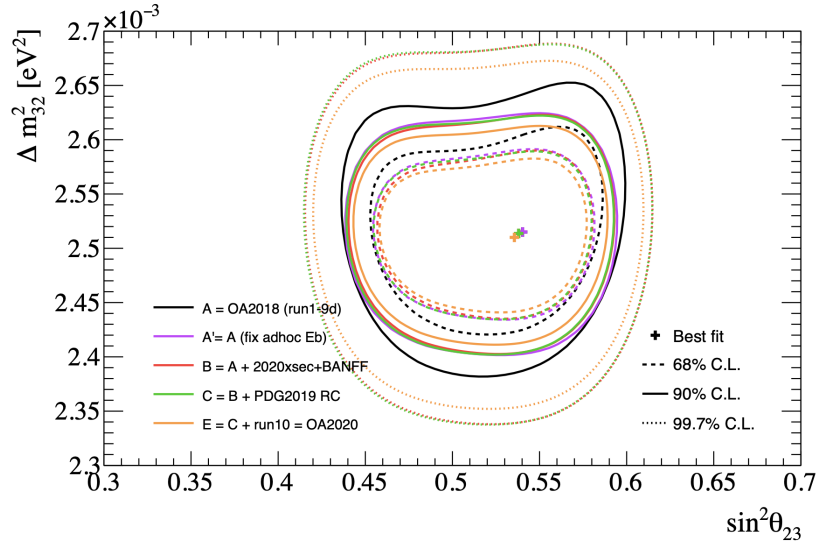
(a) Normal Hierarchy

(b) Inverted Hierarchy

Figure 6.35.: Comparison of the δ_{CP} posterior probability distribution from joint Asimov fits for the 2018 and 2020 oscillation analyses.

They also move to a higher value in $\sin^2\theta_{13}$, which is expected from the change in prior.

6.8.2. Impact of Non-Uniform Near Detector Binning

Although the uncertainty on the predicted SK event rate was not reduced by using non-rectangular fit binning at the near detector, this was a single bin analysis which did not take into account any shape information. The full impact of the non-uniform binning on the potential sensitivity is seen in the posterior probability distributions from joint Asimov fits. An identical joint Asimov fit was run, but with the uniform near detector fit binning. All other aspects of the systematics and samples for both the near and far detector were the

(a) Normal Hierarchy
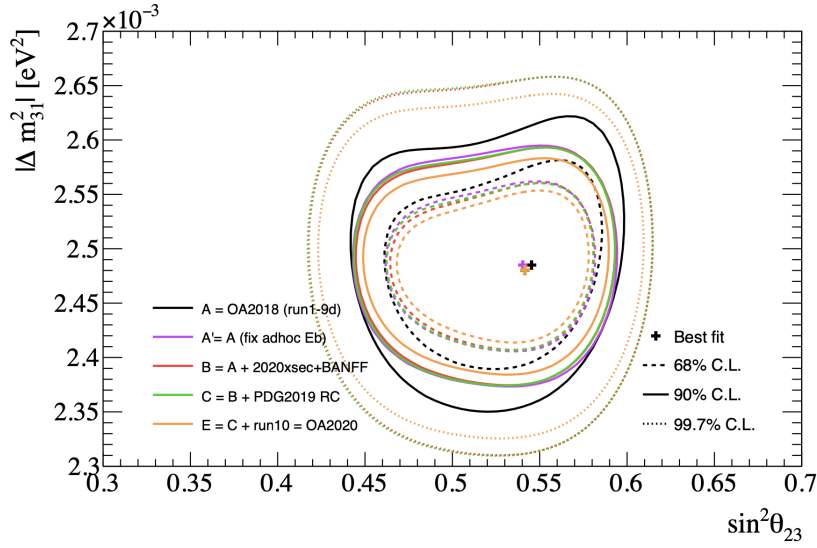
(b) Inverted Hierarchy

Figure 6.36.: Asimov sensitivity for the disappearance parameters with incremental updates from the 2018 oscillation analysis [134]. Figures from [144].

same, and both fits used the 574 merged near detector systematic binning.

The credible intervals for the appearance parameters are compared in Figure 6.38, for both the normal and inverted hierarchies. There is a better constraint on $\sin^2\theta_{23}$ for the fit using non-uniform near detector fit binning for both hierarchies. This is due to the dip in the FHC ν_μ prediction being slightly narrower in the posterior prediction for the non-uniform fits. The constraint on Δm^2_{23} is also slightly stronger for the non-uniform fits, particularly for the normal hierarchy, but this is a smaller effect.

The posterior distributions for the appearance parameters are shown in Figure 6.40. The contours for the uniform near detector binning were produced from a fit with fewer steps,

(a) Normal Hierarchy
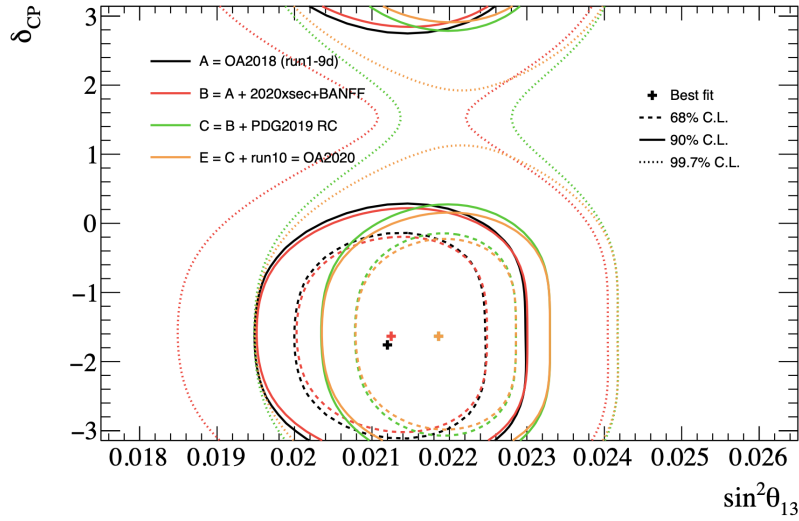

(b) Inverted Hierarchy

Figure 6.37.: Asimov sensitivity for the appearance parameters with incremental updates from the 2018 oscillation analysis [134]. Figures from [144].

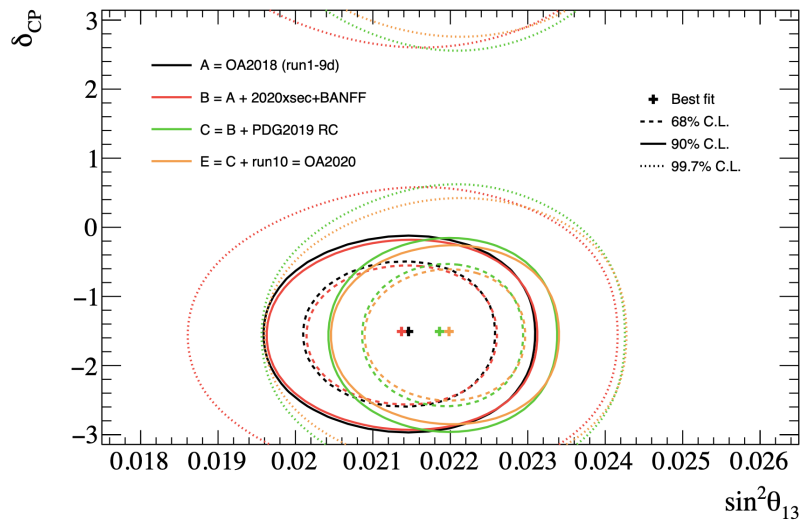causing the contours to be less smooth. For $\sin^2\theta_{13}$, the constraint is slightly tighter, but this is a very small effect. The constraint on δ_{CP} is unchanged. This is confirmed by the 1D posterior distributions for δ_{CP}, shown in Figure 6.40.

6.8.3. Impact of E_b

As discussed in Section 5.4.1.1, the binding energy was the largest individual systematic uncertainty in the previous T2K oscillation analysis. Simulated data studies have shown that the effect of varying this systematic has a significant impact on the oscillation parameter
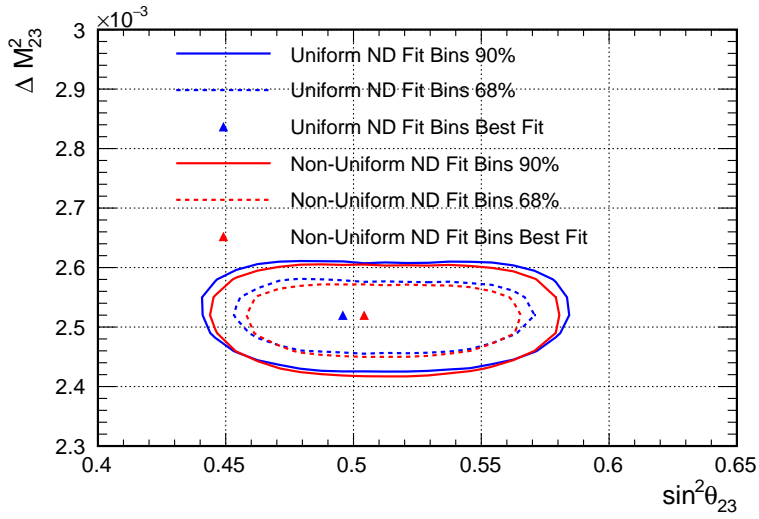
(a) Normal Hierarchy
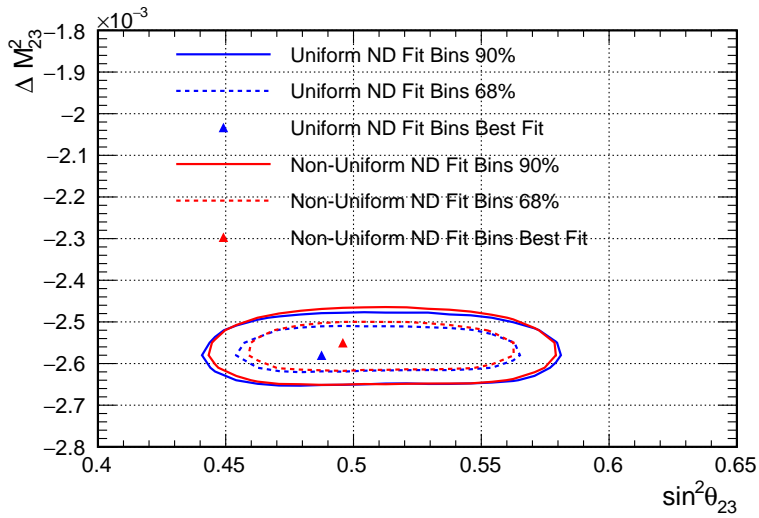
(b) Inverted Hierarchy

Figure 6.38.: Comparison of the disappearance parameter credible intervals from joint Asimov fits using uniform and non-uniform near detector fit binning. The solid (dashed) lines show the 90% (68%) credible intervals.

contours [135], and this was also seen in the piecewise updates in Section 6.8.1.

To inspect the impact of the new treatment of the binding energy on the full oscillation analysis, the $E_b\nu^{16}O$ parameter was fixed to the post near detector fit value and the same oscillation parameter contours were produced. Fixing the parameter was simulated by using only steps with $E_b\nu^{16}O$ between -3.2 and -1.6 (corresponding to a range of 23.8–25.4 MeV), a thin region around the peak in Figure 6.29c at -2.41. The region was chosen to be as small as possible while still having enough steps to produce stable contours. The number of steps contained in this region was ∼4 million.

(a) Normal Hierarchy

(b) Inverted Hierarchy

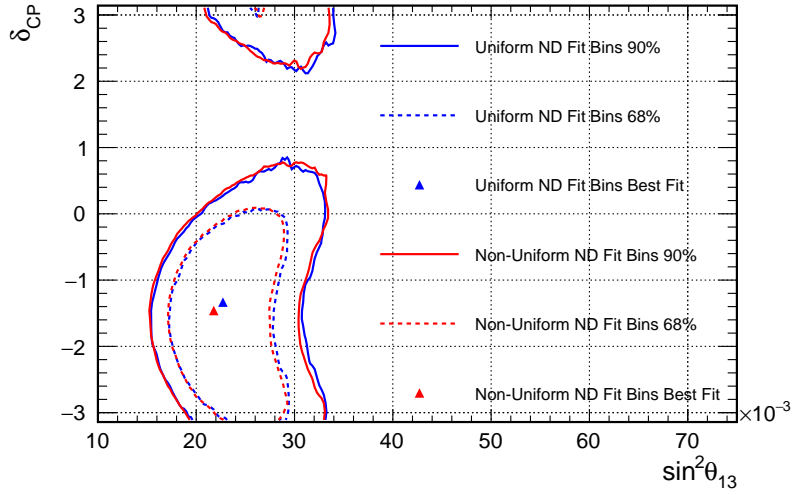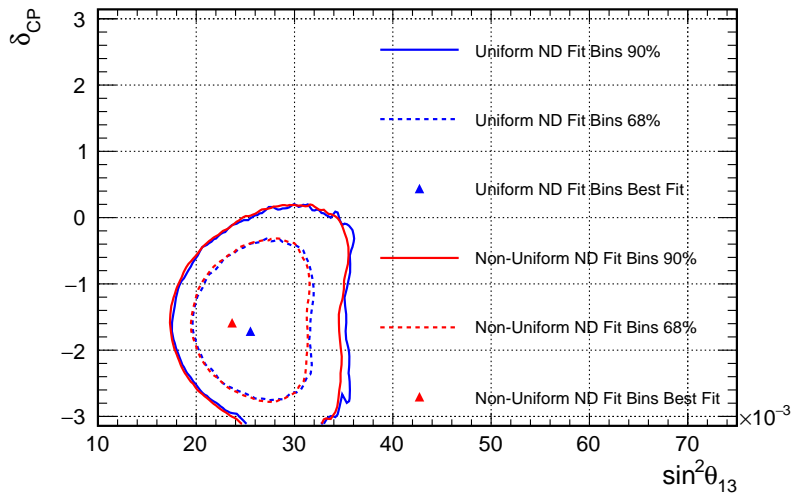Figure 6.39.: Comparison of the appearance parameter credible intervals from joint Asimov fits using uniform and non-uniform near detector fit binning. The solid (dashed) lines show the 90% (68%) credible intervals.

Only $E_b \nu^{16}O$ was fixed to ensure enough steps were selected, but the correlations mean the other E_b parameters will be somewhat fixed. Furthermore, as there is no ^{12}C in SK, and there is a larger POT for FHC than RHC, $E_b \nu^{16}O$ is the dominant E_b parameter.

The comparison of the posterior distributions for the disappearance parameters is shown in Figure 6.41. The E_b fixed contours are not as smooth as much fewer steps were used to produce them, but the shape and size is very consistent with the full fit. Any changes are much smaller than for fixing the previous E_b implementation in Figure 6.36.

The appearance contours, shown in Figure 6.42, are also very compatible. There are

(a) Normal Hierarchy
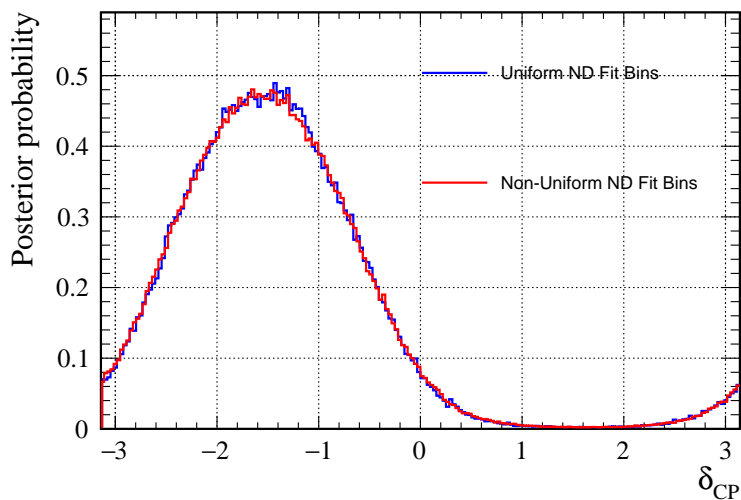
(b) Inverted Hierarchy

Figure 6.40.: Comparison of the $\delta_{\rm CP}$ posterior probability distribution from from joint Asimov fits using uniform and non-uniform near detector fit binning.

slight differences due to the lack of steps for the $E_{\rm b}$ fixed contours causing more jaggedness, but the overall shape and size aren't changed. As for the disappearance parameters, any differences are much smaller than for fixing the previous implementation in Figure 6.37. The 1D posteriors for $\delta_{\rm CP}$, shown in Figure 6.43 are also unchanged.

The fact that fixing the binding energy does not change the credible intervals demonstrates it is no longer the significant uncertainty it was in previous analyses. In 2018, the binding energy could not be fit at the near detector, and so a very large effective parameter was applied at SK. With the new treatment described in Section 5.4.1.1, the binding energy can be fit at the near detector, reducing its impact on the fit. It has now gone from being a
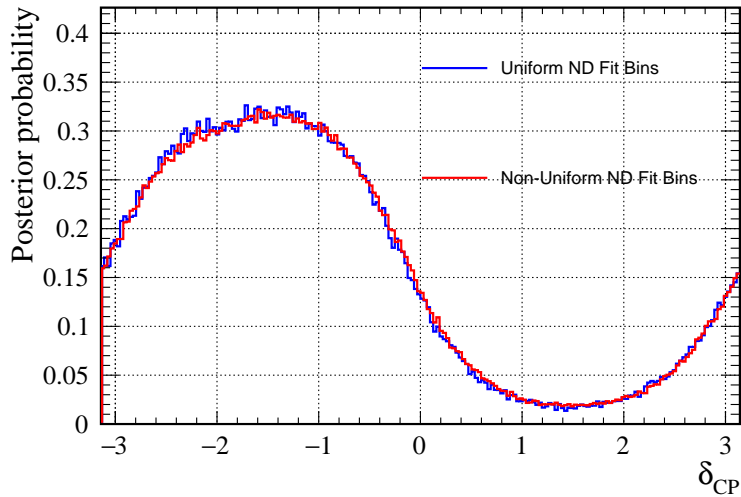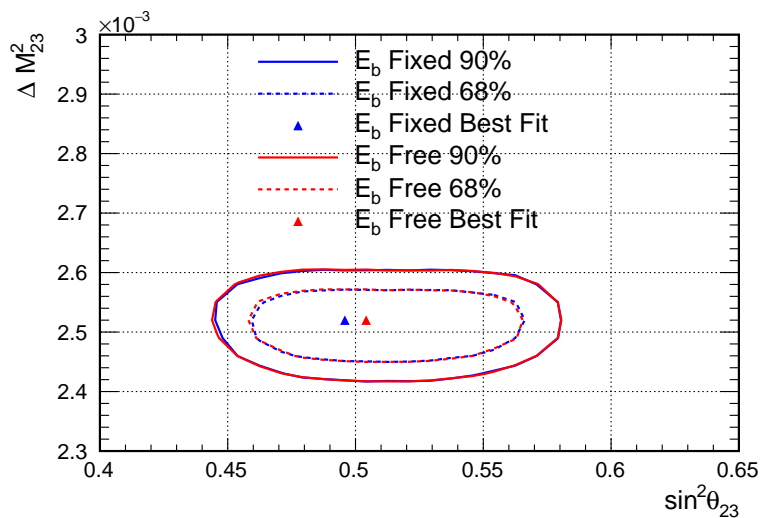
(a) Normal Hierarchy
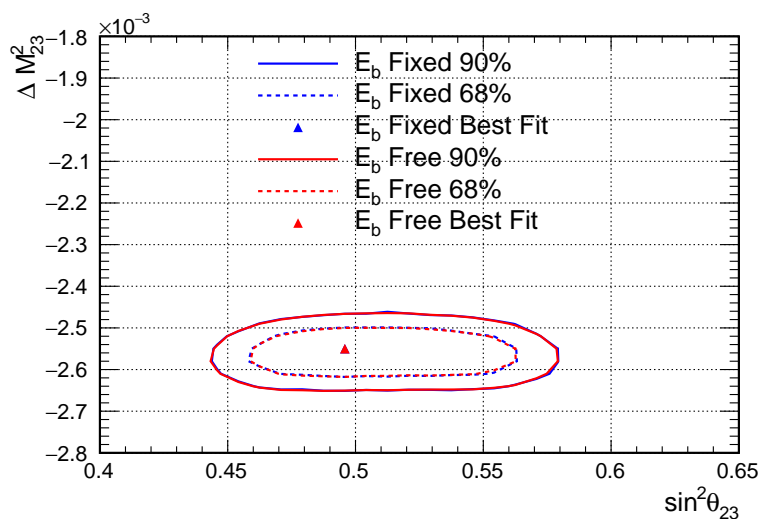
(b) Inverted Hierarchy

Figure 6.41.: Comparison of the disappearance parameter credible intervals from joint Asimov fits using steps with $-3.2 < E_b\nu^{16}O < -1.6$, and using all steps. The solid (dashed) lines show the 90% (68%) credible intervals.

dominant, to a sub-dominant systematic uncertainty in the oscillation analysis.

6.9. Summary

This chapter has presented the results of the near detector fit for the 2020 T2K oscillation analysis. The nominal MC and prior predictions have significant discrepancies to the data. The log-likelihood scans, parameter variations, and Asimov fit indicate that the fitter is behaving as expected, and is validated with the other near detector fitting group.
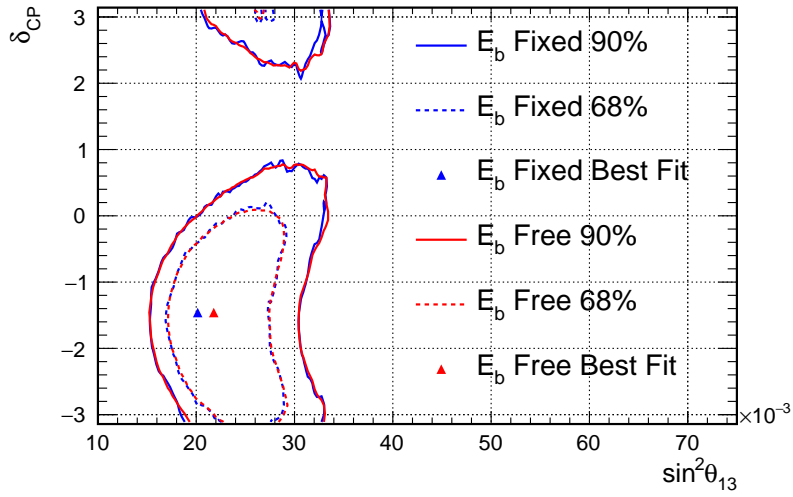
(a) Normal Hierarchy
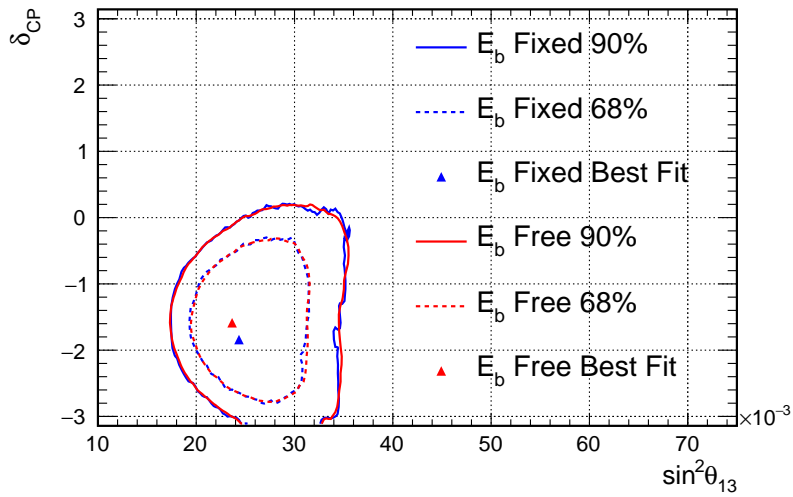
(b) Inverted Hierarchy

Figure 6.42.: Comparison of the disappearance parameter credible intervals from joint Asimov fits using steps with $-3.2 < E_b\nu^{16}O < -1.6$, and using all steps. The solid (dashed) lines show the 90% (68%) credible intervals.

The datafit results do not show any concerning deviations from the nominal central values, and most parameters are within the prefit uncertainty band. For the majority of samples, the posterior predictions are in much better agreement with the data than the prefit predictions. The uncertainties in the overall event rates are reduced significantly for all samples of both the near and far detector posterior predictions.

The posterior predictive p-values are low for a number of samples. However, this does not necessarily indicate a bad goodness of fit. The Bayesian p-values calculated in this analysis are used to show which samples and regions of phase space the fit has performed well for,
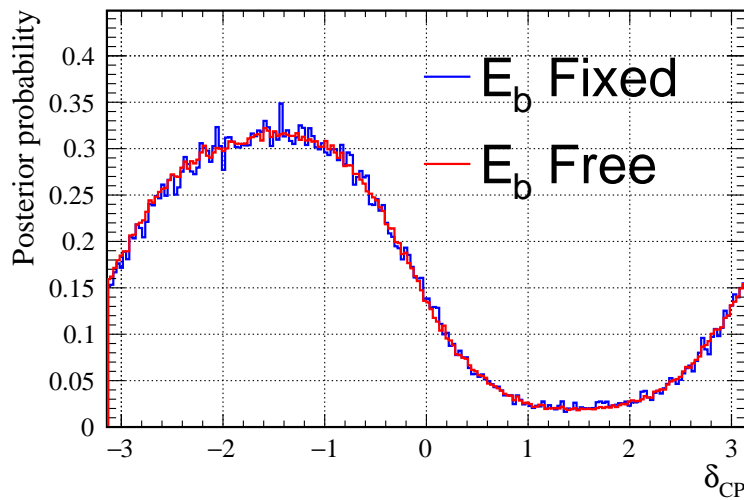
(a) Normal Hierarchy
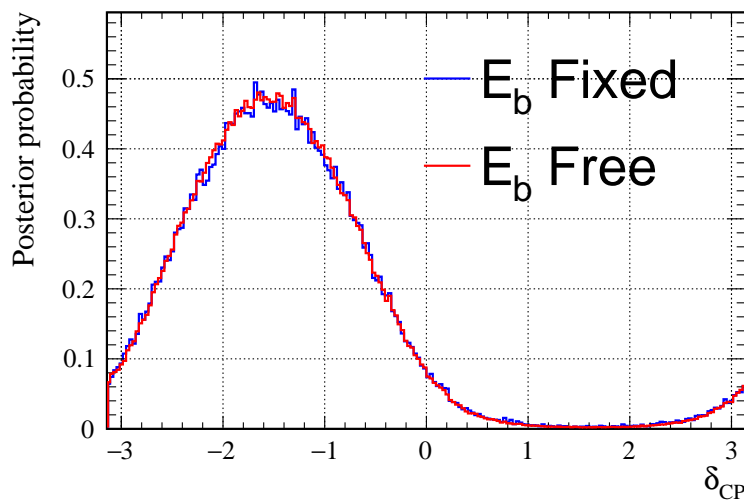
(b) Inverted Hierarchy

Figure 6.43.: Comparison of the δ_{CP} posterior probability distribution from joint Asimov fits using steps with $-3.2 < E_{\mathrm{b}}\nu^{16}\mathrm{O} < -1.6$, and using all steps.

rather than to determine if the overall goodness of fit is acceptable. The other near detector fitting group construct a more traditional p-value, which indicated that the goodness of fit had improved since previous analyses.

Similar results were obtained by using a detector covariance produced with the same detector binning as the fit binning. This shows that the bin merging process successfully reduced the number of fit parameters without reducing the accuracy of the fit.

There were larger differences in the postfit results for using uniform fit binning. This manifested as a slight shape change in the prediction at SK, but did not significantly change the uncertainty in the overall rates. However, the joint near and far detector Asimov fits

show that the non-uniform binning causes a small but not insignificant improvement in the sensitivity to $\sin^2\theta_{23}$.

Overall, there is significant improvement in the sensitivity to $\sin^2\theta_{23}$, $\sin^2\theta_{13}$, and Δm^2_{23} since the last analysis, and a slight improvement for δ_{CP}. This increase in constraint comes mostly from the updates to the cross-section model, particularly the treatment of the binding energy uncertainty. In the previous analysis, the binding energy was the dominant systematic, causing significant inflation to the oscillation parameter contours. With the new binding energy implementation, the uncertainty has very little impact on the oscillation sensitivities, showing that it is no longer a dominant systematic at T2K.

The increase in sensitivity to oscillation parameters from the non-uniform binning, and the reduction of impact of the binding energy with its new treatment are the main improvements from the work of this thesis.

7. Future Near Detectors for Long Baseline Neutrino Oscillation Experiments

For T2K, the statistical error is still the largest uncertainty on oscillation measurements. This will not be the case for the future long baseline neutrino oscillation experiments, HK and DUNE, and so they aim to perform 5σ measurements of δ_{CP}. However, as the statistical error becomes less significant, it is ever more important that systematic uncertainties are reduced. To achieve the target sensitivity, systematic errors will need to be reduced to the 1–2% level.

As discussed in Section 5.4.1.1, cross-section uncertainties are the dominant systematic, and these depend heavily on theoretical nuclear models. This is because the target nucleon resides inside a nucleus, and so nuclear effects and final state interactions (FSI) alter the measured kinematics of final state particles. As discussed in Section 2.3.3, FSI effects cause the neutrino energy to be reconstructed incorrectly, and interactions to be misclassified, contributing a large systematic uncertainty. To reduce these uncertainties, tensions in nuclear models must be resolved, which can only be done with improved measurements of the multiplicity and momentum distributions of final state particles [145].

The NEUT, GENIE, and NuWro [146] neutrino event generators use cascade models tuned to hadron-nucleus scattering measurements to simulate final-state particles leaving the target nucleus. However, these measurements are sparse, as shown in Figure 7.1, and so semi-empirical parametrisations are used to extrapolate to the relevant momentum ranges and target nuclei. The parametrisations are different in the three generators, leading to significant differences in the multiplicity of final state protons, as shown in Figure 7.2. Below 100 MeV, the proton momentum distributions diverge considerably, but this is below the detection threshold of current detectors.

It is therefore crucial that future near detectors are able to accurately measure final state
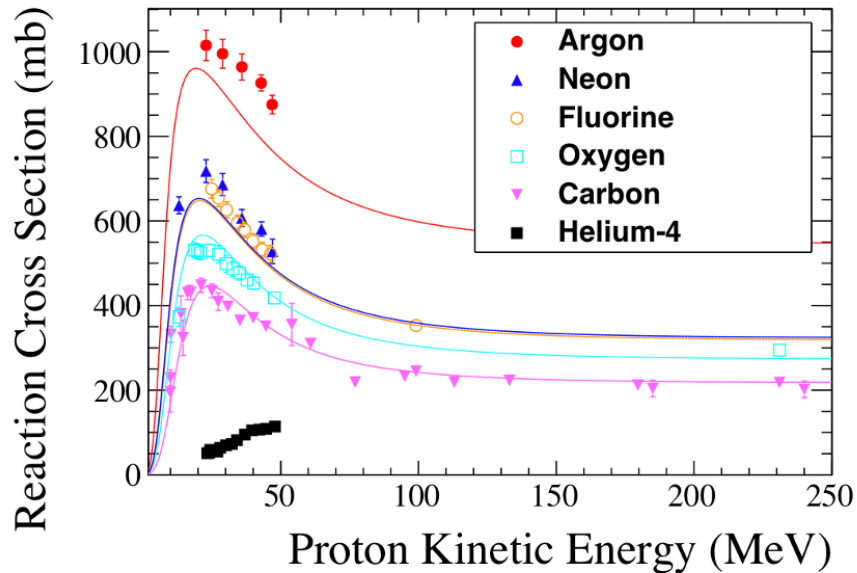
Figure 7.1.: Cross-section measurements of protons on different nuclei. Figure from [147].

particles, particularly at low momentum, to distinguish between nuclear models and reduce the effects of FSI.

7.1. High Pressure Time Projection Chamber

A high pressure time projection chamber (HPTPC) is a proposed near detector for future long baseline neutrino oscillation experiments. It is designed to be able to probe the low momentum region of parameter space to resolve nuclear model tensions and reduce neutrino interaction cross-section uncertainties.

Gas TPCs have lower momentum thresholds for detecting secondary particles as low energy hadrons travel further from the interaction point in gas than in denser detectors. The proton detection threshold is ∼110 MeV in water Cerenkov detectors, and ∼400 MeV in liquid argon TPCs. These are both too high to resolve model discrepancies.

The main disadvantage of using gas as an active target is the reduction in number of events due to the lower density, but this effect can be reduced by increasing the pressure. This, combined with the Mega-Watt beams future experiments will utilise, mean there can be enough detected events using a gaseous target.

Having the TPC filled with the active target allows 4-π angular coverage of final state particles, further adding to the HPTPC's ability to distinguish between interaction models.
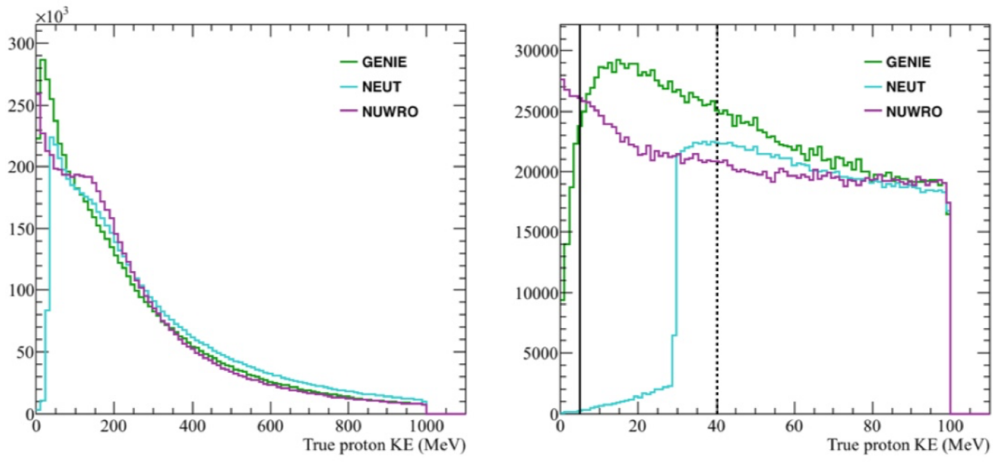
Figure 7.2.: Predicted proton energy distributions at the DUNE far detector using GENIE, NEUT and NuWro. The dashed and solid lines show the expected reconstruction thresholds in liquid and 10 atm gaseous argon detectors respecctively. Figure from [148].

An HPTPC is part of the planned near detector complex at DUNE, and T2K has explored the possibility of using an HPTPC as a long term near detector upgrade.

7.1.1. Single Transverse Variables

When the true momentum vector of the final state lepton and hadrons are projected into the plane transverse to the original neutrino's trajectory, any momentum imbalance is due to final state interactions and nuclear effects. By making accurate measurements of the multiplicity and momentum distributions of secondary particles, an HPTPC can probe the missing momentum, and better characterise events affected by FSI. This can be done using single transverse kinematics variables (STV) [149]:

- δp_T: represents the imbalance in the three momentum in the transverse plane. It is zero in the absence of FSI.

- $\delta \phi_T$: characterises how 'back-to-back' the transverse components in the final states are. It is zero for a completely balanced final state.

- $\delta \alpha_T$: describes how the hadronic system has been 'accelerated' or 'decelerated' by FSI effects. This is $< \pi/2$ if the proton transverse momentum is larger than expected for the observed lepton transverse momentum, and $> \pi/2$ if the proton transverse momentum is smaller than expected.
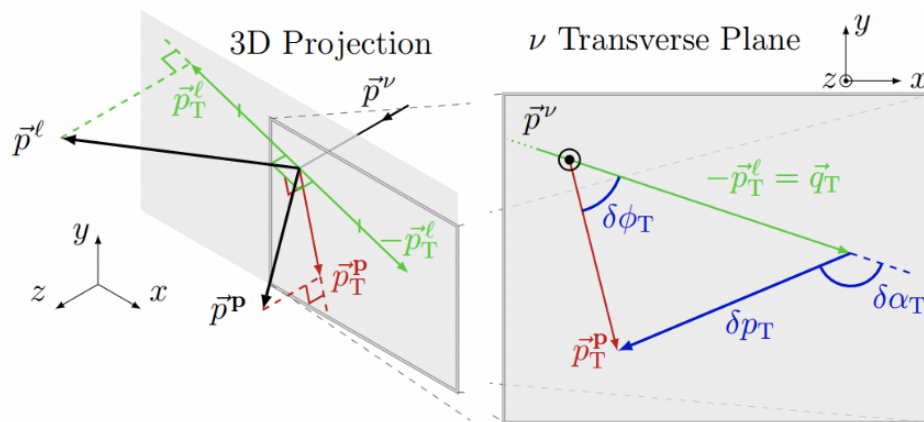
Figure 7.3.: Definition of the single transverse variables in the plane transverse to the neutrino direction. Figure from [150].

The geometric definitions of these variables in the plane transverse to the neutrino direction are shown in Figure 7.3.

These kinematics are also less dependent on the true incoming neutrino energy, which is not precisely known for individual events in accelerator neutrino experiments. Using these transverse variables can therefore provide insight into final state interactions and nuclear effects [151].

7.1.2. Sensitivity Studies

HPTPC simulation was produced to explore the potential sensitivity improvement from using an HPTPC. This was done by smearing the true kinematics of a subset of ND280 FGD1 MC events. For each event, for both the momentum and angle of the final state lepton, a random number was drawn from a Gaussian with mean equal to the true value of the variable in the ND280 MC, and width equal to the assumed resolution of the HPTPC. The assumed resolutions are shown in Table 7.1.

The selection of events is similar to that described in Section 5.2, but using the full FGD1 volume. The assumed detection thresholds were different for the ND280 and HPTPC selections, as shown in Table 7.2. The thresholds are the only differences between the ND280 and HPTPC MC.

The selected events are then divided by pion and proton multiplicity, giving seven samples in total:

- **CC 0π $0p$:** 1 muon above threshold, 0 charged pions above threshold, 0 protons above threshold.

	Value	Fractional Resolution
p_μ	< 200 MeV	0.036
	200–400 MeV	0.043
	400–600 MeV	0.053
	600–800 MeV	0.070
	800–1000 MeV	0.090
	1000–1200 MeV	0.093
	1200–1400 MeV	0.110
	1400–1600 MeV	0.120
	1600–1800 MeV	0.125
	> 1800 MeV	0.130
$\cos\theta_\mu$	< 1, > −1	0.040

Table 7.1.: Assumed HPTPC p_μ and $\cos\theta_\mu$ resolutions.

Particle	Detection Threshold (MeV)	
	ND280	HPTPC
$\mu^\pm$	100	15
$\pi^\pm$	120	16
p	450	60

Table 7.2.: Assumed detection thresholds used in these studies for the HPTPC and ND280 MC.

- **CC 0π 1p:** 1 muon above threshold, 0 charged pions above threshold, 1 proton above threshold.

- **CC 0π Np:** 1 muon above threshold, 0 charged pions above threshold, > 1 protons above threshold.

- **CC 1π 1p:** 1 muon above threshold, 1 charged pion above threshold, 0 protons above threshold.

- **CC 1π 1p:** 1 muon above threshold, 1 charged pion above threshold, 1 proton above threshold.

- **CC 1π Np:** 1 muon above threshold, 1 charged pion above threshold, > 1 protons above threshold.

- **CC Other:** 1 muon above threshold, > 1 charged pions above threshold.

The cross-section and flux systematics from the 2015 T2K oscillation analysis, described in [152] and [153] respectively, were applied. The main difference to the interaction model

described in Section 5.4.1 is that a Relativistic Fermi Gas (RFG), rather than Spectral Function (SF), nuclear model was used. The flux model was tuned using data from a thinner replica target experiment than the model described in Section 5.4.2. Detector systematics have not been developed for the HPTPC, so weren't applied to either the HPTPC or ND280 MC.

The total number of events in each sample, for each detector, are shown in Table 7.3.

	Events	
Sample	**ND280**	**HPTPC**
CC $0\pi\ 0p$	3165.52	645.38
CC $0\pi\ 1p$	3038.75	4956.65
CC $0\pi\ Np$	491.36	2634.29
CC $1\pi\ 0p$	1296.50	833.26
CC $1\pi\ 1p$	1094.97	1719.41
CC $1\pi\ Np$	94.69	516.67
CC Other	1104.51	1348.12
Total	10286.30	12653.78

Table 7.3.: Number of events in each sample for the ND280 and HPTPC MC.

The CC $0\pi\ 0p$ and CC $1\pi\ 0p$ samples have more events for ND280, as higher multiplicity events will be misclassified as having zero protons due to the higher detection thresholds. Overall, there is a greater number of interactions in the HPTPC sample, due to the lower thresholds.

The p_μ–$\cos\theta_\mu$ distributions for the two detectors are shown in Figures 7.4 and 7.5.

The distributions of CC $0\pi\ 0p$ and CC $0\pi\ 1p$ samples are similar for the two detectors, but the normalisations are higher for ND280. The $1p$ and Np samples have significantly more backward going low momentum events in the HPTPC distributions than ND280. This is because if a proton is detected in an event with a low momentum lepton, the proton momentum is also likely to be low, and so more will be below the ND280 detection threshold. The shape of the CC Other distributions are similar, but there is a slightly higher normalisation for HPTPC.

The full potential sensitivity of the HPTPC to resolving nuclear model tensions cannot be achieved with lepton kinematics alone. To investigate the impact of using hadronic information in interactions, $\pm1\sigma$ parameter variations were run on the HPTPC MC in different combinations of single transverse variables, and lepton, proton, and pion momentum and angle. The process was the same as that described in Section 6.3, but events were binned in different variables.

(a) CC 0π $0p$

(b) CC 1π $0p$

(c) CC 0π $1p$

(d) CC 1π $1p$

(e) CC 0π Np

(f) CC 1π Np

(g) CC Other
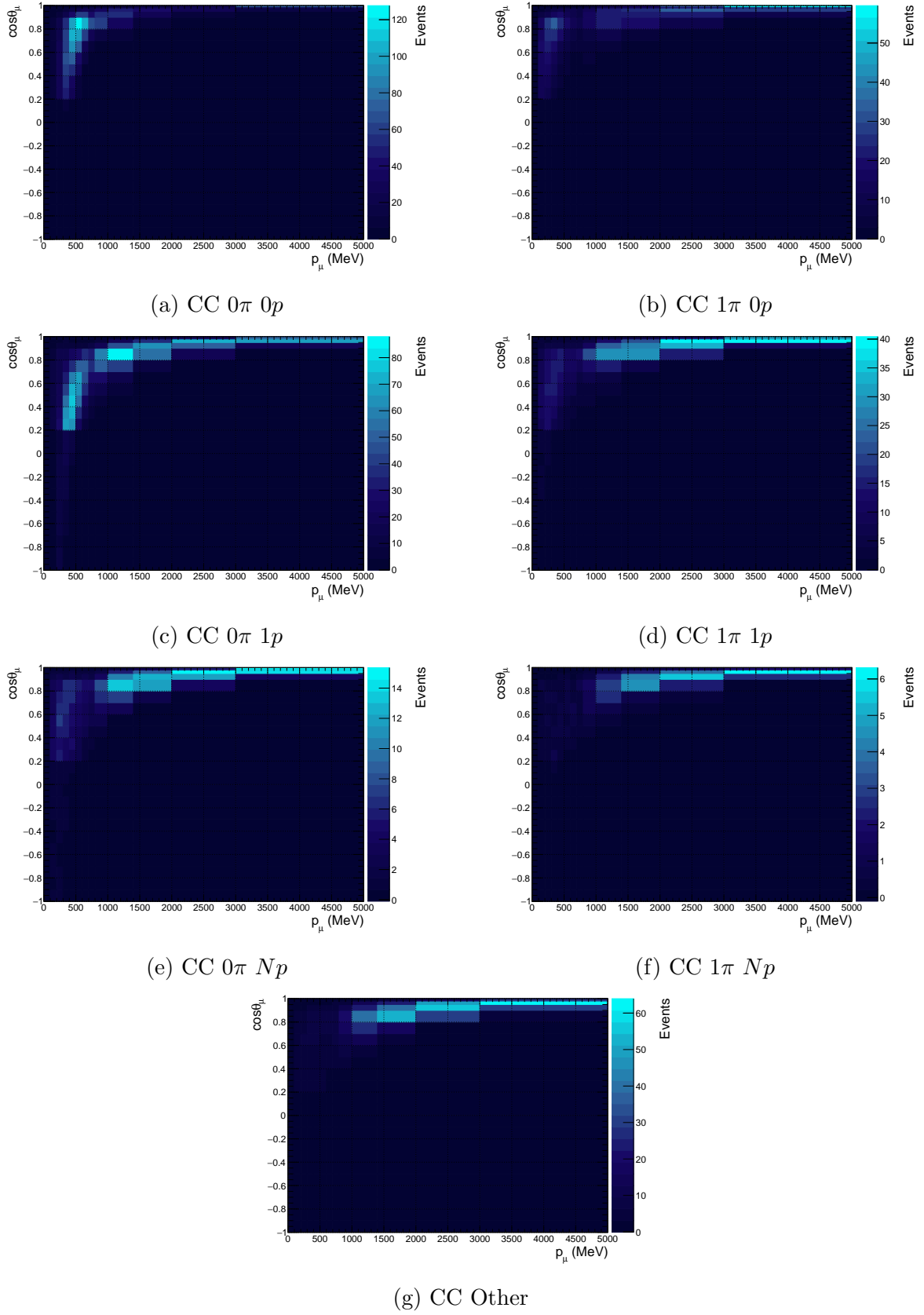
Figure 7.4.: True p_μ–$\cos\theta_\mu$ distributions for the ND280 MC.

(a) CC 0π $0p$

(b) CC 1π $0p$

(c) CC 0π $1p$

(d) CC 1π $1p$

(e) CC 0π Np

(f) CC 1π Np

(g) CC Other

Figure 7.5.: True p_μ–$\cos\theta_\mu$ distributions for the HPTPC MC.

The transverse variables are calculated for each event from the smeared true lepton momentum and angle. For each sample, the combination of interaction parameter and kinematic variables which caused the largest changed in the total number of events is shown in Table 7.4.

Sample	Parameter	Kinematic Variables
CC 0π $0p$	M_A^{RES}	p_μ–$\cos\theta_\mu$
CC 0π $1p$	2p2h Shape C	p_p–$\cos\theta_\mu$
CC 0π Np	M_A^{QE}	p_p–$\cos\theta_\mu$
CC 1π $0p$	M_A^{RES}	p_π–$\cos\theta_\mu$
CC 1π $1p$	C_A^5	δp_T–$\cos\theta_\mu$
CC 1π Np	C_A^5	p_p–$\cos\theta_\mu$
CC Other	CC DIS	p_μ–$\cos\theta_\mu$

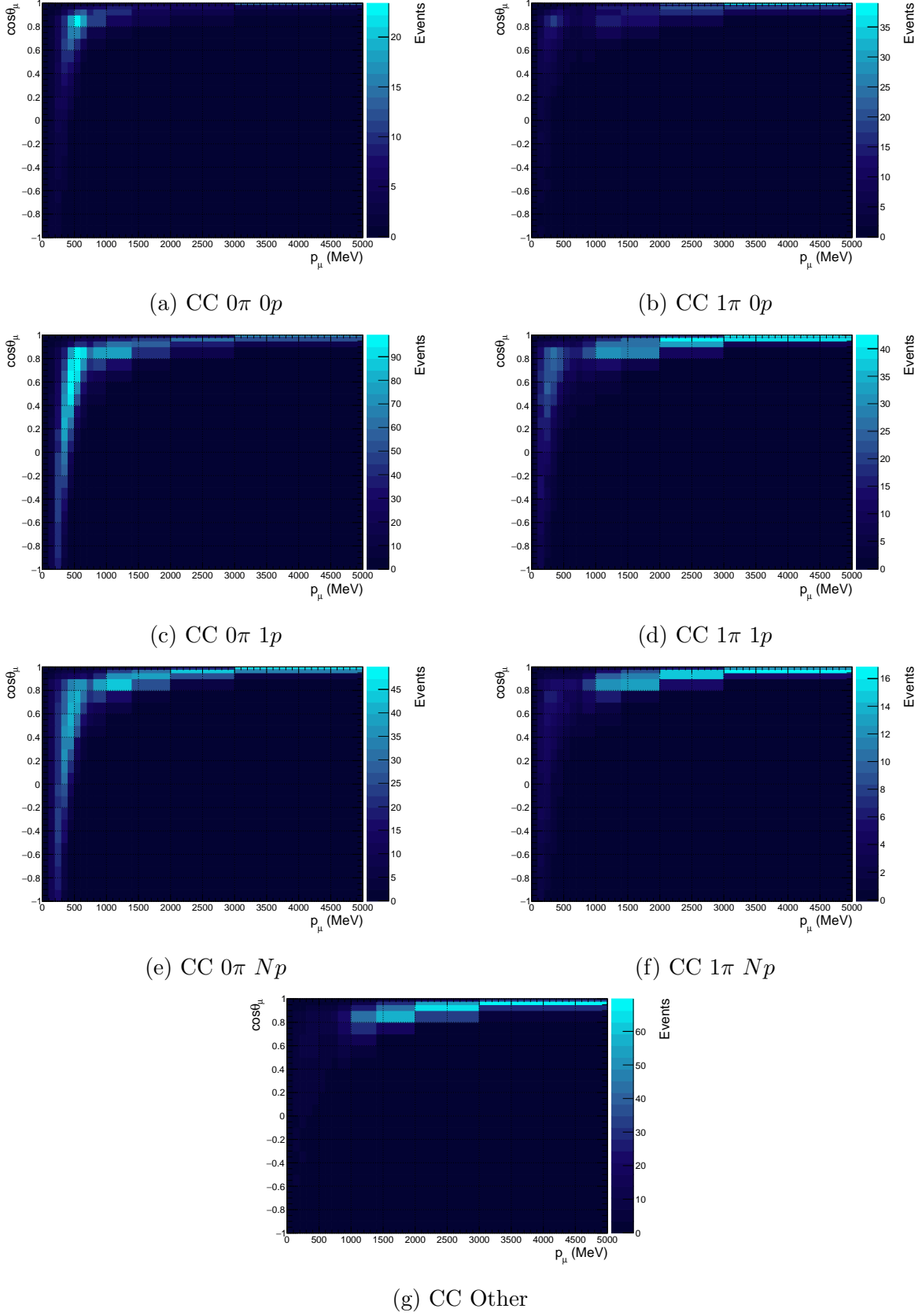Table 7.4.: Combination of kinematic variable pair and interaction parameter which caused the biggest change in the total event rate for each sample in the $\pm 1\sigma$ parameter variations.

Interestingly, for each sample one of the pair of kinematic that variables that cause the biggest change in the event rates in the parameter variations is either the lepton momentum or angle. This is likely because these variables can be measured with the best resolution by the detectors.

The distributions in these pairs of kinematics for the HPTPC MC samples are shown in Figure 7.6.

The samples binned in p_p and p_π have hard cuts at low momentum, beyond which there are very few events. However, these are significantly below the ND280 detection thresholds. These samples, along with the δp_T binned CC 1π $1p$ sample, have much narrower momentum distributions than those binned in p_μ. The 'gaps' in the distributions at $\cos\theta_\mu \sim 0.8$ are binning effects, whereby bins in the peak regions are finer, so have fewer events. The binnings have not been fully tuned to the new kinematic variables.

7.1.2.1. Asimov Fits

Three Asimov fits were ran to compare the sensitivities of the two detectors. The ND280 and HPTPC nominal MC distributions were fit in p_μ–$\cos\theta_\mu$, and the HPTPC MC was also fit in the combination of single transverse variables for each sample shown in Figure 7.6.

The fit results are shown in Figures 7.7, 7.8 and 7.9.

As expected, the postfit values of all parameters are very close to their nominal values. The slight differences are due to marginalisation effects, which are larger than for the full analysis fits due to the lower statistics.

(a) CC 0π 0p: p_μ–$\cos\theta_\mu$

(b) CC 1π 0p: p_π–$\cos\theta_\mu$

(c) CC 0π 1p: p_p–$\cos\theta_\mu$

(d) CC 1π 1p: δp_T–$\cos\theta_\mu$

(e) CC 0π Np: p_p–$\cos\theta_\mu$

(f) CC 1π Np: p_p–$\cos\theta_\mu$

(g) CC Other: p_μ–$\cos\theta_\mu$
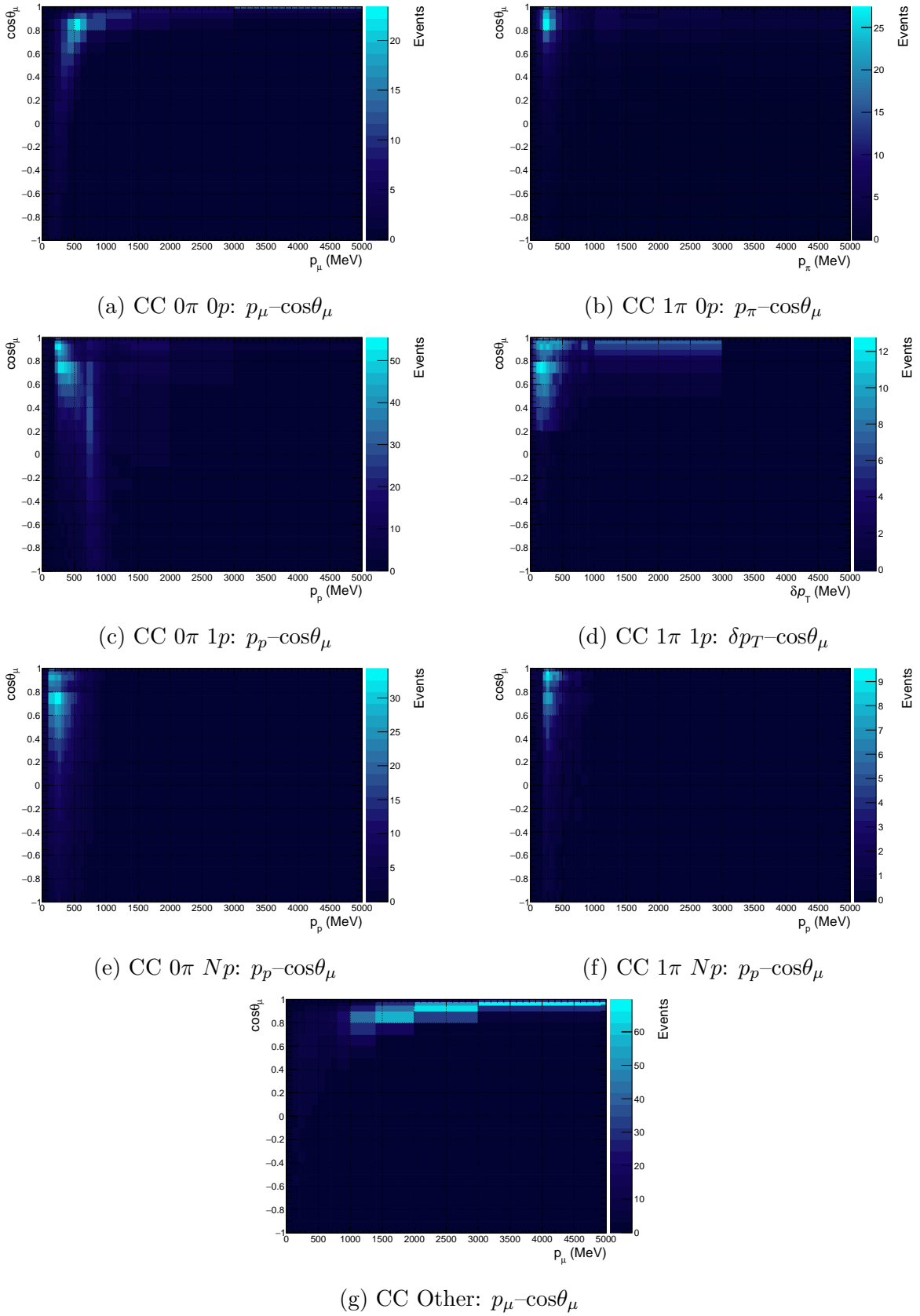
Figure 7.6.: Distributions of HPTPC MC in different kinematic variables. For each sample, the pair of kinematics shown are those that had the largest change in event rates in $\pm 1\sigma$ parameter variations.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$

Figure 7.7.: ND280 flux parameters for ND280 p_μ–$\cos\theta_\mu$, HPTPC p_μ–$\cos\theta_\mu$, and HPTPC mixed STV fits.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
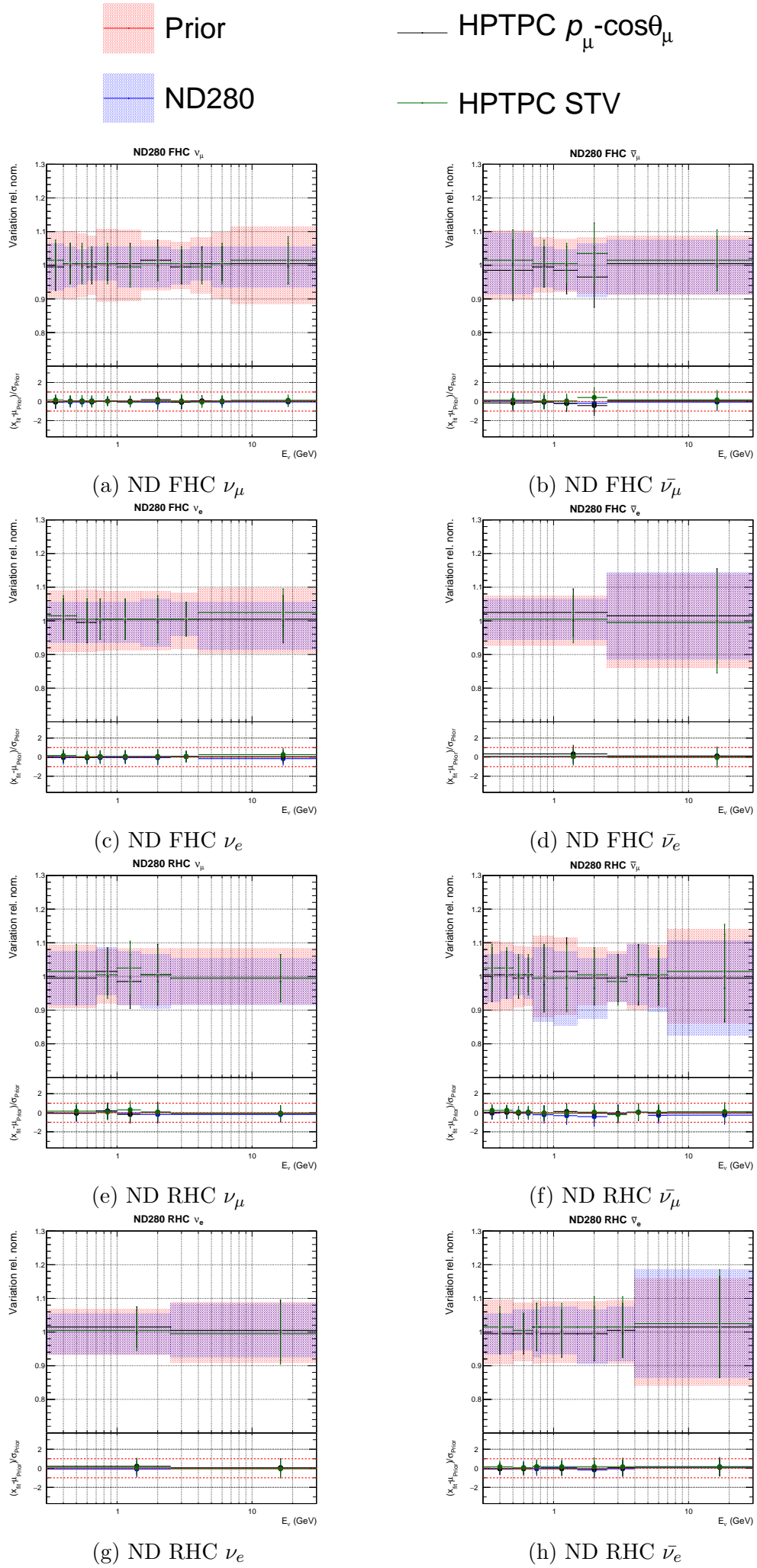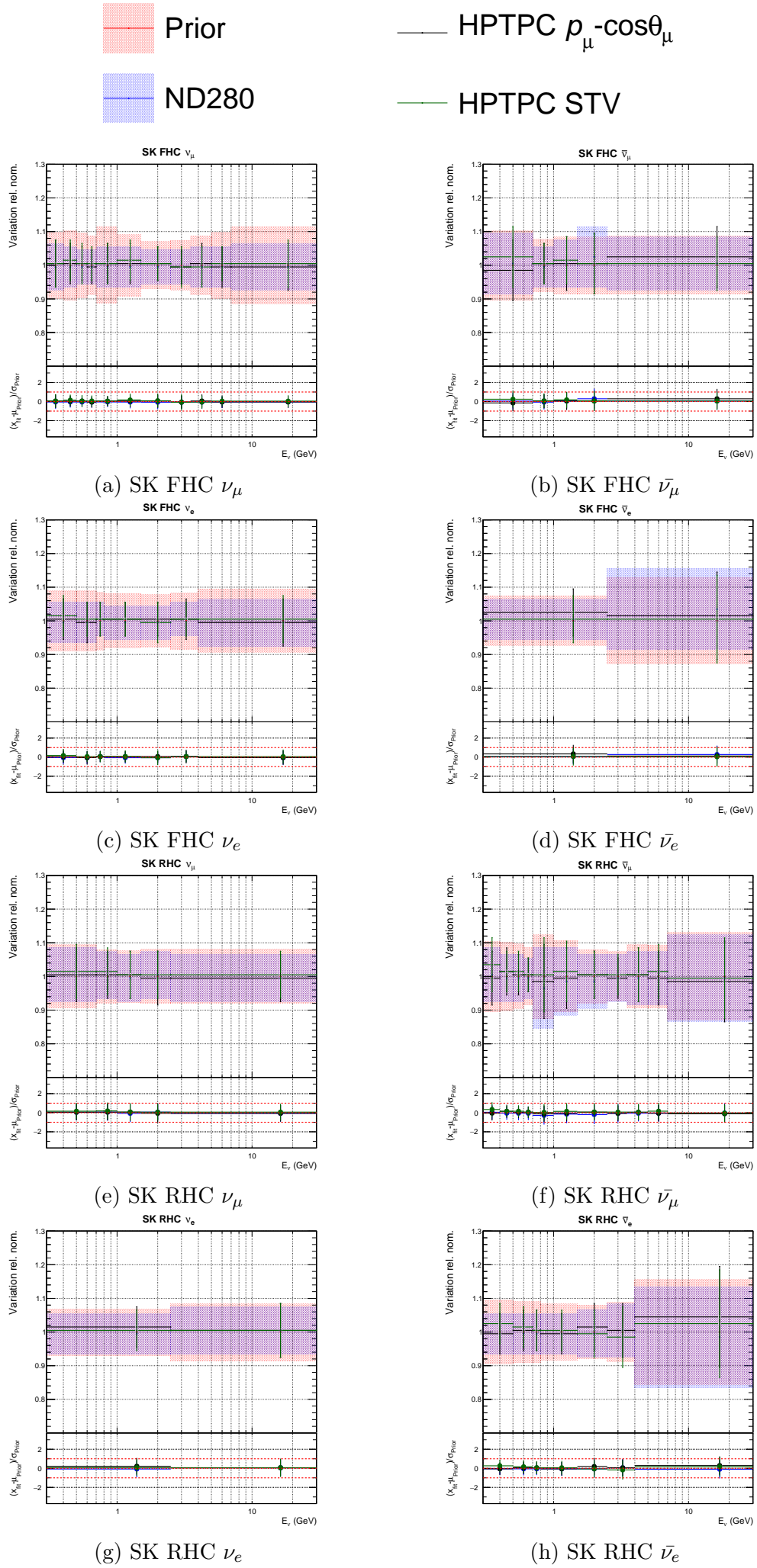
Figure 7.8.: SK flux parameters for ND280 p_μ–$\cos\theta_\mu$, HPTPC p_μ–$\cos\theta_\mu$, and HPTPC mixed STV fits.

(a) CC 0π

(b) CC 1π, ν_e, CC DIS, CC coh., NC

(c) π FSI
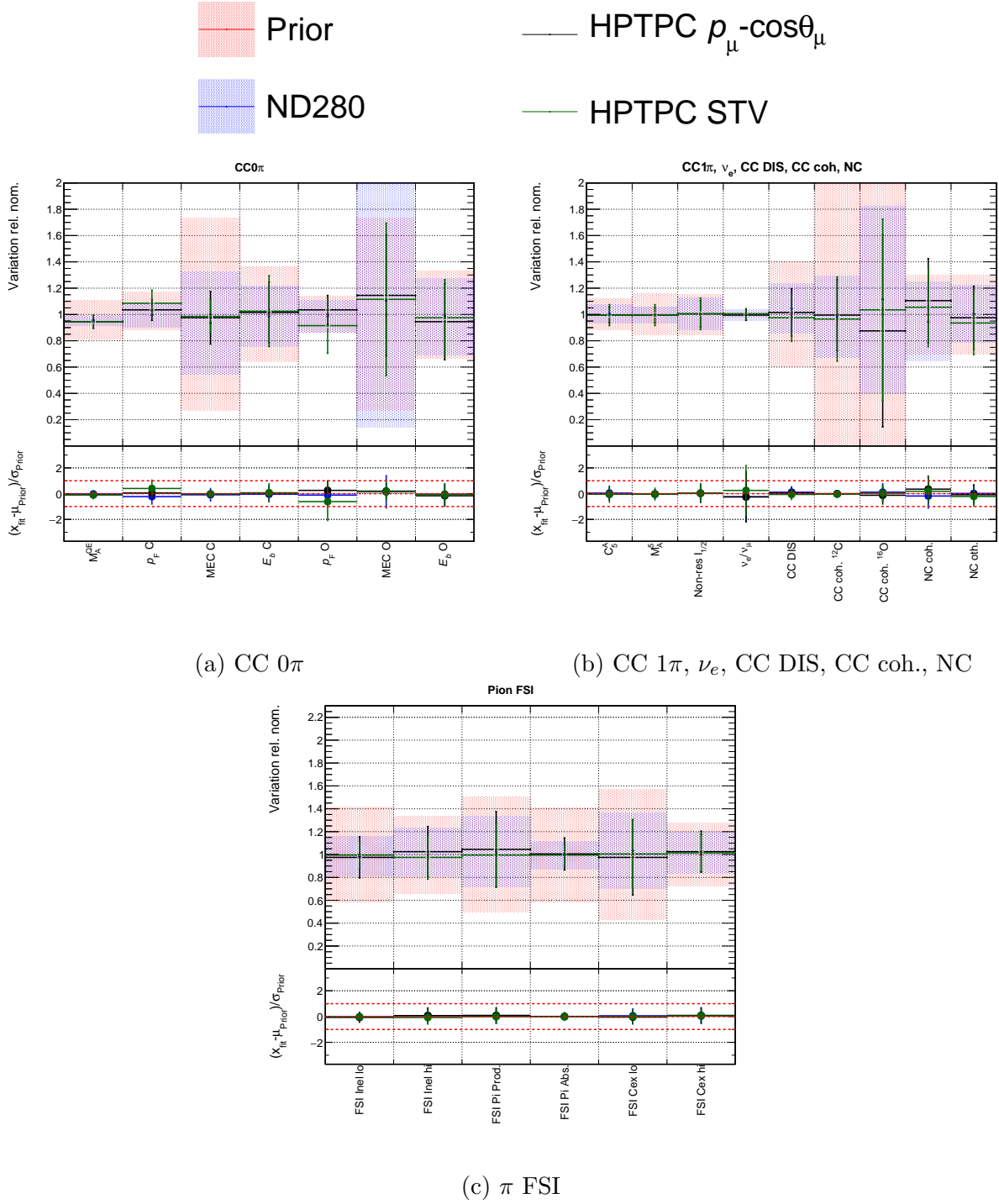
Figure 7.9.: Interaction parameters for ND280 p_μ–$\cos\theta_\mu$, HPTPC p_μ–$\cos\theta_\mu$, and HPTPC mixed STV fits.

The majority of flux parameters have a smaller postfit uncertainty for the HPTPC than ND280. The increase in number of events from lowering the detection thresholds allows a slight increase in the constraint on these systematics.

There is also a further reduction in the majority of postfit flux uncertainties for fitting in STVs rather than just p_μ–$\cos\theta_\mu$. This is a smaller effect however, and is not consistent across all energies.

7. Future Near Detectors for Long Baseline Neutrino Oscillation Experiments

Several of the interaction parameters also have a general trend of the HPTPC seeing better constraint than ND280, and the STVs being a slight improvement on just the lepton variables.

The size of the M_A^{QE}, Fermi momentum ^{12}C, and binding energy ^{12}C uncertainties are very similar in the three fits. The MEC ^{12}C uncertainty has a stronger constraint for the HPTPC, but only a small improvement from the STVs. The ^{16}O parameters have very little constraint in any of the fits as only FGD1 MC is used.

Of the 1 π parameters, the $I_{1/2}$ non-resonant background is the only uncertainty to reduce for the HPTPC compared to ND280. The STVs do not improve the constraint any further. The CC DIS and CC coh. ^{12}C parameters have small improvements in constraint for HPTPC, but this is not reduced any further by the STVs. The CC coh. ^{16}O has a similar postfit uncertainty for all three fits, as the constraint comes entirely from the prior. The NC parameters have similar postfit uncertainties for each of the fits.

The FSI parameters all have similar constraints for the HPTPC and ND280, but a significant improvement from using the STVs.

Overall, the majority of parameters have a smaller uncertainty for the HPTPC STV fit than the ND280 p_μ–$\cos\theta_\mu$ fit.

These results show that lowering detection thresholds, and using more kinematic information than just the lepton variables, can improve the constraint of systematic uncertainties in the fit. However, this is not a complete study, for a number of reasons. Firstly, without detector systematics in the fit the results are not entirely valid. These could correlate with the other parameters differently for the two detectors, altering the uncertainties in each fit.

There is also a low sample size used in this study compared to the full T2K oscillation analysis. Using the entire runs 2–9 MC to produce HPTPC-like events would give a better comparison of the potential sensitivities.

Furthermore, the HPTPC events are truth-smeared ND280 FGD1 events, and so the target is carbon. Future studies aiming for a more complete comparison of the sensitivity of the detectors should use a full simulation of an HPTPC with different target gasses. The target volume of an HPTPC would also be larger than that of ND280, increasing the target mass and therefore the number of interactions, which is not accounted for in this study.

Further improvements to the fitting framework could also help achieve the full potential of an HPTPC. With the increase in statistics future experiments will benefit from, events could be binned in more than two kinematic variables. This would allow better characterisation of interactions, and so systematics could be applied to the target events more accurately. The

systematics themselves could also be generated to be more appropriate for use with transverse kinematics. The systematics used in this study were designed to depend on the leptonic information in the event. As the lower detection thresholds allow better measurements of the hadronic side of interactions, models constructed to use this information would better use the full constraining power of an HPTPC.

However, this study does show that just by lowering detection thresholds and using different kinematic variables, systematic uncertainties gain a significant improvement in constraint compared to ND280. This is not just due to the increase in interactions, but because the extra interactions detected are at low momentum. By reducing cross-section systematics, different interaction model components can be distinguished, and in particular by probing the low momentum regions an HPTPC will allow nuclear FSI model tensions to be resolved. This will be crucial for reducing systematic uncertainties sufficiently to measure δ_{CP} to 5σ in future long baseline neutrino oscillation experiments. The full constraining power of an HPTPC will be further improved on these results by the larger target mass and fitting in more than two kinematic variables. Doing a full scale simulation of the detector would give a better indication of the HPTPC's potential sensitivity.

7.2. Summary

This chapter has presented an overview of sensitivity studies for a proposed near detector for future long baseline neutrino oscillation experiments. For the next generation of experiments to achieve their target sensitivity, systematic uncertainties will need to be reduced to the 1–2% level. This will require tensions in nuclear models to be resolved.

Using a gas target will allow an HPTPC to probe the low momentum region of parameter space where nuclear models currently diverge. Increasing the pressure of the gas, combined with future Mega-Watt beams, means there can be enough detected events using a gaseous target.

The 4-π angular coverage and lower momentum thresholds of an HPTPC allow accurate measurement of the momentum and multiplicity of secondary particles. This allows STVs to be used to better characterise events affected by FSI, and so nuclear model tensions can be better resolved.

HPTPC MC events were produced by smearing the true kinematics of ND280 FGD1 MC events and using estimated improvements in detection thresholds. Asimov fits show an improvement in sensitivity from the estimated thresholds for an HPTPC compared to

ND280, and for using a combination of different STVs for different samples compared to using lepton kinematics only.

Although a larger dataset, and full HPTPC simulation should be used for more comprehensive sensitivity studies, these results show that lowering detection thresholds and using different kinematic variables significantly improve the constraint on systematic uncertainties compared to ND280. This will allow nuclear models to be better distinguished in future long baseline neutrino oscillation experiments, which will be crucial for measuring δ_{CP} to 5σ.

8. Conclusions

This thesis has presented a Bayesian analysis of T2K near detector data, using Markov Chain Monte Carlo methods to constrain systematic uncertainties for oscillation measurements. Parametrised models of the interaction, flux, and detector systematics are fit to near detector data, and the results are propagated to the far detector by a joint ND280 and SK fit. The constraint on systematics in the full joint fit comes almost entirely from near detector data. The uncertainties on predicted SK event rates are reduced from 12–14% to 2–3% by this analysis.

Since the last oscillation analysis, there have been several updates to the near detector fit. This has included improved interaction, flux and detector models, new near detector samples, non-rectangular fit binning, and doubling the amount of data. Implementing these updates into the near detector framework has been the work of this thesis.

In particular, the treatment of the binding energy systematic has been overhauled for this analysis. Having a parameter directly shift the kinematics of an MC event, rather than just reweighting, has not been done before at T2K. Many studies and validations of the parameter were performed as part of this thesis. The new implementation has reduced the binding energy from being a dominant systematic in the previous oscillation analysis, to being sub-dominant in this analysis.

Updating the near detector fitting framework to accommodate non-uniform binning and tuning this binning was another large aspect of this thesis. The new binning has improved the sensitivity to the disappearance parameters, Δm^2_{23} and $\sin^2\theta_{23}$, significantly. In the future, as more data is taken, the impact of non-uniform binning will increase.

Although the Bayesian p-value for this result was low for several samples, the frequentist p-value, calculated by the other near detector fitting group on T2K, has increased significantly since the last analysis. It is not expected that the two p-value calculations give the same results, as they answer different questions. The Bayesian p-value is a harsh test by construction, used for determining for which samples and regions of phase space the fit is performing well, whereas the frequentist version is a more traditional p-value. The frequen-

tist p-value also uses a more accurate treatment of detector systematics, and so is used as the main indicator of goodness of fit for the analysis. The improvement in frequentist p-value is largely due to the improvements to the fit and input models since the last analysis.

The full effect of all the updates to the fit implemented for this analysis was a significant improvement in the sensitivity of the measurement of oscillation parameters, particularly for the disappearance parameters where the 90% credible intervals for this analysis are similar to the 68% credible intervals for the previous analysis.

In future analyses, using Principle Component Analysis to reduce the number of fitted detector parameters will allow the full fit binning to be used as the detector binning in joint fits. This will allow more accurate application of ND280 detector systematics. Although the merged detector binning produced very similar results to the full fit binning as detector binning in near detector only fits, the Bayesian p-value was not as high for several samples. Furthermore, as was shown by the improvement in $\sin^2\theta_{23}$ sensitivity for the non-uniform binning, the SK prediction uncertainty not reducing does not necessarily mean the oscillation parameter sensitivity will not improve. Ultimately, fitting each underlying detector systematic individually, allowing some to directly shift event kinematics like the binding energy parameter, would be a more accurate implementation.

In the longer term future, sensitivity studies show that using new technologies, such as an HPTPC, in combination with fitting in Single Transverse Variables would enable better distinguishing of final state interaction models, which diverge at low momentum. With increased statistics, it will be possible to fit in more than two dimensions, allowing further improvements in distinguishing interaction types and therefore providing a better distinction of signal and background. This will provide a better constraint on systematics, which will be even more vital with the higher statistics future long baseline neutrino oscillation experiments will benefit from. This will ultimately improve the sensitivity to oscillation parameters, allowing even more precise measurements of their values to be made.

Appendices

A. Detector Binning

The result of the fits with different detector binnings, used to tune the final detector binning, are shown in Figures A.1 and A.2 for the flux parameters, and Figure A.3 for the cross-section parameters. To avoid tuning the detector binning on data, fake data was produced by setting the cross-section parameters to their best fit values from the 2017 analysis and reweighting the runs 2–6 nominal MC. The MC was then fitted to this fake data.
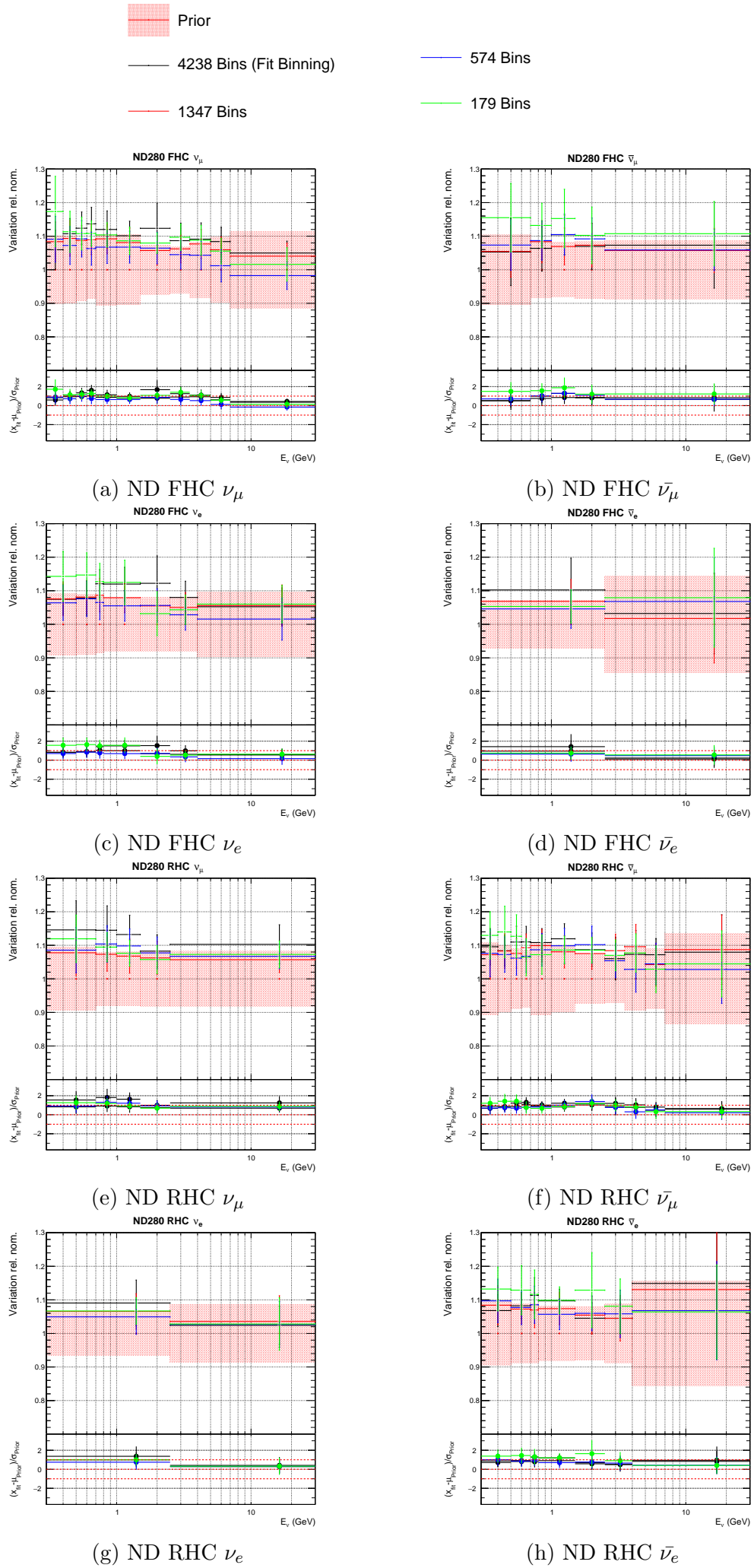
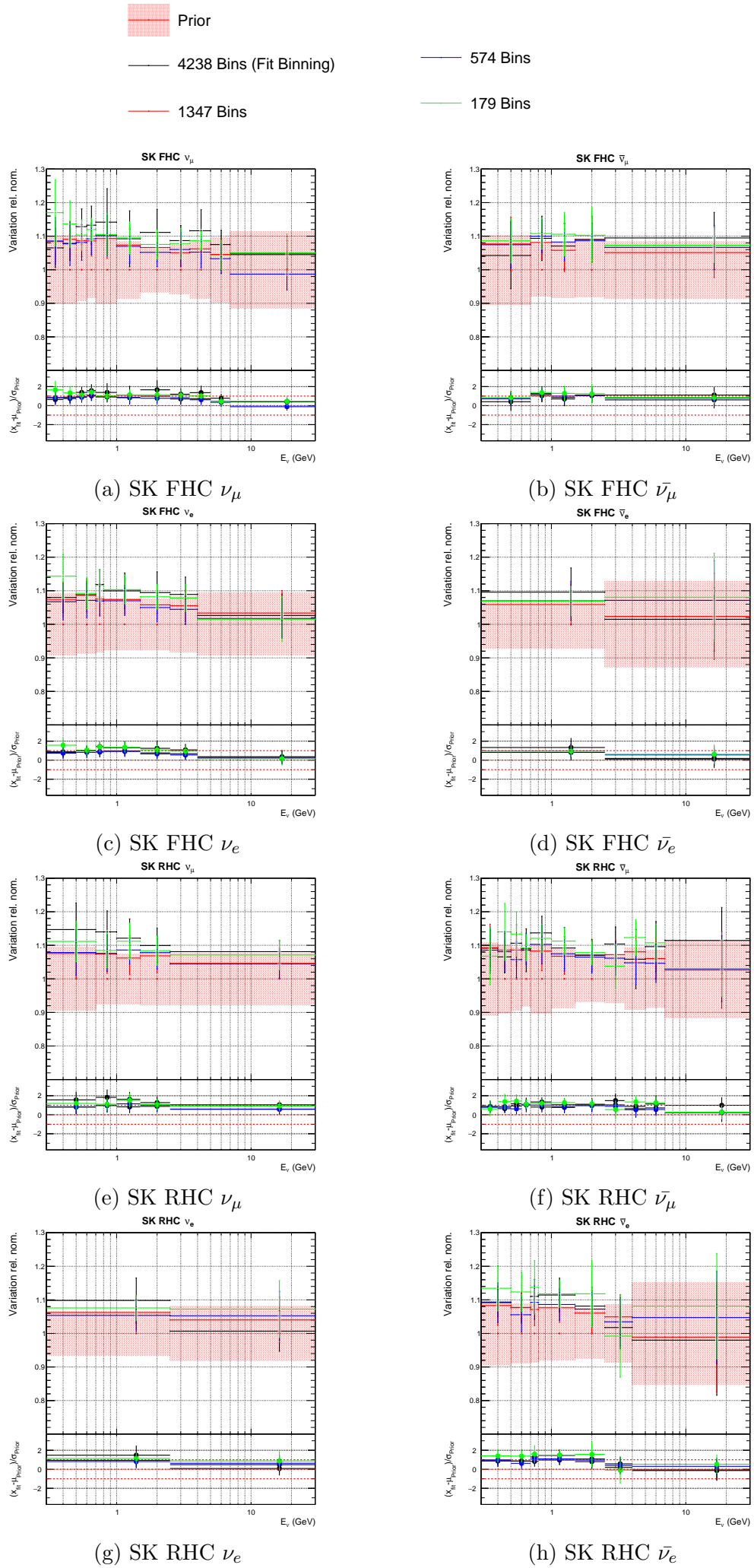Figure A.1.: ND280 flux parameters for fake data fits using different detector binnings.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$

Figure A.2.: SK flux parameters for fake data fits using different detector binnings.

(a) CC 0π

(b) CC 1π, ν_e, CC DIS, and CC coh.

(c) NC

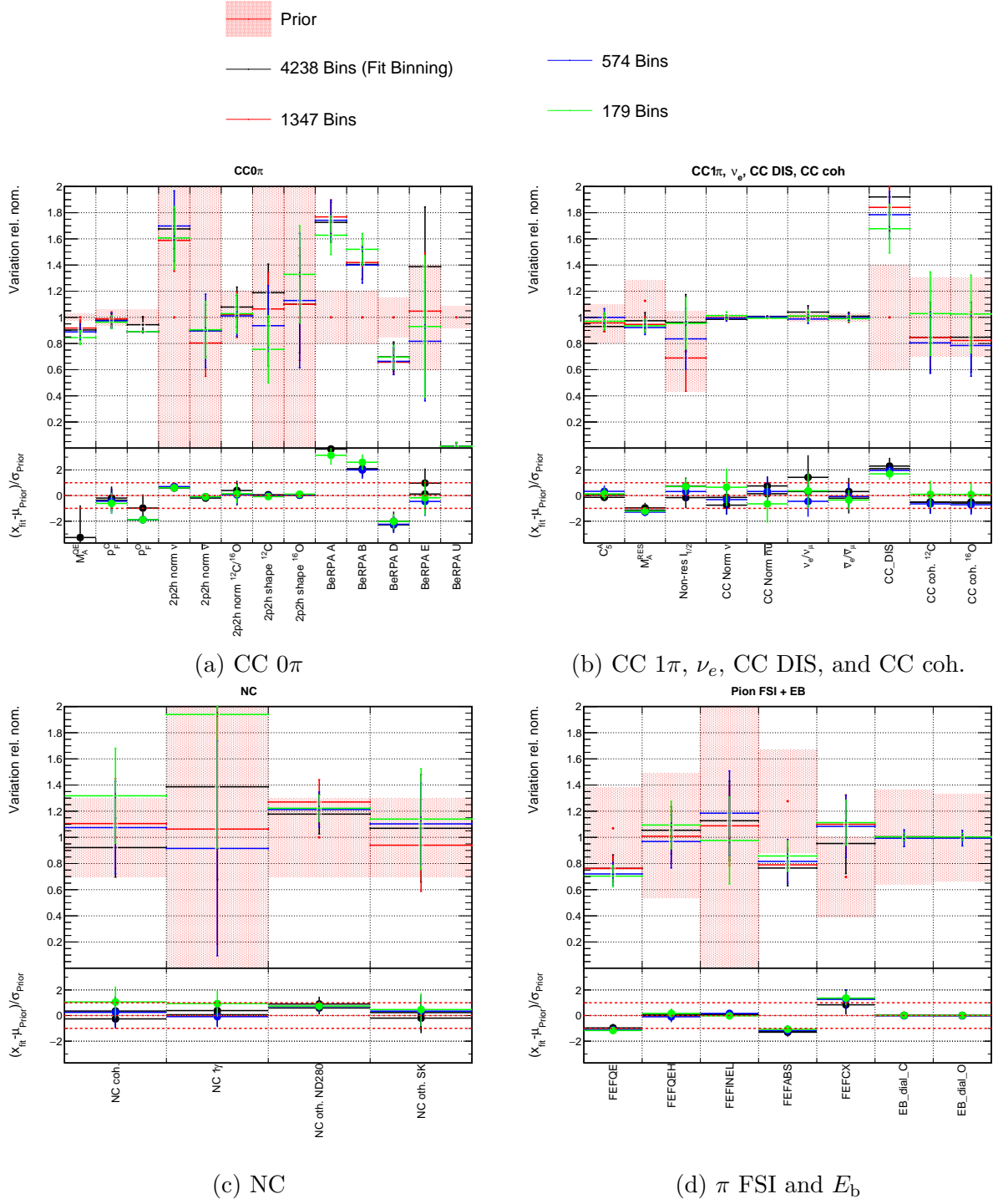(d) π FSI and E_b

Figure A.3.: Interaction parameters for fake data fits using different detector binnings.

B. Selection Binning

Runs 2–8 data and MC were used to tune the algorithm for producing the non-uniform rectangular sample binning. The distribution of events binned using the original uniform binning is shown for all samples in Figure B.1. The distribution of events binned using the final non-uniform rectangular binning is shown for all samples in Figure B.2. For the non-uniform rectangular binning, the bins containing the largest amount of events are in the peak regions at high angle and ~500 MeV.

Templates for the non-uniform fit binning for each sample are presented in this section, as it is not feasible to express the bin edges in text for non-uniform binning. Figure B.3 shows the x-axis range reduced to 0–5000 MeV so that the smaller bins in the peak can be seen. Figure B.4 shows the full distributions out to 30 GeV, for each sample.

The uniform binning used to validate against the other near detector fitting group is as follows:

- **FHC ν_μ CC 0π:**

  $\boldsymbol{p_\mu}$ **(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 θ_μ: -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.

- **FHC ν_μ CC 1π:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0, 300, 350, 400, 500, 600, 650, 700, 750, 800, 900, 1000, 1100, 1200, 1500, 2000, 3000, 5000, 30000.

  cos θ_μ: -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.

- **FHC ν_μ CC Other:**

  $\boldsymbol{p_\mu}$ **(MeV/c):** 0, 300, 400, 500, 600, 650, 700, 750, 800, 900, 1000, 1100, 1250, 1500, 1750, 2000, 3000, 5000, 30000.
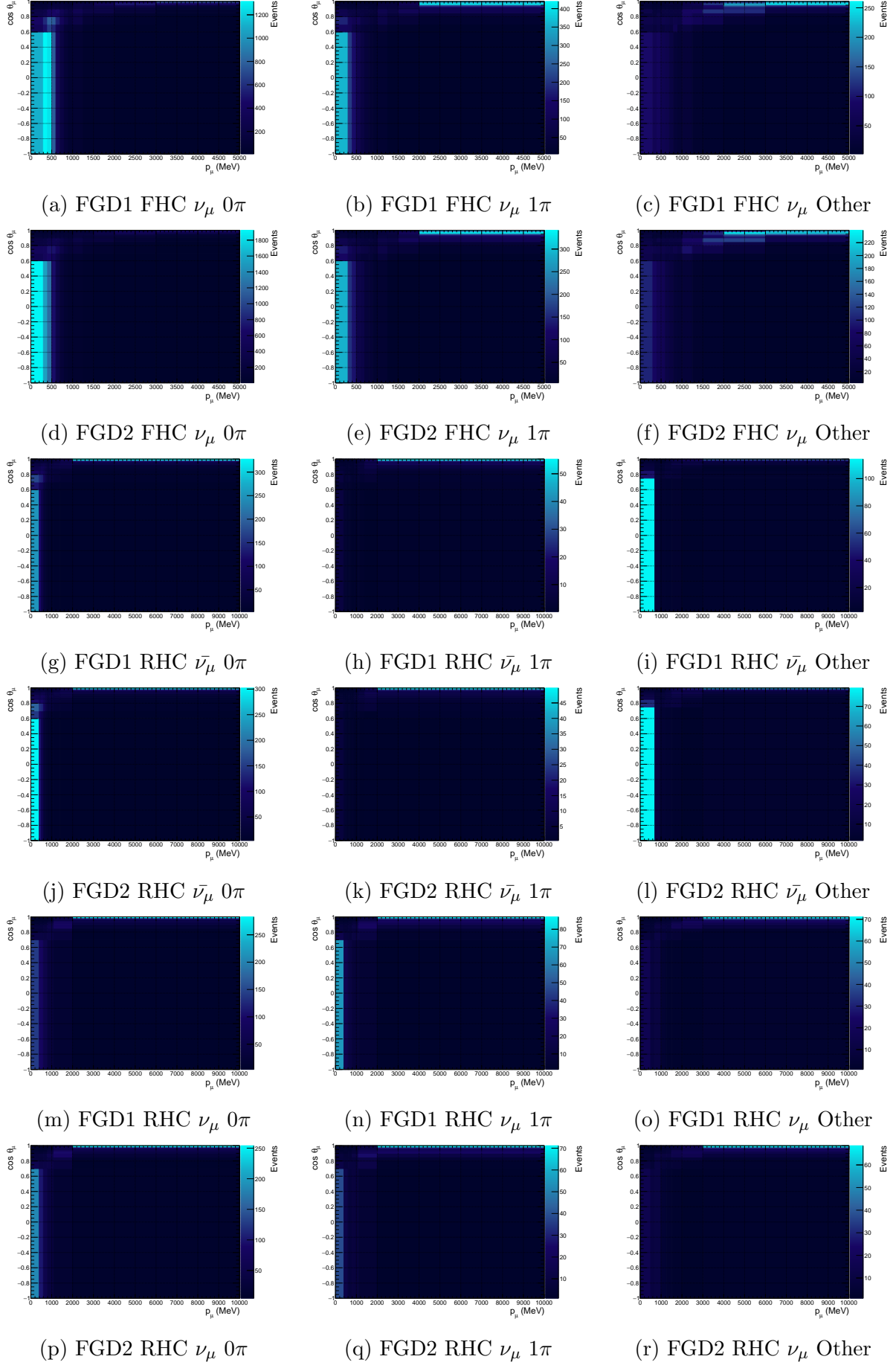
211

(a) FGD1 FHC ν_μ 0π (b) FGD1 FHC ν_μ 1π (c) FGD1 FHC ν_μ Other

(d) FGD2 FHC ν_μ 0π (e) FGD2 FHC ν_μ 1π (f) FGD2 FHC ν_μ Other

(g) FGD1 RHC $\bar{\nu}_\mu$ 0π (h) FGD1 RHC $\bar{\nu}_\mu$ 1π (i) FGD1 RHC $\bar{\nu}_\mu$ Other

(j) FGD2 RHC $\bar{\nu}_\mu$ 0π (k) FGD2 RHC $\bar{\nu}_\mu$ 1π (l) FGD2 RHC $\bar{\nu}_\mu$ Other

(m) FGD1 RHC ν_μ 0π (n) FGD1 RHC ν_μ 1π (o) FGD1 RHC ν_μ Other

(p) FGD2 RHC ν_μ 0π (q) FGD2 RHC ν_μ 1π (r) FGD2 RHC ν_μ Other

Figure B.1.: Uniform rectangular binning of MC events for T2K runs 2–8.

(a) FGD1 FHC ν_μ 0π (b) FGD1 FHC ν_μ 1π (c) FGD1 FHC ν_μ Other

(d) FGD2 FHC ν_μ 0π (e) FGD2 FHC ν_μ 1π (f) FGD2 FHC ν_μ Other

(g) FGD1 RHC $\bar{\nu}_\mu$ 0π (h) FGD1 RHC $\bar{\nu}_\mu$ 1π (i) FGD1 RHC $\bar{\nu}_\mu$ Other

(j) FGD2 RHC $\bar{\nu}_\mu$ 0π (k) FGD2 RHC $\bar{\nu}_\mu$ 1π (l) FGD2 RHC $\bar{\nu}_\mu$ Other

(m) FGD1 RHC ν_μ 0π (n) FGD1 RHC ν_μ 1π (o) FGD1 RHC ν_μ Other

(p) FGD2 RHC ν_μ 0π (q) FGD2 RHC ν_μ 1π (r) FGD2 RHC ν_μ Other
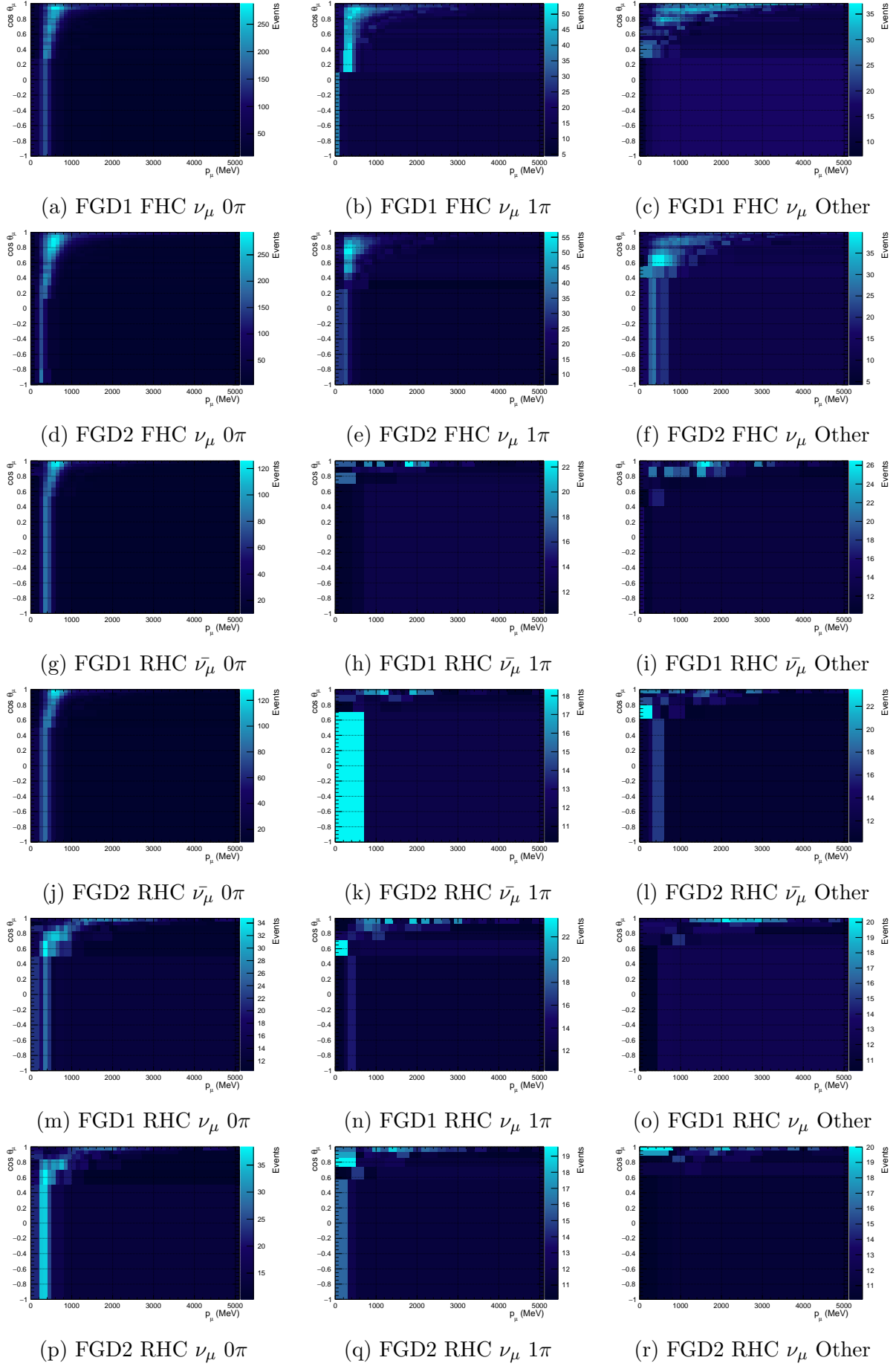
Figure B.2.: Non-uniform rectangular binning of MC events for T2K runs 2–8.

(a) FGD1 FHC ν_μ 0π (b) FGD1 FHC ν_μ 1π (c) FGD1 FHC ν_μ Other

(d) FGD2 FHC ν_μ 0π (e) FGD2 FHC ν_μ 1π (f) FGD2 FHC ν_μ Other
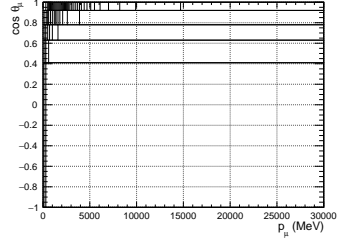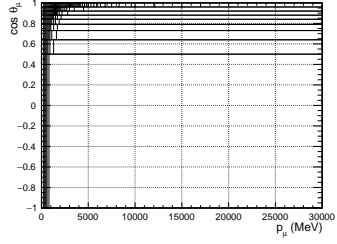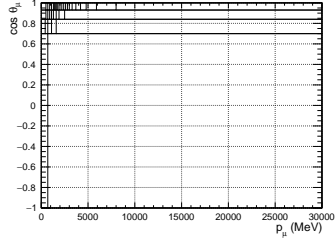
(g) FGD1 RHC $\bar{\nu}_\mu$ 0π (h) FGD1 RHC $\bar{\nu}_\mu$ 1π (i) FGD1 RHC $\bar{\nu}_\mu$ Other

(j) FGD2 RHC $\bar{\nu}_\mu$ 0π (k) FGD2 RHC $\bar{\nu}_\mu$ 1π (l) FGD2 RHC $\bar{\nu}_\mu$ Other

(m) FGD1 RHC ν_μ 0π (n) FGD1 RHC ν_μ 1π (o) FGD1 RHC ν_μ Other

(p) FGD2 RHC ν_μ 0π (q) FGD2 RHC ν_μ 1π (r) FGD2 RHC ν_μ Other
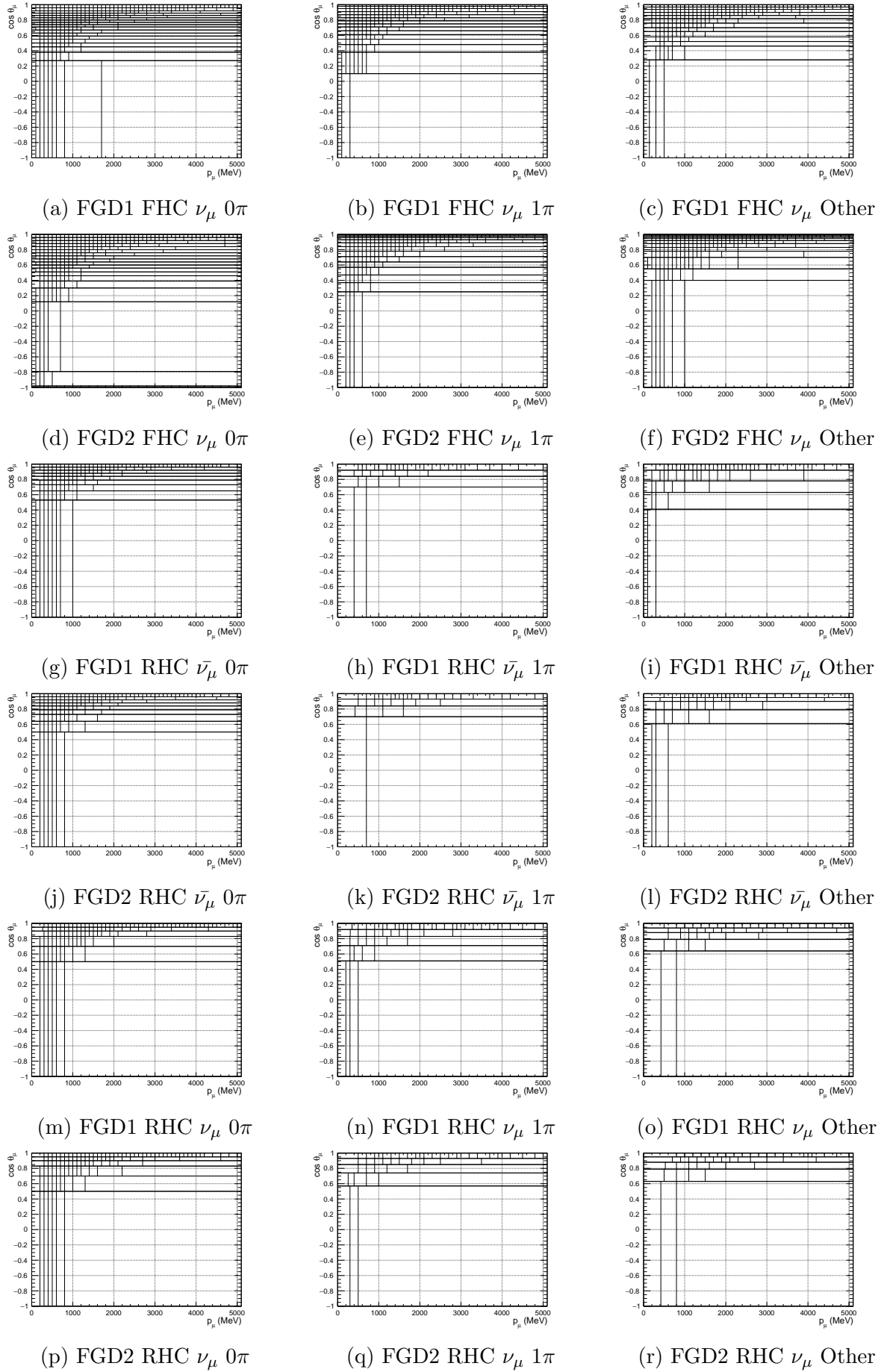
Figure B.3.: Non-uniform rectangular binning used in this analysis for each sample. The x-axis is reduced to better show the smaller bins at low momentum and high angle.
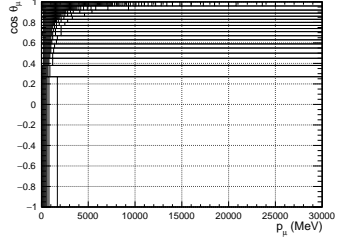
(a) FGD1 FHC ν_μ 0π (b) FGD1 FHC ν_μ 1π (c) FGD1 FHC ν_μ Other
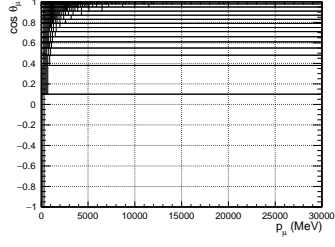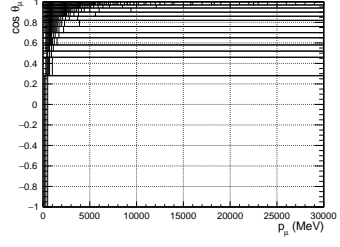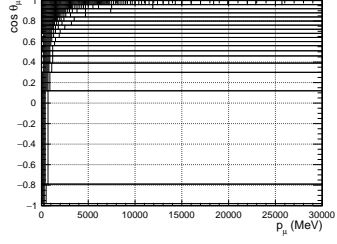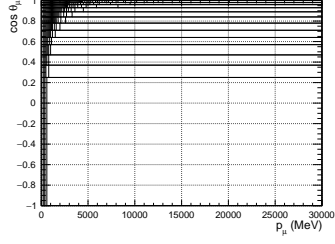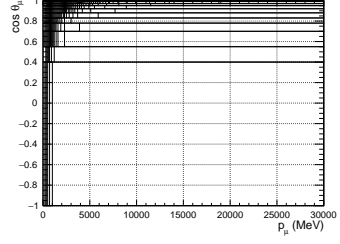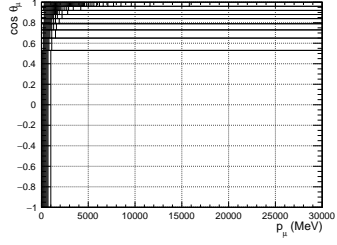
(d) FGD2 FHC ν_μ 0π (e) FGD2 FHC ν_μ 1π (f) FGD2 FHC ν_μ Other

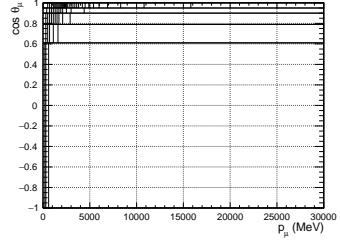(g) FGD1 RHC $\bar{\nu}_\mu$ 0π (h) FGD1 RHC $\bar{\nu}_\mu$ 1π (i) FGD1 RHC $\bar{\nu}_\mu$ Other
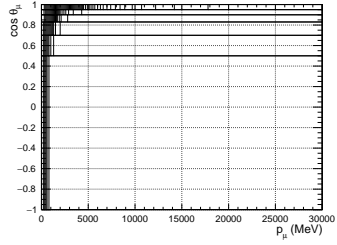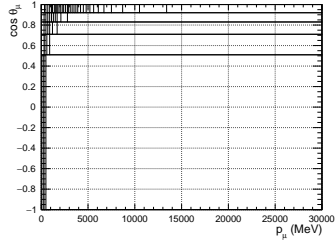
(j) FGD2 RHC $\bar{\nu}_\mu$ 0π (k) FGD2 RHC $\bar{\nu}_\mu$ 1π (l) FGD2 RHC $\bar{\nu}_\mu$ Other
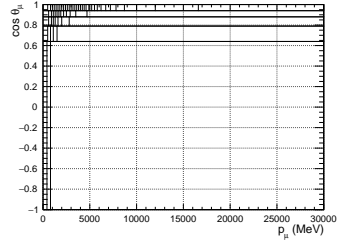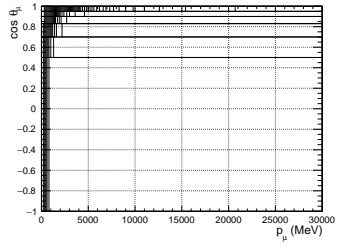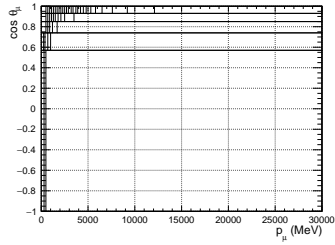
(m) FGD1 RHC ν_μ 0π (n) FGD1 RHC ν_μ 1π (o) FGD1 RHC ν_μ Other
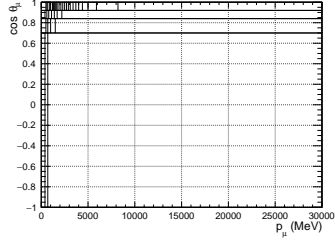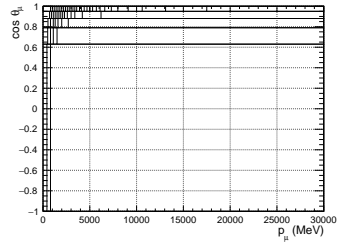
(p) FGD2 RHC ν_μ 0π (q) FGD2 RHC ν_μ 1π (r) FGD2 RHC ν_μ Other

Figure B.4.: Non-uniform rectangular binning used in this analysis for each sample.

215

cos θ_μ: -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.

- **RHC $\bar{\nu}_\mu$ CC 0π:**

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

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

- **RHC $\bar{\nu}_\mu$ CC 1π:**

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

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

- **RHC $\bar{\nu}_\mu$ CC Other: p_μ (MeV/c):** 0, 600, 800, 1000, 1250, 1500, 2000, 4000, 30000.

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

- **RHC ν_μ CC 0π:**

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

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

- **RHC ν_μ CC 1π:**

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

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

- **RHC ν_μ CC Other:**

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

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

C. Full Prefit Correlation Matrices

The full flux correlation matrix for all samples is shown in Figure C.1.
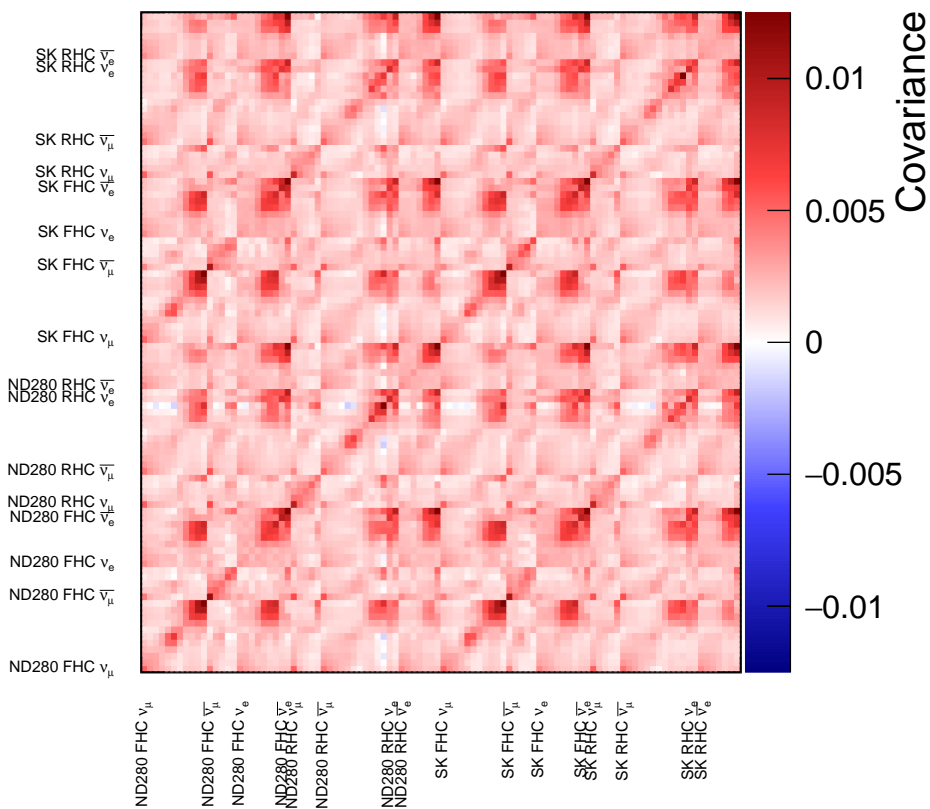

Figure C.1.: The full flux covariance matrix for all samples.

The full 574-bin detector correlation matrix for all samples is shown in Figure C.2. The bunching of the RHC samples at the end of each axis shows the detector bin merging process targeted bins and samples with fewer events, keeping more bins the well populated FHC samples.

The full 3071-bin detector correlation matrix for all samples is shown in Figure C.3. Again, the bunching of the RHC samples at the end of each axis shows the detector bin merging process targeted bins and samples with fewer events, keeping more bins in the well populated FHC samples.
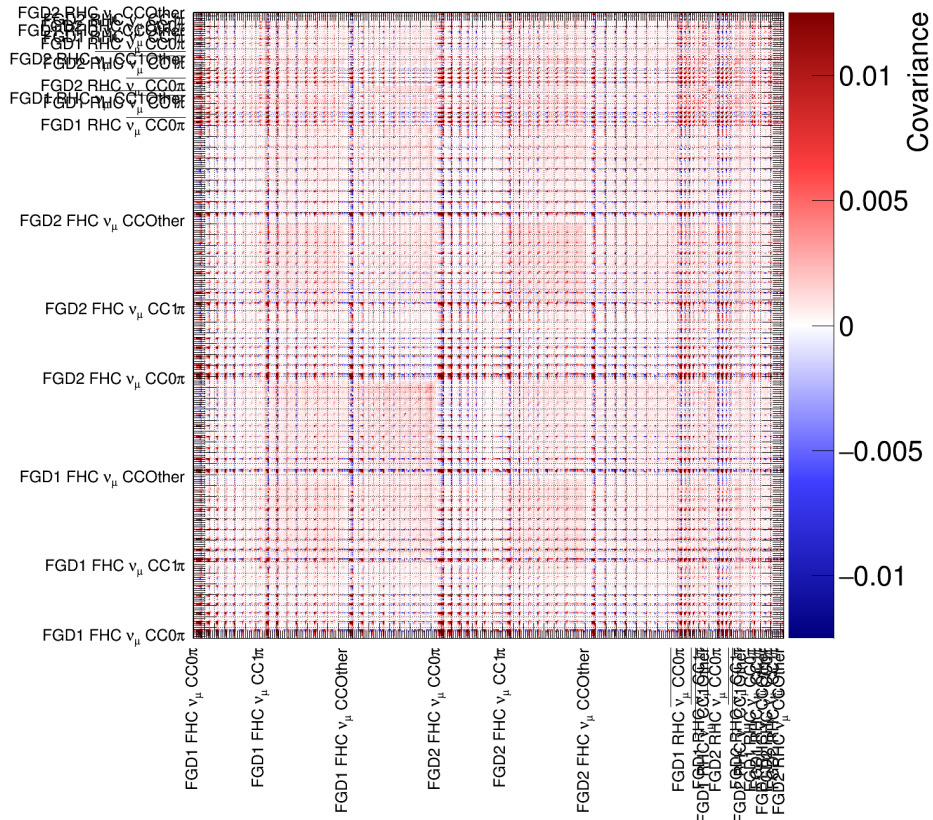
Figure C.2.: The ND280 detector covariance matrix for all samples with 574 merged bins, produced using runs 2–9 MC.
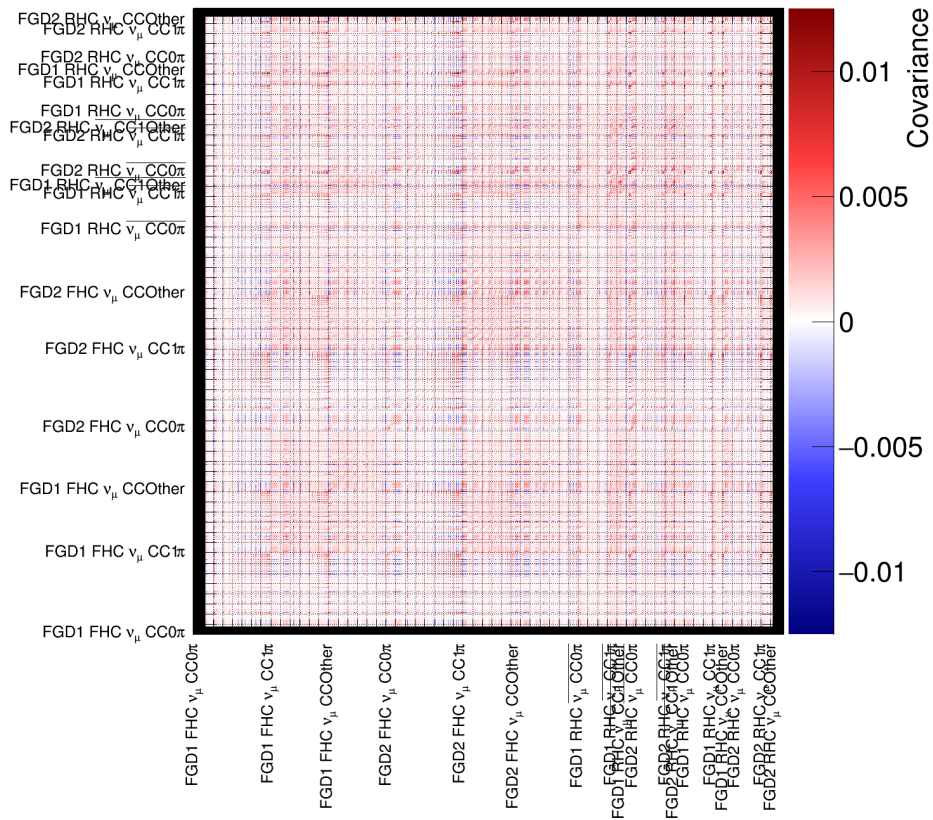

Figure C.3.: The ND280 detector covariance matrix for all samples with the full 3071-bin non-uniform fit binning, produced using runs 2–9 MC.

D. Nominal MC Distributions

The 2D nominal, uniformly-binned runs 2–9 MC distributions are shown in Figure D.1.

The projection of the non-uniformly binned nominal MC distributions onto the p_μ axis are shown in Figures D.2, D.3, and D.4, along with the interaction mode breakdown and data.

The projection of the non-uniformly binned nominal MC distributions onto the $\cos\theta_\mu$ axis are shown in Figures D.5, D.6, and D.7, along with the interaction mode breakdown and data.

(a) FGD1 FHC ν_μ 0π

(b) FGD1 FHC ν_μ 1π

(c) FGD1 FHC ν_μ Other

(d) FGD2 FHC ν_μ 0π

(e) FGD2 FHC ν_μ 1π

(f) FGD2 FHC ν_μ Other

(g) FGD1 RHC $\bar{\nu}_\mu$ 0π

(h) FGD1 RHC $\bar{\nu}_\mu$ 1π

(i) FGD1 RHC $\bar{\nu}_\mu$ Other

(j) FGD2 RHC $\bar{\nu}_\mu$ 0π

(k) FGD2 RHC $\bar{\nu}_\mu$ 1π

(l) FGD2 RHC $\bar{\nu}_\mu$ Other

(m) FGD1 RHC ν_μ 0π

(n) FGD1 RHC ν_μ 1π

(o) FGD1 RHC ν_μ Other

(p) FGD2 RHC ν_μ 0π

(q) FGD2 RHC ν_μ 1π

(r) FGD2 RHC ν_μ Other
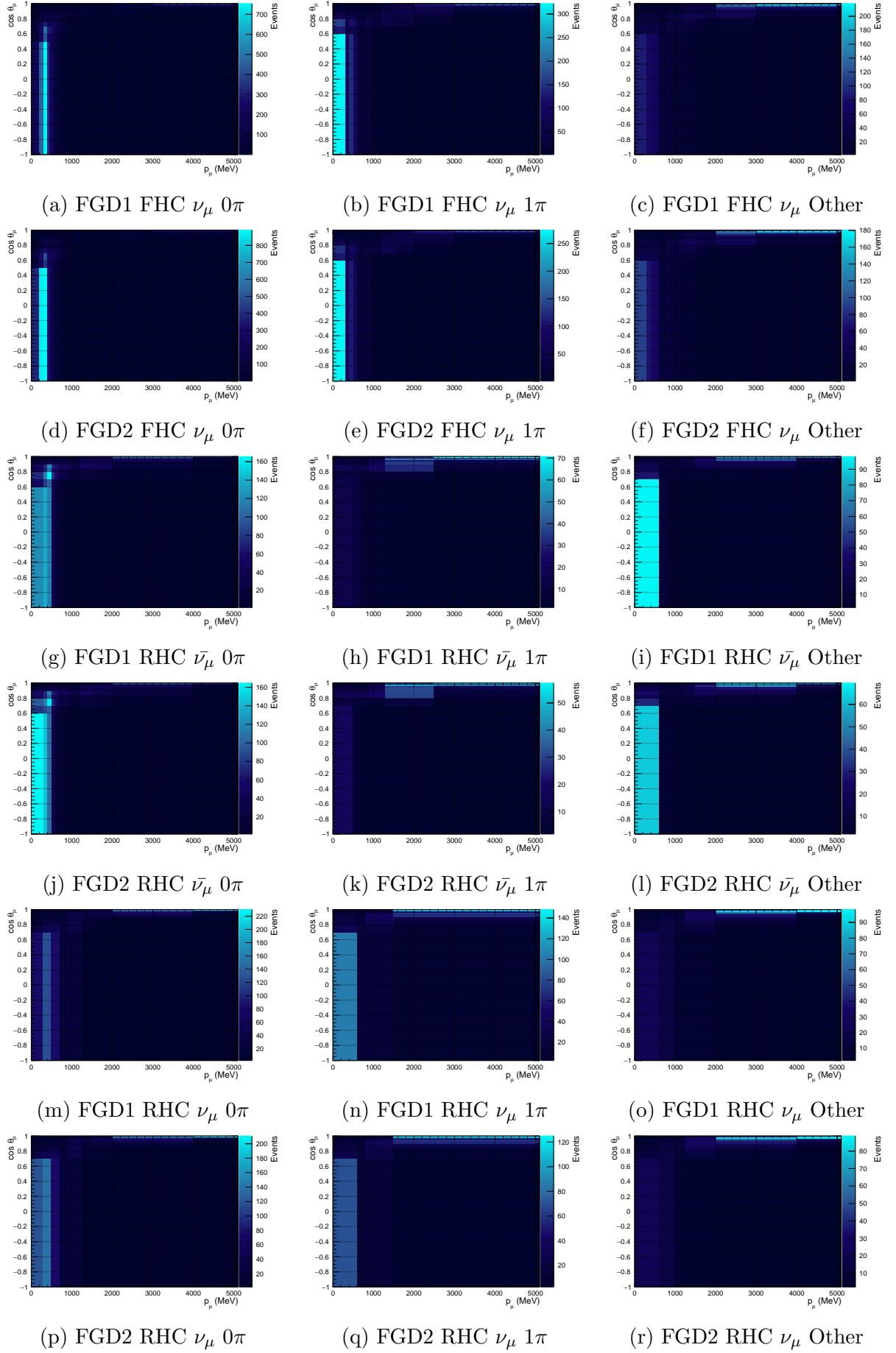
Figure D.1.: p_μ–cosθ_μ distributions for the nominal MC binned uniformly.
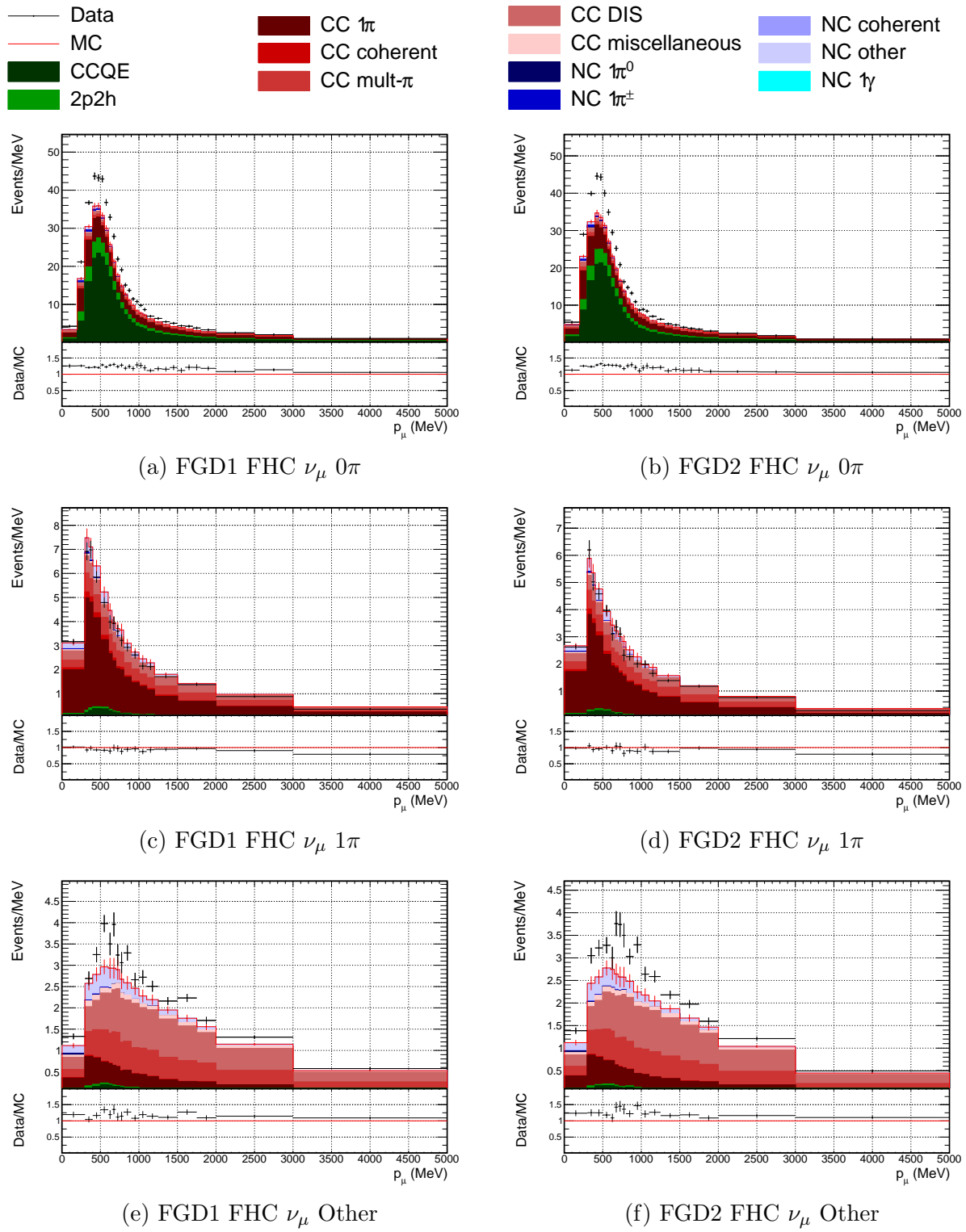
Legend:
- Data
- MC
- CCQE
- 2p2h
- CC 1π
- CC coherent
- CC mult-π
- CC DIS
- CC miscellaneous
- NC 1π⁰
- NC 1π±
- NC coherent
- NC other
- NC 1γ

(a) FGD1 FHC ν_μ 0π

(b) FGD2 FHC ν_μ 0π

(c) FGD1 FHC ν_μ 1π

(d) FGD2 FHC ν_μ 1π

(e) FGD1 FHC ν_μ Other

(f) FGD2 FHC ν_μ Other

Figure D.2.: p_μ projections of data and nominal MC broken down by interaction mode for FHC selections.

(a) FGD1 RHC $\bar{\nu}_\mu$ 0π

(b) FGD2 RHC $\bar{\nu}_\mu$ 0π

(c) FGD1 RHC $\bar{\nu}_\mu$ 1π

(d) FGD2 RHC $\bar{\nu}_\mu$ 1π

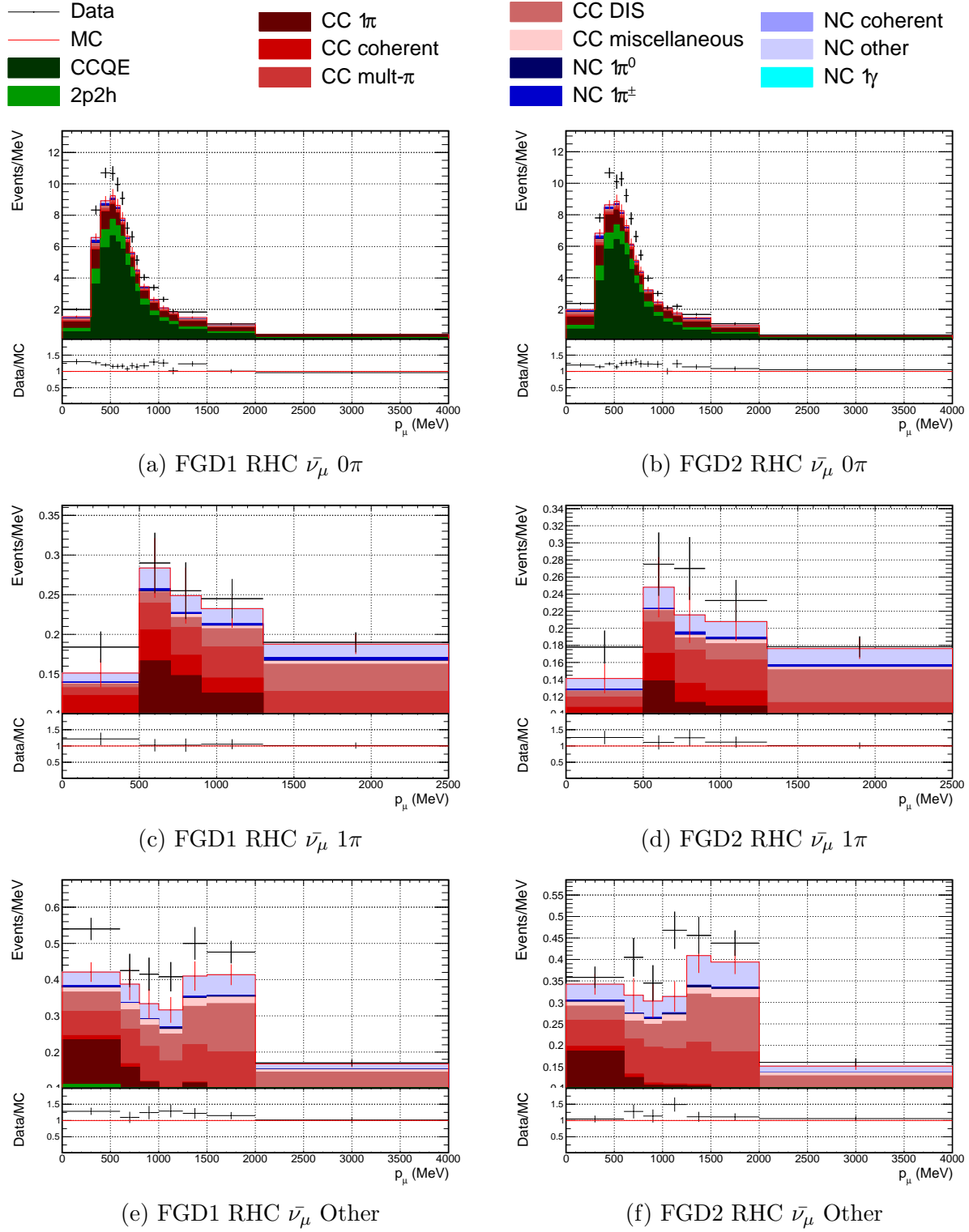(e) FGD1 RHC $\bar{\nu}_\mu$ Other

(f) FGD2 RHC $\bar{\nu}_\mu$ Other

Figure D.3.: p_μ projections of data and nominal MC broken down by interaction mode for RHC $\bar{\nu}_\mu$ selections.
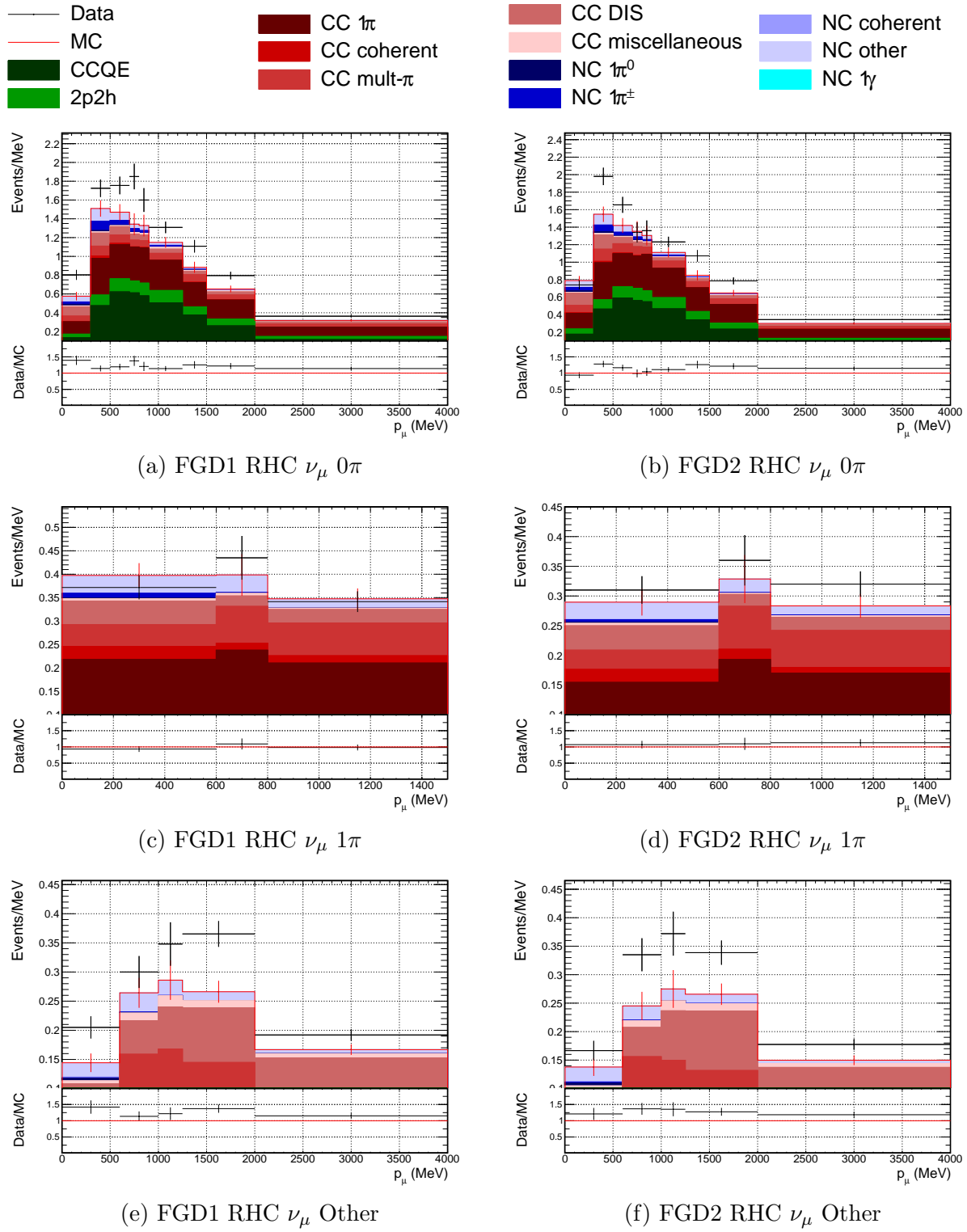
Legend:
- Data
- MC
- CCQE
- 2p2h
- CC 1π
- CC coherent
- CC mult-π
- CC DIS
- CC miscellaneous
- NC 1π⁰
- NC 1π±
- NC coherent
- NC other
- NC 1γ

(a) FGD1 RHC ν_μ 0π

(b) FGD2 RHC ν_μ 0π

(c) FGD1 RHC ν_μ 1π

(d) FGD2 RHC ν_μ 1π

(e) FGD1 RHC ν_μ Other

(f) FGD2 RHC ν_μ Other

Figure D.4.: p_μ projections of data and nominal MC broken down by interaction mode for RHC ν_μ selections.

(a) FGD1 FHC ν_μ 0π

(b) FGD2 FHC ν_μ 0π

(c) FGD1 FHC ν_μ 1π

(d) FGD2 FHC ν_μ 1π
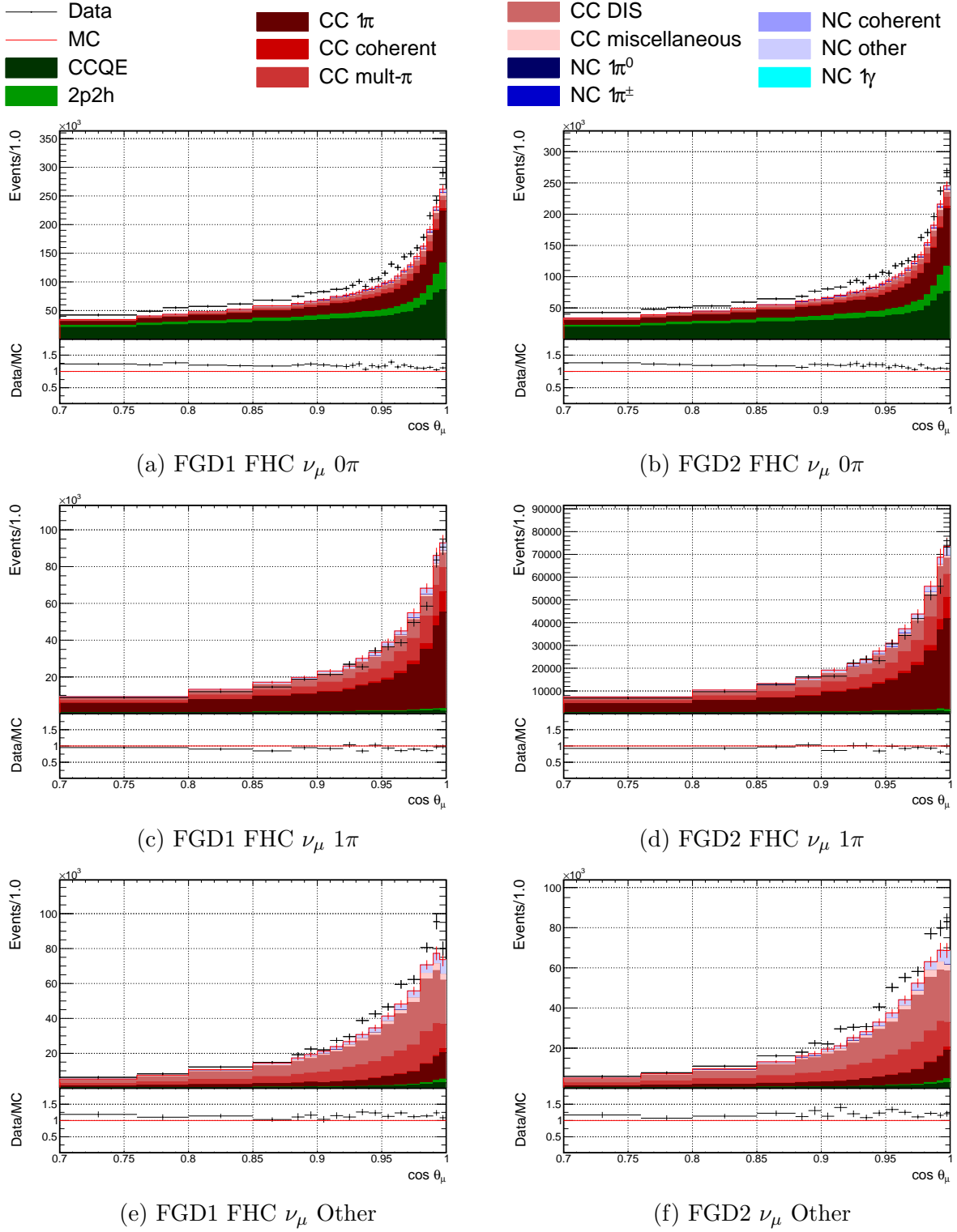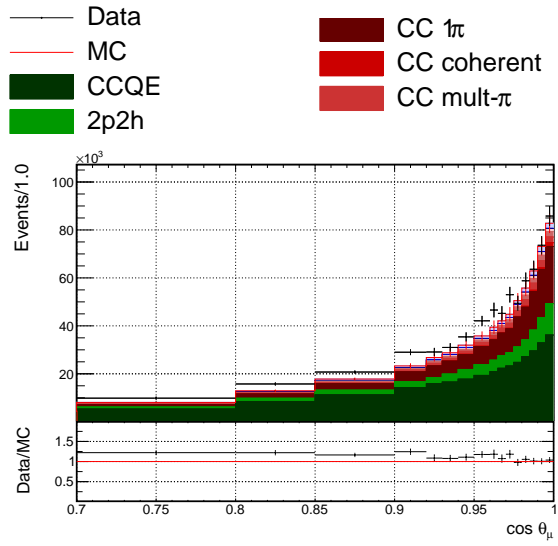
(e) FGD1 FHC ν_μ Other

(f) FGD2 ν_μ Other
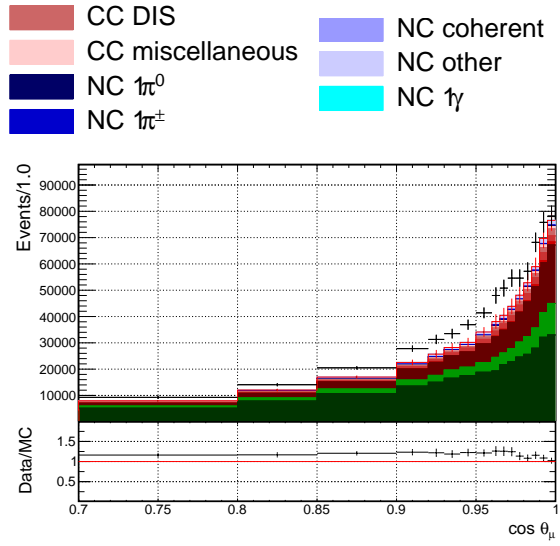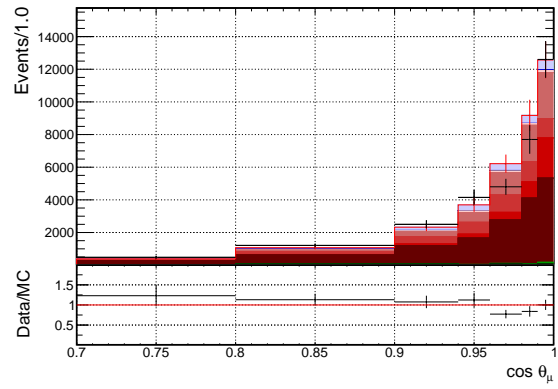
Figure D.5.: $\cos\theta_\mu$ projections of data and nominal MC broken down by interaction mode for FHC selections.

Data	CC 1π	CC DIS
MC	CC coherent	CC miscellaneous
CCQE	CC mult-π	NC 1π⁰
2p2h		NC 1π±

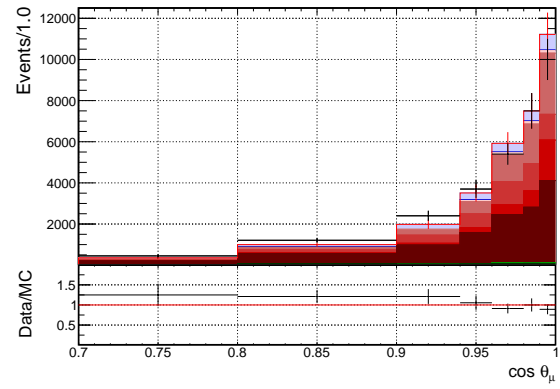|	
NC coherent	
NC other	
NC 1γ	

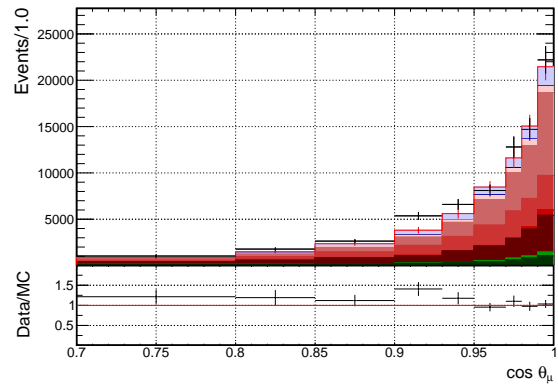(a) FGD1 RHC $\bar{\nu}_\mu$ 0π
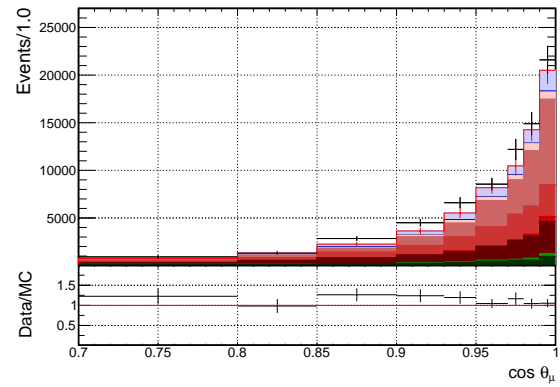
(b) FGD2 RHC $\bar{\nu}_\mu$ 0π

(c) FGD1 RHC $\bar{\nu}_\mu$ 1π

(d) FGD2 RHC $\bar{\nu}_\mu$ 1π

(e) FGD1 RHC $\bar{\nu}_\mu$ Other

(f) FGD2 RHC $\bar{\nu}_\mu$ Other

Figure D.6.: $\cos\theta_\mu$ projections of data and nominal MC broken down by interaction mode for RHC $\bar{\nu}_\mu$ selections.

(a) FGD1 RHC ν_μ 0π

(b) FGD2 RHC ν_μ 0π

(c) FGD1 RHC ν_μ 1π

(d) FGD2 RHC ν_μ 1π

(e) FGD1 RHC ν_μ Other
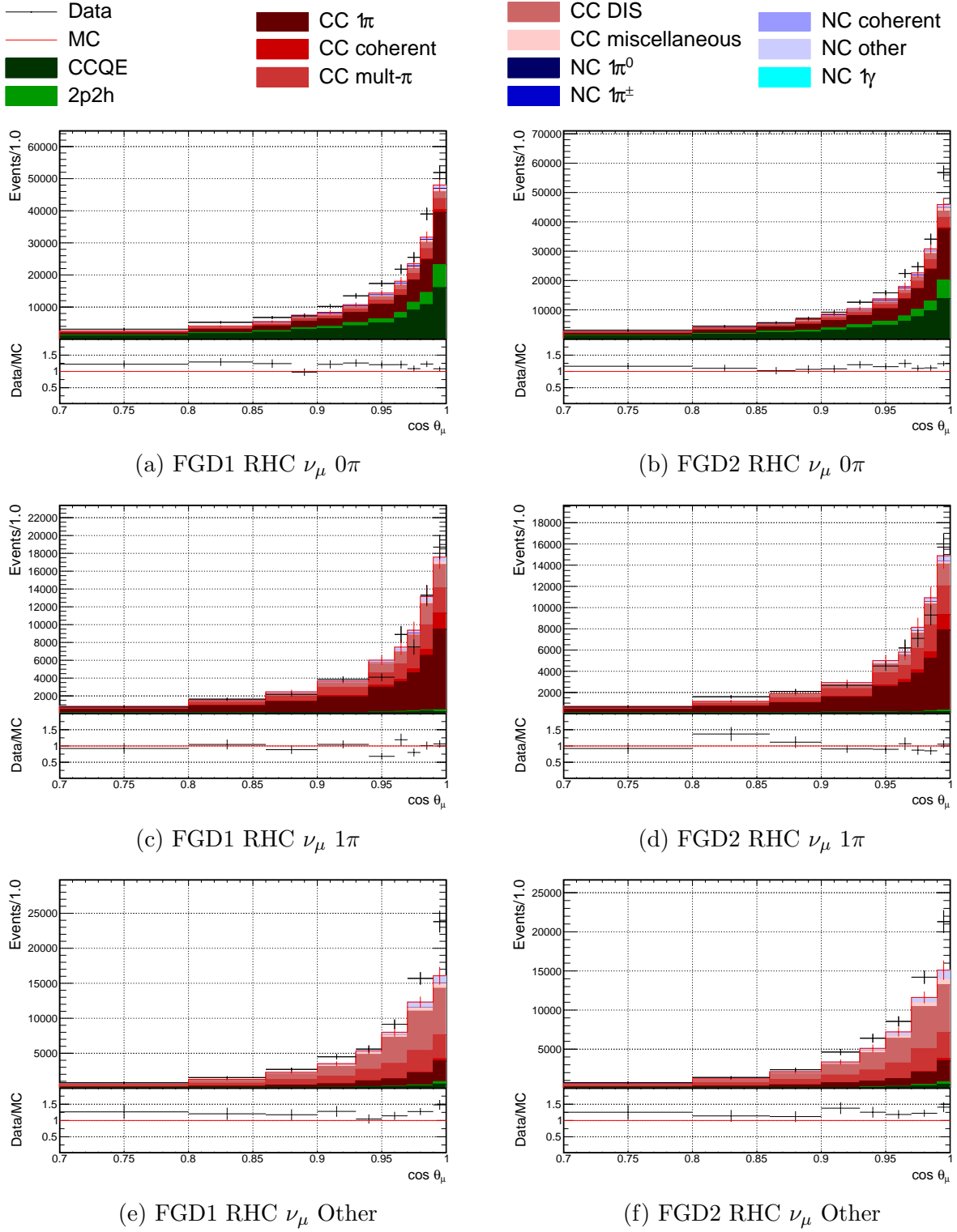
(f) FGD2 RHC ν_μ Other

Figure D.7.: $\cos\theta_\mu$ projections of data and nominal MC broken down by interaction mode for RHC ν_μ selections.

E. Prior Predictive Distributions

The p_μ and $\cos\theta_\mu$ projections of the prior predictive distributions are shown in Figures E.1–E.6.
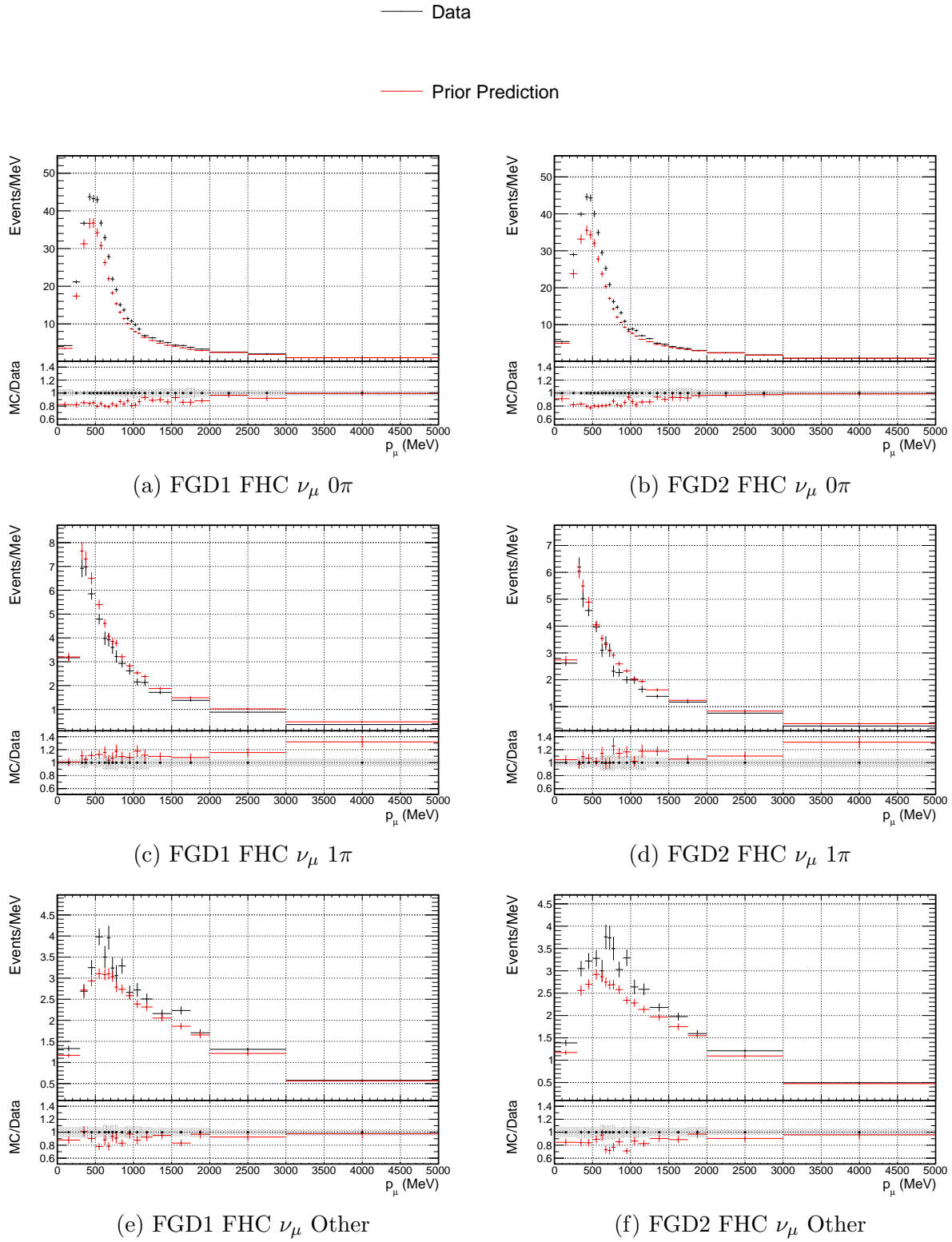
Figure E.1.: p_μ projections of the prior predictive distributions and data for FHC ν_μ selections.

(a) FGD1 RHC $\bar{\nu}_\mu$ 0π

(b) FGD2 RHC $\bar{\nu}_\mu$ 0π

(c) FGD1 RHC $\bar{\nu}_\mu$ 1π

(d) FGD2 RHC $\bar{\nu}_\mu$ 1π

(e) FGD1 RHC $\bar{\nu}_\mu$ Other

(f) FGD2 RHC $\bar{\nu}_\mu$ Other

Figure E.2.: p_μ projections of the prior predictive distributions and data for RHC $\bar{\nu}_\mu$ selections.

(a) FGD1 RHC ν_μ 0π

(b) FGD2 RHC ν_μ 0π

(c) FGD1 RHC ν_μ 1π

(d) FGD2 RHC ν_μ 1π

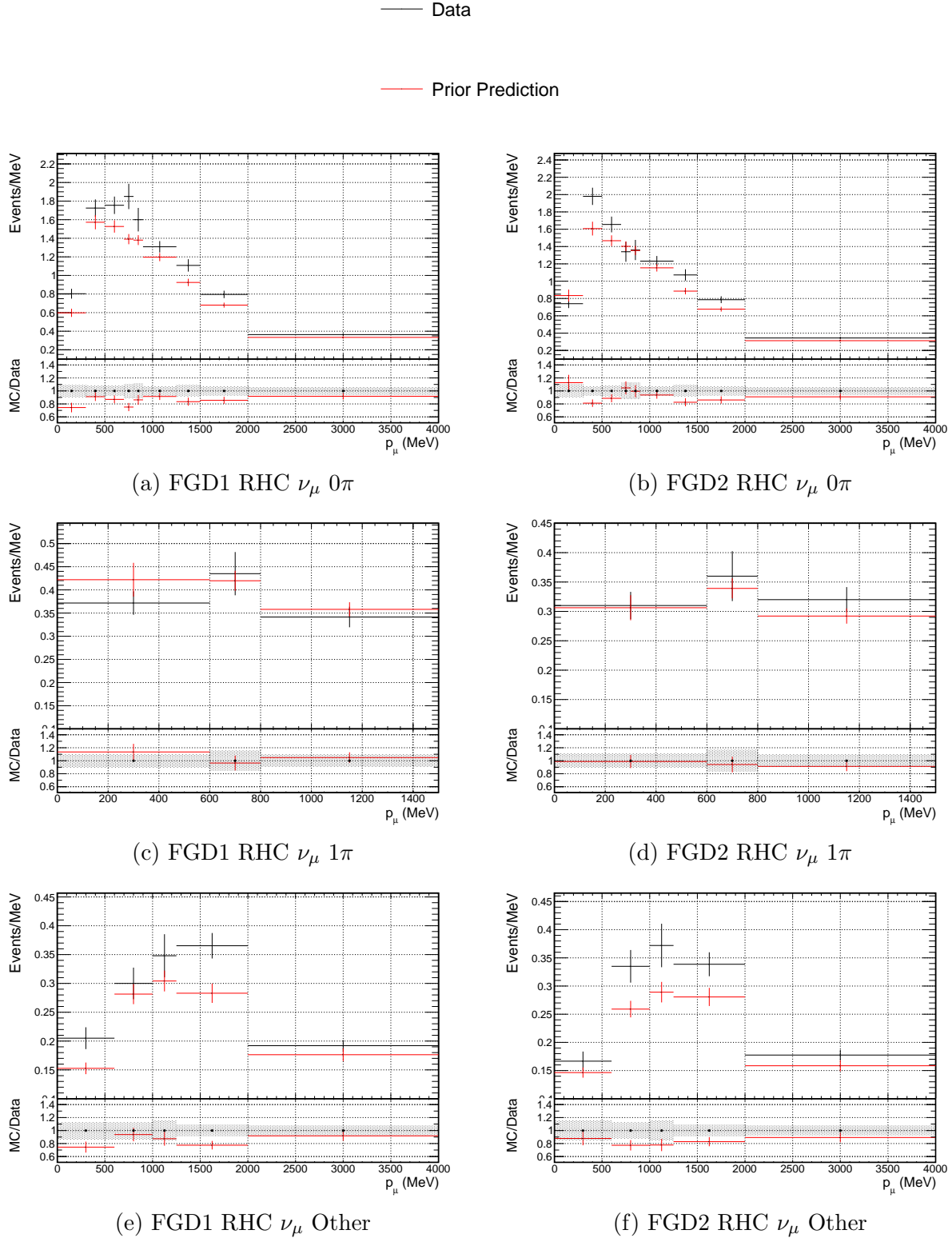(e) FGD1 RHC ν_μ Other

(f) FGD2 RHC ν_μ Other

Figure E.3.: p_μ projections of the prior predictive distributions and data for RHC ν_μ selections.
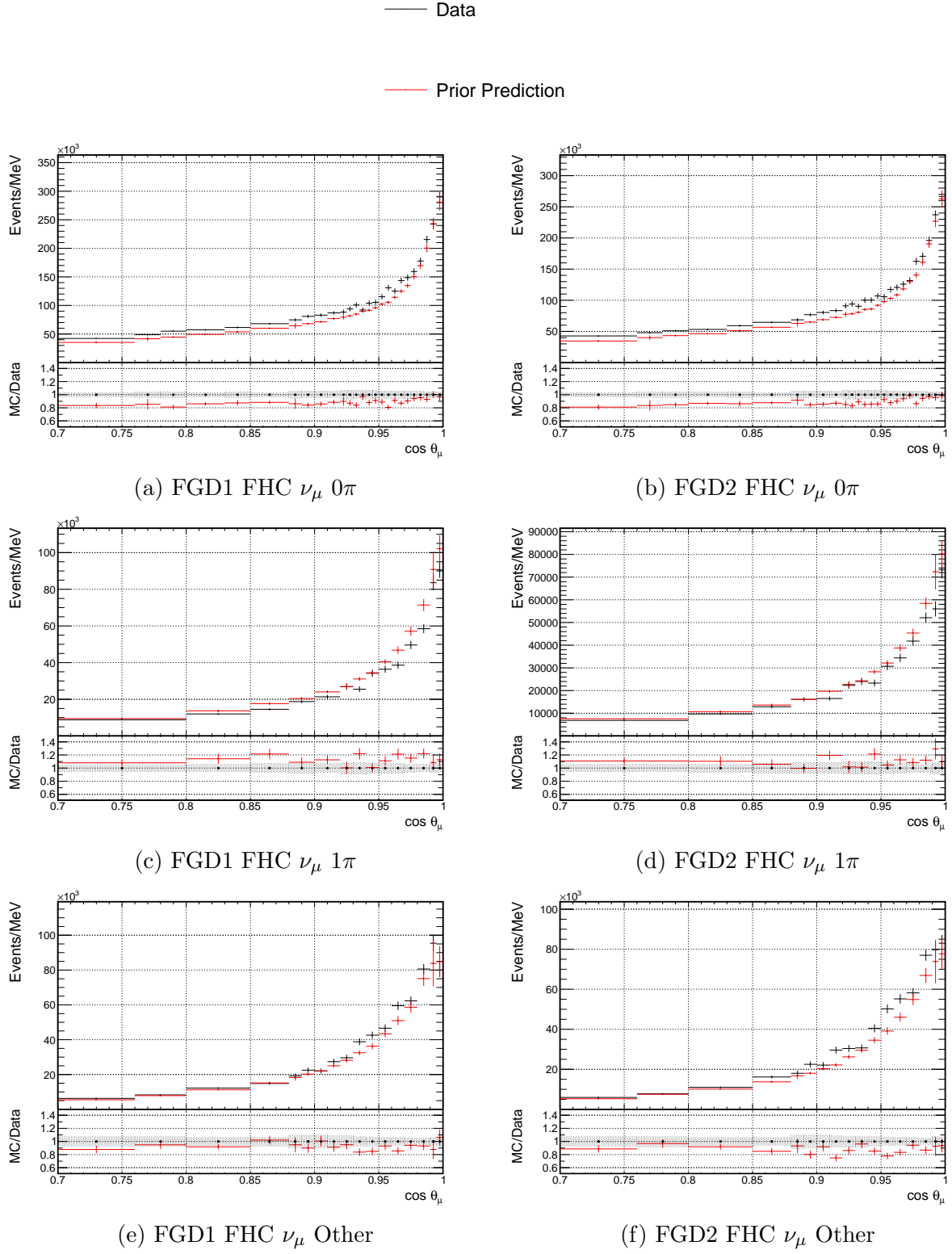
Figure E.4.: $\cos\theta_\mu$ projections of the prior predictive distributions and data for FHC ν_μ selections.

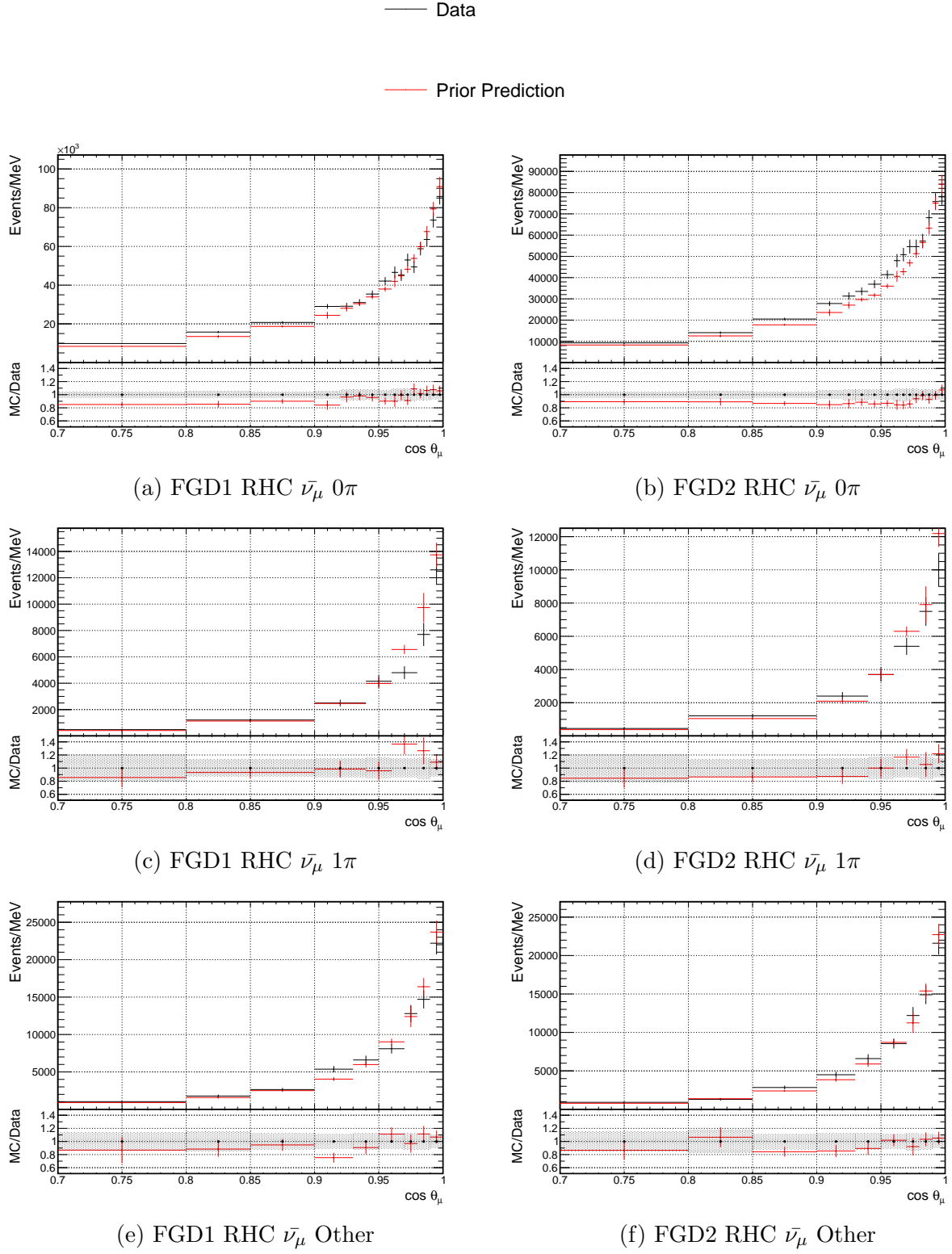Figure E.5.: $\cos\theta_\mu$ projections of the prior predictive distributions and data for RHC $\bar{\nu}_\mu$ selections.
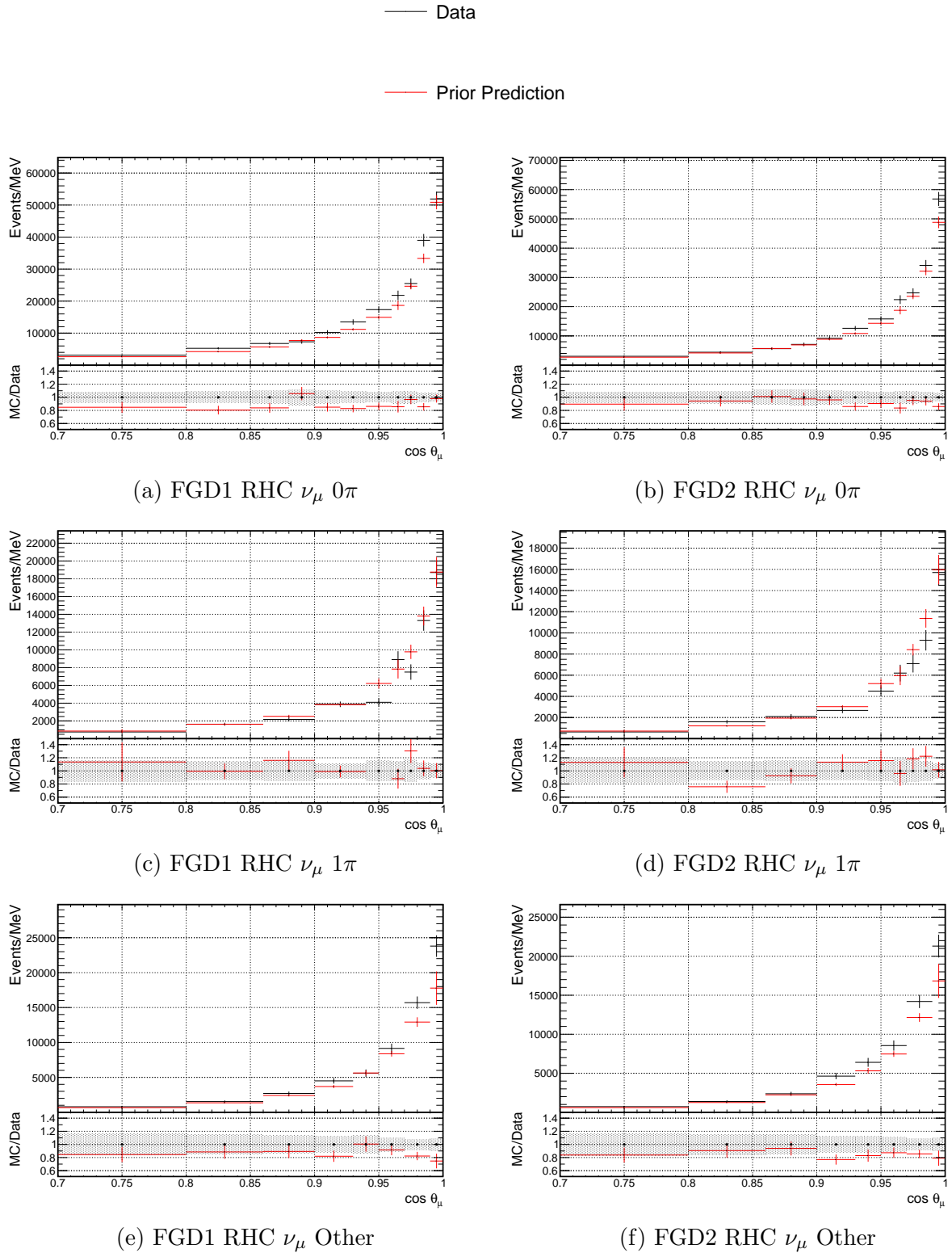
Figure E.6.: cosθ_μ projections of the prior predictive distributions and data for RHC ν_μ selections.

F. Full Fit Results

The results for the Asimov fit are shown in Figures F.1, F.2 and F.3.

The results for the data fit are shown in Figures F.4, F.5 and F.6.

The results for the Asimov fits comparing different fit and detector binnings are shown in Figures F.7, F.8 and F.9.

The results for the data fits comparing different fit and detector binnings are shown in Figures F.10, F.11 and F.12.

The joint near and far, and near detector only fits are compared in Figures F.13, F.14 and F.15.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$

Figure F.1.: ND280 flux parameters for the Asimov fit.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$

Figure F.2.: SK flux parameters for the Asimov fit.

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI

Figure F.3.: Interaction parameters for the Asimov fit.

Prior

Postfit

(a) ND FHC ν_μ
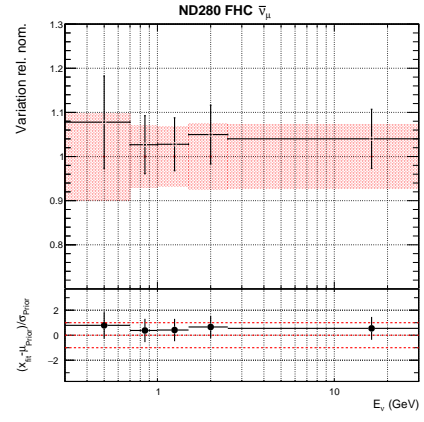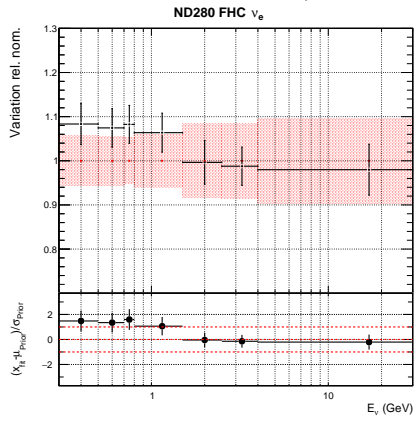
(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

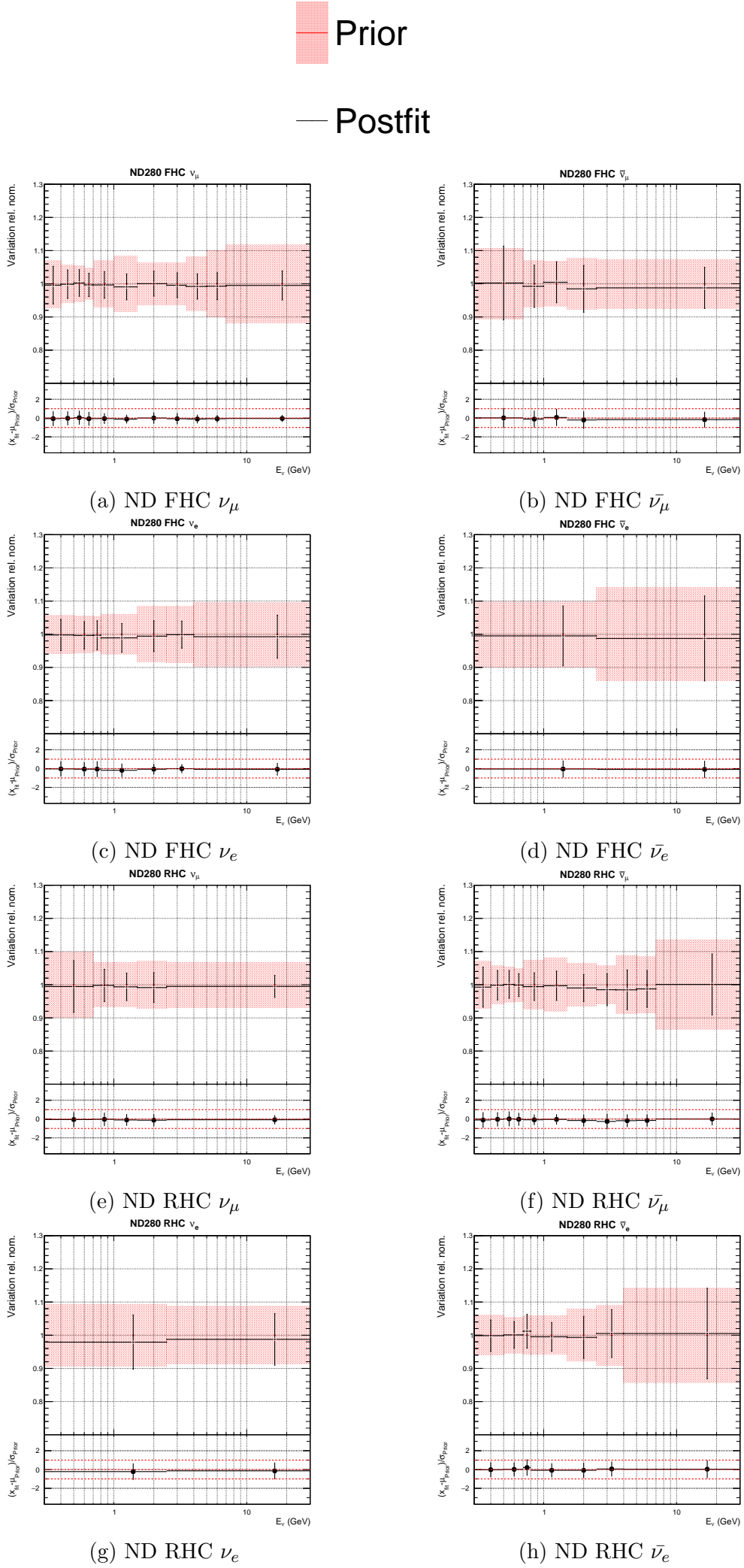(e) ND RHC ν_μ
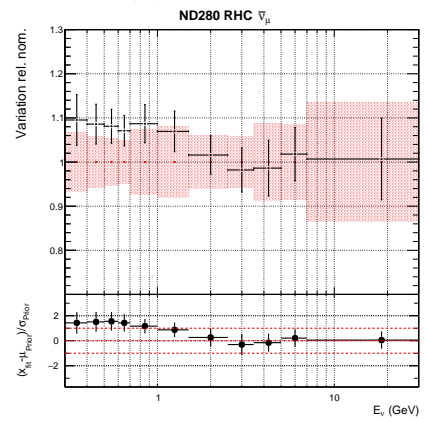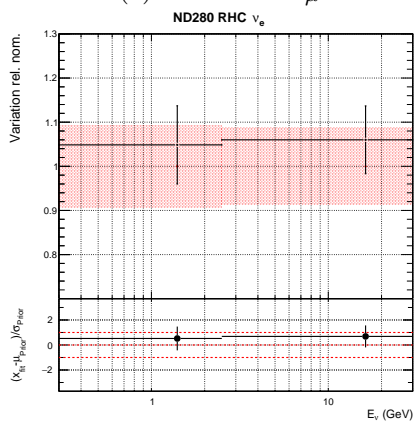
(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e
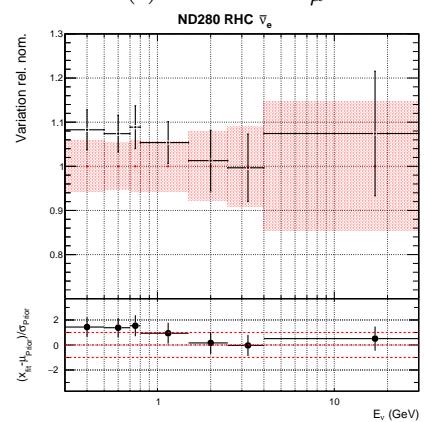
(h) ND RHC $\bar{\nu}_e$

Figure F.4.: ND280 flux parameters for the data fit.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
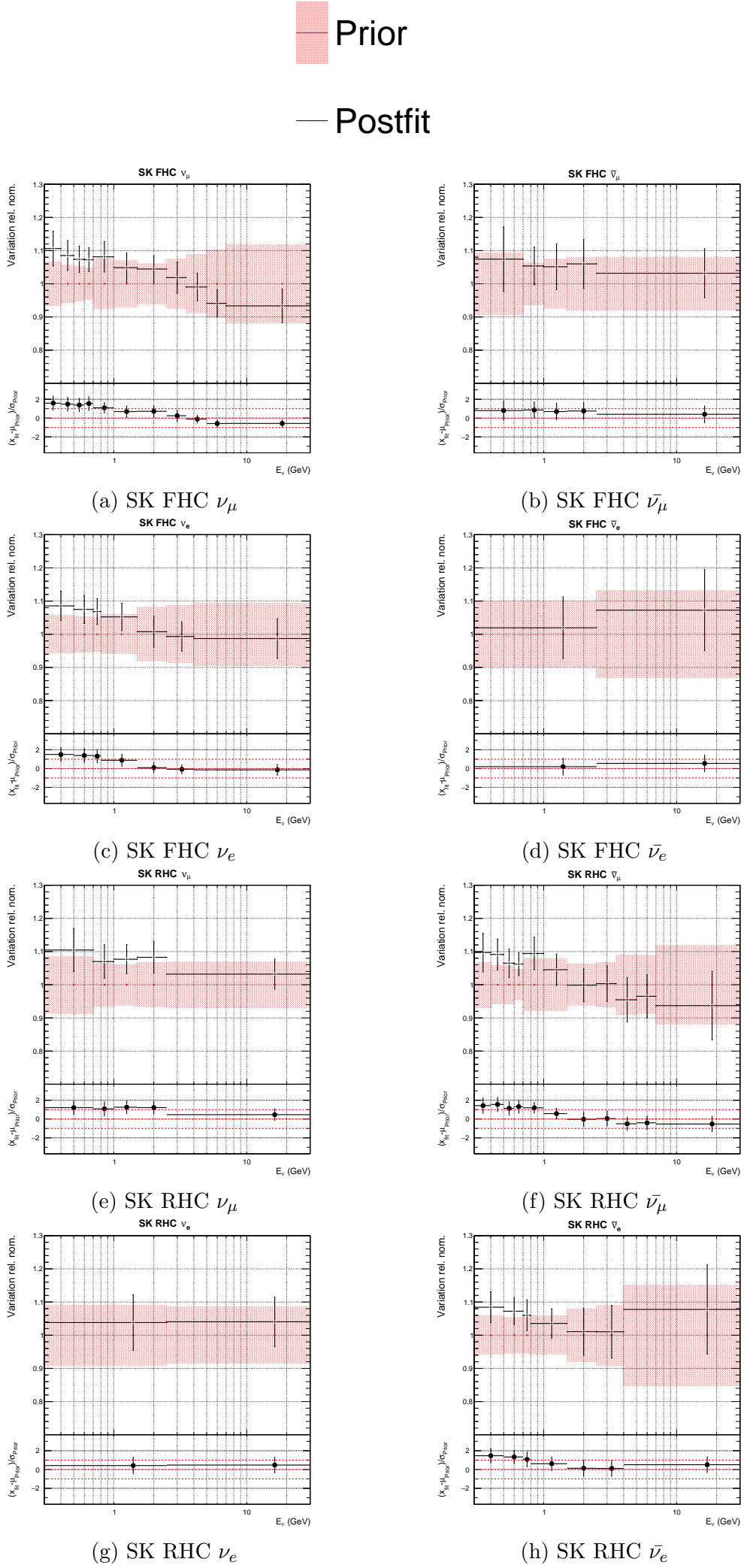
Figure F.5.: SK flux parameters for the data fit.

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC coh.

(d) NC

(e) π FSI
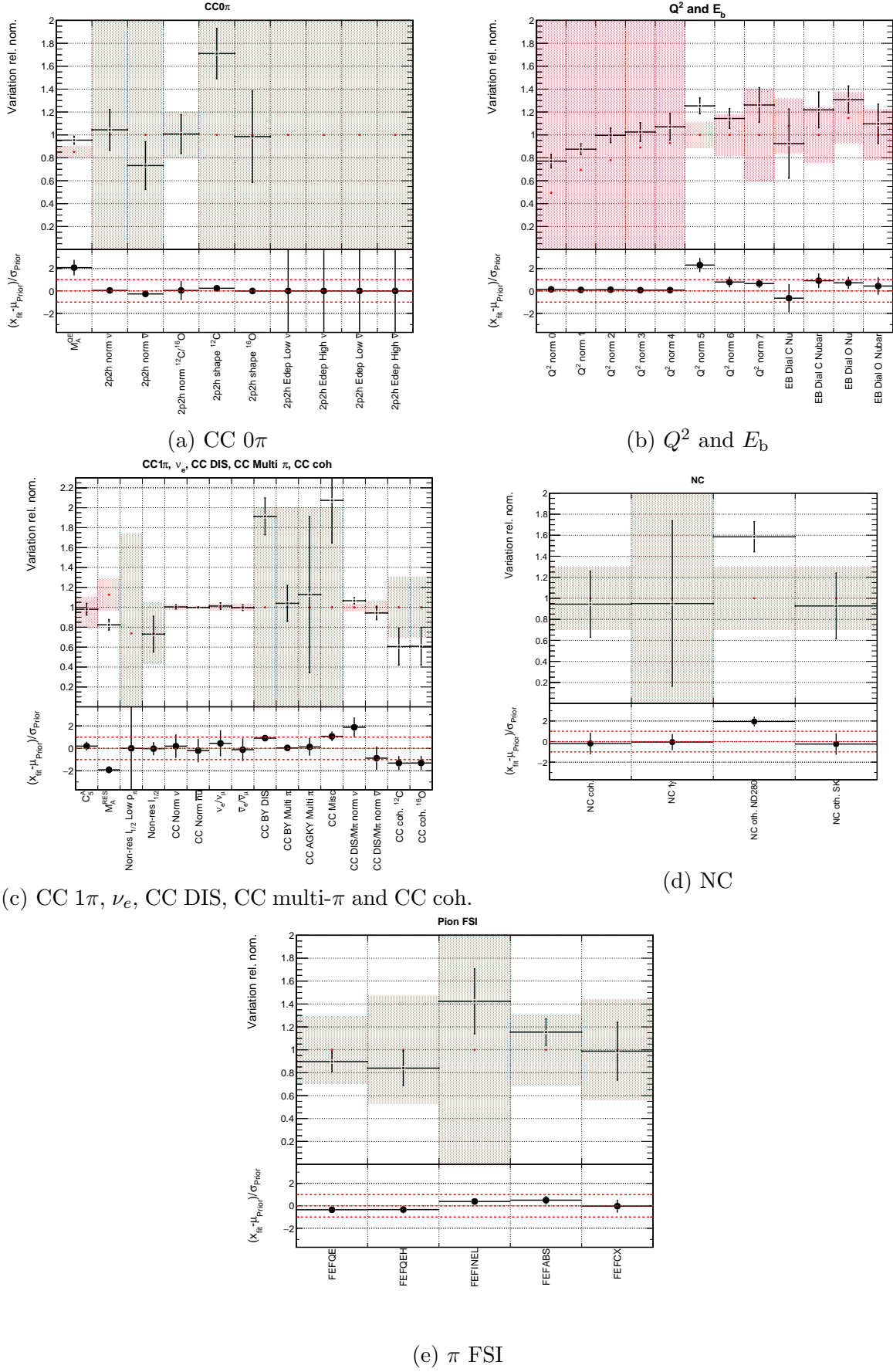
Figure F.6.: Interaction parameters for the data fit.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
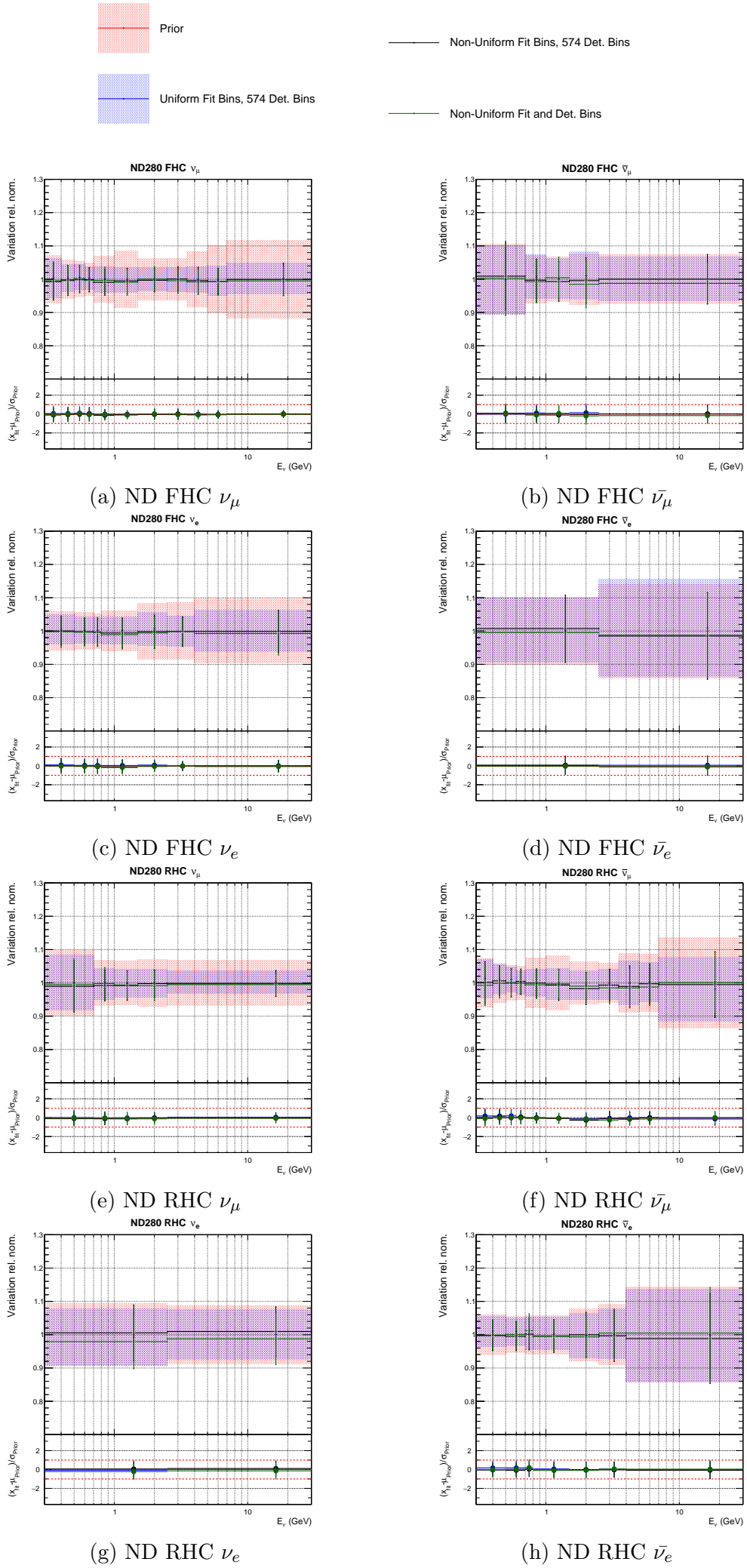
Figure F.7.: Comparison of ND280 flux parameters for the Asimov fits with different fit and detector binnings.
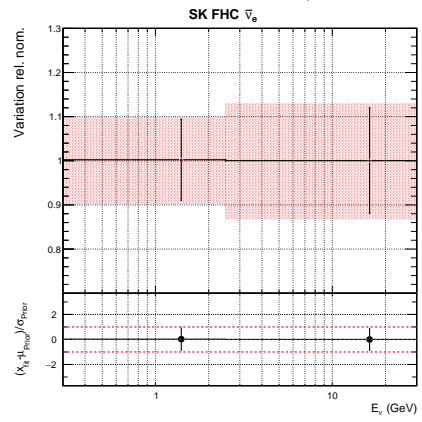
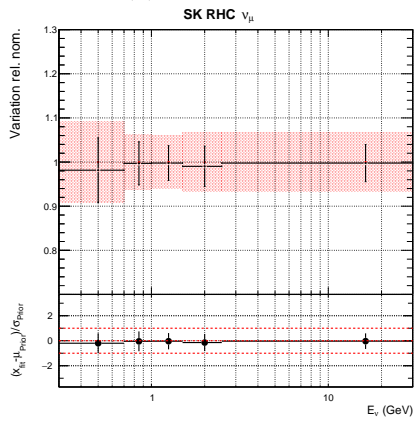(a) SK FHC ν_μ (b) SK FHC $\bar{\nu}_\mu$

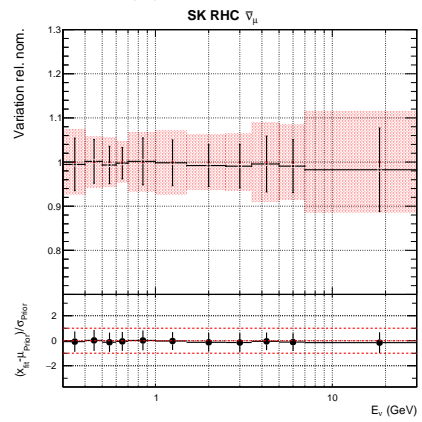(c) SK FHC ν_e (d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ (f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e (h) SK RHC $\bar{\nu}_e$
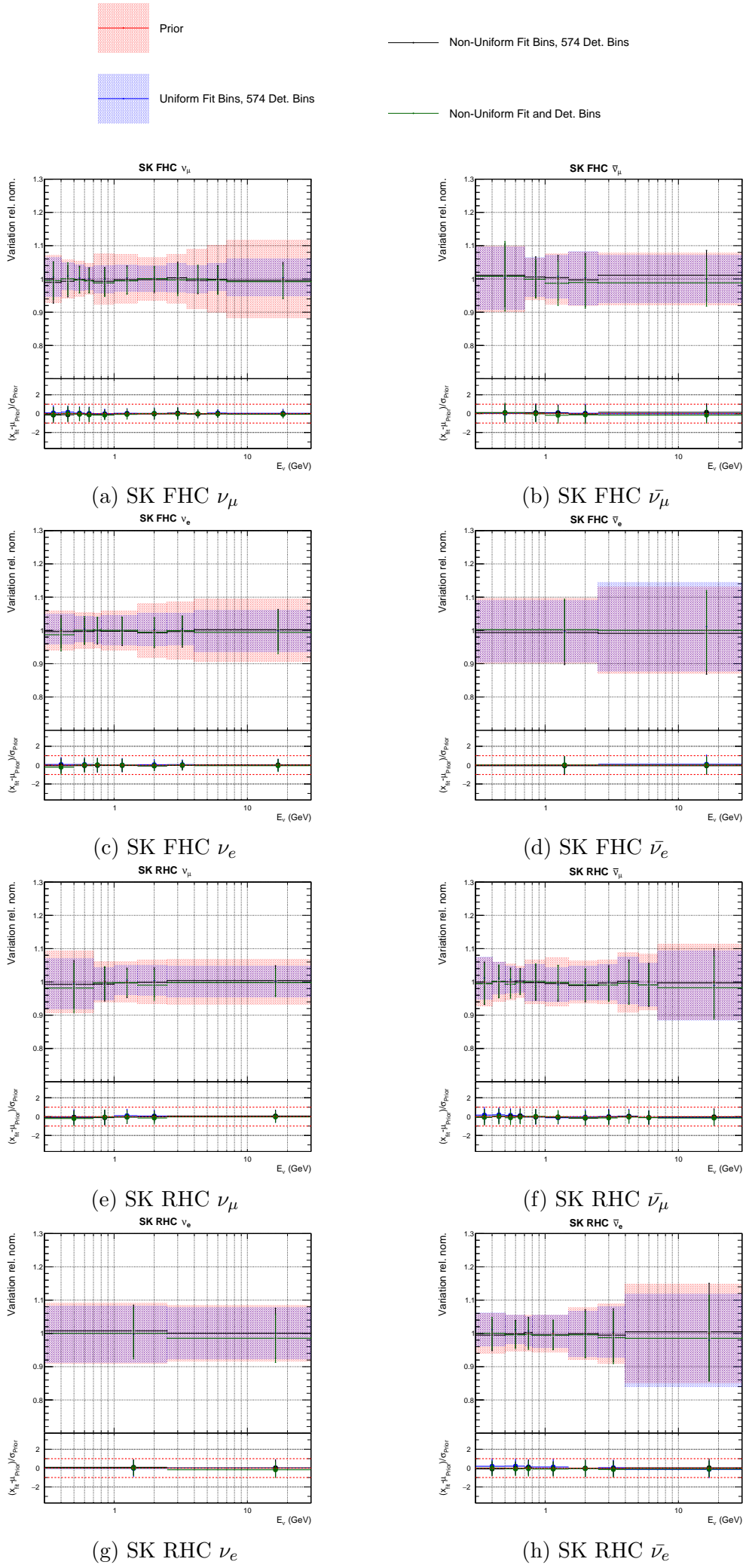
Figure F.8.: Comparison of SK flux parameters for the Asimov fits with different fit and detector binnings.

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
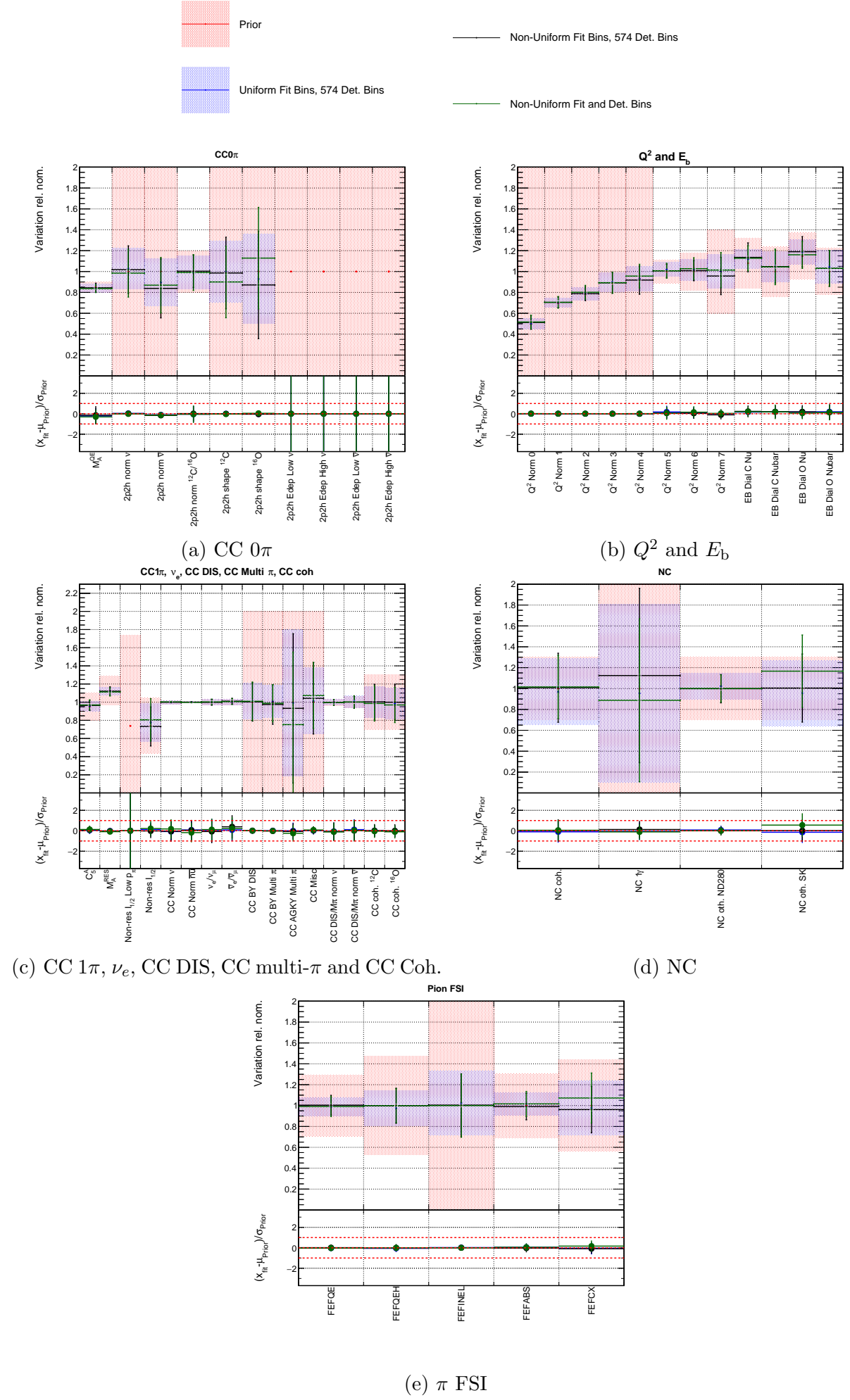
Figure F.9.: Comparison of interaction parameters for the Asimov fits with different fit and detector binnings.

(a) ND FHC ν_μ (b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e (d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ (f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e (h) ND RHC $\bar{\nu}_e$
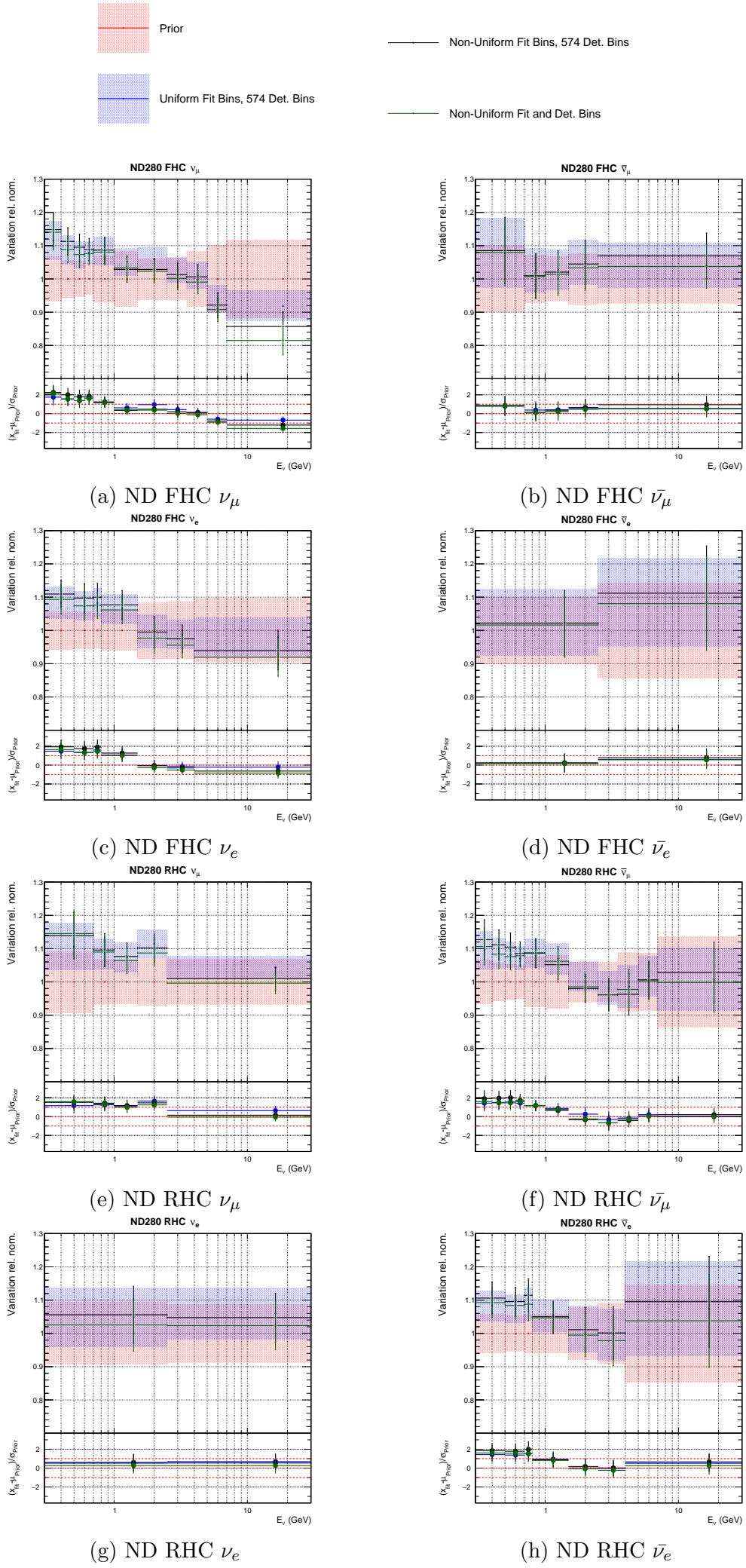
Figure F.10.: Comparison of ND280 flux parameters for the data fits with different fit and detector binnings.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

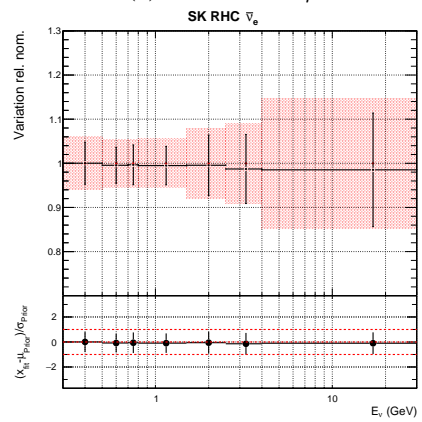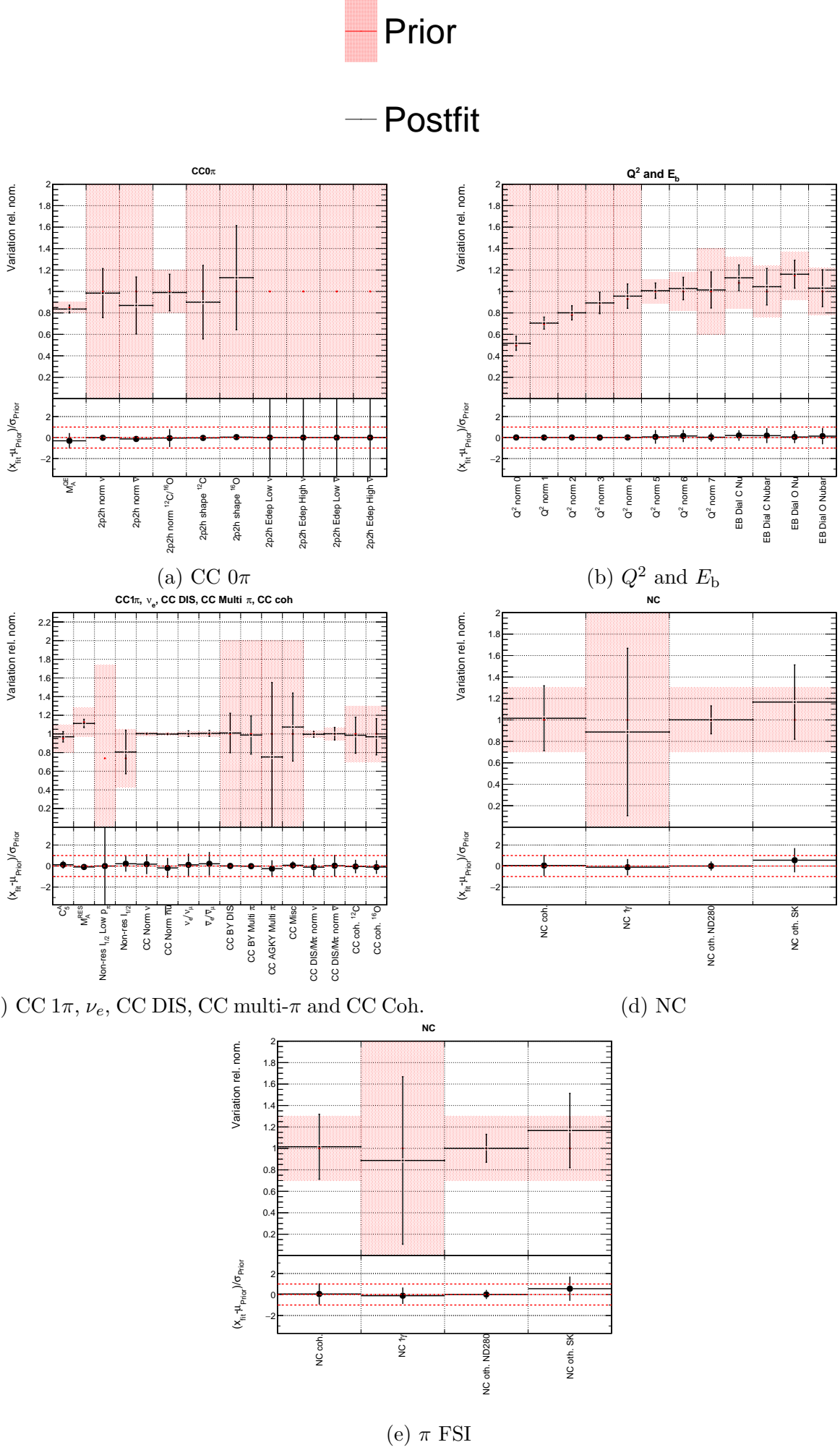(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
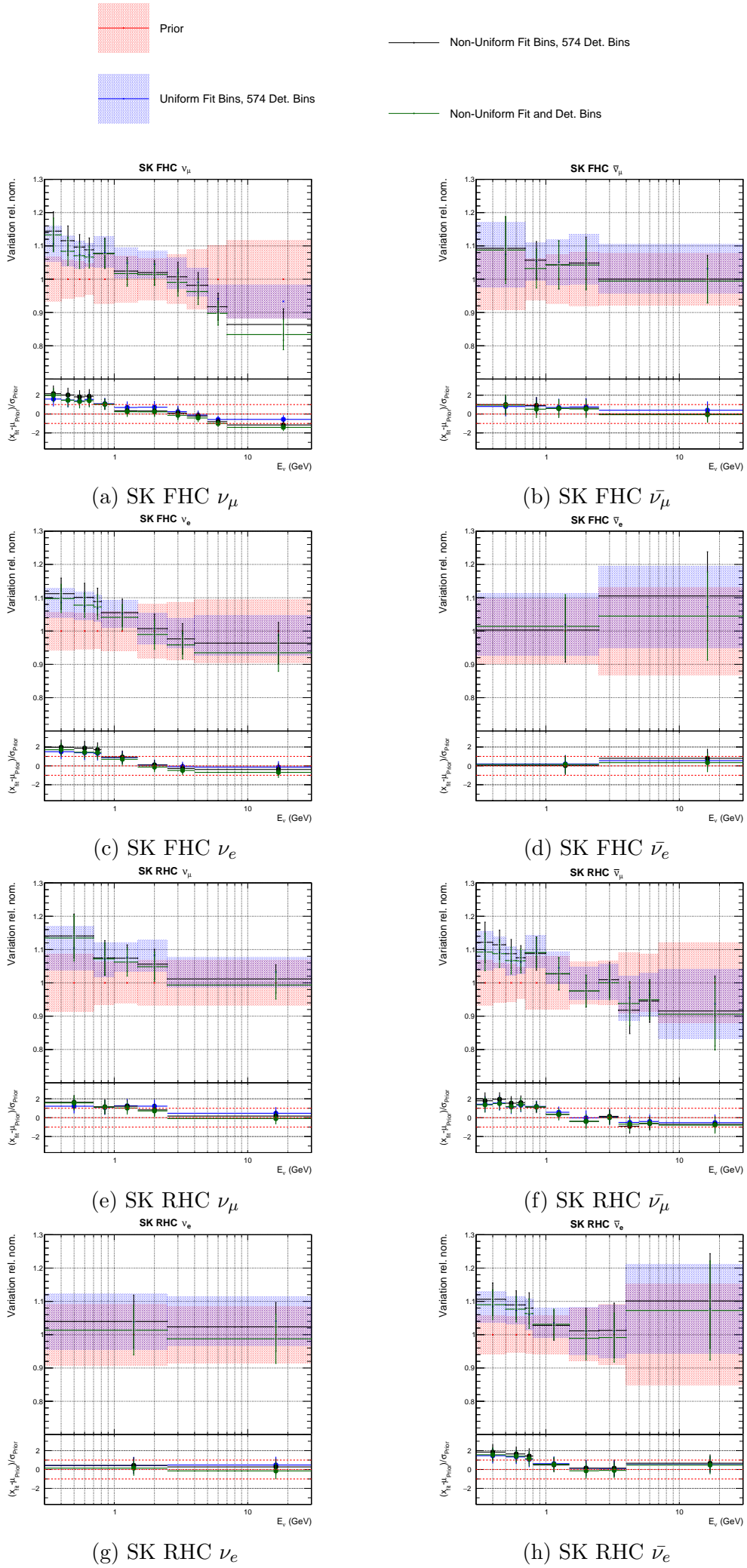
Figure F.11.: Comparison of SK flux parameters for the data fits with different fit and detector binnings.

(a) CC 0π

(b) Q^2 and $E_{\rm b}$

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
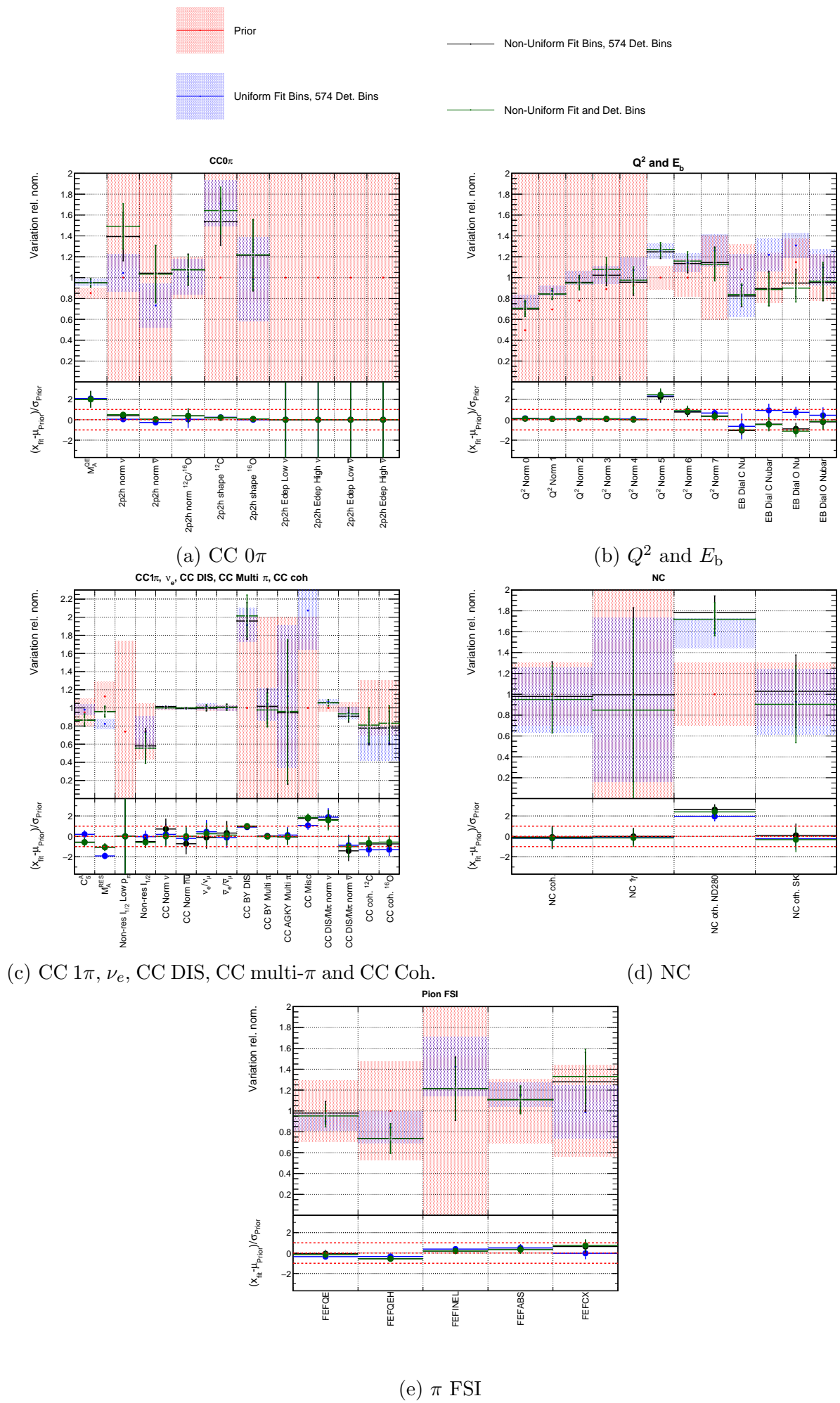
Figure F.12.: Comparison of interaction parameters for the data fits with different fit and detector binnings.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
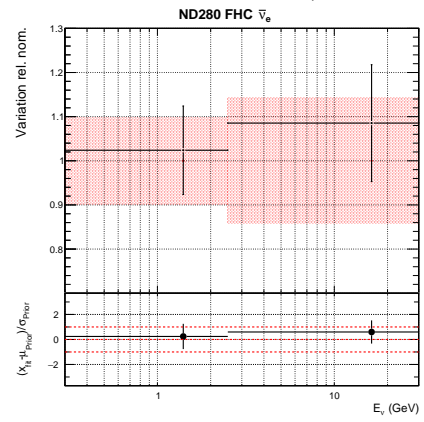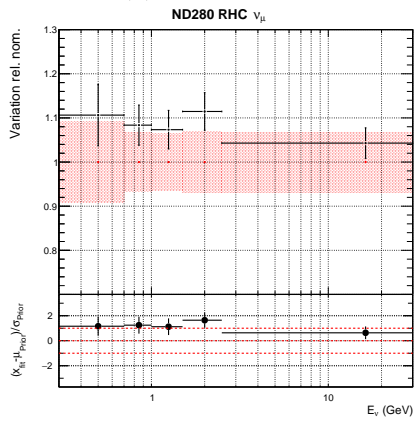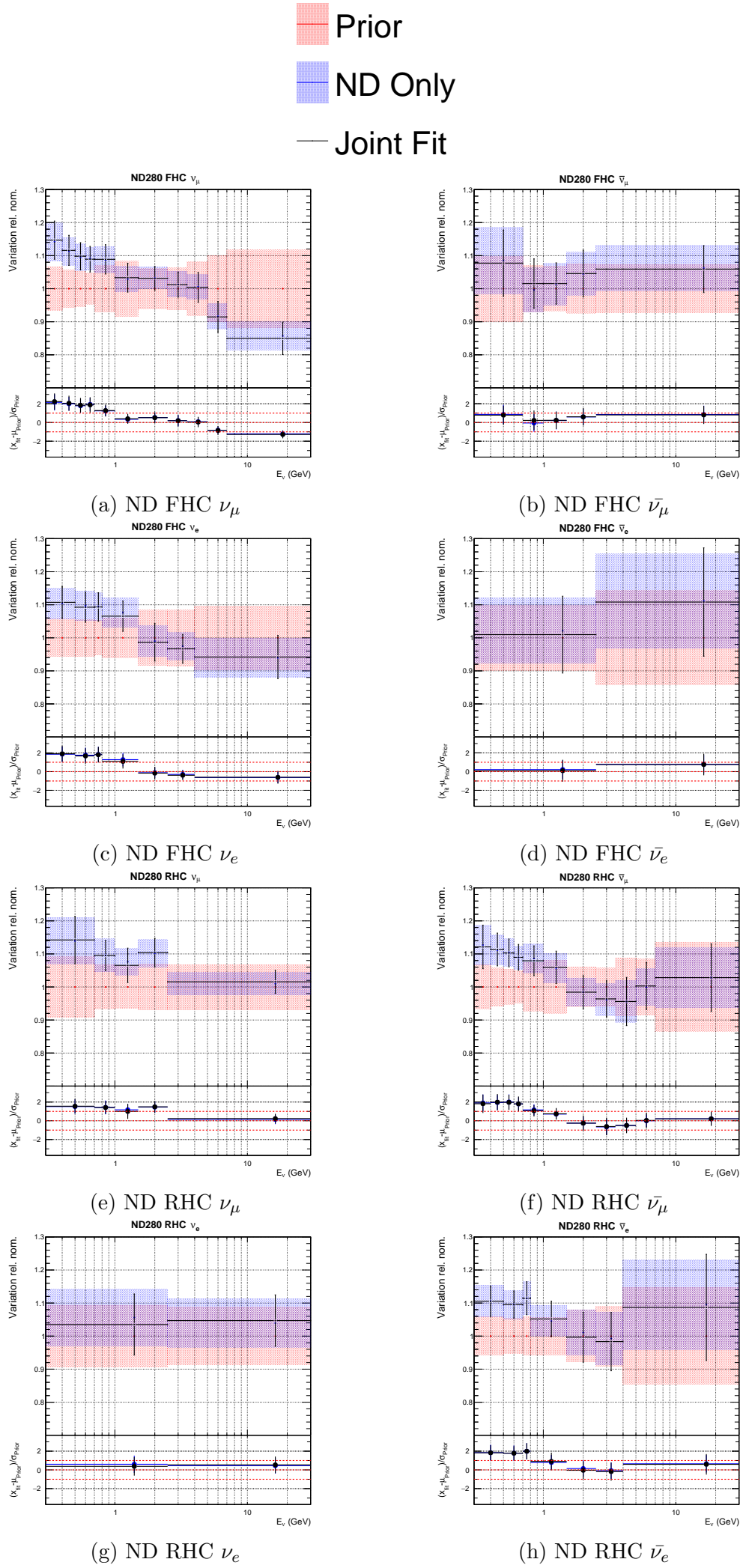
Figure F.13.: ND280 flux parameters for the joint and near detector only fits.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
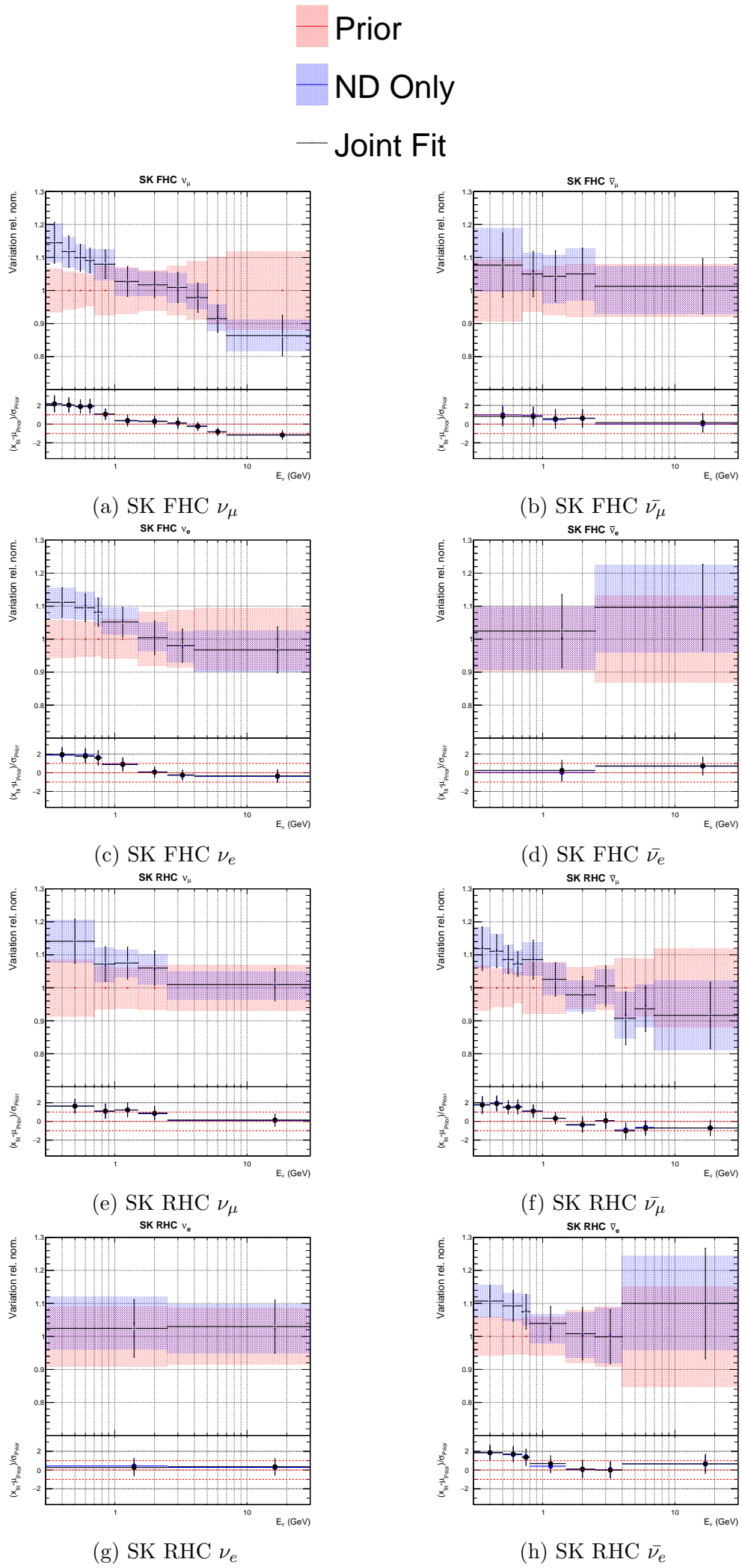
Figure F.14.: SK flux parameters for the joint and near detector only fits.

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC coh.

(d) NC

(e) π FSI
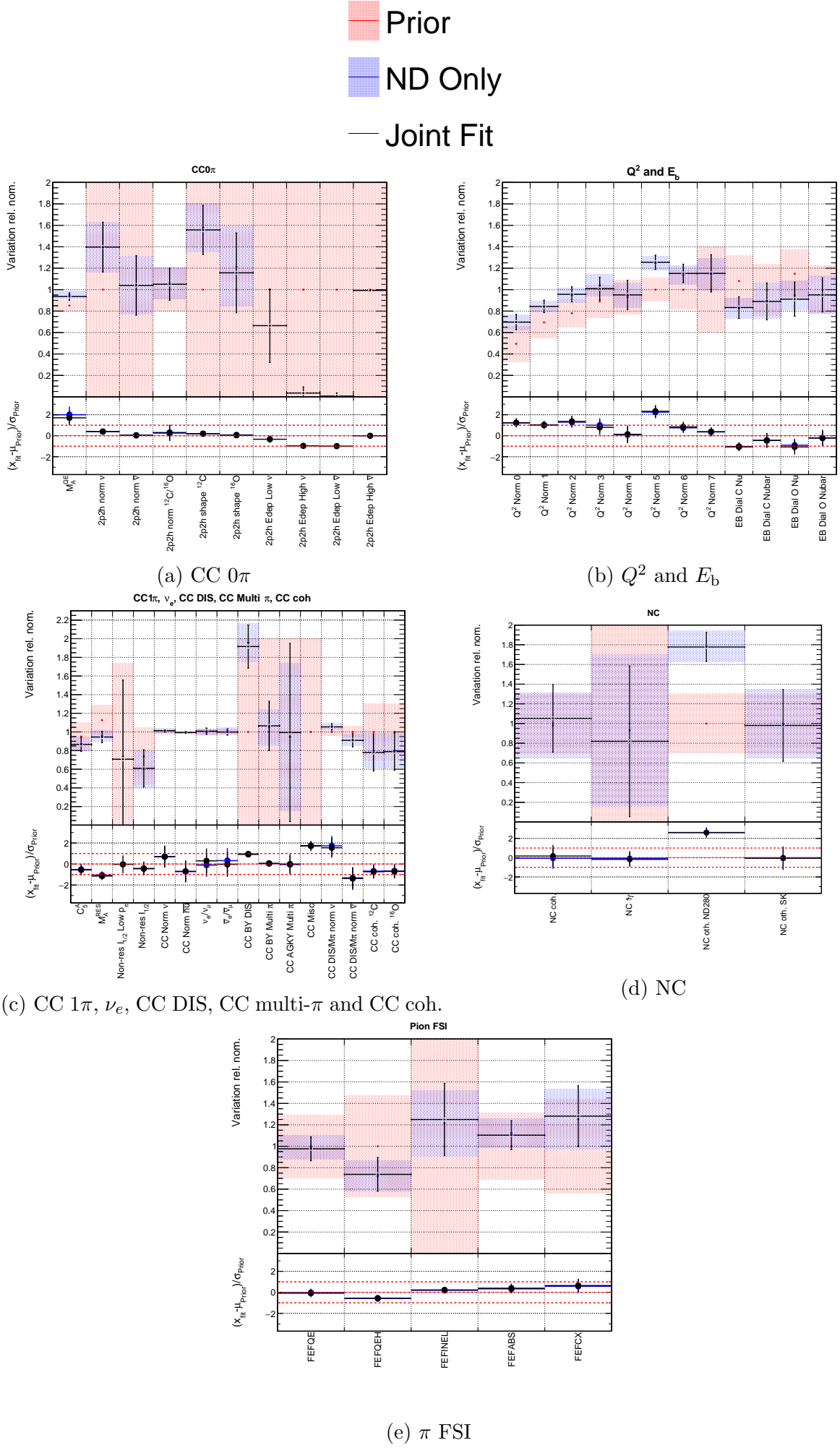
Figure F.15.: Interaction parameters for the joint and near detector only fits.

G. Full Postfit Correlation Matrices

The post-Asimov-fit correlation matrix for the flux and interaction parameters is shown in Figure G.1.
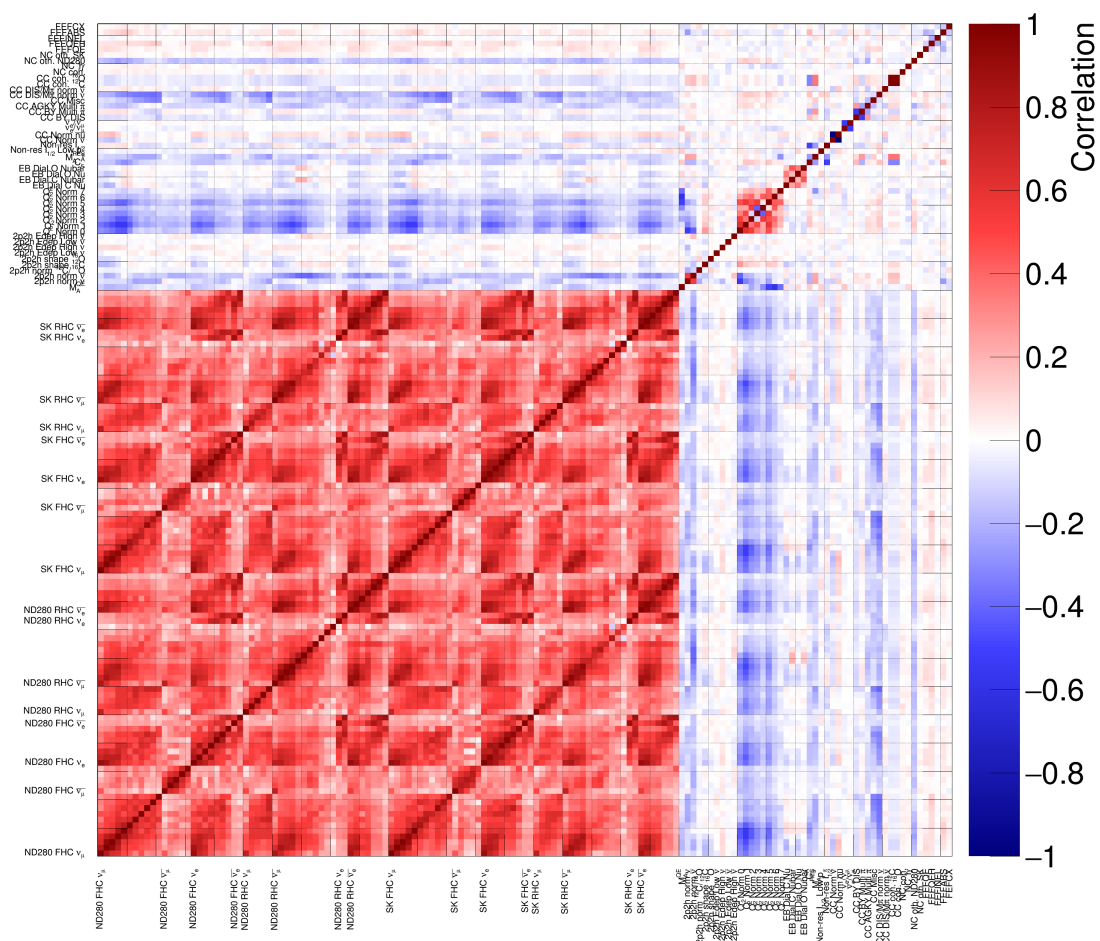
Figure G.1.: Asimov postfit correlation matrix for flux and interaction parameters.

The post-data-fit correlation matrix for the flux and interaction parameters is shown in Figure G.2.
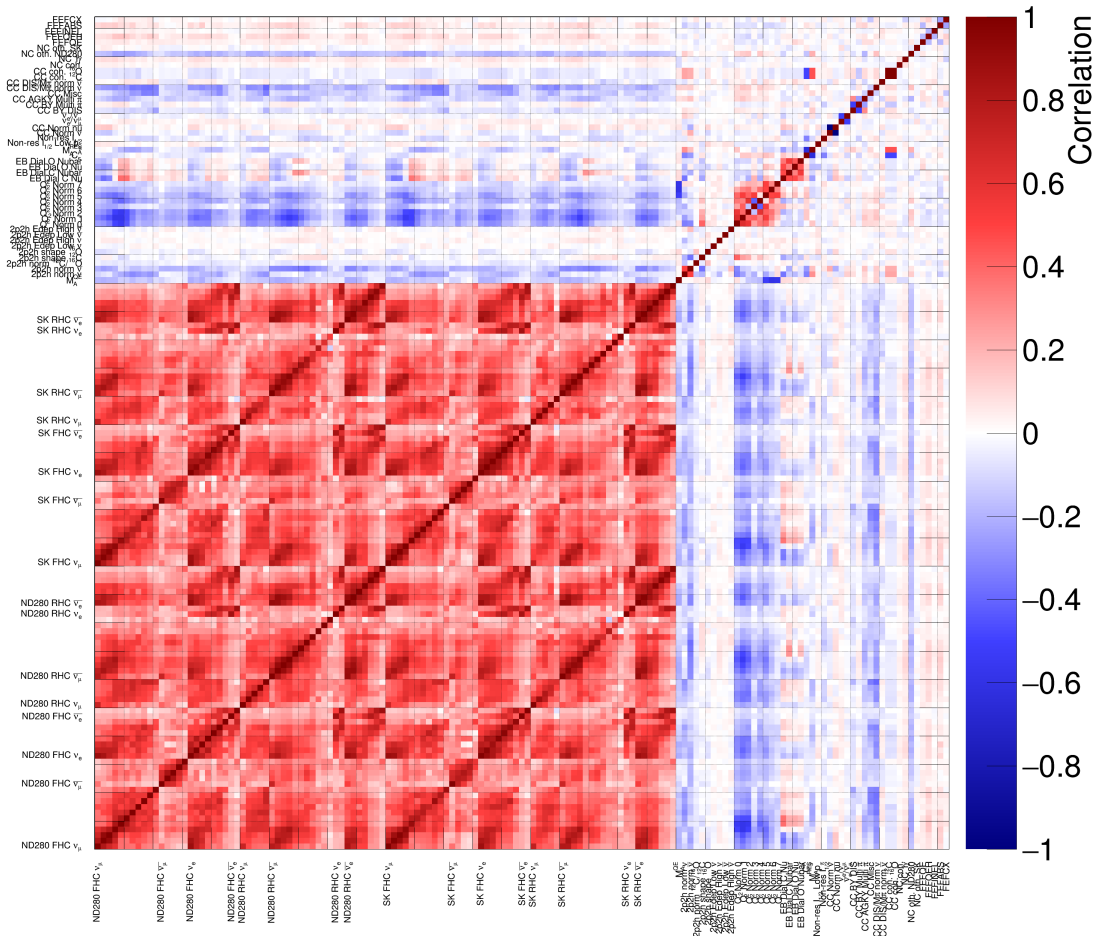
Figure G.2.: Data postfit correlation matrix for flux and interaction parameters.

H. Posterior Predictive Distributions

The p_μ and $\cos\theta_\mu$ projections of the posterior predictions for each sample are shown in Figures H.1–H.6, along with the prior predictions and data.
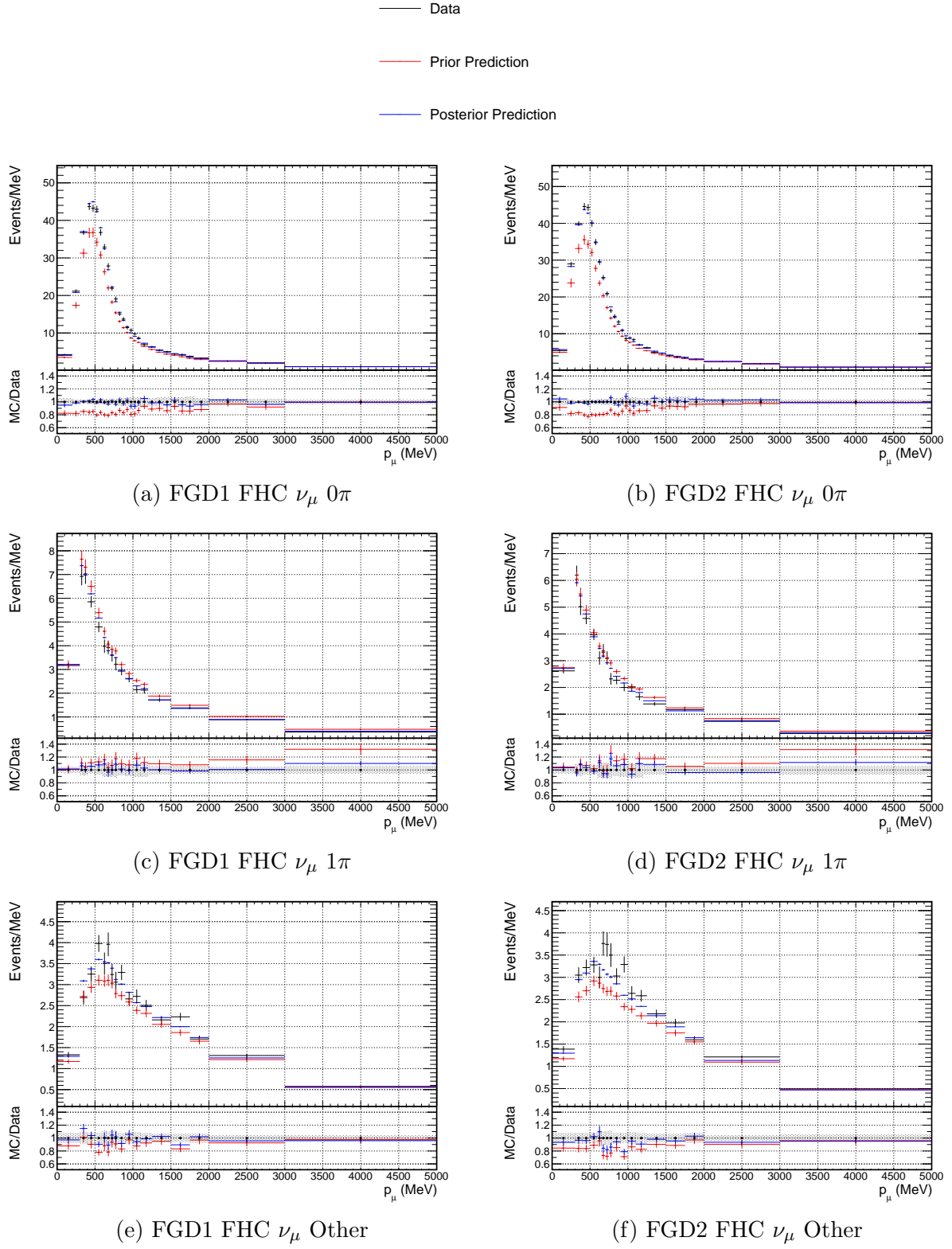
(a) FGD1 FHC ν_μ 0π

(b) FGD2 FHC ν_μ 0π

(c) FGD1 FHC ν_μ 1π

(d) FGD2 FHC ν_μ 1π

(e) FGD1 FHC ν_μ Other

(f) FGD2 FHC ν_μ Other

Figure H.1.: p_μ projections of the prior and posterior predictive distributions and data for FHC ν_μ selections.

(a) FGD1 RHC $\bar{\nu}_\mu$ 0π

(b) FGD2 RHC $\bar{\nu}_\mu$ 0π

(c) FGD1 RHC $\bar{\nu}_\mu$ 1π

(d) FGD2 RHC $\bar{\nu}_\mu$ 1π

(e) FGD1 RHC $\bar{\nu}_\mu$ Other

(f) FGD2 RHC $\bar{\nu}_\mu$ Other

Figure H.2.: p_μ projections of the prior and posterior predictive distributions and data for RHC $\bar{\nu}_\mu$ selections.

(a) FGD1 RHC ν_μ 0π

(b) FGD2 RHC ν_μ 0π

(c) FGD1 RHC ν_μ 1π

(d) FGD2 RHC ν_μ 1π

(e) FGD1 RHC ν_μ Other

(f) FGD2 RHC ν_μ Other

Figure H.3.: p_μ projections of the prior and posterior predictive distributions and data for RHC ν_μ selections.

(a) FGD1 FHC ν_μ 0π

(b) FGD2 FHC ν_μ 0π

(c) FGD1 FHC ν_μ 1π

(d) FGD2 FHC ν_μ 1π

(e) FGD1 FHC ν_μ Other

(f) FGD2 FHC ν_μ Other
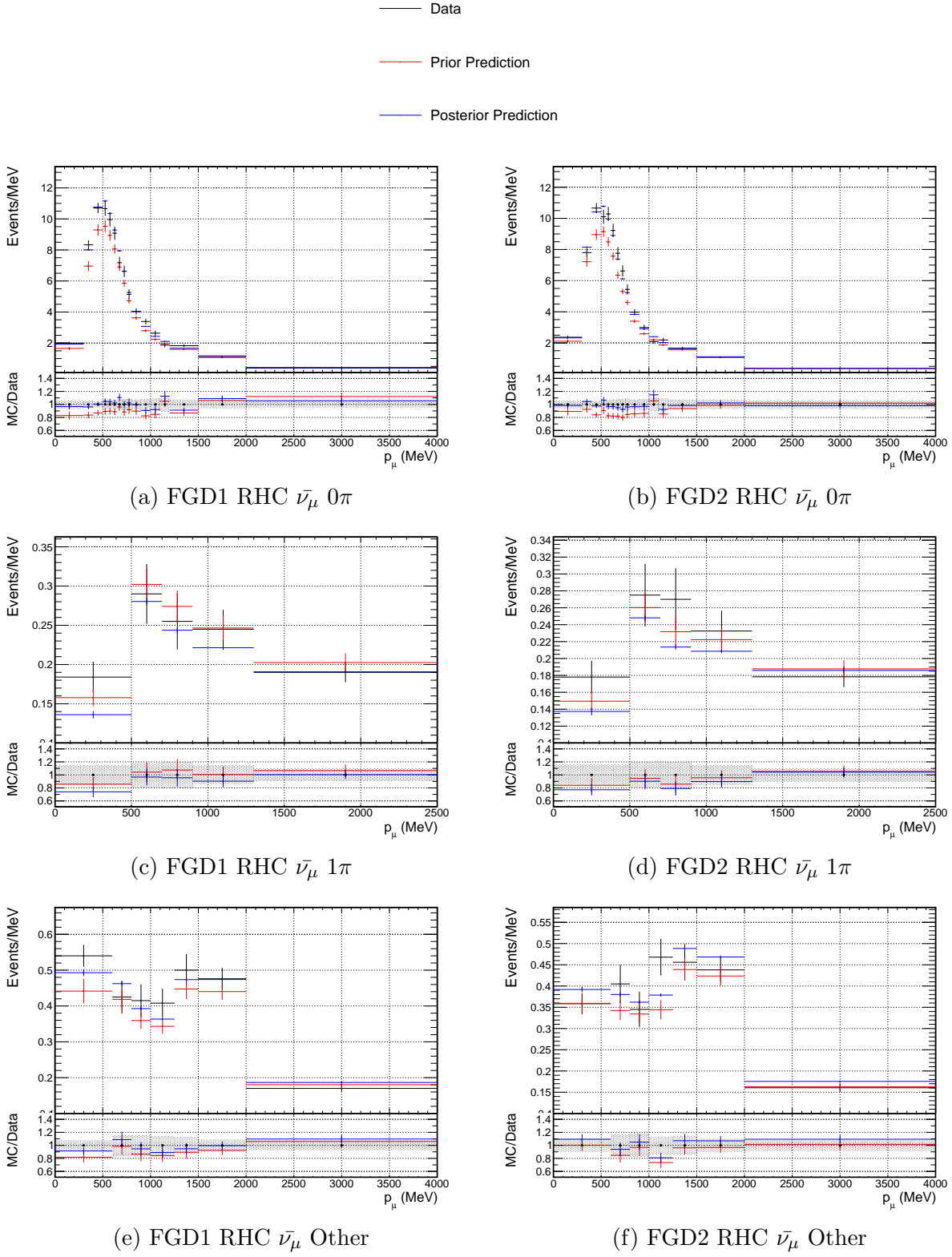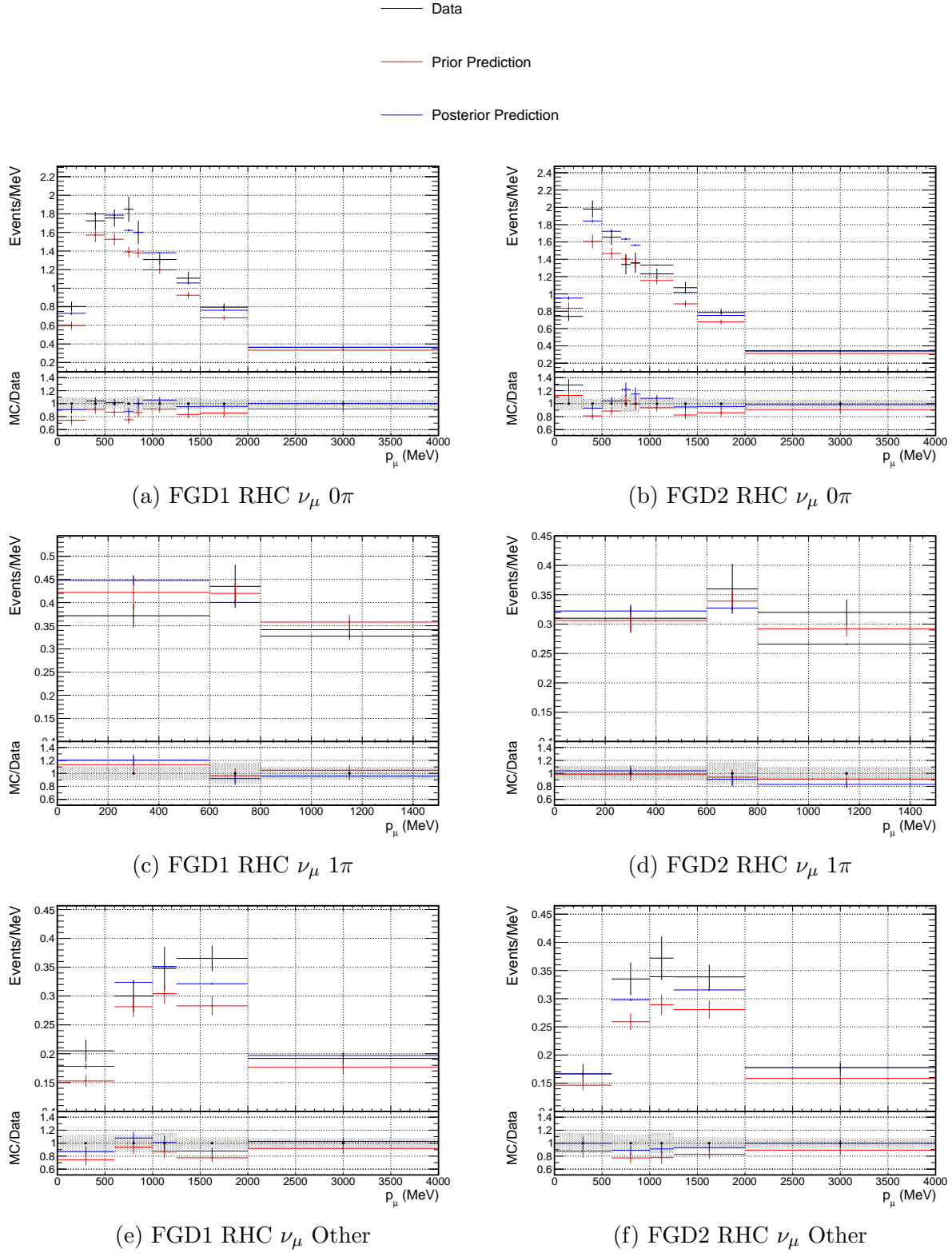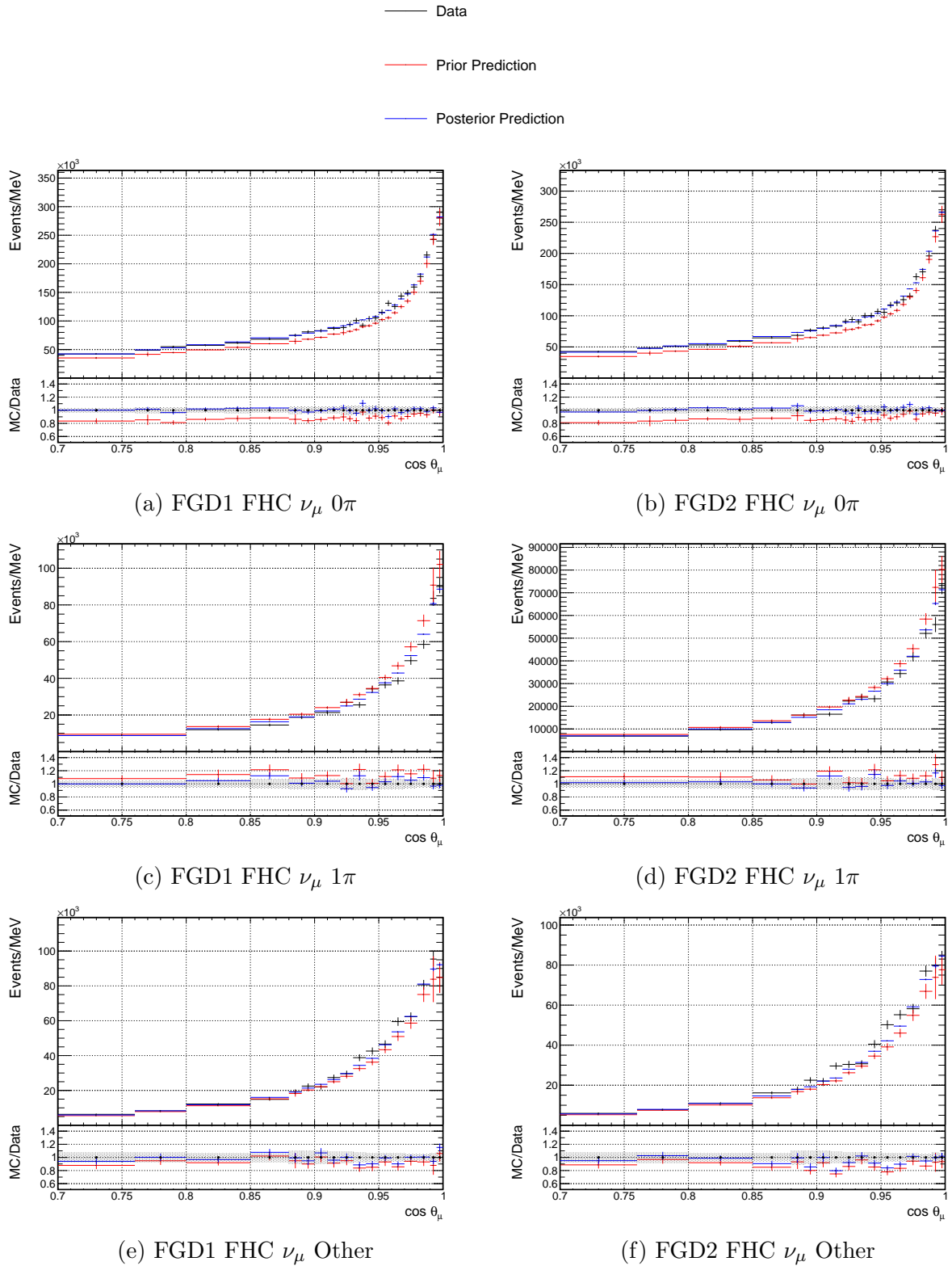
Figure H.4.: $\cos\theta_\mu$ projections of the prior and posterior predictive distributions and data for FHC ν_μ selections.

Figure H.5.: cosθ_μ projections of the prior and posterior predictive distributions and data for RHC $\bar{\nu}_\mu$ selections.

Data

Prior Prediction

Posterior Prediction

(a) FGD1 RHC ν_μ 0π

(b) FGD2 RHC ν_μ 0π

(c) FGD1 RHC ν_μ 1π

(d) FGD2 RHC ν_μ 1π

(e) FGD1 RHC ν_μ Other

(f) FGD2 RHC ν_μ Other
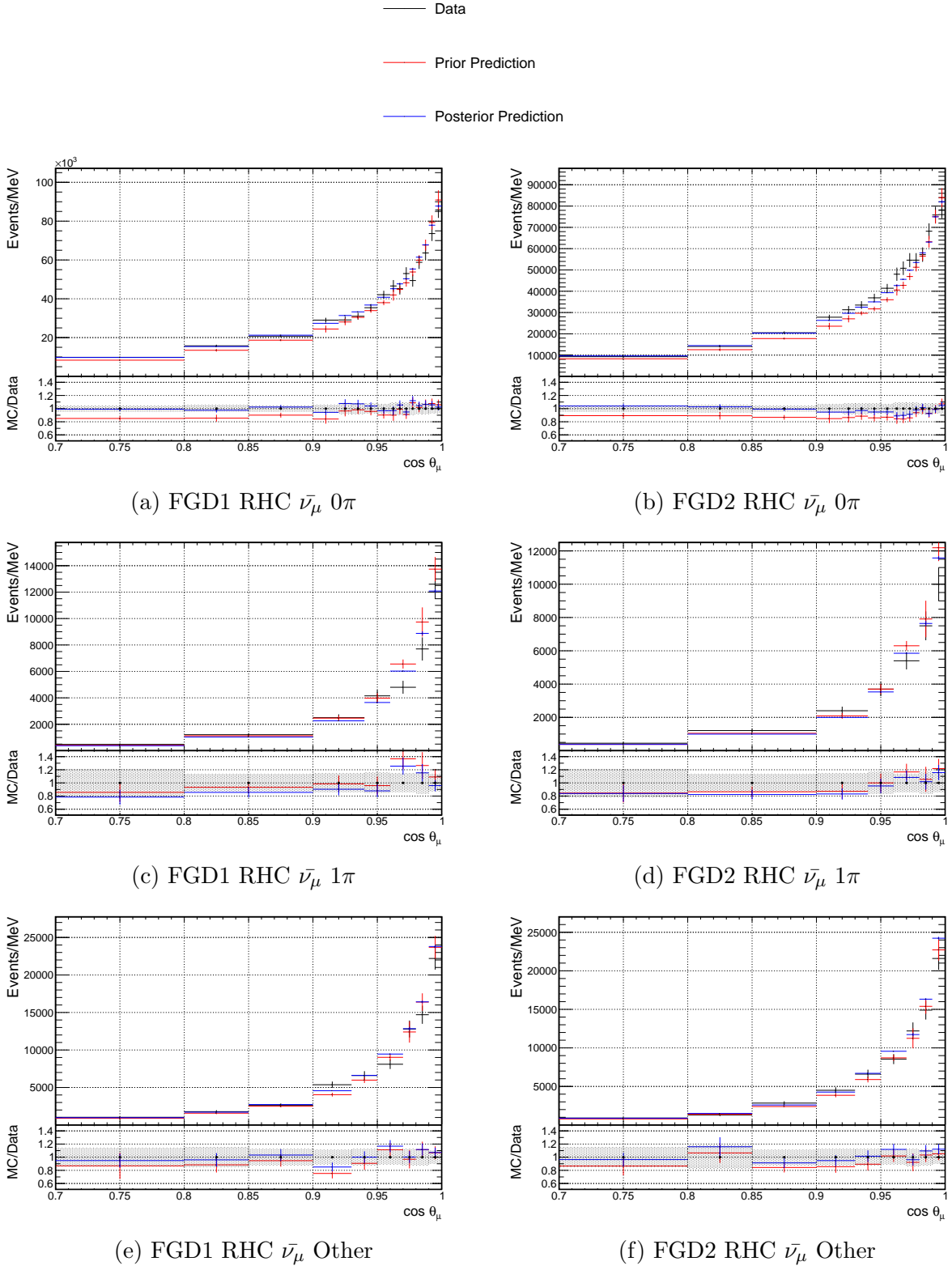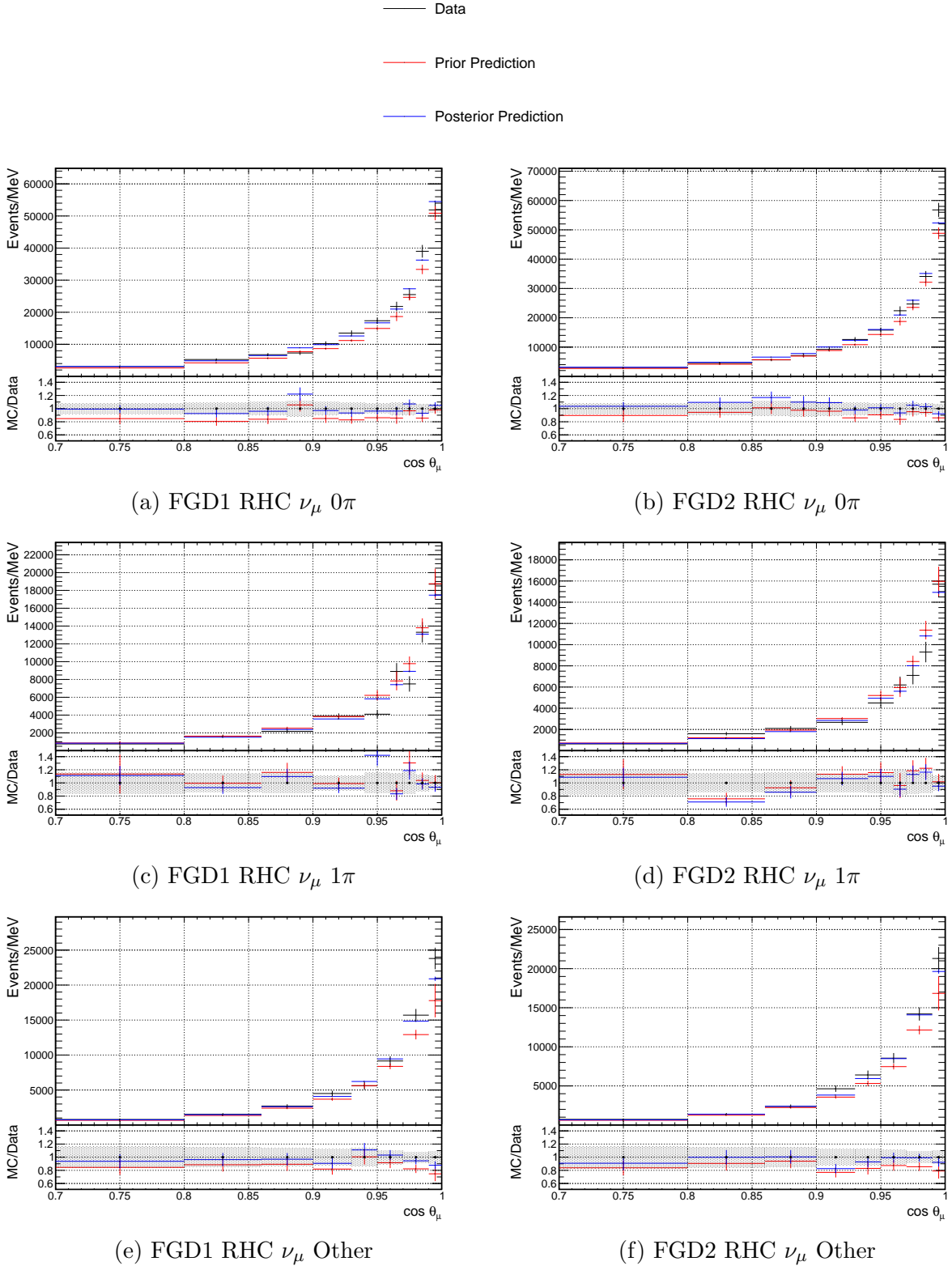
Figure H.6.: $\cos\theta_\mu$ projections of the prior and posterior predictive distributions and data for RHC ν_μ selections.

I. Alternative Fit Studies

A number of compatibility studies were performed, comparing subsets of data and using alternative fit models.

The data-subset results presented here were not used in the full oscillation analysis, but are part of the validations of the input models and fitting framework. The alternative model fits were used to inform choices in the interaction model used for the final fits.

These fits were all run using the uniform-rectangular fit binning and 574 merged detector bins.

I.1. FGD1 and FGD2 Only Fits

As discussed in Section 3.2.2.1, FGD1 and FGD2 have different target materials. FGD2 has water layers, and so contains a significant amount of ^{16}O, whereas FGD1 has only plastic scintillator (C_8H_8) layers. The constraint on ^{16}O only parameters therefore comes only from FGD2.

The location of the FGDs also causes differences in reconstruction between the two subdetectors. FGD1 has two TPCs downstream of it, and so performs more accurate reconstruction for forward-going tracks, whereas FGD2 has two TPCs upstream of it, and so performs more accurate reconstruction for backward going-tracks.

To investigate the compatibility of the FGD1 and FGD2 samples, fits were run using each independently. There is approximately the same amount of data in the FGD1 and FGD2 sample, so the constraining power of each is expected to be similar.

The results for the flux parameters are shown in Figures I.1 and I.2, along with the full FGD1 and FGD2 fit. The postfit values of the FGD1 and FGD2 only fits are consistently within the postfit uncertainties of both each other, and the full fit. The high pulls at lower energies are seen in all fits. The gradient of the decrease in pull with increasing energy shows some small discrepancies, but this is to be expected given the high correlations between the flux parameters.

I. Alternative Fit Studies

The SK flux parameters have very similar behaviour to the ND flux parameters, as is expected as they are only constrained by their prior correlations.

The interaction parameters, shown in Figure I.4, have larger differences. M_A^{QE} is pushed higher for FGD1 than FGD2, with the full fit value lying between two. The 2p2h ν normalisation parameter is also higher for FGD1, with the FGD2 value remaining close to nominal as it is for the full fit. The 2p2h shape parameters have large differences. For FGD2, 2p2h shape C is pushed to its upper bound at $\sim$2, whereas it is only pushed to $\sim$1.3 for FGD1. This is interesting as both FGDs are able to constrain the parameter. For the full fit, the 2p2h shape C postfit value lies between the FGD1 and FGD2 values, despite not affecting any FGD1 events. This is likely driven by the prior correlations between 2p2h shape C and O.

The shape of the increase in Q^2 parameters is slightly different between the FGD1 and FGD2 fits, but all are within uncertainty of each other.

The E_b parameters are significantly higher for FGD1. As discussed in Section 4.4.1, as these parameters are non-Gaussian it is important to inspect the full distributions as well as the extracted values. These are shown in Figure I.4. For $E_b \nu$ C, there is clearly a strong peak much higher than for both the FGD2 and full fits. For $E_b \bar{\nu}$ C, there is again a higher peak for FGD1, but the full fit lies between the two subdetector fits. For $E_b \nu$ O there is more overlap between the fits, which is more expected as there is no ^{16}O in FGD1. However, for $E_b \bar{\nu}$ O, there are different peaks for the different fits, driven by the correlations between the E_b parameters.

The CC DIS BY parameter is higher for FGD1 than FGD2, with the full fit falling between the two. The other CC 1π and CC Other targetting parameters are more consistent for the two fits. NC 1γ and NC Other show some differences, but this is not too concerning as there are low statistics for the interaction modes they affect.

For FSI, the charge exchange parameter has the largest discrepancy, but this is within the postfit uncertainty.

Overall there is fairly good compatibility between the FGD samples. Where there are discrepancies, the full fit value tends to fall between the FGD1 and FGD2 only results. CC DIS BY is the only parameter affecting both ^{12}C and ^{16}O events which are not within the uncertainty for the two fits.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
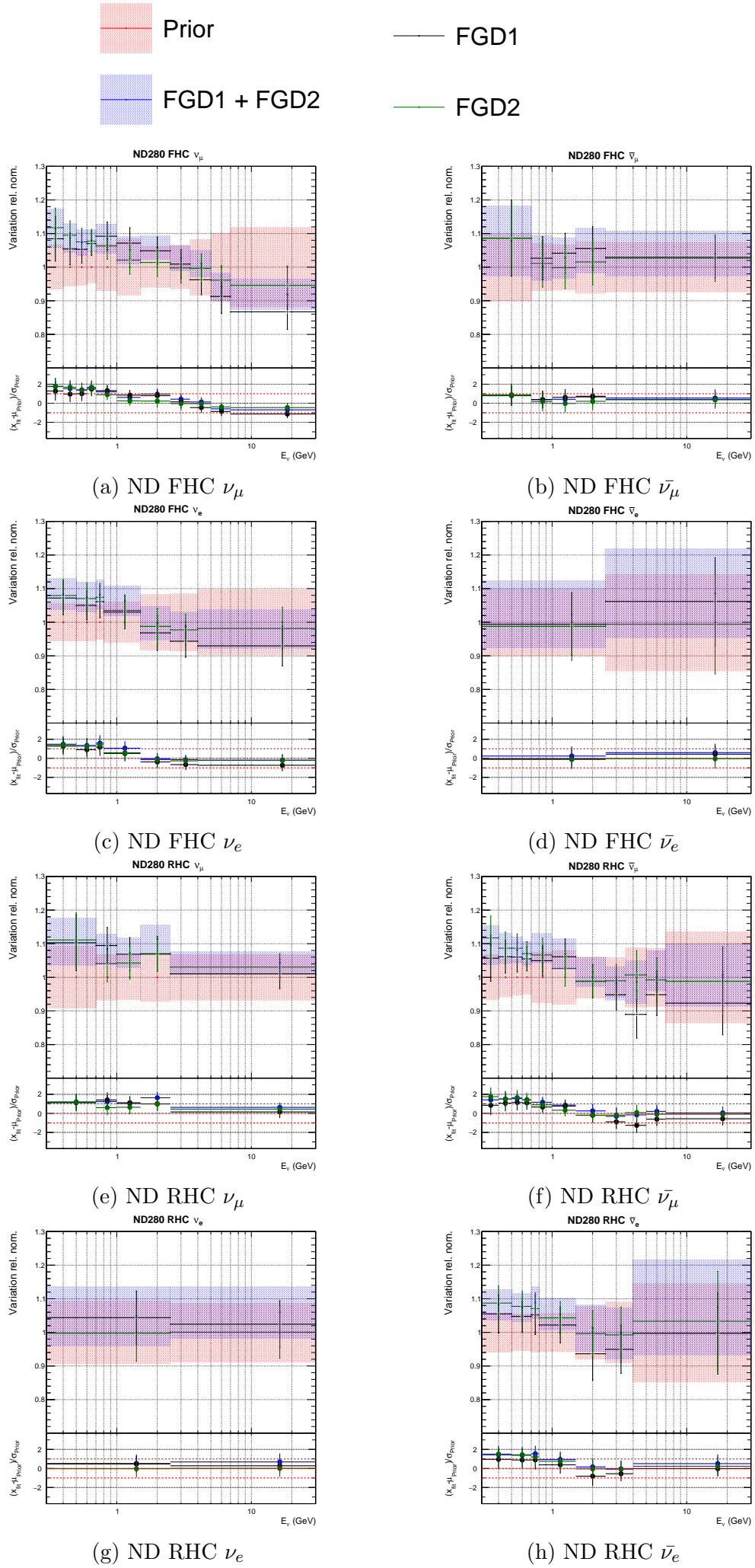
Figure I.1.: ND280 flux parameters for the FGD1 and 2 only fits.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
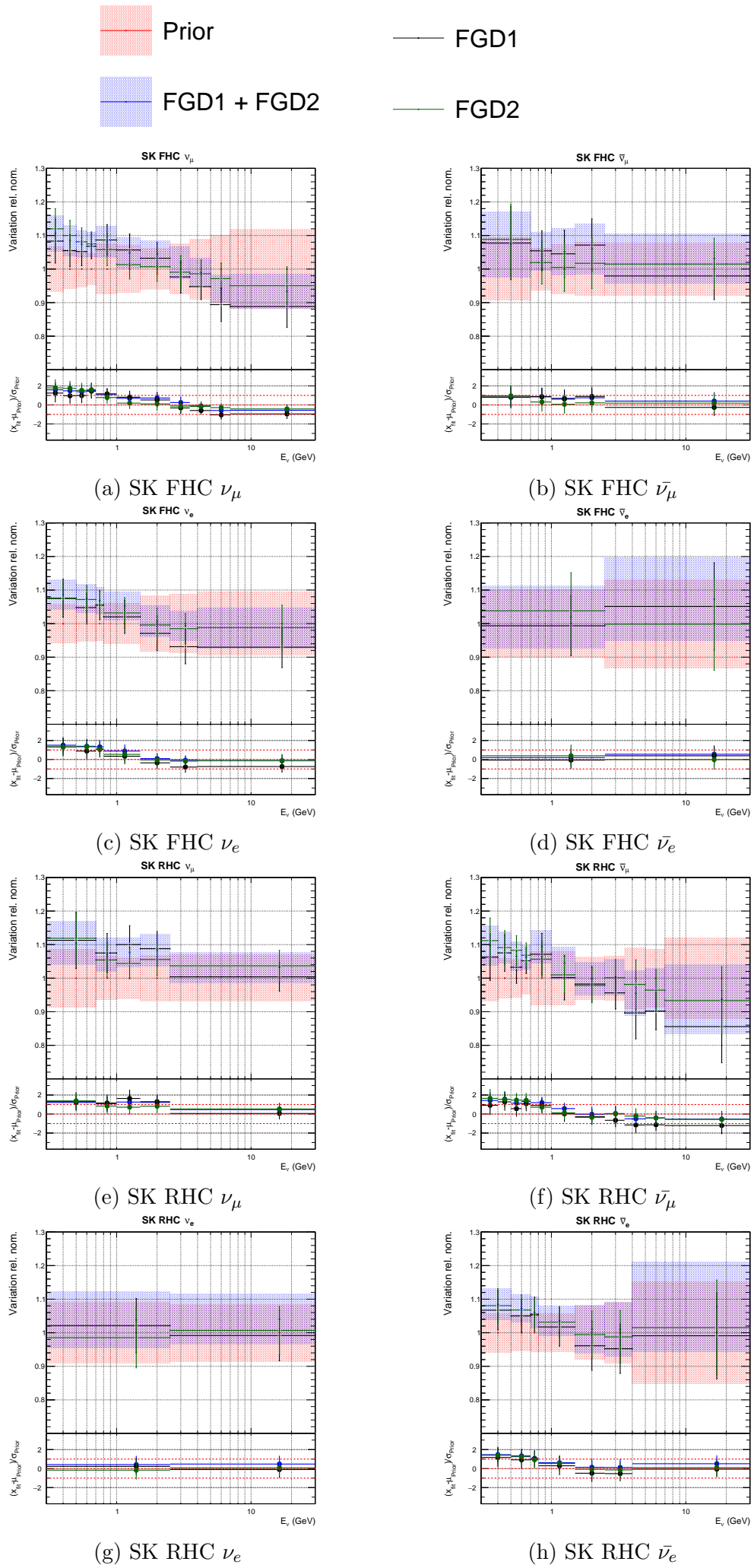
Figure I.2.: SK flux parameters for the FGD1 and 2 only fits.

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
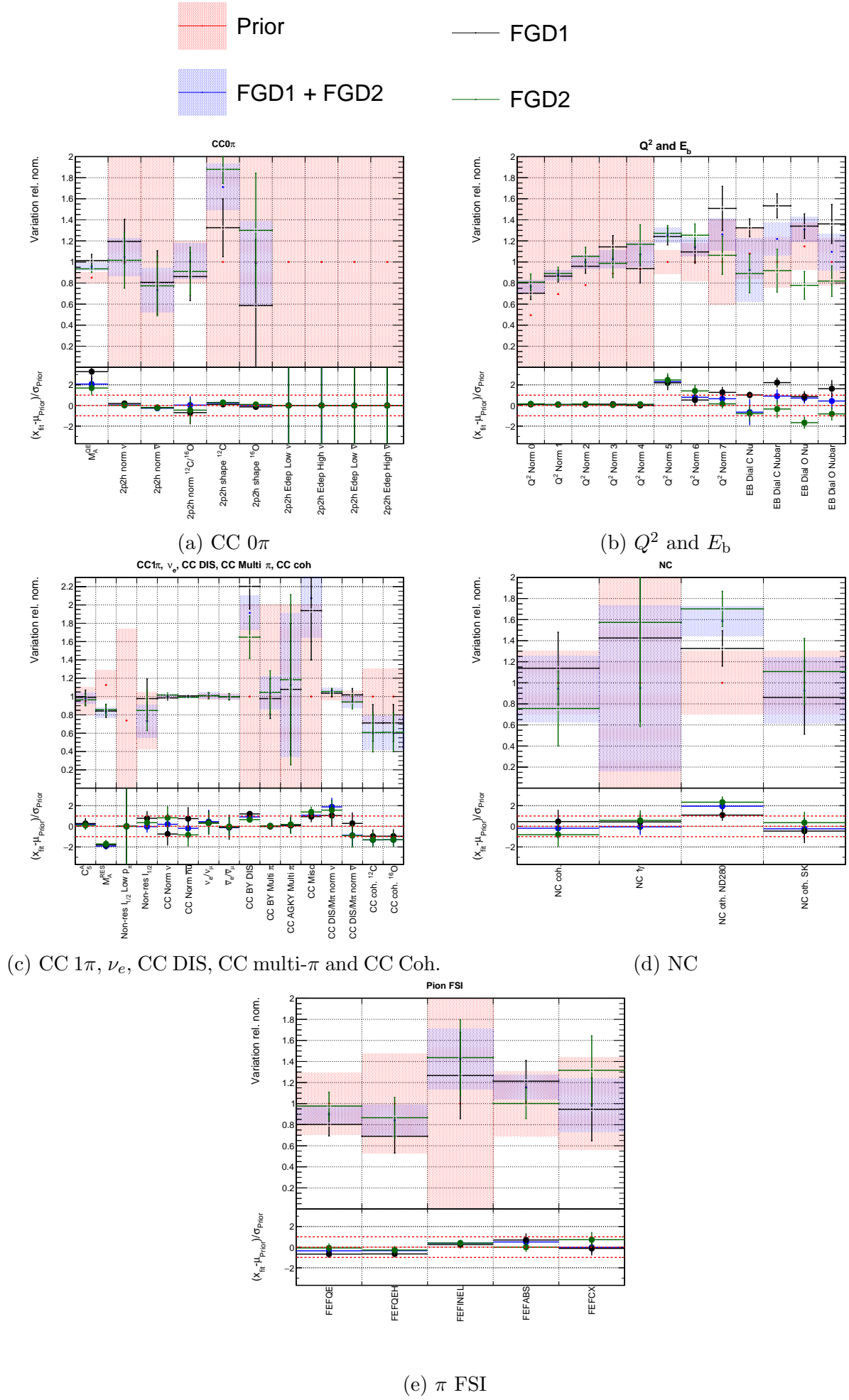
Figure I.3.: Interaction parameters for the FGD1 and 2 only fits.

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
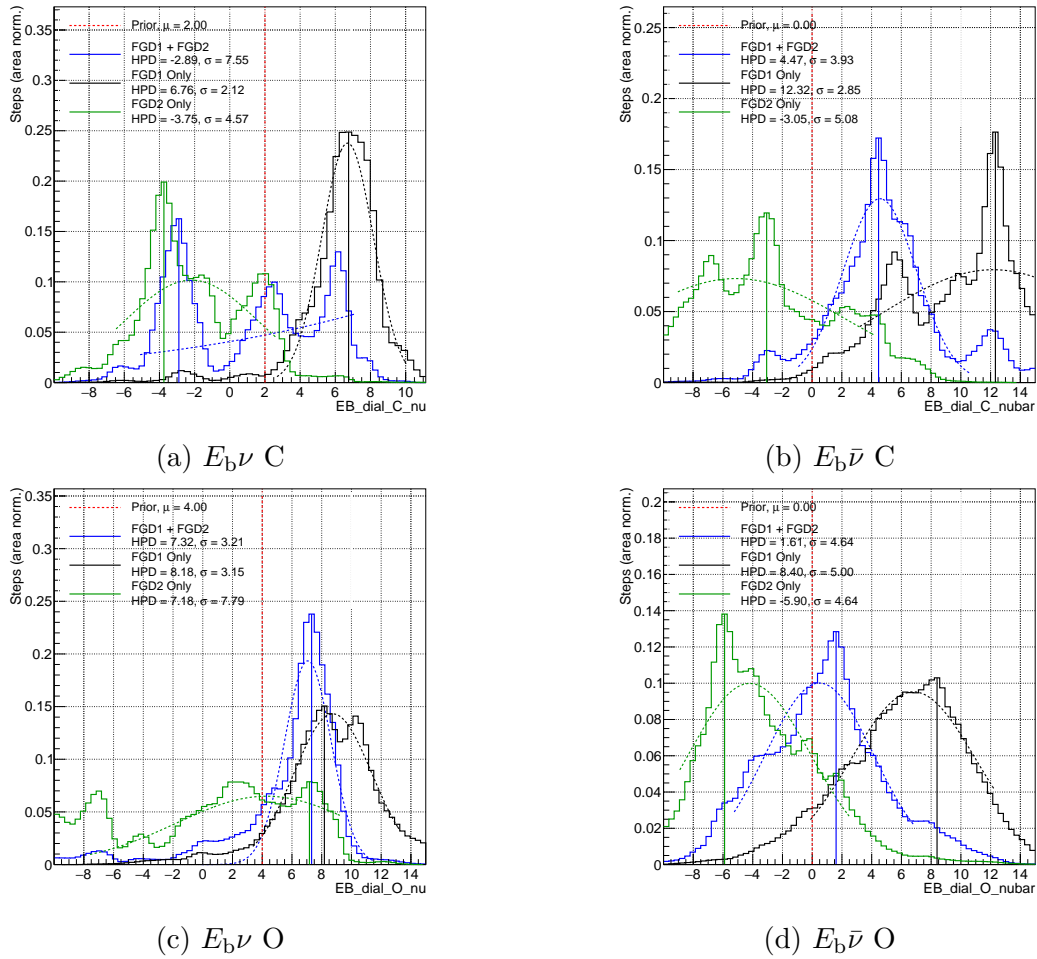
Figure I.4.: Posterior distributions from the binding energy parameters for the FGD1 and FGD2 only fits.

I.2. FHC and RHC Only Fits

For runs 2–4, 8, data was taken in FHC mode, and for runs 5–7, 9, data was taken in RHC mode. To see the compatibility of the model fit to neutrino and anti-neutrino data, fits were running using runs 2–4+8 and runs 5–7+9 separately. As the amount of FHC data is ∼3× larger than for RHC, it is expected there will be larger constraints for the FHC only fit.

The results for the flux parameters are shown in Figures I.5 and I.6. The shape of the pulls are very similar for the three fits, but for RHC only the parameters are consistently closer to nominal than for FHC only, with the full fit value lying between the two. In the FHC fit, the RHC flux parameters are constrained via the prior correlations with the FHC flux parameters, and vice versa.

The results for the interaction parameters are shown in Figure I.7. M_A^{QE} is not pushed as high for FHC as RHC, but both are within the postfit uncertainty of the full fit value. The

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
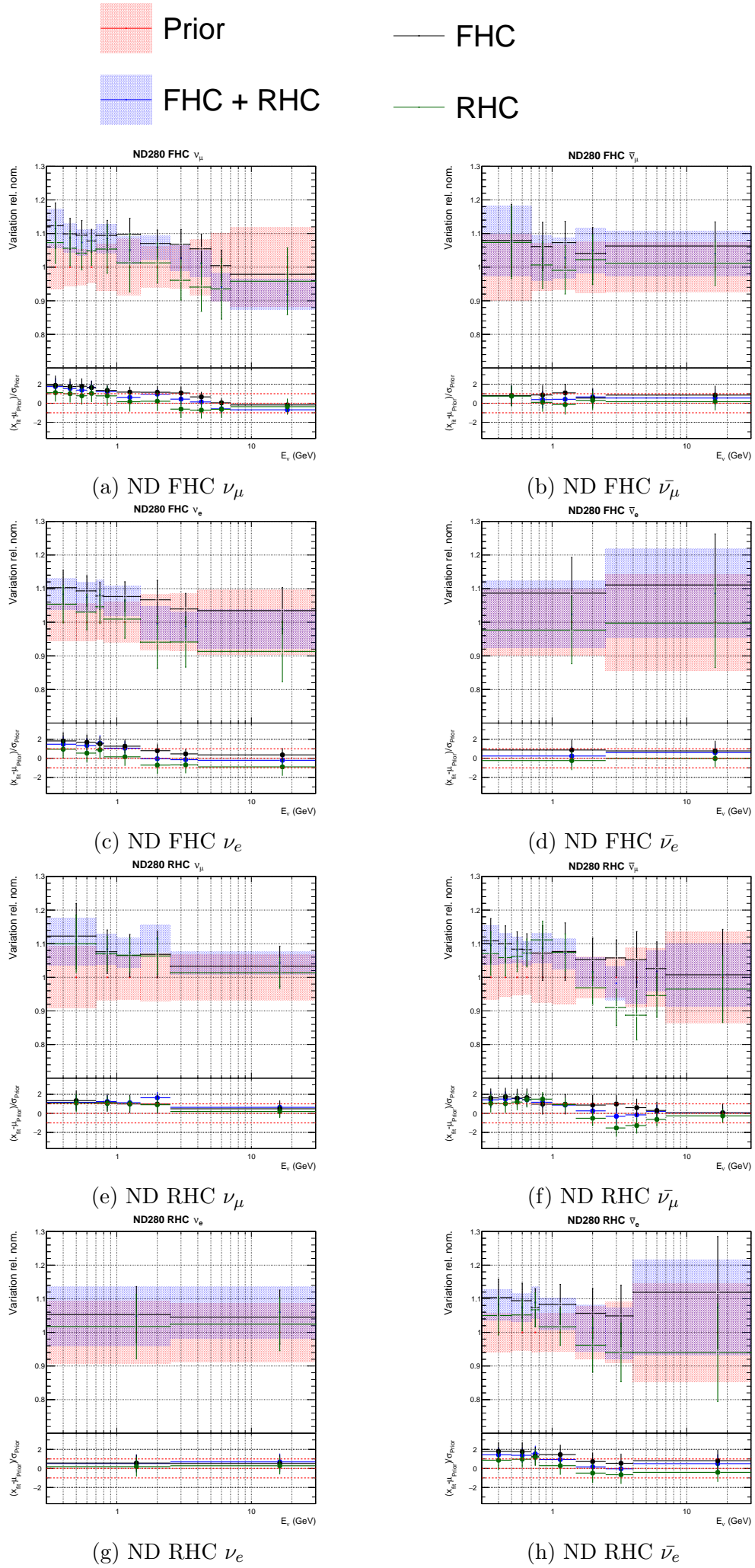
Figure I.5.: ND flux parameters for the FHC and RHC only fits.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
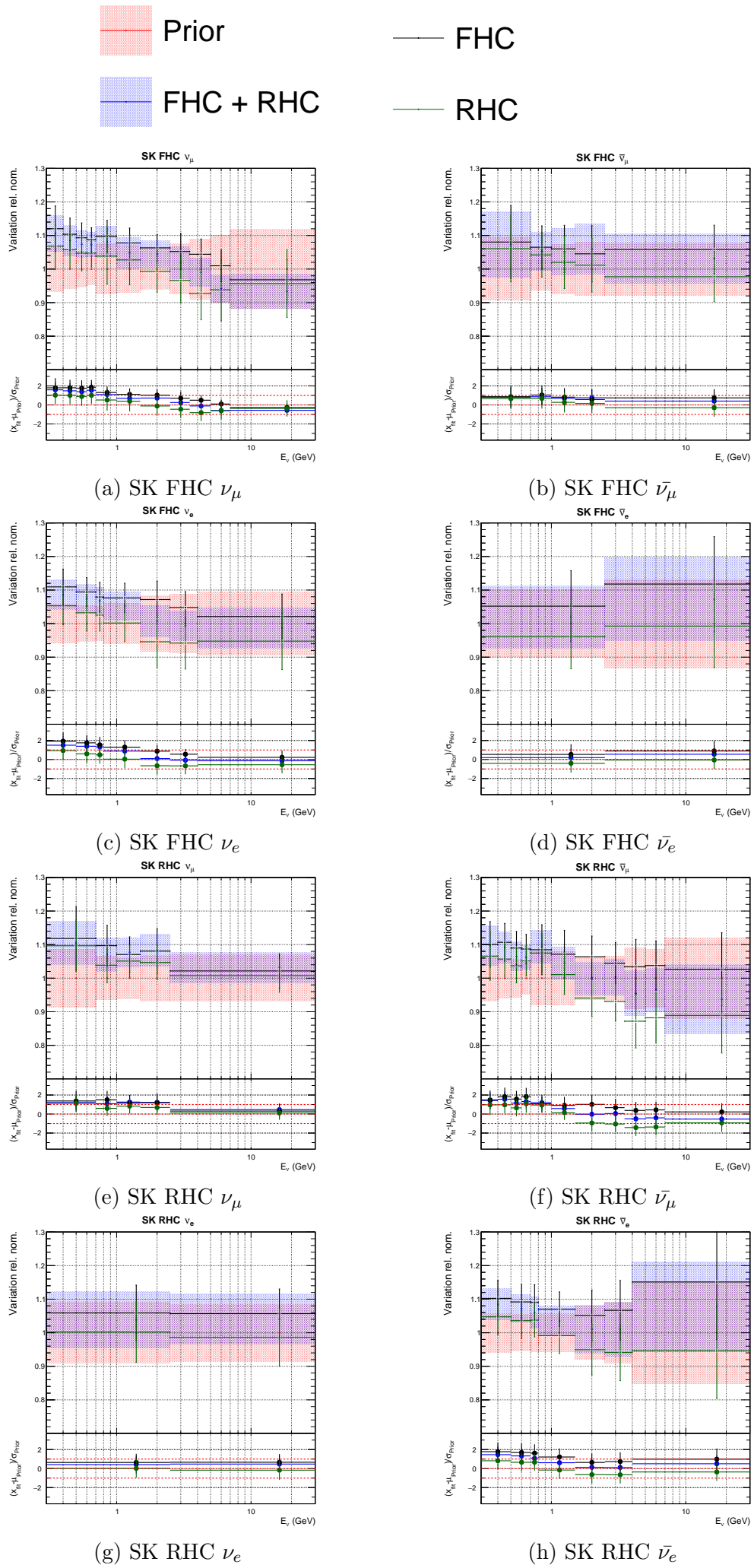
Figure I.6.: SK flux parameters for the FHC and RHC only fits.

2p2h ν normalisation gets a small constraint in the RHC fit from the ν in $\bar{\nu}$ samples, but the 2p2h $\bar{\nu}$ normalisation gets no constraint in the FHC only fit. The 2p2h C to O normalisation and shape parameters for the FHC and RHC only fits are 1σ either side of the full fits.

The Q^2 normalisations are closer to nominal for RHC, but the shape of the increase with increasing Q^2 is similar.

The E_b distributions are shown in Figure I.7 as they are non-Gaussian and so the single extracted values do not show the whole story. The distributions for $E_b\nu$ C parameter are similar for the FHC and full fits, as would be expected. In the RHC fit, it is only constrained through the prior correlations to the FHC E_b parameters and the wrong-sign background samples. Similarly, $E_b\bar{\nu}$ C is very similar for RHC only and the full fit. As there are no FHC wrong-sign samples, $E_b\bar{\nu}$ C is only constrained through the prior uncertainty and correlations in FHC only fit, and so the postfit value is close to the prior central value. The E_b O parameters are both much more consistent across the three fits.

M_A^{RES} is much closer to nominal for RHC, but the full fit favours the FHC value. The BY parameters are both also closer to nominal for RHC, with the full fit values lying between the FHC and RHC only values. The NC parameters are all consistent for the fits, and of the FSI parameters, high energy quasi-elastic is the only systematic for which the postfit values are outside 1σ of each other, with the full fit favouring the FHC value.

Overall these fits show some tension, with the postfit values being generally closer to nominal for RHC only. However, this is likely because there are lower statistics for RHC, so the contribution to the likelihood from the prior uncertainty becomes more significant.

I.3. New and Old Data Only Fits

The previous oscillation analysis [4] used data from runs 2–6. The addition of runs 7, 8 and 9 for this analysis approximately doubles both the FHC and RHC data. To investigate the compatibility of the new and old data, fits were run using data from just runs 2–6, and just runs 7–9. This will show if unexpected results are coming from the changes in the model or the extra data. There is approximately the same amount of data in each set, and so the constraint is expected to be similar.

The results for the flux parameters are shown in Figures I.9 and I.10. They are mostly consistent between the three fits, with the postfit parameters lying within the uncertainty of each other. For runs 7–9, the fluxes tend to be slightly closer to nominal, but the overall shape of the pulls in energy are similar.

I. Alternative Fit Studies

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
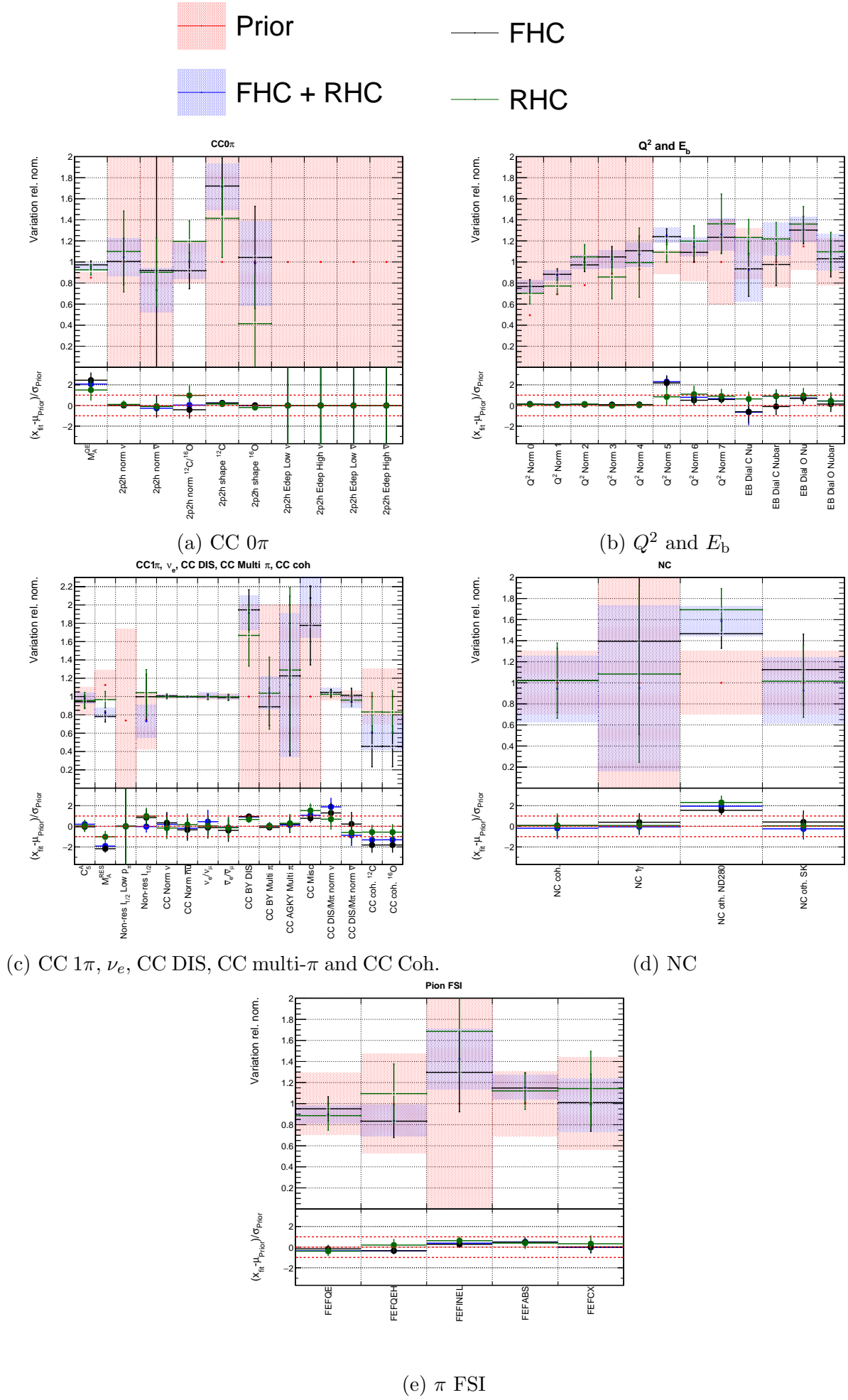
Figure I.7.: Interaction parameters for the FHC and RHC only fits.

(a) $E_{\mathrm{b}}\nu$ C

(b) $E_{\mathrm{b}}\bar{\nu}$ C

(c) $E_{\mathrm{b}}\nu$ O

(d) $E_{\mathrm{b}}\bar{\nu}$ O
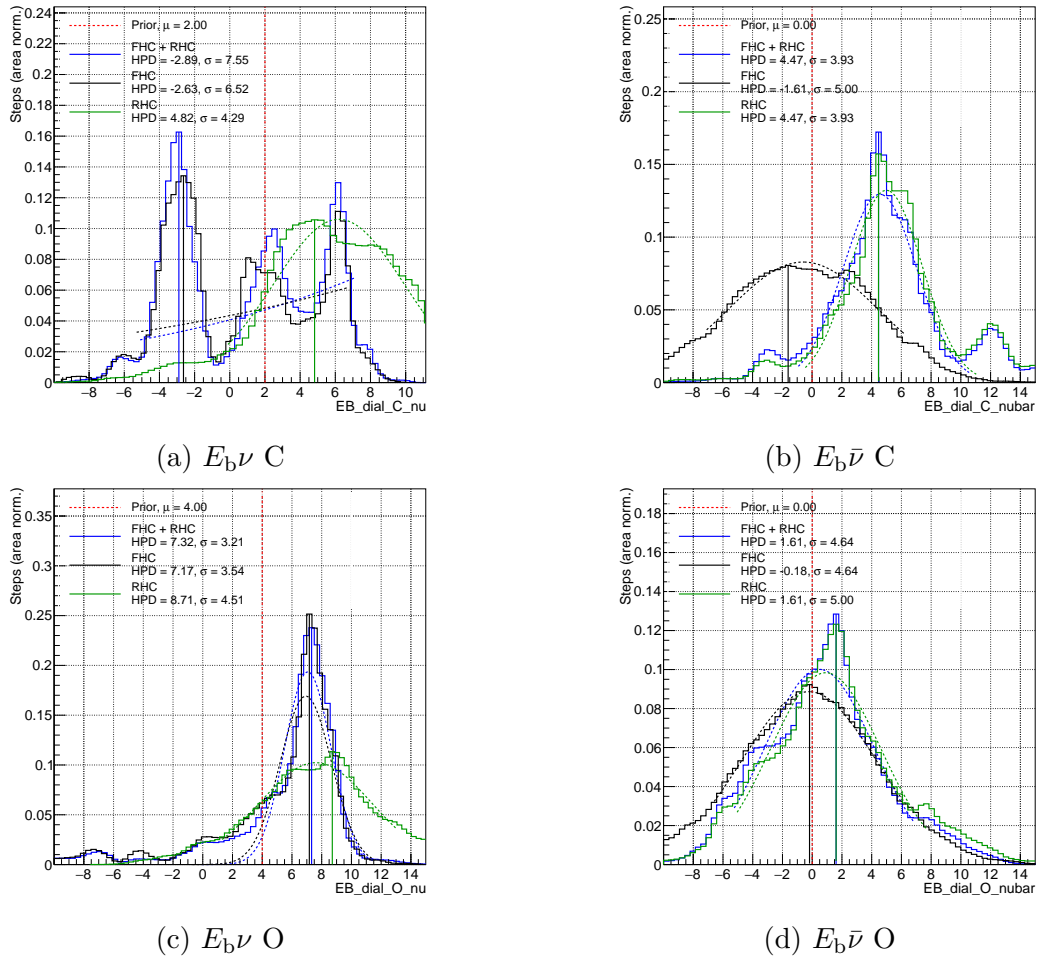
Figure I.8.: Posterior distributions from the binding energy parameters for the FHC and RHC only fits.

The interaction parameters, shown in Figure I.11, show more differences between the fits. M_A^{QE} is very consistent, but the 2p2h normalisations are both closer to nominal for runs 2–6, as is 2p2h shape C. Interestingly, the full fit value for 2p2h shape C to O is beyond both the subset fits. The Q^2 normalisations are all very similar for the three fits.

The E_{b} distributions are shown in Figure I.12, as they are non-Gaussian and so the single extracted values alone are not sufficient. The $E_{\mathrm{b}}\nu$ C and $E_{\mathrm{b}}\nu$ O parameters both have large peaks at higher values for runs 2–6. The full fit prefers the runs 7–9 value for C, and the runs 2–6 value for O. The full fit distributions for the $E_{\mathrm{b}}\bar{\nu}$ C and $E_{\mathrm{b}}\bar{\nu}$ O parameters are both similar to those for runs 7–9. The distributions for runs 2–6 have more peaks, but center around similar values.

As runs 2–6 have a larger proportion of FHC data, it would be expected that more of the full fit constraint on ν parameters would be coming from runs 2–6, and more of the full fit constraint on $\bar{\nu}$ parameters would be coming from runs 7–9. However, there does not seem

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
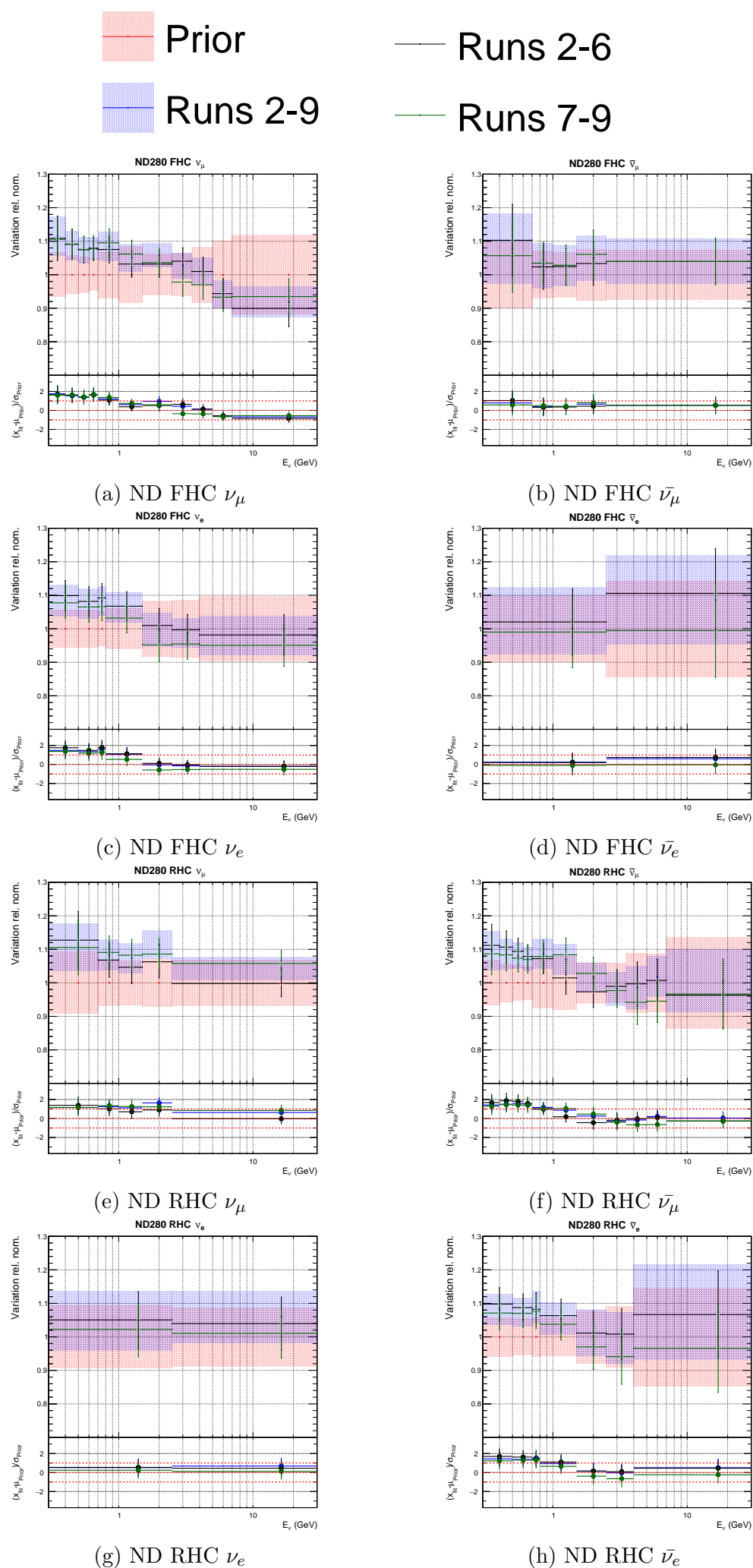
Figure I.9.: ND280 flux parameters for the new and old data only fits.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
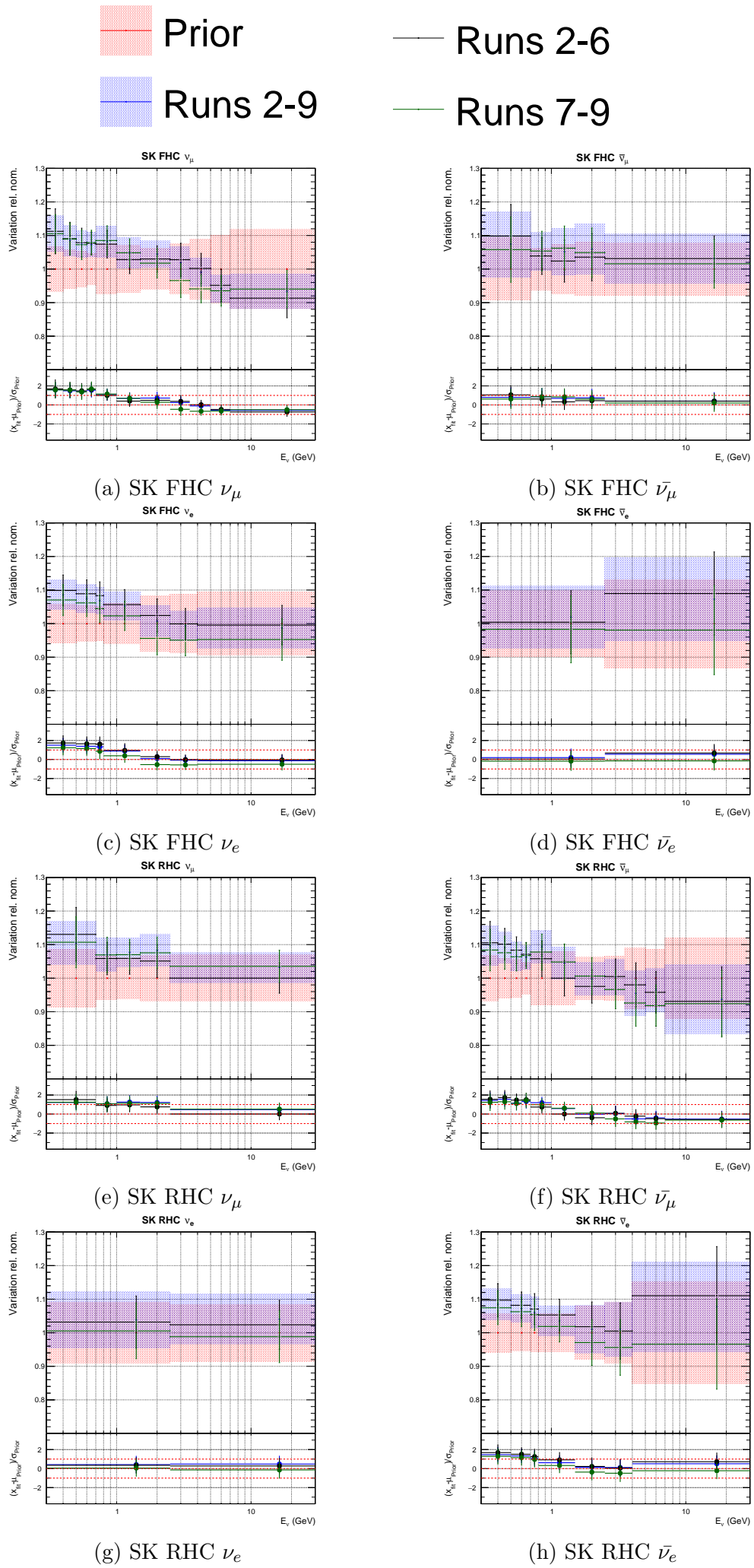
Figure I.10.: SK flux parameters for the new and old data only fits.

(a) CC 0π

(b) Q^2 and $E_{\rm b}$

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
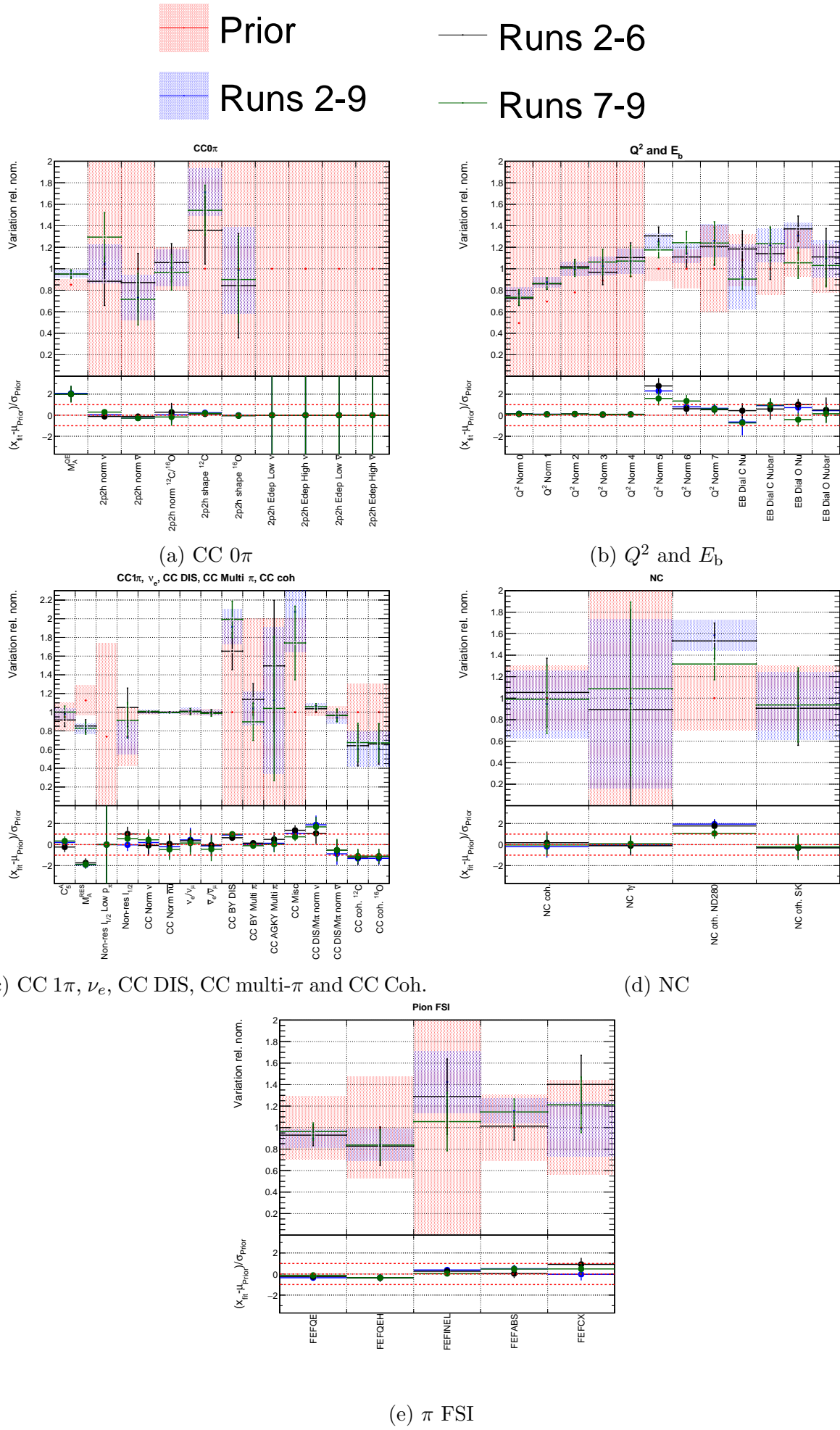
Figure I.11.: Interaction parameters for the new and old data only fits.

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
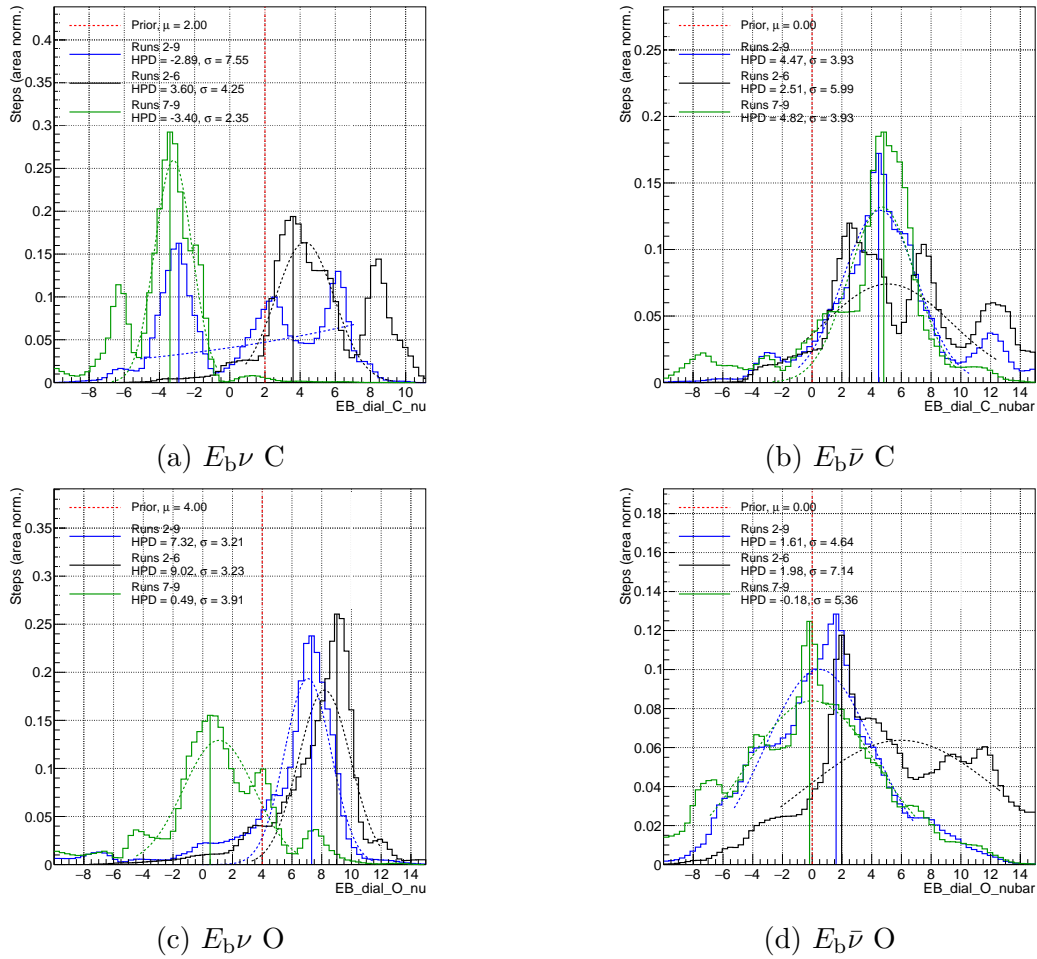
Figure I.12.: Posterior distributions from the binding energy parameters for the new and old data only fits.

to be a general trend of the full fit parameters for ν or $\bar{\nu}$ being closer to the runs 2–6 or 7–9 values. This is probably because the proportions of FHC/RHC data in each set are vastly different, and so other factors dominate.

For runs 7–9, CC DIS BY is higher and CC BY multi-π, CC AGKY mult., and CC misc. are lower than for runs 2–6, with the full fit values lying between the two. The rest of the CC 1π and CC Other targetting parameters are fairly consistent. The inelastic and charge exchange FSI parameters are both closer to nominal for runs 7–9.

Overall, there is some tension between these fits, but nearly all parameters are within 1σ of the each other. In most cases, the full fit value lies between the runs 2–6 and 7–9 values.

I.4. Flat MAQE Prior and Less Q^2 Freedom

Two interaction models were proposed for this analysis. As M_A^{QE} affects events with high Q^2, one model had only five Q^2 normalisations (up to 0.25 GeV), but with a flat prior on M_A^{QE}. The other model invoked a tight prior uncertainty of 0.06 GeV² on M_A^{QE}, but had extra high Q^2 freedom from the three additional high Q^2 normalisations. Fits were ran with both these models.

The comparison of the results for the flux parameters are shown in Figures I.13 and I.14. The shape of the pulls in energy are similar for the two fits, but the fluxes are generally slightly lower for the fit with eight Q^2 normalisations.

Similar results are seen for the ND280 and SK flux parameters, as would be expected.

The interaction parameters are shown in Figure I.16. M_A^{QE} is pushed even higher when it does not have a prior uncertainty. The 2p2h normalisations are lower for having eight Q^2 normalisations, and the five low Q^2 normalisations themselves are slightly higher. The low Q^2 parameters are likely moved by their co-correlations with the fluxes, which control all regions of Q^2.

The comparison of the 1D distributions for the E_b parameters are shown in Figure I.16. There are peaks in similar locations for $E_b\nu$ C, but the lower peak is larger for having five Q^2 normalisations, whereas the higher peak is larger for having eight. The other E_b parameters have similar shaped distributions, but with small changes in the relative sizes of peaks causing changes in the constraints, particularly for $E_b\bar{\nu}$ C and $E_b\nu$ O.

The CC 1π, NC, and π FSI parameters are all unchanged.

The eight Q^2 parameters give better control over high Q^2 events and allows the M_A^{QE} prior uncertainty to be tuned to external data. For this reason, and as the fit results were similar, it was decided the eight Q^2 normalisation with 0.6 GeV² M_A^{QE} prior uncertainty would be used in the final fits. It should be noted that these fits were run before other inputs to the fits were finalised, and so the results aren't the same as those in Section 6.5, but the conclusions should not be different.

I.5. Fixing the 2p2ph Energy Dependence Parameters

The 2p2h energy dependence parameters were found to have no constraint from the near detector fit, as shown in Figure I.17. This causes large marginalisation effects, and also can prevent the other near detector fitting group's gradient descent fit from converging. It was

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
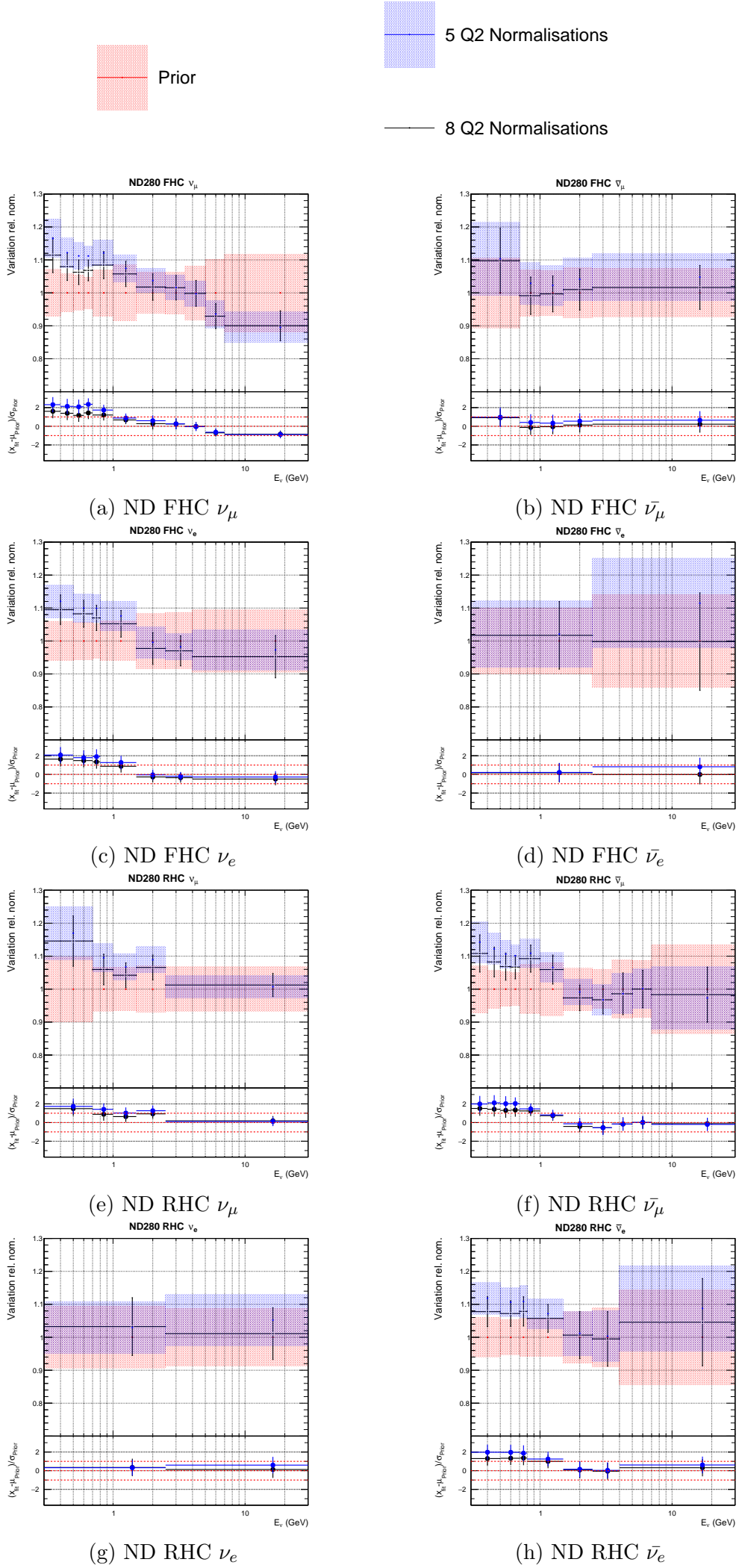
Figure I.13.: Comparison of ND280 flux parameters for the five and eight Q^2 normalisation models.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
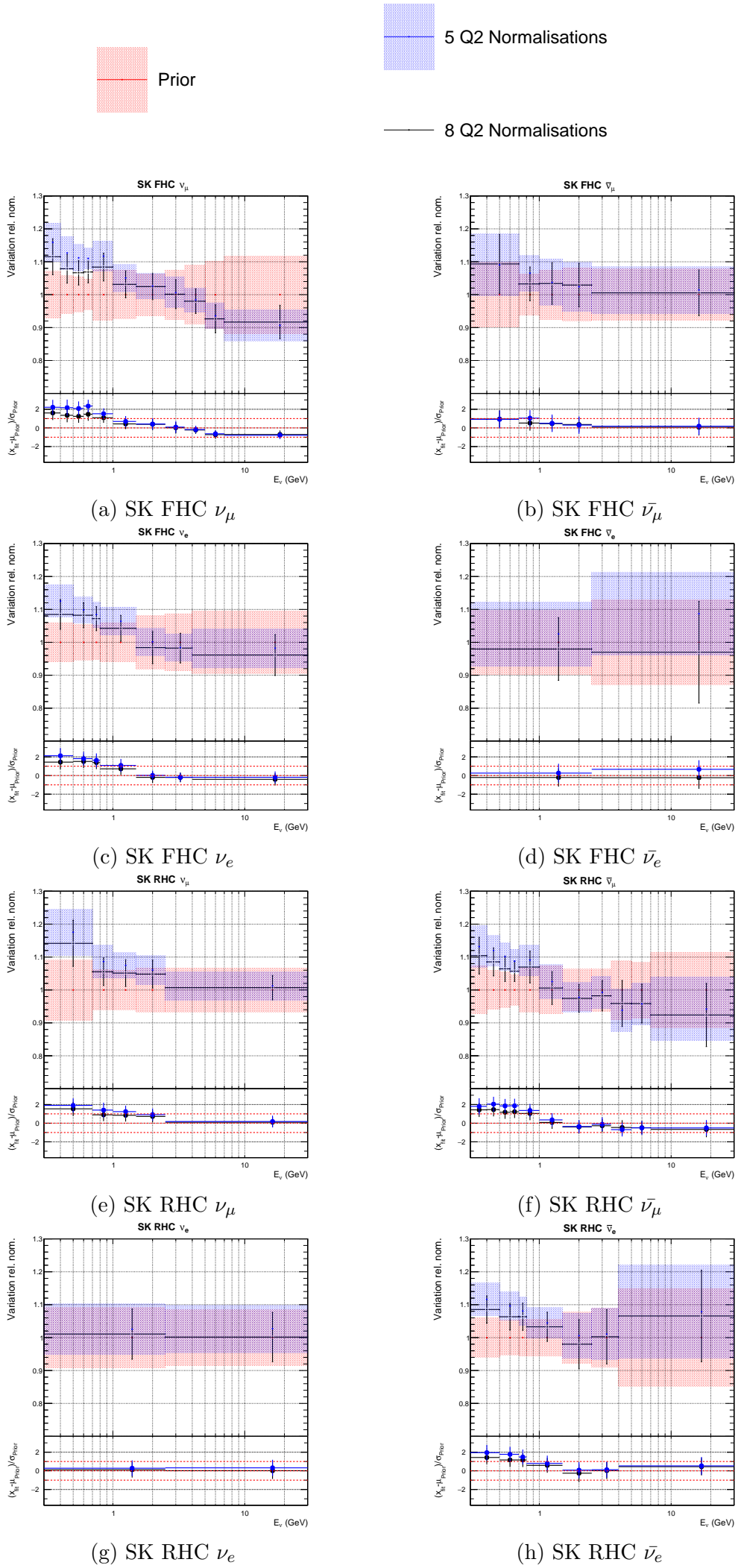
Figure I.14.: Comparison of SK flux parameters for the five and eight Q^2 normalisation models.

(a) CC 0π

(b) Q^2 and $E_{\rm b}$

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI

Figure I.15.: Comparison of interaction parameters for the five and eight Q^2 normalisation models.

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
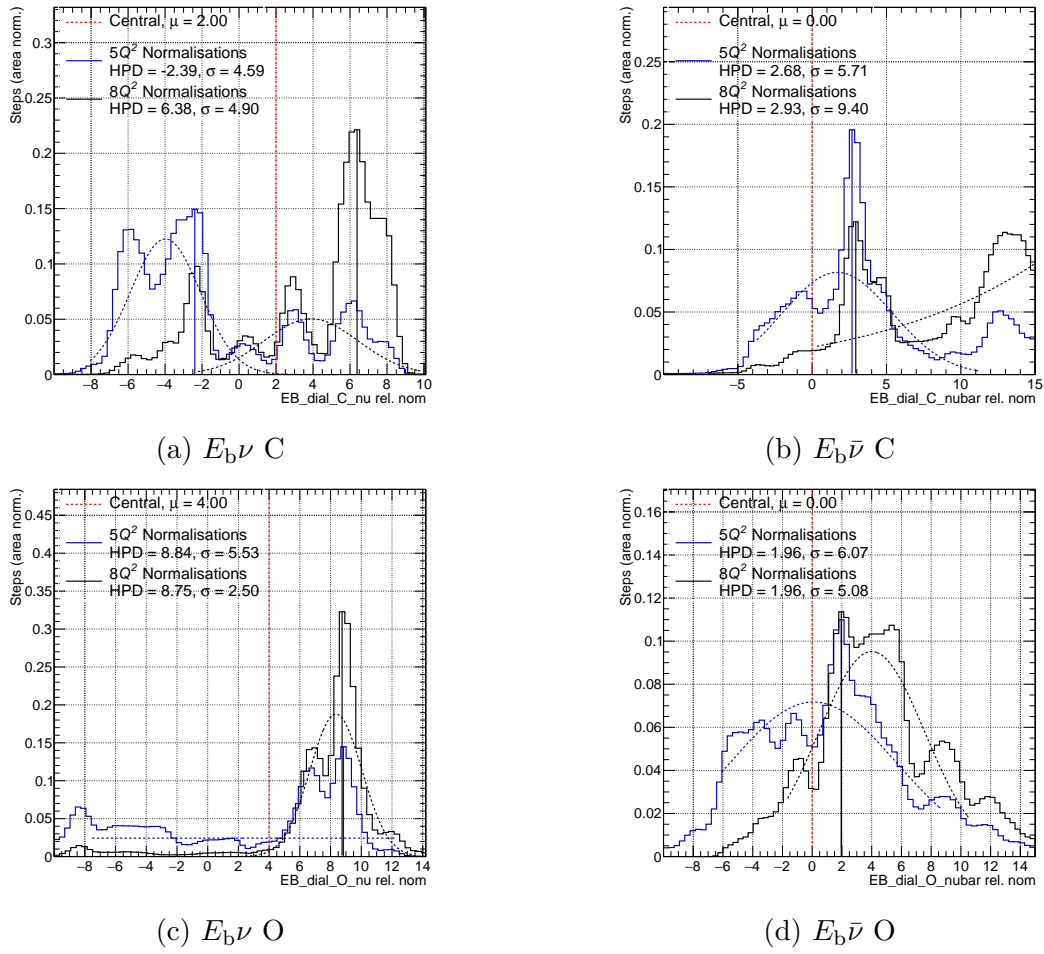
Figure I.16.: Posterior distributions from the binding energy parameters for the five and eight Q^2 normalisation models.

therefore decided that the parameters would not be fitted in the near detector analysis. To make sure this did not have any unintended consequences, fits were run with them free and fixed at their prior central values.

The results are in Figures I.18, I.19, I.20, and I.21. As expected, there is very little change in the postfit parameter values. The 2p2h energy dependence parameters were therefore fixed in this analysis, but are free for the full joint fits.

(a) Low Energy ν

(b) Low Energy $\bar{\nu}$

(c) High Energy ν
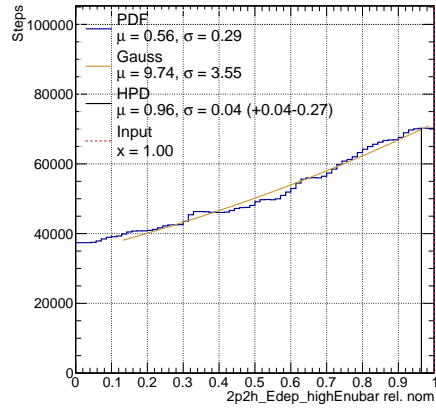
(d) High Energy $\bar{\nu}$

Figure I.17.: Posterior distributions for the 2p2h energy dependence parameters.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
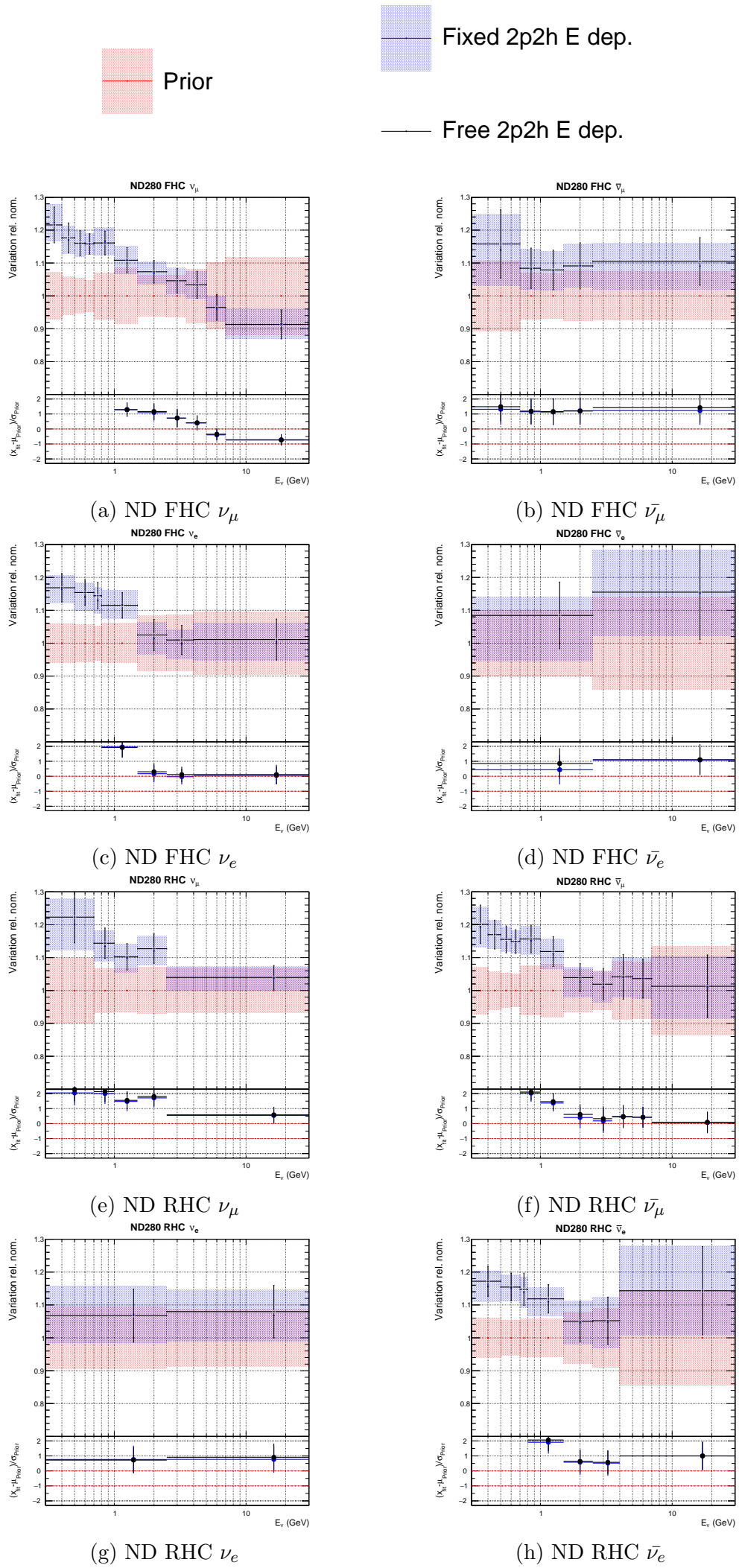
Figure I.18.: Comparison of ND280 flux parameters with 2p2h energy dependence fixed and free.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
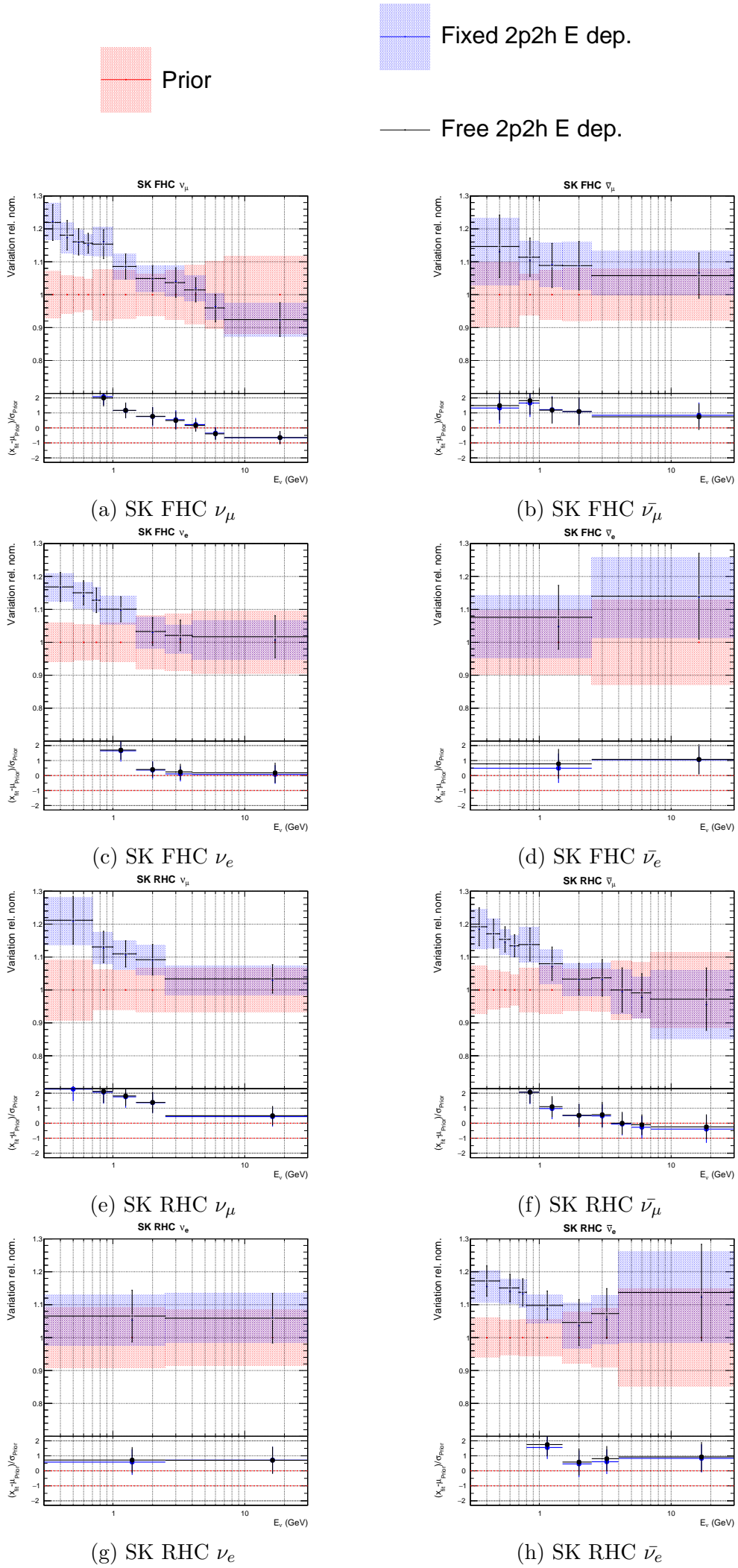
Figure I.19.: Comparison of SK flux parameters with 2p2h energy dependence fixed and free.

(a) CC 0π

(b) Q^2 and E_b

(c) CC 1π, ν_e, CC DIS, CC multi-π and CC Coh.

(d) NC

(e) π FSI
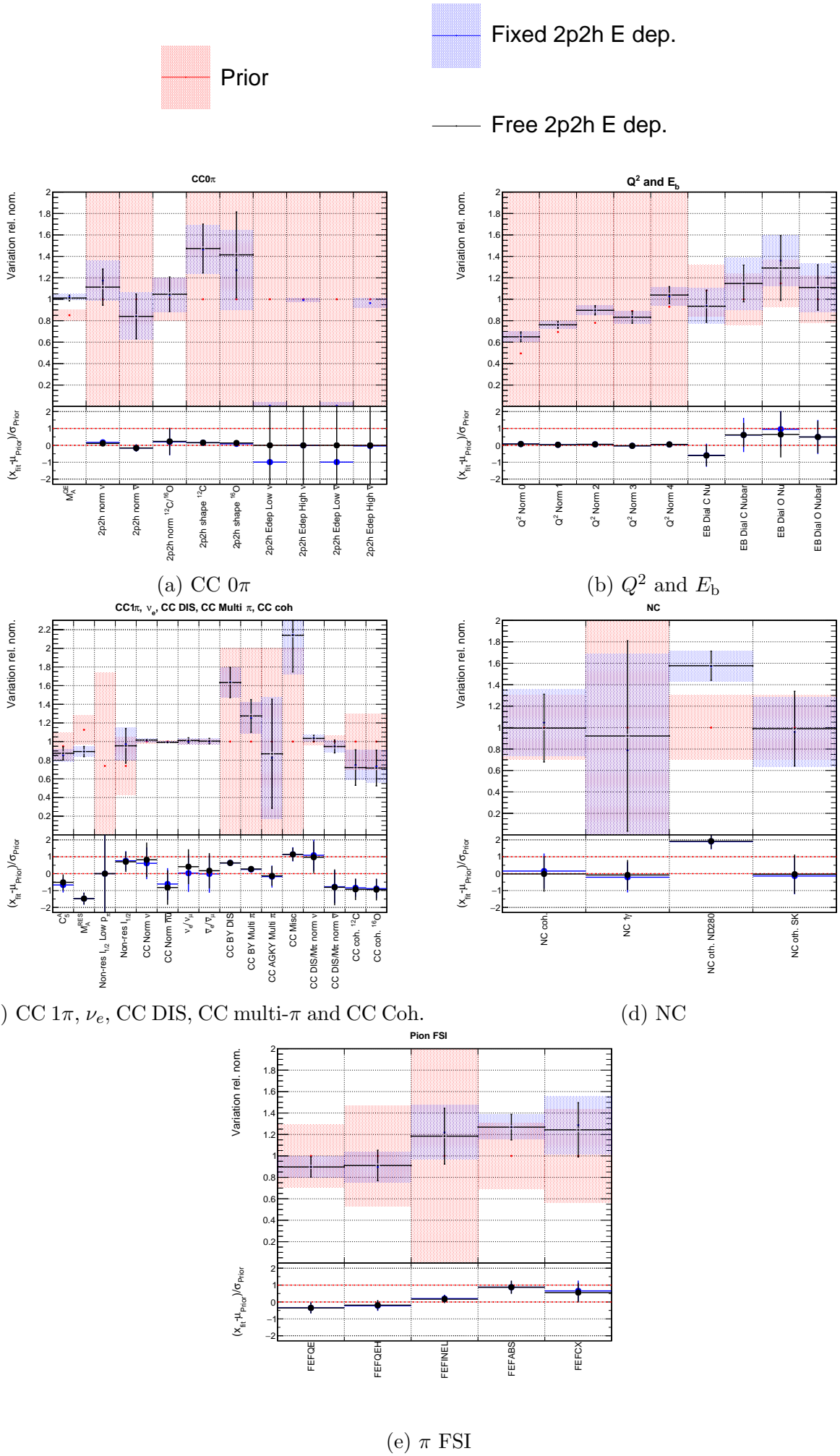
Figure I.20.: Comparison of interaction parameters with 2p2h energy dependence fixed and free.

(a) $E_b\nu$ C

(b) $E_b\bar{\nu}$ C

(c) $E_b\nu$ O

(d) $E_b\bar{\nu}$ O
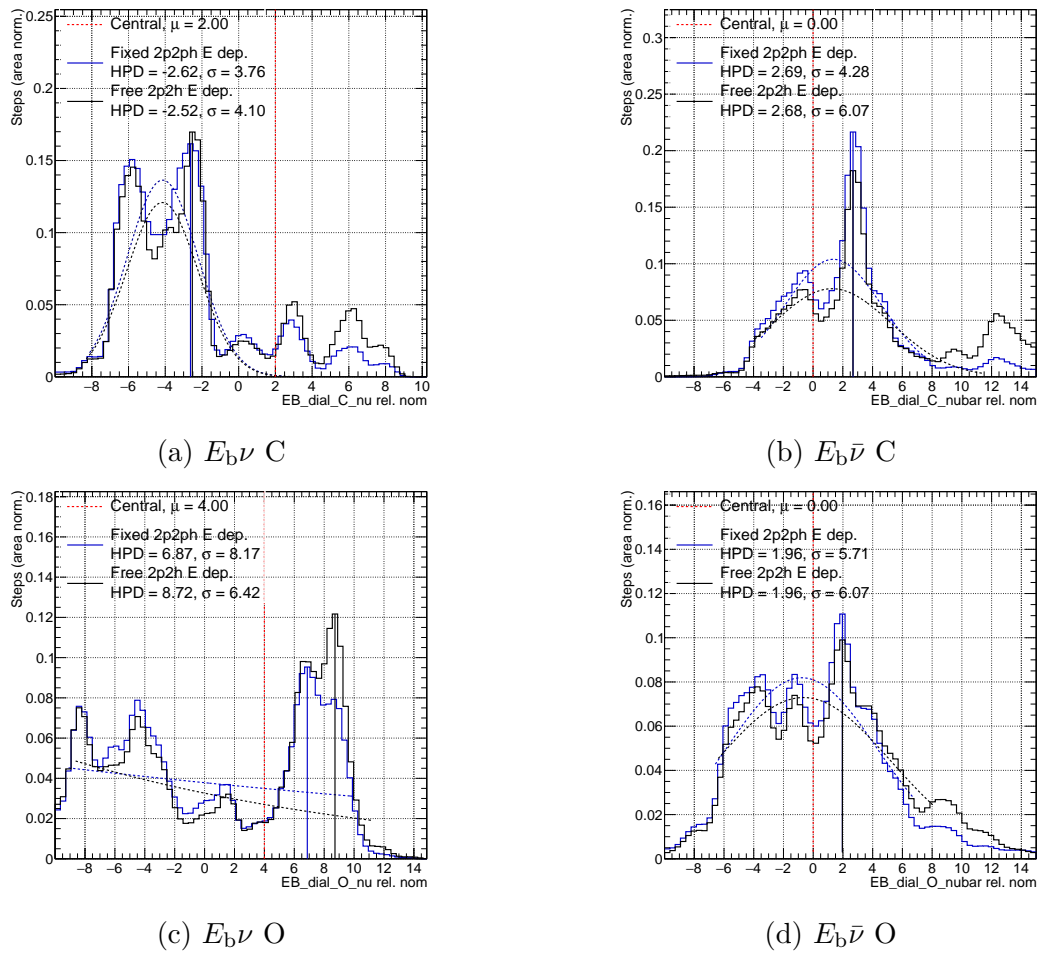
Figure I.21.: Posterior distributions from the binding energy parameters with 2p2h energy dependence fixed and free.

J. RHC Multi Pi Samples Validation

Various fits were run to validate the implementation of the RHC samples being divided by π, rather than track, multiplicity. Data fits using only runs 2–4 and 8 were run, as the FHC sample should be unaffected by the changes. The results are shown in Figures J.1, J.2 and J.3, and as expected, there is no difference using the RHC multi-π or RHC multi-track sample implementations.

To see the full impact of the change in samples, Asimov and data fits were run with the multi-π and track implementations, using FHC and RHC data (runs 2–8). The Asimov fit results, shown in Figures J.5, J.5 and J.6, are very similar for the two samples. The slight differences are due to marginalisation effects, and the two fits are entirely compatible. There is a very slight reduction in uncertainties using the RHC multi-π samples, showing a small improvement in sensitivity.

The data fits using both FHC and RHC data have more differences. The flux parameters, shown in Figures J.7 and J.8, are pulled further from nominal at low energies. The oscillatory shape of the pulls in energy are similar for the two selections, and postfit values are consistently within 1σ of each other. The ND and SK flux parameters have similar behaviour.

The interaction parameters, shown in J.9, also have fairly significant differences. The 2p2h normalisations for ν and $\bar{\nu}$ are both closer to nominal using the mulit-π samples, whereas the 2p2h shape parameter on C is pulled about 1σ further. The CC 1 π parameters are very compatible for the two selections, but the CC DIS parameter is closer to nominal for the multi-π samples. The NC Other (NC 1π and NC DIS) is the only NC parameter to show any differences, being pulled $>1\sigma$ higher. For the π FSI parameters, the high energy quasi-elastic and π absorption parameters are closer to nominal, whereas the charge exchange parameter is pulled further away for the multi-π samples.

Overall the fits are largely compatible, with parameters mostly being within 1σ for the two selections, and often being closer to nominal for multi-π.

These validation fits were run using an intermediate cross-section model made up of the

one used for the 2017 oscillation analysis, which is described in more detail in [140], plus normalisations for the CC ν and $\bar{\nu}$ cross-section, Coulomb corrections, and binding energy dials.

(a) ND FHC ν_μ (b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e (d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ (f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e (h) ND RHC $\bar{\nu}_e$

Figure J.1.: Comparison of ND280 flux parameters for RHC multi-π and multi-track data fits using FHC only data.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$

Figure J.2.: Comparison of SK flux parameters for RHC multi-π and multi-track data fits using FHC only data.

(a) CC 0π

(b) CC 1π, ν_e, CC DIS, and CC coh.

(c) NC

(d) π FSI and E_b

Figure J.3.: Comparison of interaction parameters for RHC multi-π and multi-track data fits using FHC only data.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$
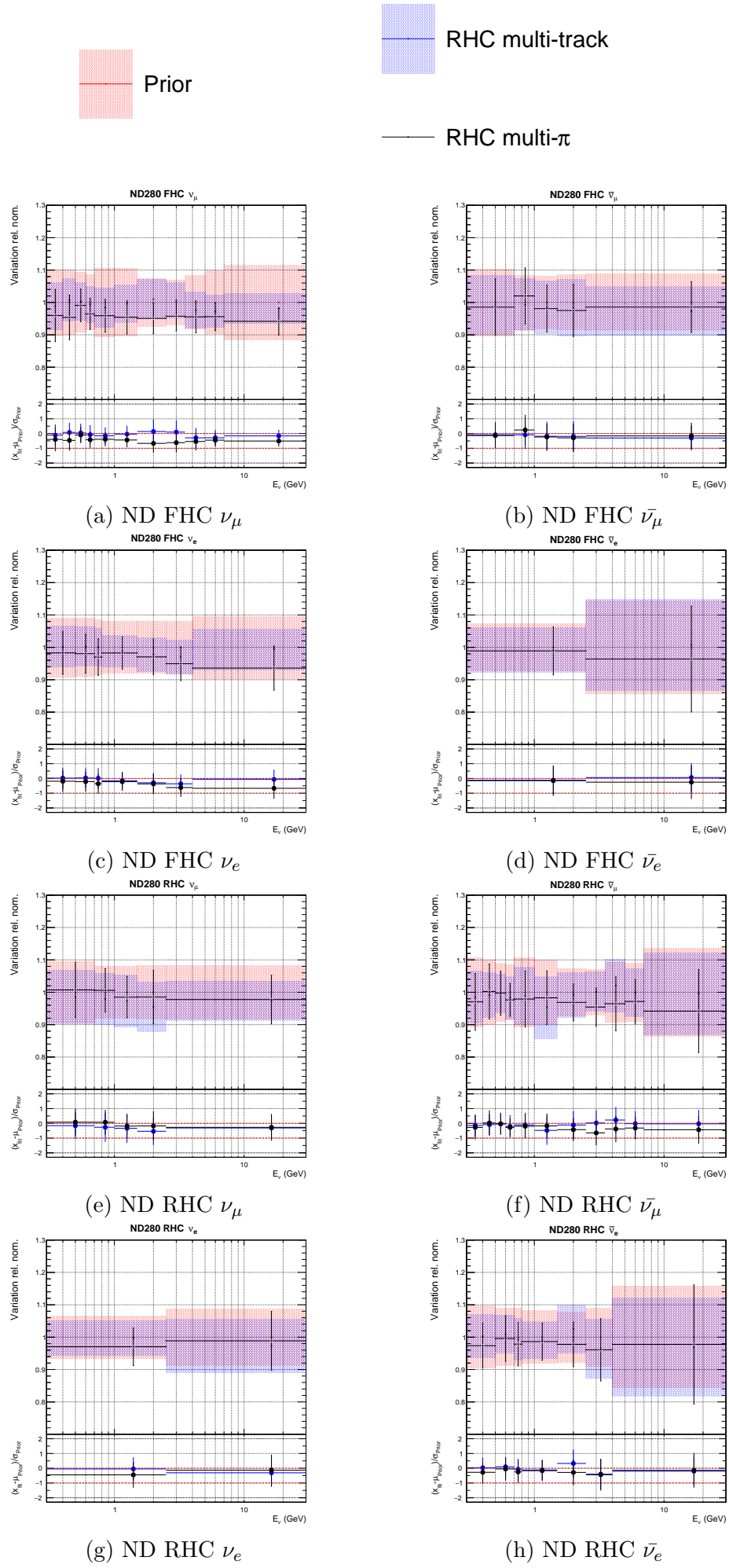
Figure J.4.: Comparison of ND280 flux parameters for RHC multi-π and multi-track Asimov fits using FHC and RHC data.

Legend: Prior — RHC multi-track — RHC multi-π

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK RHC ν_e

(h) SK RHC $\bar{\nu}_e$
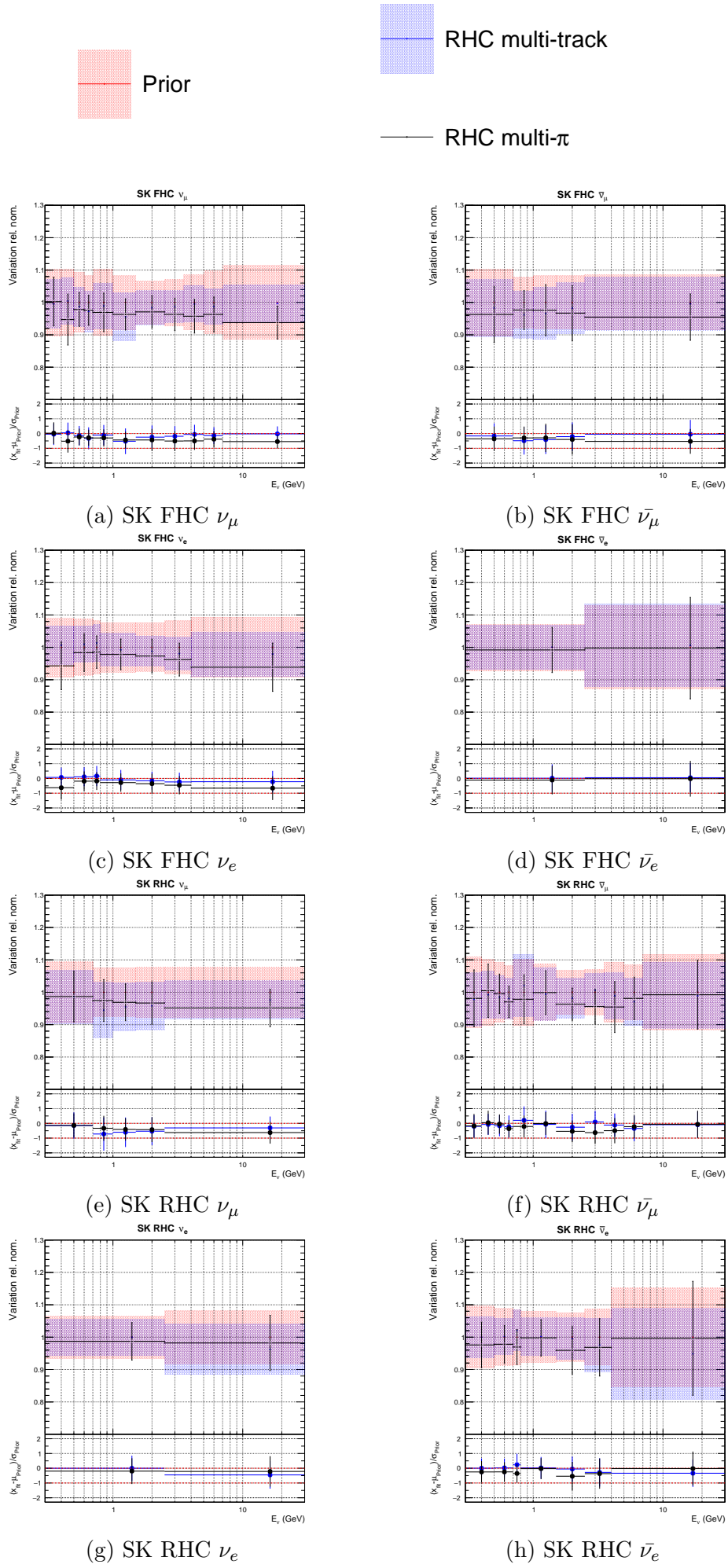
Figure J.5.: Comparison of SK flux parameters for RHC multi-π and multi-track Asimov fits using FHC and RHC data.

(a) CC 0π

(b) CC 1π, ν_e, CC DIS, and CC coh.

(c) NC

(d) π FSI and $E_{\rm b}$

Figure J.6.: Comparison of interaction parameters for RHC multi-π and multi-track Asimov fits using FHC and RHC data.

(a) ND FHC ν_μ

(b) ND FHC $\bar{\nu}_\mu$

(c) ND FHC ν_e

(d) ND FHC $\bar{\nu}_e$

(e) ND RHC ν_μ

(f) ND RHC $\bar{\nu}_\mu$

(g) ND RHC ν_e

(h) ND RHC $\bar{\nu}_e$

Figure J.7.: Comparison of ND280 flux parameters for RHC multi-π and multi-track data fits using FHC and RHC data.

(a) SK FHC ν_μ

(b) SK FHC $\bar{\nu}_\mu$

(c) SK FHC ν_e

(d) SK FHC $\bar{\nu}_e$

(e) SK RHC ν_μ

(f) SK RHC $\bar{\nu}_\mu$

(g) SK FHC ν_e

(h) SK RHC $\bar{\nu}_e$

Figure J.8.: Comparison of SK flux parameters for RHC multi-π and multi-track data fits using FHC and RHC data.

(a) CC 0π

(b) CC 1π, ν_e, CC DIS, and CC coh

(c) NC

(d) π FSI and E_b

Figure J.9.: Comparison of interaction parameters for RHC multi-π and multi-track data fits using FHC and RHC data.

Bibliography

[1] A. D. Sakharov. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5, 1967.

[2] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π Decay of the K^0 Meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[3] M. Fukugita and T. Yanagida. Baryogenesis without Grand Unification. *Phys. Lett. B*, 174:45–47, 1986.

[4] K. Abe et al. Search for CP Violation in Neutrino and Antineutrino Oscillations by the T2K Experiment with 2.2×10^{21} Protons on Target. *Phys. Rev. Lett.*, 121:171802, Oct 2018.

[5] K. Abe et al. Hyper-Kamiokande Design Report. 2018.

[6] R. Acciarri et al. Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report, Volume 4 The DUNE Detectors at LBNF. 2016.

[7] S. Fukuda et al. Determination of Solar Neutrino Oscillation Parameters using 1496 days of Super-Kamiokande-I Data. *Physics Letters B*, 539(3):179 – 187, 2002.

[8] Q. R. Ahmad et al. Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by B^8 Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87(7), Jul 2001.

[9] J. Chadwick. Bakerian Lecture: The Neutron. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Science*, volume 142 of *1*, 1933.

[10] C. D. Ellis and N. F. Mott. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Science*, volume 147 of *502*, 1934.

[11] W. Pauli. In *Rapports du SeptiAlme Conseil de Physique Solvay, GauthieraARVillars, Paris*, 1993.

Bibliography

[12] E. Fermi. *Zeitschrift fur Physik*, 88(161), 1934.

[13] C. L. Cowan, F. Reines, et al. *Science*, 124(103), 1956.

[14] R. Davis. Attempt to Detect the Antineutrinos from a Nuclear Reactor by the $Cl^{37}(\nu, e^-)A^{37}$ Reaction. *Phys. Rev.*, 97:766–769, 1955.

[15] G. Danby et al. Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos. *Phys. Rev. Lett.*, 9:36–44, 1962.

[16] M. M. Block et al. Neutrino Interactions in the CERN Heavy Liquid Bubble Chamber. *Phys. Rev. Lett.*, 12 (3):281–285, 1964.

[17] K. Kodama et al. Observation of Tau Neutrino Interactions. *Physics Letters B*, 504 (3):218–224, 2001.

[18] M. L. Perl et al. Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation. *Phys. Rev. Lett.*, 35:1489–1492, 1975.

[19] M. Lattanzi. *Journal of Physics: Conference Series*, 718, 032008, 2016.

[20] The ALEPH Collaboration et al. Precision Electroweak Measurements on the Z Resonance. *Physics Reports*, 427:257–454, 2006.

[21] J. Bahcall and A. Serenelli. New Solar Opacities, Abundances, Helioseismology, and Neutrino Fluxes. *Astrophys. J.*, 621:85–L88, 2005.

[22] R. Davis, D. S. Harmer, and Hoffman K. C. Search for Neutrinos from the Sun. *Phys. Rev. Lett.*, 20:1205–1209, 1968.

[23] B. Pontecorvo. Neutrino Experiments and the Problem of Conservation of Leptonic Charge. *Sov. Phys. JETP*, 26:984, 1968.

[24] N. Cabibbo. *Phys. Rev. Lett.*, 10:531, 1963.

[25] M. Kobayashi and T. Maskawa. *Progress of Theoretical Physics*, 49:652, 1973.

[26] Hirata K. S. et al. *Phys. Rev. Lett.*, 63:16, 1989.

[27] W. Hampel et al. *Physics Letters B*, 447:127, 1999.

[28] J. N. Abdurashitov et al. *Phys. Rev. C*, 60, 1999.

[29] Q. R. Ahmad et al. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89, 2002.

[30] S. Basu J.N. Bahcall, M.H. Pinsonneault. *Astrophys. J.*, 555:990, 2001.

[31] F. Reines et al. *Phys. Rev. Lett.*, 15, 1965.

[32] T. J. Haines et al. *Phys. Rev. Lett.*, 57, 1986.

[33] K. Hirata et al. *Physics Letters B*, 205:416, 1988.

[34] T. Kajita et al. Establishing Atmospheric Neutrino Oscillations with Super-Kamiokande. *Nuclear Physics B*, 908:14–29, 2016.

[35] K. Eguchi et al. First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance. *Phys. Rev. Lett.*, 90, 2003.

[36] S. Abe et al. Precision Measurement of Neutrino Oscillation Parameters with Kam-LAND. *Phys. Rev. Lett.*, 100:221803, Jun 2008.

[37] J. K. Ahn et al. *Phys. Rev. Lett.*, 108, 2012.

[38] F. P. An et al. *Phys. Rev. Lett.*, 108, 2012.

[39] Y. Abe et al. *Journal of High Energy Physics 2016*, 163, 2016.

[40] M. Apollonio et al. Chooz Collaboration. *Eur. Phys. J.*, 27, 2003.

[41] P. Novella. The Antineutrino Energy Structure in Reactor Experiments. *Advances in High Energy Physics*, 2015, 2015.

[42] M. H. Ahn et al. *Phys. Rev. D*, 74, 2006.

[43] D. G. Michael et al. *Phys. Rev. Lett.*, 97, 2006.

[44] P. Adamson et al. *Phys. Rev. Lett.*, 116, 2016.

[45] R. Acquafredda et al. *Journal of Instrumentation*, 4:P04018, 2009.

[46] A. Aguilar et al. *Phys. Rev. D*, 64, 2001.

[47] K. Eitel et al. *Nuclear Physics B-Proceedings Supplements*, 91(1-3):191–197, 2001.

[48] S. Amerio et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 527:329, 2004.

[49] A. A. Aguilar-Arevalo et al. *Phys. Rev. Lett.*, 110, 2013.

[50] A. A. Aguilar-Arevalo et al. *Phys. Rev. Lett.*, 121, 2018.

Bibliography

[51] V. Gribov and B. Pontecorvo. *Physics Letters B*, 28:493, 1969.

[52] X. Qian et al. Neutrino Mass Hierarchy. *Progress of Particle and Nuclear Physics*, 83, 2015.

[53] M. Ettore and L. Maiani. Teoria Simmetrica dell'Elettrone e del Positrone. *Il Nuovo Cimento*, 14 (4):171–84, 1937.

[54] S. Mikheyev and A. Smirnov. *Soviet Journal of Nuclear Physics*, 42:913, 1985.

[55] K. Hagiwara, N. Okamura, and K. Senda. The Earth Matter Effects in Neutrino Oscillation Experiments from Tokai to Kamioka and Korea. *Journal of High Energy Physics*, 2011 (9):82, 2011.

[56] M. E. Shaposhnikov S. Y. Khlebnikov. The Statistical Theory of Anomalous Fermion Number Non-Conservation. *Nuclear Physics B*, 308:885–912, 1988.

[57] A. Riotto S. Pascoli, S.T. Petcov. *Phys. Rev. D*, 75:083511, 2007.

[58] A. Riotto S. Pascoli, S.T. Petcov. *Phys. Rev. B*, 774:1, 2007.

[59] L. Alvarez-Ruso, Y. Hayato, and J. Nieves. *New Journal of Physics*, 16, 2014.

[60] J. G. Morfin, J. Nieves, and J. Sobczyk. Recent Developments in Neutrino/Antineutrino - Nucleus Interactions. *Advances in High Energy Physics*, 6, 2012.

[61] A. A. Aguilar-Arevalo et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 599:28, 2009.

[62] P. A. N. Machado and O. S. Palamara. The Short-Baseline Neutrino Program at Fermilab. *Annual Review of Nuclear and Particle Science*, 69:363–387, 2019.

[63] G. Bellini, J. Benziger, D. Bick, S. Bonetti, et al. First Evidence of *pep* Solar Neutrinos by Direct Detection in Borexino. *Phys. Rev. Lett.*, 108(5):051302, 2012.

[64] C. Kraus et al. SNO with Liquid Scintillator: SNO+. *Progress in Particle and Nuclear Physics*, 57(1):150–152, 2006.

[65] A. Karle, J. Ahrens, J. N. Bahcall, et al. IceCube—The Next Generation Neutrino Telescope at the South Pole. *Nuclear Physics B-Proceedings Supplements*, 118:388–395, 2003.

[66] M. Ageron, J. A. Aguilar, et al. ANTARES: The First Undersea Neutrino Telescope. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 656(1):11–38, 2011.

[67] I. Alekseev, V. Belov, et al. Search for Sterile Neutrinos at the DANSS Experiment. *Physics Letters B*, 787:56–63, 2018.

[68] Y. J. Ko, B. R. Kim, et al. Sterile Neutrino Search at the NEOS Experiment. *Phys. Rev. Lett.*, 118(12):121802, 2017.

[69] J. Ashenfelter, A. B. Balantekin, et al. First Search for Short-Baseline Neutrino Oscillations at HFIR with PROSPECT. *Phys. Rev. Lett.*, 121(25):251802, 2018.

[70] J. Haser et al. Search for eV Sterile Neutrinos at a Nuclear Reactor—the Stereo Project. In *Journal of Physics: Conference Series*, volume 718, page 062023. IOP Publishing, 2016.

[71] G. Ban, W. Beaumont, et al. SoLid: Search for Oscillations with Lithium-6 Detector at the SCK-CEN BR2 Reactor. *Nuclear and Particle Physics Proceedings*, 273:2690–2692, 2016.

[72] M. Tanabashi et al. (Particle Data Group). *Phys. Rev. D*, 98:030001, 2018 and 2019 update.

[73] G. Ranucci et al. Status and Prospects of the JUNO Experiment. In *Journal of Physics: Conference Series*, volume 888, page 012022. IOP Publishing, 2017.

[74] C. E. Aalseth, E. Aguayo, et al. The Majorana Experiment. *Nuclear Physics Section B Proceedings Supplements*, 217(1):44, 2011.

[75] B. Mong et al. nEXO-Neutrinoless Double Beta Decay Experiment. In *XIII International Conference on Heavy Quarks and Leptons*, volume 274, page 074. SISSA Medialab, 2017.

[76] A. Gando, Y. Gando, et al. Precision Analysis of the Xe-136 Two-Neutrino $\beta\,\beta$ Spectrum in KamLAND-Zen and its Impact on the Quenching of Nuclear Matrix Elements. *Phys. Rev. Lett.*, 122(19):192501, 2019.

[77] M. Drewes. The Phenomonology of Right Handed Neutrinos. *International Journal of Modern Physics E*, 22(08):1330019, Aug 2013.

Bibliography

[78] K. Abe et al. The T2K Experiment. *Nuclear Instruments and Methods in Physics Research A*, 659:106–135, 2011.

[79] P. Adamson et al. *Phys. Rev. Lett.*, 116:151806, 2014.

[80] K. Abe et al. Observation of Electron Neutrino Appearance in a Muon Neutrino Beam. *Phys. Rev. Lett.*, 112:061802, Feb 2014.

[81] T. Fukuda et al. Proposal for Precise Measurement of Neutrino-Water Cross-Section in NINJA Physics Run. *Proposal for J-PARC and KEK*, 2017.

[82] T. Ovsiannikova et al. *Physics of Particles and Nuclei*, 48:1014, 2017.

[83] M. Antonova et al. *Journal of Instrumentation*, 12:C07028, 2017.

[84] Y. Yamazaki et al. Accelerator Technical Design Report for J-PARC. 2003.

[85] L. MyeongJae et al. COMET Muon Conversion Experiment in J-PARC. *Frontiers in Physics*, 6:133, 11 2018.

[86] K. Suzuki, S. Aoki, A. Ariga, T. Ariga, F. Bay, C. Bronner, A. Ereditato, M. Friend, M. Hartz, T. Hiraki, and et al. Measurement of the Muon Beam Direction and Muon Flux for the T2K Neutrino Experiment. *Progress of Theoretical and Experimental Physics*, 2015(5):53C01–0, May 2015.

[87] T. Vladisavljevic. *Predicting the T2K Neutrino Flux and Measuring Oscillation Parameters*. Springer Nature, 2020.

[88] K. Matsuoka et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 624:591, 2010.

[89] Kenji Yasutome. T2K Integrated POT, Feb 2020.

[90] D. Beavis et al. Long Baseline Neutrino Oscillation Experiment, E889, Physics Design Report. Chapter III A:BNL 52459, 1995.

[91] Kirsty Duffy. First Antineutrino Oscillation Results from T2K, 2016.

[92] Pablo Fernandez. Recent Results from the Long-Baseline (LBL) Neutrino Oscillation Experiments, journal = SciPost Physics Proceedings. pages 24–28, 09 2019.

[93] G. Battistoni et al. The FLUKA Code: Description and Benchmarking. *AIP Conference Proceedings*, 896 (1):31–49, 2007.

[94] N. Abgrall et al. Measurements of Cross Sections and Charged Pion Spectra in Proton-Carbon Inter- actions at 31 GeV/c. *Phys. Rev. C*, 84:034604, 2011.

[95] R. Brun, F. Carminiti, and S. Giani. GEANT: Detector Description and Simulation Tool. *Technical Report No. CERN-W5013*, 1994.

[96] K. Abe. et al. The T2K Neutrino Flux Prediction. *Phys. Rev. D*, 87:012001, Jan 2013.

[97] C. Zeitnitz and T. A. Gabriel. The GEANT-CALOR Interface and Benchmark Calculations of ZEUS Test Calorimeters. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 349 (1):106 – 111, 1994.

[98] K. Abe et al. Proposal for an Extended Run of T2K to 20×10^{21} POT. 09 2016.

[99] K. Tatsuya et al. INGRID and MUMON Measurements, Apr 2020.

[100] P.-A. Amaudruz et al. The T2K Fine-Grained Detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 696:1 – 31, 2012.

[101] N. Abgrall et al. Time Projection Chambers for the T2K Near Detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 637(1):25 – 46, 2011.

[102] M. Pimiä. *First Results of the UA1 Experiment*, pages 43–66. Springer US, Boston, MA, 1983.

[103] Y. Hayato. *Acta Physica Polonica B*, 40, 2009.

[104] S. Kasuga et al. A Study on the Electron-Muon Identification Capability of a Water Cherenkov Detector and the Atmospheric Neutrino Problem. *Physics Letters B*, 374:238–242, 1996.

[105] R. Brun et al. GEANT: Detector Description and Simulation Tool. *CERN Program Library*, Long Writeup W5013, 1993.

[106] S. Fukuda et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 418, 2003.

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

Bibliography

[108] N. Metropolis et al. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21:1087–1092, 1953.

[109] W. K. Hastings et al. Monte Carlo Sampling Methods Using Markov Chains and their Applications. *Biometrika*, 57:97–109, 1970.

[110] J. Dunkley et al. Fast and Reliable Markov Chain Monte Carlo Technique for Cosmological Parameter Estimation. *Monthly Notices of the Royal Astronomical Society*, 356:925–936, 2005.

[111] Andrew Gelman, Yuri Goegebeur, Francis Tuerlinckx, and Iven Van Mechelen. Diagnostic Checks for Discrete Data Regression Models using Posterior Predictive Simulations. *Appl. Statist.*, 49 Par 2:247–268, 2000.

[112] Andrew Gelman, Xiao-Li Meng, and Hal Stern. Posterior Predictive Assessment of Model Fitness via Realized Discrepancies. *Statistica Sinica*, 6:733–807, 1996.

[113] A. Gelman. Understanding Posterior p-values. *Electronic Journal of Statistics*, 2013.

[114] S. Bienstock et al. Constraining the Flux and Cross Section Models with Data from the ND280 Detector using FGD1 and FGD2 for the 2017 Joint Oscillation Analysis. *T2K*, 2017.

[115] F. James. MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1. 1994.

[116] P. Bartet et al. ν_μ CC Event Selections in the ND280 Tracker Using Run 2+3+4 Data. *T2K*, 2015.

[117] V. Berardi et al. CC $\bar{\nu}_\mu$ Event Selection in the ND280 Tracker using Run 5c Anti-Neutrino Beam Data. *T2K*, 2015.

[118] R. Brun and F. Rademakers. ROOT - An Object Orientated Data Analysis Framework. *Nucl. Instrum. Meth. in Phys. Res. A*, 389:81–86, 1997.

[119] O. Benhar and A. Fabrocini. Two-Nucleon Spectral Function in Infinite Nuclear Matter. *Phys. Rev. C*, 62:034304, 2000.

[120] S. Bolognesi et al. NIWG Model and Uncertainties for 2019-2020 Oscillation Analysis. *T2K*, 2020.

[121] L. Aliaga et al. Design, Calibration, and Performance of the MINERνA Detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 743:130–159, 2014.

[122] R. Gran, J. Nieves, F. Sanchez, and M. J. Vicente Vacas. Neutrino-Nucleus Quasi-Elastic and 2p2h Interactions up to 10 GeV. *Phys. Rev. D*, 88:113007, Dec 2013.

[123] M. Martini, M. Ericson, G. Chanfray, and J. Marteau. Unified Approach for Nucleon Knock-Out and Coherent and Incoherent Pion Production in Neutrino Interactions with Nuclei. *Phys. Rev. C*, 80:065501, Dec 2009.

[124] D. Rein and L. M. Sehgal. Neutrino-Excitation of Baryon Resonances and Single Pion Production. *Annals of Physics*, 133, 1981.

[125] C. Wilkinson, P. Rodrigues, S. Cartwright, L. Thompson, and K. McFarland. Re-analysis of Bubble Chamber Measurements of Muon-Neutrino Induced Single Pion Production. *Phys. Rev. D*, 90:112017, Dec 2014.

[126] D. Rein and L. M. Sehgal. Coherent π0 Production in Neutrino Reactions. *Nuclear Physics B*, 223(1):29 – 44, 1983.

[127] M. Gluck, E. Reya, and A. Vogt. Dynamical Parton Distributions Revisited. *The European Physical Journal C - Particles and Fields*, 5:461–470, 1998.

[128] A. Bodek and U. Yang. A Unified Model for Inelastic $e - N$ and $\nu - N$ cross-sections at all Q^2. *AIP Conf. Proc.*, 792:257–260, 2005.

[129] T. Yang, C. Andreopoulos, H. Gallagher, and P Kehayias. A Hadronization Model for the MINOS Experiment. *AIP Conf. Proc.*, 967:269–275, 2007.

[130] E. Wang, L. Alvarez-Ruso, Y. Hayato, K. Mahn, and J. Nieves. Photon Emission in Neutral Current Interactions at the T2K Experiment. *Phys. Rev. D*, 92:053005, Sep 2015.

[131] M. Day and K. S. McFarland. Differences in Quasielastic Cross Sections of Muon and Electron Neutrinos. *Phys. Rev. D*, 86:053003, Sep 2012.

[132] L. L. Salcedo, E. Oset, M. J. Vicente-Vacas, and C Garcia-Recio. Computer Simulation of Inclusive Pion Nuclear Reactions. *Nuclear Physics A*, 484(3):557 – 592, 1988.

[133] G. Rowe, M. Salomon, and R. H. Landau. Energy-Dependent Phase Shift Analysis of Pion-Nucleon Scattering below 400 MeV. *Phys. Rev. C*, 18:584–589, Jul 1978.

Bibliography

[134] K. Abe, R. Akutsu, A. Ali, et al. Constraint on the Matter–Antimatter Symmetry-Violating Phase in Neutrino Oscillations. *Nature*, 580:339–344, 2020.

[135] S. Bienstock et al. Assessing the Effect of Cross-Section Model Uncertainties on the T2K Oscillation Analyses with Fake Data Studies using the BANFF, MaCh3, P-Theta and VALOR Fit Frameworks. *T2K*, 2018.

[136] J. A. Caballero. General Study of Superscaling in QuasiElastic $(e, e^{'})$ and (ν, μ) Reactions Using the Relativistic Impulse Approximation. *Phys. Rev. C*, 74:015502, Jul 2006.

[137] Isaac Asimov. Franchise. In *Isaac Asimov: The Complete Stories, Vol. 1*. Broadway Books, 1990.

[138] L. Munteanu et al. Constraining the Flux and Cross Section Models with Data from ND280 using FGD1 and FGD2 for the 2020 Oscillation Analysis. *T2K*, 2020.

[139] L. Berns et al. Flux Prediction and Uncertainty with NA61/SHINE 2010 Replica-Target Data. *T2K*, 2020.

[140] Bolognesi, S. and others. NIWG Model and Uncertainties for 2017 Oscillation Analysis. *T2K*, 2017.

[141] P. Gueye et al. Coulomb Distortion Measurements by Comparing Electron and Positron Quasi-Elastic Scattering off ^{12}C and ^{208}Pb. *Phys. Rev. C*, 60:044308, 1999.

[142] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic Formulae for Likelihood-Based Tests of New Physics. *The European Physical Journal C*, 71(2), Feb 2011.

[143] R. Wendell et al. Super-Kamiokande Data Quality, MC, and Systematics in Run10. *T2K*, 2020.

[144] L. Berns et al. Five Sample Joint Oscillation Analysis with T2K Run1–10 Data. *T2K*, 2020.

[145] L. Alvarez-Ruso et al. Status and Challenges of Neutrino–Nucleus Scattering. *Progress in Particle and Nuclear Physics*, 100, 2018.

[146] T. Golan et al. NuWro: The Wrocław Monte Carlo Generator of Neutrino Interactions. *Nuclear Physics B - Proceedings Supplements 2012*, 499:229–232, 2010.

[147] C. Andreopoulos et al. Proposal to Measure Hadron Scattering with a Gaseous High Pressure TPC for Neutrino Oscillation Measurements. *CERN-SPSC-2017-030*, 2017.

[148] DUNE HPgTPC Working Group. High-Pressure Argon gas TPC Option for the DUNE Near Detector. *DUNE*, 2018.

[149] X.-G. Lu et al. Measurement of Nuclear Effects in Neutrino Interactions with Minimal Dependence on Neutrino Energy. *American Physical Society*, 94, 2016.

[150] Stephen Dolan. Probing Nuclear Effects at the T2K Near Detector Using Transverse Kinematic Imbalance. *NuPhys 2015*, 2015.

[151] X.-G. Lu and L. Pickering. Theoretical Predictions of Transverse Kinematic Imbalance in Neutrino-Nucleus Interactions. *NuPhys 2015*, 2015.

[152] M. Dunkman et al. Updated Recommendation of the 2014-5 Oscillation Parameters. *T2K*, 2015.

[153] M. Friend et al. Flux Prediction and Uncertainty Updates with NA61 2009 Thin Target Data and Negative Focussing Mode Predictions. *T2K*, 2015.