Study of the νμ interactions via charged current in the T2K near detector

Alfonso Andrés García Soto

Study of the ν_μ interactions via charged current in the T2K near detector

Alfonso Andrés García Soto

Tesis presentada para optar al grado de

DOCTOR EN FÍSICA

Supervisada por:

Dr. Federico Sánchez Nieto (Director)

Dr. Enrique Fernández Sánchez (Tutor)

y evaluada por:

Dr. Juan Antonio Caballero Carretero (Presidente)

Dra. Lluïsa-Maria Mir (Secretaria)

Dr. Kenneth Long (Vocal)

Cerdanyola (Barcelona), 15 de Mayo de 2017

A quien me ha iluminado
durante esta travesía...

Acknowledgments

When you start a research career you might be disoriented and I have realized that a good guidance helps enormously in focusing your efforts. In my opinion, Dr. Federico Sanchéz represents the figure of a perfect guide. Apart from his deep knowledge about physics, he has really taught me how to confront problems in this field with enthusiasm. Te estaré eternamente agradecido...

If T2K is one of the leading experiments in high energy physics is mainly because it is composed by extraordinary physicists. Besides, its co-working environment makes the research really excited for young students. I have met astonishing researchers as Dr. Anthony Hillairet and Dr. Sara Bolognesi who are one of the main reason to achieve this goal. Furthermore, I really appreciate the opportunity that Dr. Hirohisa Tanaka gave me by letting me know TRIUMF.

I would like to specially thanks people from NuMu ans XSEC groups and specially to Dr. Melody Ravonel, Dr. Justyna Lagoda, Dr. Kendall Mahn, Dr. Callum Wilkinson, Dr. Francesca Di Lodovico and Dr. Andrea Longhin. Their advices have made this analysis much more rigorous. In particular, I must mention Dr. Anselmo Cervera and Dr. Alexander Izmaylov who have to put up with me answering plenty of emails patiently.

Definitely, I have known amazing people in this collaboration. My Italian mates makes my stays in Japan really enjoyable. Grazie mille Ciro, Davide, Enrico, Francesco and Lorenzo. I wont forget our favorite Ueno spots. Moreover, I have to thanks Jiae because she was the best host ever during my stay in Vancouver.

Back in Spain, I have to particularly thanks several people from IFAE. Thorsten, Raquel, Javi and Stefania, because they were great mentors during first years. All the mates that I have met in the module during our Letras/Science lunches: Bruno, David, Emanuele, Fabian, Ivan, Joern and Stefano.

También debo mencionar las maravillosas personas que he conocido en Barcelona. Personas que ahora considero familia, como mis compis de piso: Johny, Oihana, Marta, Rober, Rocio y Tamara. Personas que me han echo pasar buenos momentos y que considero grandes amigos: Alba, Alberto, Andrea, Anna, Ari, Carlos, Laura, Maribel, Marta, Martin y Victor.

Desde Granada, mis padres (Lourdes y Javier), mis hermanas (Rocío y Lourditas) y mi otro hermano (Jose) me han permitido encarar todo este recorrido con fuerza porque su apoyo ha sido increíble siempre. Gracias por no faltar nunca.
Por último, Alba, ella ha sido mi luz.

Study of the ν_μ interactions via charged current in the T2K near detector

Alfonso Andrés García Soto

Institut de Física d'Altes Energies
Universitat Autònoma de Barcelona
Barcelona, Spain
2017

ABSTRACT

T2K is a long baseline experiment located in Japan, which aims to measure neutrino oscillation. An accelerator is used to produce neutrino, which are detected in a far detector (Super-Kamiokande) and a near detector complex (both INGRID and ND280). It was design to be sensitive to the $\nu_\mu \rightarrow \nu_\mu$ and $\nu_\mu \rightarrow \nu_e$ channels, which mainly depends on θ_{23} and θ_{13} parameters from the PMNS matrix respectively. Currently, T2K provides competitive results for all those factors and the first hint for the violation of CP symmetry in the leptonic sector has just been released.

ND280 plays a fundamental role in the oscillation analysis. It is used to constrain the systematic uncertainties associated to both the flux and cross section by measuring the rate of ν_μ and ν_e interactions. Fine grained detectors (FGDs) are narrow scintillators layers (made of plastic) designed to have very good vertexing capability, which makes then a perfect target. They are surrounded by Time Projection Chambers (TPCs), whose performance identifying particles is outstanding. Besides, Electromagnetic Calorimeters (ECals) cover both FGDs and TPCs and they have been designed to tag MIP tracks. Everything is embedded in a magnet, which allows us to measure both momentum and charge of the tracks with a high precision.

The analysis described in this thesis study ν_μ-nucleus interactions via Charged Current (CC) happening in the FGD. From this interaction a muon is produced plus hadrons, which are generated within the nuclear media. This analysis aims to reconstruct the outgoing muon, increasing the angular acceptance with respect previous analysis. The goal is to provide a sample of ν_μ CC interaction that mimic the features of such interactions when they are detected in the far detector. Consequently, the extrapolation of systematic uncertainties from near to far detector becomes more robust.

Such sample provides the perfect opportunity to report a double differential cross section for ν_μ CC interaction on plastic as function of the muon momentum and angle.

Estudio de las interacciones ν_μ via corrientes cargadas en el detector cercano de T2K

Alfonso Andrés García Soto

Institut de Física d'Altes Energies
Universitat Autònoma de Barcelona
Barcelona, Spain
2017

RESUMEN

T2K es un experimento de largo recorrido situado en Japón que pretende medir oscillación de neutrinos. Los neutrinos se producen usando un acelerador y son detectados en un detector lejano (Super-Kamiokande) y un complejo de detectores cercanos (INGRID y ND280). Fue diseñado para medir con precisión los canales $\nu_\mu \to \nu_\mu$ y $\nu_\mu \to \nu_e$, los cuales dependen principalemente de los parámetros θ_{23} y θ_{13} en la matriz PMNS respectivamente. En la actualidad, T2K aporta resultados competitivos para esos factores y los primeros indicios de violación de la simetría CP en el sector leptónico acaban de ser publicados.

ND280 desempea un papel fundamental en el análisis de oscilaciones. Se usa para reducir los errores sistemáticos asociados al flujo y la sección eficaz midiendo la cantidad de interacciones de ν_μ y ν_e. Los detectores de alta granulación (FGDs) son finas capas de centelladores (hechas de plástico) diseñados para localizar bien vértices, haciendolos perfectos blancos. Están rodeados por cámaras de proyección temporal (TPCs), cuyo rendimiento identificando partículas es espectacular. Además, calorímetros electromagnéticos (ECals) cubren ambos FGDs y TPCs y han sido diseñados para identificar trazas MIP. Todo está situado dentro de un un imán, el cual permite medir con mucha precisión el momento y carga de las trazas.

El análisis descrito en esta tesis estudia las interacciones ν_μ-nucleo mediante corrientes cargadas (CC) que ocurren en el FGD. En estas interacciones un muón y hadrones, generados en el medio nuclear, son producidos. Este análisis pretende reconstruir el muón saliente, incrementando la aceptancia angular de análisis anteriores. El objetivo es aportar una muestra de interacciones ν_μ CC que imite las características de estas interacciones cuando son detectadas en el detector lejano. De esta forma la extrapolación de los errores sistemáticos del detector cercano al lejano será más robusta.

Esta muestra proporciona la oportunidad perfecta para mostrar una sección eficaz doble diferencial de interacciones ν_μ CC en plástico en función del momento y ángulo del muón.

Contents

List of Figures

List of Tables

Outline

This dissertation details the study of ν_μ interactions via charged current. This analysis aims to characterize such interactions using the near detector of T2K experiment.

In the first block, a brief summary of the current knowledge about the neutrino is presented. To begin with, a short historical introduction is included to put the study of neutrino oscillation in context. Then, theoretical aspects related to neutrino interactions will be discussed, including an overview of the experimental status. Finally, the intrinsic connexion between neutrino oscillations and interactions will be explained.

The second block describes the different subsystems that constitute the T2K experiment. Firstly, the main features of the accelerator and beamline are summarizes. Secondly, Super-Kamiokande and INGRID detectors are briefly described. Finally, the different components of the near detector (both at software and hardware level) are explained in more detail.

In the third block of this dissertation, the core of the main analysis is explained. To begin with, the event selection is described. Afterward, the different uncertainties affecting this study are evaluated. Finally, the methodology used to extract a cross section is detail together with the final results. Furthermore, several prospects and interpretation of the results are given.

The last block includes several appendices that provide additional information for several sections. Particularly interesting is the measurement of the electric field distortion in the TPC. Such study was perform in parallel to the main analysis, although they are not independent.

Part I

Physics of Neutrinos

Chapter 1

Historical Background

During the first decades of 20th century, it was widely accepted that the nucleus were composed by electrons (e^-) and protons (p). However, using this picture, physicists were not able to give a convincing explanation for two phenomena:

- The kinetic energy of the e^- coming out from a β decay was not fixed but a continuous spectra was observed. Therefore, there was some missing energy that jeopardized a fundamental principle: the energy conservation.
- The spin of the Nitrogen nucleus was found to be integer, while a nucleus made of e^- and p has half-integer spin.

In 1930, W. Pauli wrote a letter trying to address both problems. He postulated the existence of an extra particle in the nucleus, called "neutron" (n), with 1/2 spin and a mass similar to e^-. Under this hypothesis, both phenomena could be explained. However, in 1932, J. Chadwick discovered the n and it was found to be as heavy as the p [1]. Such a heavy particle could explain the problem of the spin (assuming the nucleus as a bound state of p and n [2]), but not the spectra of the e^- in the β decay.

Two years later, a theory for β decay was developed by E. Fermi taking into account the emission of both an e^- and a neutral light particle (called "neutrino" ν) [3]. Although the model was very simplistic (based on electromagnetic interactions), it allowed to relate β decay with ν reactions. This was a critical step forward because it changed the view of the ν as an active particle rather than a passive one. Nevertheless, the probability of interaction was estimated to be $\sigma < 10^{-44}$ cm^2 [4].

The first attempt to detect ν was based on B. Pontecorvo's idea, observing the presence of Argon in a Chlorine tank placed next to a nuclear reactor [5]. However, in 1955, an experiment placed next to Brookhaven reactor (designed by R. Davis) claimed no evidences of ν [6]. The following year, another experiment (projected by F. Reines and C. Cowan), which used a liquid scintillator placed close to the reactor at Savannah River plant, found the first empirical evidence of the ν (measuring the time delay

between the light produced by an $e^+ - e^-$ annihilation and the n capture) [7]. The negative result from R. Davis was the first evidence of the leptonic number conservation including ν. Nowadays, it is well known that $\bar{\nu}$ are produced in reactors, which are unlikely to interact with Chlorine producing Argon.

Using cosmic rays, muons (μ) and pions (π) (predicted by H. Jukawa in 1935 [8]) were observed for the first time in 1936 [9] and 1947 [10] respectively. In 1956, T. Lee and C. Yang suggested that weak interactions could violate parity and their hypothesis could be tested observing either β, π and μ decays [11]. Their hypothesis was validated one year later in two different experiments designed by C.Wu [12] and R. Garwin [13]. Their discovery changed drastically the understanding of weak interactions. Soon after, L. Landau [14], A. Salam [15] and T. Lee and C. Yang [16] developed a model which explained the parity violation observed in weak interactions: the two component ν theory. In that framework the ν was assumed to have zero mass and helicity either negative (left-handed, ν_L) or positive (right-handed, ν_R). In 1958, the results from M. Goldhaver's experiment confirmed that theory and it only found ν_L [17].

In 1958, R. Feynman and M. Gell-Mann and R. Marshak and G. Sudarshan developed the V-A theory, in which the weak interaction was built including just the left-handed component of all fermions (not only ν) [18] [19]. V-A theory was able to predict results for a wide variety of processes such as π decay [20] or elastic scattering of ν and e^- [21]. Moreover, it was mentioned that the interaction should be mediated by a charged vector meson with high mass.

Nevertheless, V-A theory did not include the flavor concept in the ν. In fact, in 1962, while studying the products from π decay, an experiment placed at Brookhaven showed that the ν associated to μ or e^- was not the same particle (ν_μ and ν_e) [22]. After that result, the flavor concept and its conservation was established for leptons.

In the following decade, a third lepton was discovered at SLAC accelerator, called tau (τ) [23]. By that time, the Standard Model of the electroweak interaction already existed (see Sec. 2). Consequently, the new lepton should have an associated neutrino (ν_τ), which was discovered in 2000 by DONUT experiment [24]. In fact, the same year, data from the LEP experiment (an $e^- - e^+$ collider) established that the number of light ν ($m_\nu < m_{Z^0}/2$) should be three [25].

Chapter 2

The Standard Model

The Standard Model is the theory that has unified the weak and electromagnetic interactions. It was proposed by S. Weinberg [26] and A. Salam [27] in 1968 (at the beginning including only one flavor).

In this model, the V-A and QED theories come naturally from a $SU(2) \times U(1)$ symmetry that is spontaneously broken. In fact, it predicts the existence of one massless boson mediator of electromagnetic interactions (photon) and two charged massive bosons mediators of the V-A reactions ($W^{\pm}$).

Furthermore, it predicts a new class of weak interaction mediated by a neutral massive boson (Z^0). That interaction was discovered in 1973 at the Gargamelle experiment, in which ν_μ (created using an accelerator) were detected using a bubble chamber [28]. Then, such interactions were studied intensively, allowing free SM parameters to be accurately set: the Weinberg angle θ_W and Fermi constant G_F [29].

The key ingredient of the SM is the symmetry breaking, usually called Higgs mechanism. Such a process is responsible for the bosons masses, whose values are related to θ_W, α and G_F. In the case of quarks and leptons, this mechanism is also responsible for their masses, but it does not provide their value. In particular, in the case of the ν, the SM model can neither predict their masses. However, ν must be massless under the hypothesis of no ν_R, which have not been found up to day.

In this framework, ν are chargeless and colorless particles, so they are only affected by the weak force. On the one hand, when the interaction is mediated by $W^{(+,-)}$ boson, it is called charged current interaction (CC). On the other hand, for Z^0 boson, the interaction is referred as neutral current (NC). The coupling between such bosons and point like fermions is described using the following currents [30]:

$$\begin{aligned} J^{\mu}_{NC} &= \bar{u}_i \gamma^{\mu} \frac{g_i^V - g_i^A \gamma^5}{2} u_i \\ J^{\mu}_{CC} &= \bar{u}'_i \gamma^{\mu} \frac{1 - \gamma^5}{2} u_i \end{aligned} \tag{2.1}$$

Where $\gamma^{\mu,5}$ are the gamma matrices and $g_i^{V,A}$ are the vector and axial vector couplings for a fermion i (which depends on its weak isospin, electric charge and θ_W). Besides, u represents the spinors (Dirac field) for those fermions. Both currents have the characteristic V-A structure, which allows parity violation.

Besides, in this model quark mixing is allowed in CC interactions. Fig. 2.1 summarizes the current knowledge of the elementary particles within the SM.

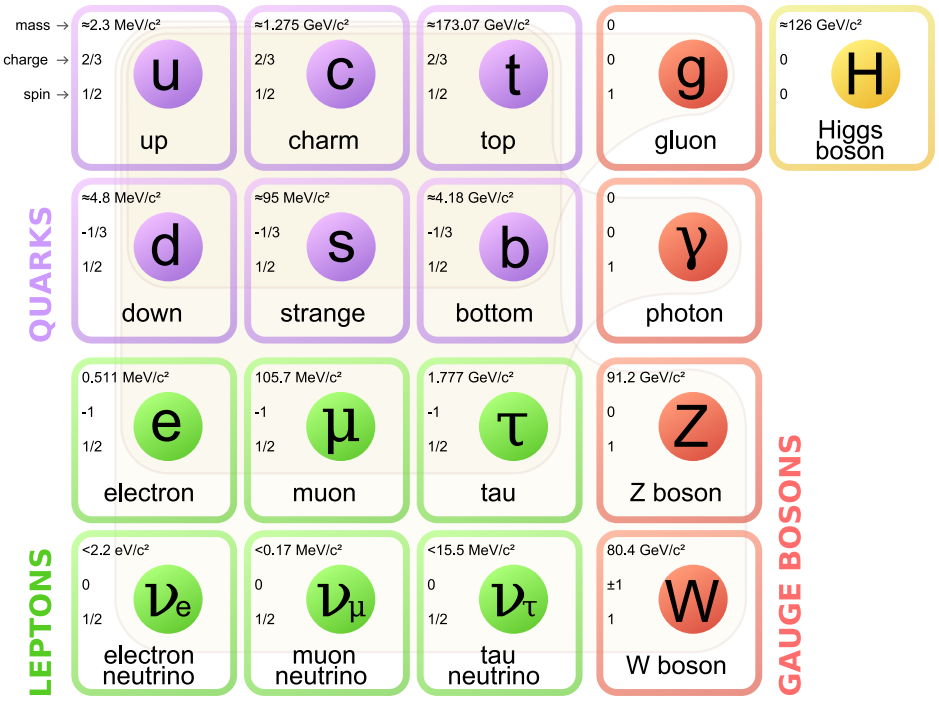

Figure 2.1: Electroweak Standard Model scheme after symmetry breaking (Source: Wikipedia).

Chapter 3

Neutrino Oscillations

ν oscillations were mentioned for the first time by B. Pontecorvo just after parity violation and the theory of two component ν appeared in 1957 [31]. At that time, only one flavor was known and he suggested a $\nu - \bar{\nu}$ oscillation, which violated the leptonic number.

One year later, a second paper was published by B. Pontercorvo about ν oscillation trying to explain a possible violation of the leptonic number in the experiment developed by R. Davis at Brookhaven [32]. By then, the theory of two component ν was confirmed and only ν_L ($\bar{\nu}_R$) were observed. B. Pontercorvo suggested the existence of an extra component ν_R, which could explain such violation if an oscillation $\bar{\nu}_R \to \nu_R$ was happening (assuming ν to be massive). However, the experiment concluded that lepton number was conserved, as no ν_R interactions were observed.

In the following decade, the V-A theory and the flavor concept for ν_e and ν_μ were established. In 1962, Maki, Nakagawa and Sakata proposed a ν oscillation framework in which the lepton number was conserved [33]. They discussed about possible transitions $\nu_e \to \nu_\mu$.

During 1967, the sterile concept for ν was introduced by B. Pontecorvo [34]. His hypothesis was that ν_e and ν_μ could oscillate between them but also with non-interacting particles (sterile ν). In that paper, he suggested that such oscillations could be measured detecting ν coming from the sun.

Using a similar detector and concept, R. Davis designed a new underground experiment placed in the Homestake Gold Mine using a different ν source, the sun instead of a reactor. In 1970, a large deficit of ν coming from the sun was observed when comparing its rate with the flux predictions for solar ν [35].

In the 70s, several papers studying ν oscillation were published. They were trying to explain the solar ν problem [36] [37] [38]. The idea was to construct a unitary mixing matrix ($U^\dagger U = 1$), analogous to the one used in the quark sector, in which lepton number was conserved. Therefore, the $\nu_{e,\mu}$ could be interpreted as a superposition of massive Dirac particles $\nu_{1,2}$ weighted by a mixing angle θ.

Once the τ lepton was discovered, such mixing matrix was generalized for the three families (not including sterile ν). Since then, it is known as the PMNS matrix:

$$|\nu_l(t)\rangle = \sum_{i=0}^{3} U_{li}^{*} |\nu_i(t)\rangle \tag{3.1}$$

On the one hand, $|\nu_l(t)\rangle$ are the ν flavor eigenstates of the Hamiltonian H_f at a time t. Using Schrodinger equation, such state can be written as a plane wave solution:

$$|\nu_l(t)\rangle = e^{-iH_f t} |\nu_l(0)\rangle \tag{3.2}$$

On the other hand, $|\nu_i(t)\rangle$ are the mass eigenstates with energy E_i and momentum $\vec{p}$ of the Hamiltonian, which in vacuum is diagonal and it fulfills the following relation (assuming $|\vec{p}| \gg m_i$):

$$H |\nu_i\rangle = E_i |\nu_i\rangle = \sqrt{|\vec{p}|^2 + m_i^2} |\nu_i\rangle = (|\vec{p}| + m_i^2/2|\vec{p}|) |\nu_i\rangle \tag{3.3}$$

Using previous relations, the probability of transition for a ν from a flavor α to β after traveling a time t can be derived:

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) &= |\langle \nu_\beta(0)|\nu_\alpha(t)\rangle|^2 = |\sum_{i=0}^{3} U_{\beta i} U_{\alpha i}^{*} e^{-i(|\vec{p}|+m_i^2/2|\vec{p}|)t}|^2 = |\sum_{i=0}^{3} U_{\beta i} U_{\alpha i}^{*} e^{-i(m_i^2/2|\vec{p}|)t}|^2 \\ &= |\sum_{i=0}^{3} U_{\beta i} U_{\alpha i}^{*} e^{-i[(m_i^2 - m_1^2/2 - m_2^2/2])/2|\vec{p}|]t}|^2 \end{aligned} \tag{3.4}$$

In the ultrarelativistic limit, the distance that a ν traveled L is equivalent to t. Therefore, the probability can be rewritten as follows (assuming $E = |\vec{p}|$):

$$P(\nu_\alpha \to \nu_\beta) = |\sum_{i=0}^{3} U_{\beta i} U_{\alpha i}^{*} e^{-i[(m_i^2 - m_1^2/2 - m_2^2/2])/2E]L}|^2 \tag{3.5}$$

In the three flavor scheme (not including sterile ν), the standard way of building the PMNS matrix U is done using four parameters: three angles (θ_{12}, θ_{13} and θ_{23}) and one complex phase (δ).

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -\sin\theta_{13} e^{i\delta} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.6}$$

Thus, in this framework, the $3-\nu$ mixing can be parametrized using 7 fundamental factors: 3 mixing angles, one complex phase and 3 masses. In particular, the δ factor is related to the CP symmetry in the lepton sector. It can be demonstrated that if $\delta \neq 0$, CP is not conserved. For m_i, the actual factors that affect the oscillation probability are their squared differences (using Eq. 3.5), which can be defined through two factors:

$$\begin{aligned} \Delta m^2_{21} &= m^2_2 - m^2_1 \\ \Delta m^2_{32} &= m^2_3 - \frac{m^2_1 + m^2_2}{2} \end{aligned} \tag{3.7}$$

As it appears in the literature, the standard way has been to define $m_1 < m_2$ so $\Delta m^2_{21} > 0$. Regarding the other masses, two different configurations are possible: $m_1 < m_2 < m_3$ so $\Delta m^2_{32} > 0$ (called normal hierarchy NH) or $m_3 < m_1 < m_2$ so $\Delta m^2_{32} < 0$ (called inverted hierarchy IH). A pictorial view of the mass ordering can be seen in Fig.3.1. It is important to stress out that from the study of such probabilities, nothing can be said about the scale of the ν masses, apart from their differences.

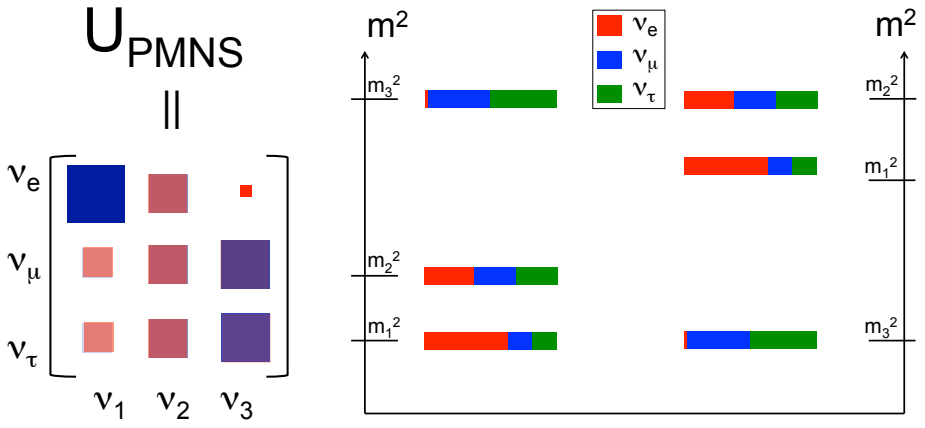

Figure 3.1: Schematic representation of the PMNS matrix (left) and the two possible configurations for mass ordering (right). Colors indicate the flavor content of each mass state.

In order to measure such factors, the basic idea in all experiments has been to compare the rate of ν, with certain flavor and E, at a production point and after flying certain fixed distance L. Using that concept, the value of $P(\nu_\alpha \to \nu_\beta)$ can be calculated for different mixing combinations and energies.

Mixing combinations are commonly broken down into two channels: disappearance and appearance. In the disappearance channel, the survival probability for certain flavor is measured (for instance, in the case of ν_μ disappearance, the rate of ν_μ is measured both at the creation point and after flying

a certain distance, $P(\nu_\mu \rightarrow \nu_\mu)$). Meanwhile, in the appearance channel, the oscillation probability from one channel into other is measured (for instance, in the case of ν_e appearance from ν_μ, the rate of ν_μ is measured at the creation point and after flying a certain distance the rate of ν_e is measured , $P(\nu_\mu \rightarrow \nu_e)$).

It is important to notice that Eq. 3.5 was obtained assuming that the ν are propagated through vacuum. However, in the presence of matter, the Hamiltonian includes a potential that takes into account the interaction of the ν with matter. On the one hand, NC interactions do not affect the probability pattern because their potential is diagonal and symmetric, so they just add an overall phase shift to Eq. 3.5. On the other hand, CC interactions only involve ν_e (ν_μ and ν_τ do not interact with e^- via CC), so the probability pattern will vary depending on its associated potential, which is related to the matter density ($V_{CC} = \pm G_F n_e$, where + (−) sign applies to ν_e ($\bar{\nu}_e$)) [39].

Depending on the configuration of the experiment, different sensitivities to measure the factors ruling the ν oscillations can be achieved. Different configurations are possible changing L, E, matter effects and the mixing channel under study. Experiments are usually classified depending on the ν source: atmospheric, accelerator, reactor and solar.

3.1 Solar Neutrino Experiments

Stars are one of the main ν factories in the universe. Those ν are mainly produced in the fusion reactions in which ν_e are released. In the case of the sun, the flux prediction of ν has greatly evolved. Currently, there have been very sophisticated models that calculate the flux of ν coming from the sun with high accuracy [40]. It is interesting to notice that such ν have a maximum energy of ∼15 MeV.

In order to detect solar ν, several detector technologies have been used apart from the R. Davis experiment in Homestake. In the 90s, two experiments (plus an extension) were developed using Gallium tanks in which excess of Germanium was produced from $\nu_e - Ga$ CC interactions: GALLEX [42], SAGE [41] and GNO [43]. That reaction has an energy threshold lower than $\nu_e - Cl$ (233 keV instead of 814 keV), so those experiments were sensitive to a wider energy range of the ν_e flux.

In the same decade, the first water based detector was created (Kamiokande [44]). The idea was to detect the Čerenkov light emitted by e^- (produced in $\nu_e - e^-$ scattering), when they fly faster than light in the medium. This experiment was improved enlarging the tank size (Super-Kamiokande [45]).

Besides, another experiment using heavy water was constructed at the end of the decade (SNO [46]). Those detectors were able to observe NC reactions (apart from CC) so the flavor content of the solar ν could be measured.

Moreover, in the early 21st century, a detector filled with highly pure liquid scintillator have been used to measure solar ν_e in a lower energy region: Borexino [47]. All those experiments claimed a large deficit of ν_e coming from the sun with respect to models predictions without assuming oscillations.

In those experiments, in which the factor $L/E \gg 1$, the ν_e disappearance channel is very sensitive to θ_{12} and Δm^2_{12} (assuming $\theta_{13} \approx 0$ and $\Delta m^2_{32} \gg \Delta m^2_{21}$). Currently, the best fit values for those factors are shown in Fig. 3.2.

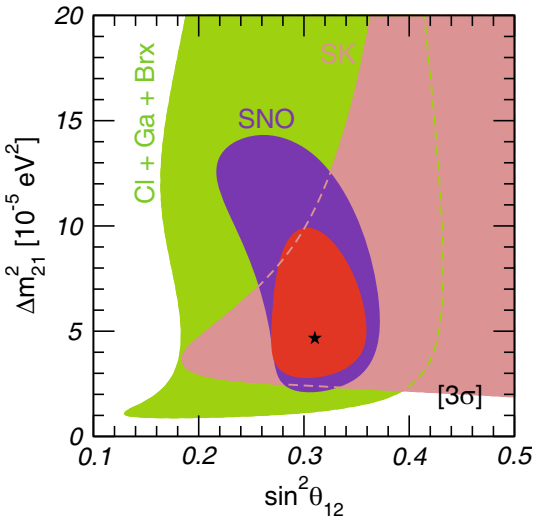

Figure 3.2: Allowed regions of θ_{12} and Δm^2_{21} for solar neutrino experiments [48].

3.2 Reactor Neutrino Experiments

A very pure flux of $\bar{\nu}_e$ is produced in nuclear reactors. Those ν appear when the products of fission reactions, which are very unstable, suffer from β decay. Although β decay is well known, in nuclear reactors thousands of different isotopes decay contribute to the total ν flux. Recent computations of such flux are available [49].

In order to observe such $\bar{\nu}_e$, all detectors use a similar concept to the one used at the Savannah River plant in the 50s. The idea has been to identify a inverse β decay in a liquid scintillator and Gd tank, which is placed 1 km far from the nuclear reactor. Such decay produces a characteristic light signal

due to the $e^+ - e^-$ annihilation and the neutron capture.

In the 90s, two experiments of the same kind were developed (Chooz [50] and Palo Verde [51]) and no evidences were found of neither excess nor deficit of $\bar{\nu}_e$. In the next decade, three new experiments were projected including a near detector in the configuration. In 2012, they published results ruling out the non oscillation hypothesis with high significance (RENO [52], Double Chooz [53] and Daya Bay [54]).

In reactor experiments, the common configuration is $L \sim 1$ km and $E \sim 3$ MeV. Studying the $\bar{\nu}_e$ disappearance channel they are very sensitive to θ_{13} and $|\Delta m^2_{32}|$ (in first approximation, Δm^2_{21} impact is negligible). The most precise measurement of those factors is shown in Fig. 3.3.

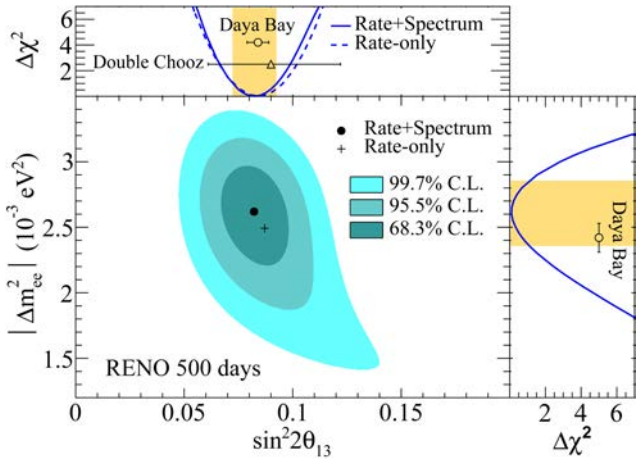

Figure 3.3: Allowed regions of θ_{13} and $|\Delta m^2_{ee}|$ (which is very similar to $|\Delta m^2_{32}|$) using the latest published results from RENO, Daya Bay and Double Chooz [55].

3.3 Atmospheric Neutrino Experiments

Cosmic rays (mainly protons) interact with the atmosphere producing hadrons (predominantly π) that decay into $\nu_{e,\mu}$. The energy range of such ν is highly wide (from the GeV range up to TeV range). In fact, it is well known that ν_μ is the dominant flavor, although ν_e component (associated to μ decay) is not negligible at energies around 1 GeV. Moreover, the contribution of both ν and $\bar{\nu}$ is found to be very similar over the whole energy range [56].

Several experiments have detected ν_μ produced in the atmosphere placing the detector underground (Kamiokande [57] and IMB [58]), actually the most important one has been Super-Kamiokande. That experiment cannot distinguish between ν and $\bar{\nu}$ but it can disentangle e and μ coming out from CC

reactions. Therefore, the main concept is to measure the rate of $\nu_\mu - \bar{\nu}_\mu$ and $\nu_e - \bar{\nu}_e$ for different zenith angles (different travel distances) and energies.

In 1998, the first statistically significant measurement was released showing a deficit of ν_μ with respect to the predictions without assuming oscillations [59], while the ν_e component showed no excess. Moreover, while studying the rates as function of the zenith angles, a clear dependency was observed. These features pointed to a possible $\nu_\mu \rightarrow \nu_\tau$ oscillation. That result was later confirmed by two other independent experiments (MACRO [60] and MINOS [61]).

In a first approximation, for atmospheric ν experiments, in which $L \sim 20 - 1000$ km and $E \sim 1 - 100$ GeV, the ν_μ disappearance pattern is driven by θ_{23} and $|\Delta m^2_{32}|$. The latest results from Super-Kamiokande and IceCube are shown in Fig. 3.4.

3.4 Accelerator Neutrino Experiments

Nowadays, one of the most widespread methods for production of ν is done through the use of accelerators. In the basic configuration, protons are accelerated up to a desired energy and later they collide onto a nuclear target. In those collisions, mesons are generated (primarily π and secondly K) which decay after flying a certain distance in ν (mainly ν_μ) and leptons .

In order to increase the flux, mesons are focused using magnetic horns. Taking into account the polarization of such horns, negative (positive) mesons can be deflected, so a very pure ν_μ ($\bar{\nu}_\mu$) beam can be obtained. Besides, a very narrow ν flux peaked at certain energy can be achieved using the angular dependency of the ν energy produced in the meson decay.

Current predictions of the ν flux produced in accelerators are calculated through very detailed models of the hadron production in proton interactions. Those models are usually tested using specific experiments in which hadron production is measured [62].

In 1999 K2K became the first experiment that measured ν oscillations using an accelerator (producing ν flux peaked at 1 GeV) [63]. It consisted of a near detector (a water tank at 300 m of the ν production) and a far detector (Super Kamiokande at 250 km). In 2005 a second experiment called MINOS started to take data [64]. Both near (1 km) and far (735 km) detectors were made of iron and scintillators layers. In this case, the ν energy peaked at 3 GeV and it was possible to invert the horns polarity enhancing $\bar{\nu}_\mu$ flux.

In 2010 T2K was a pioneer experiment in the off-axis technique [65]. It achieved a narrow peak centered at 0.6 GeV for the ν_μ flux. Its configuration was very similar to the K2K one but the near detector was different. In 2014, NOVA experiment took advantage on that technique creating a narrow band centered at 3 GeV [66]. In that experiment both near and far detectors were made of PVC cells filled with liquid scintillator.

Typically, those experiments use a configuration in which $L \sim 100 - 1000$ km and $E \sim 1 - 10$ GeV. Through the measure of the ν_μ (or $\bar{\nu}_\mu$) disappearance pattern, they are more sensitive to θ_{23} and $|\Delta m^2_{32}|$ than atmospheric experiments, because the ν flux and their flying distance are better controlled. Moreover, studying both ν_μ and $\bar{\nu}_\mu$ disappearance, the CPT symmetry can be tested, which predicts the same survival probabilities in both cases. The most precise measurements of those factors are shown in Fig. 3.4.

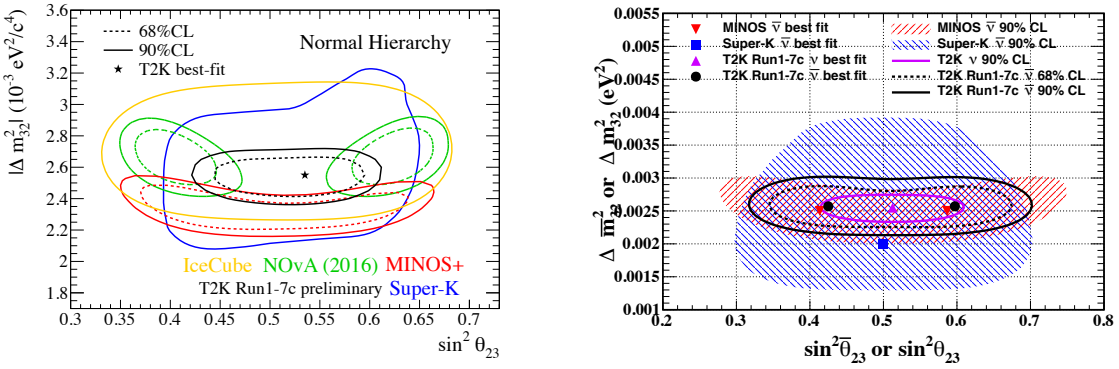

Figure 3.4: Allowed regions of θ_{23} and $|\Delta m^2_{23}|$ using the ν_μ (left) and $\bar{\nu}_\mu$ (right) disappearance channels and assuming NH.

In addition, the factor θ_{13} can be extracted from the $\nu_\mu(\bar{\nu}_\mu) \to \nu_e(\bar{\nu}_e)$ appearance channel (assuming a value of δ, θ_{23} and Δm^2_{32}). In 2014 T2K published a study in which the non oscillation hypothesis ($\theta_{13} = 0$) was excluded at 7σ level [65]. In 2016, NOVA published an analysis of the same channel showing the same tendency. It is important to notice that using this strategy, the effective angle θ_{13} depends on the δ CP phase. Therefore, if this dependency shows some tension with the factor θ_{13} measured by reactor experiments, certain sensitivity for the value of δ can be obtained. This procedure has been followed by both T2K and NOVA in order to set a constrain in the value of δ.

Finally, δ factor can also be extracted by comparing $\nu_\mu \to \nu_e$ and $\bar{\nu}_\mu \to \bar{\nu}_e$ channels. Therefore, this strategy allows to measure the value of δ without requiring any constraint from other experiments.

First studies have been developed by T2K and preliminary results, as shown in Fig. 3.5, hint to a possible CP violation ($\delta \neq 0$) at 90% C.L.

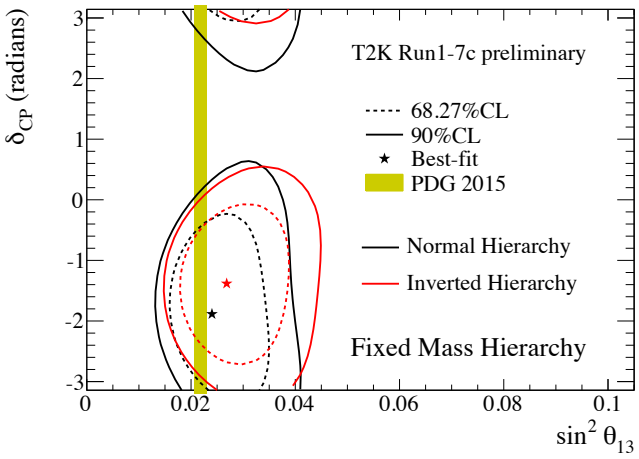

Figure 3.5: Allowed regions of θ_{13} and δ using a global fit (which includes ν_μ, $\bar{\nu}_\mu$ disappearance and ν_e, $\bar{\nu}_e$ appearance) without using reactor constraint in θ_{13}.

3.5 Energy Reconstruction

In order to measure the mixing parameters, the oscillation probability pattern must be studied for different values of the ν energy (when L is fixed), $P_{osc}(E)$. In accelerator based experiments, this is done by comparing the rate of ν in two different detectors placed at different distances with respect to the production target (near and far). Those rates can be expressed as follows:

$$\begin{aligned} N_{near}(E) &= \sigma_{near}(E)\Phi(E)\varepsilon_{near}(E) \\ N_{far}(E) &= \sigma_{far}(E)\Phi(E)P_{osc}(E)\varepsilon_{far}(E) \end{aligned} \tag{3.8}$$

Where $\Phi(E)$ is the ν flux created at the production target and $\sigma_{near,far}(E)$ ($\varepsilon_{near,far}(E)$) is the cross section (acceptance) for both near and far detectors. In the simplest configuration, in which both detectors are identical ($\sigma_{near} = \sigma_{far}(E)$), they have perfect acceptance ($\varepsilon_{near,far}(E) = 1$) and the flux is monochromatic ($\Phi(E) = \delta(E - E_0)$), then, $P_{osc}(E)$ can be easily extracted from $N_{far}(E)/N_{near}(E)$.

However, current ν sources do not produce a monochromatic flux but a spectrum of energies. Consequently, the E must be extracted from the final states of ν reactions (E_{rec}). Then, the ratio of rates

becomes more complicated:

$$\frac{N_{far}(E_{rec})}{N_{near}(E_{rec})} = \frac{\int \sigma_{far}(E)\Phi(E)\varepsilon_{far}(E)P_{osc}(E)P(E_{rec}|E)dE}{\int \sigma_{near}(E)\Phi(E)\varepsilon_{near}(E)P(E_{rec}|E)dE} \tag{3.9}$$

Where $P(E_{rec}|E)$ represents the probability that a ν with energy E has been reconstructed with energy E (usually called smearing matrix). This factor describes the performance of the detector, so Monte Carlo simulation is needed in order to extract it. Moreover, this simulation requires the prediction from a cross section theory.

Therefore, $P(E_{rec}|E)$ and $\sigma(E)$ are model dependent factors. This means that $P_{osc}(E)$ can be measured with different patterns depending on the cross section theory and parameters that have been used. Consequently, in order to make a ν oscillation measurement it is critical to understand the cross section of the ν reactions using reliable models and uncertainties.

Currently, energy reconstruction is done using two different approaches:

- **Kinematic**: E_{rec} is extracted from the kinematics of the outgoing lepton. This method is useful when $E < 2$ GeV because most of the interactions are quasielastic (see Sec. 4.1.1) and four body kinematics can be assumed. Therefore, this approach requires a channel identification. One the hand $P(E_{rec}|E)$ will be smeared by the Fermi motion of the nucleons (see Sec. 4.2.1). On the other hand, wrong identification of quasielastic interactions, as 2p2h (see Sec. 4.2.2) or resonant interactions (see Sec. 4.1.2) in which the π^+ is absorbed within the nucleus, must be taken into account by the model (see Fig. 3.6).

- **Calorimetric**: E_{rec} is extracted summing the contribution from all final states particles that are reconstructed. When $E > 2$ GeV, inelastic reactions dominate, so four body kinematics is not useful. This inclusive approach is less affected by wrong identification (background), but it must rely on the model to take into account invisible particles (usually neutral hadrons).

Several studies have been performed in order to understand the impact of $P(E_{rec}|E)$ in an oscillation measurement using both energy reconstruction approaches [67] [68] [69]. They all arrive to similar conclusions. The kinematic approach is more robust than the calorimetric one with current detector technologies. The main disadvantage of the kinematic approach is that it requires a accurate knowledge of the ν scattering with nuclei.

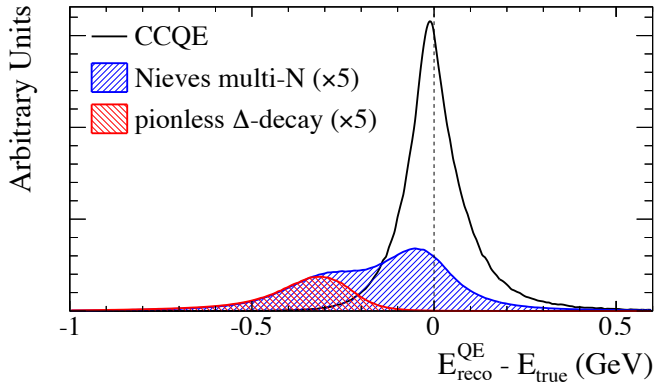

Figure 3.6: Difference between E_{rec} (using the four body kinematic formula) and E for events selected as quasielastic. Simulation is performed using `NEUT` and T2K flux. 2p2h and resonant with π^+ absorption are scaled up a factor of 5.

Chapter 4

Neutrino Interactions

In Particle Physics, the probability of interaction between particles is usually expressed using the cross section concept. It relates the number of incoming particles that are scattered with the number of targets within a certain volume:

$$\sigma = \frac{N_X}{\phi N_B} \tag{4.1}$$

Where N_X are the number of scattered events, ϕ is the incoming flux and N_B the density of targets. Using the Golden Rule, the quantity can be expressed as a function of the transition matrix $\mathfrak{M}$, which describes the evolution between initial states (A and B) and final states (X) [70]:

$$\sigma = \frac{1}{4E_A E_B (v_A + v_B)} \int \prod_X \frac{d^3\vec{p}_X}{(2\pi)^3 2E_X} (2\pi)^2 \delta^{(4)} \left(\sum_X p_X - p_A - p_B \right) |\mathfrak{M}|^2 \tag{4.2}$$

Where p, v and E are the momentum, velocity and energy of the particles respectively. This amplitude can be derived from theory, often using the Feynman rules [71]. Meanwhile, Eq. 4.1 is used by experiments to measure the cross section. Consequently, the prediction of different models for an specific interaction can be tested.

In ν experiments, a common configuration assumes the target at rest. From that configuration, momentum and angle with respect to the incoming ν of the outgoing particles are measured. Therefore, the scattering process in the lab frame is of particular interest (presented in Fig. 4.1). Its associated kinematics variables are described as follows:

$$\begin{aligned} Q^2 &\equiv -(p_\nu - p_l)^2 = 2E_\nu \left(E_l - |\vec{p}_l| \cos\theta_{\nu l}\right) - m_l^2 \\ \nu &\equiv \frac{p_t(p_\nu - p_l)}{m_t} = E_\nu - E_l \quad // \quad y \equiv \frac{m_t \nu}{p_t p_\nu} = \frac{E_\nu - E_l}{E_\nu} \\ x &\equiv \frac{Q^2}{2m_t \nu} = \frac{2E_\nu \left(E_l - |\vec{p}_l| \cos\theta_{\nu l}\right) - m_l^2}{2m_t(E_\nu - E_l)} \end{aligned} \tag{4.3}$$

Using Eq. 4.2, differential cross section as function of the energy of the outgoing particle E_l and its

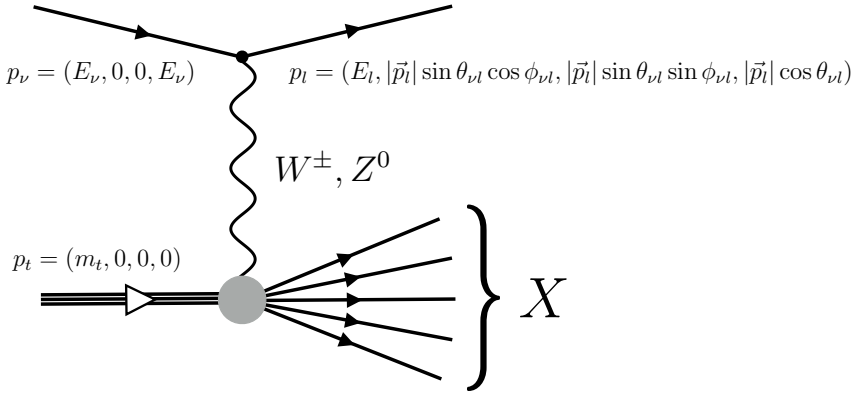

Figure 4.1: Schematic view of a interaction between a ν with a target on the lab frame. Associated four-momentum for the incoming particles and the outgoing lepton are shown.

angle with respect to the incoming particle $\theta_{\nu l}$ can be expressed as follows (neglecting the mass of the ν):

$$\frac{d\sigma}{dE_l d(\cos\theta_{\nu l})} = \frac{1}{32\pi^2 m_t} \frac{|\vec{p}_l|}{E_\nu} |\bar{\mathfrak{M}}|^2 \prod_X \frac{d^3\vec{p}_X}{(2\pi)^3 2E_X} (2\pi)^4 \delta^{(4)} \left(\sum_X p_X - p_l - p_t - p_\nu \right) \tag{4.4}$$

Where $|\bar{\mathfrak{M}}|^2$ sums (averaged) over all possible polarizations of the final (initial) state. One of the simplest processes is the scattering between ν and quarks. Their point-like structure allows the total cross section to be computed analytically (using Eq. 2.1). Assuming both a low transferred momentum and a ν energy much larger than the quarks mass, several relations are derived:

$$\begin{aligned} &\sigma_{\nu q}^{CC,NC} \propto E_\nu \\ &\sigma_{\bar{\nu} q}^{CC} = 3\sigma_{\nu q}^{CC} \\ &\sigma_{\nu q}^{NC} = (a + bsin^2\theta_W + csin^4\theta_W)\sigma_{\nu q}^{CC} \end{aligned} \tag{4.5}$$

Where a, b and c depend on the weak isospin and the electric charge of the quarks. Even though, previous relations are valid for quarks, in first approximation, they can be extrapolated for complex nucleus. Consequently, ν interaction with matter is more likely to happen via CC.

4.1 Neutrino-Nucleon Scattering

This dissertation details the interactions between ν and complex nuclei, specially, ν_μ interactions via CC. In order to define the formalism and terminology, it is useful to describe the scattering between ν and nucleons. In particular, interaction of ν_μ and a nucleon via CC (the outgoing lepton is a μ^-) will

be studied.

Assuming that the transferred momentum is much lower than the mass of W (it simplifies the definition of the boson propagator), using the Born approximation [73], $\mathfrak{M}$ can be written as follows:

$$\mathfrak{M} = \frac{G_F}{\sqrt{2}} \left[\bar{u}(p_3)\gamma_\mu(1-\gamma^5)u(p1)\right] < X|J^\mu_{CC}|N > \tag{4.6}$$

Where G_F is the Fermi constant, which characterizes the strength of the weak force (the same as the fine structure constant does for electromagnetic interactions). The last term takes into account how the W boson couples to the nucleon structure. Using Eq. 4.4, the neutrino-nucleon cross section via charged current can be expressed in the following form:

$$\frac{d\sigma}{dE_\mu d(cos\theta_{\nu\mu})} = \frac{G_F^2}{2\pi}\frac{|\vec{p}_\mu|}{E_\nu} L_{\alpha\beta}H^{\alpha\beta} \tag{4.7}$$

Where $L_{\alpha\beta}$ is the leptonic tensor (easily calculated) and $H^{\alpha\beta}$ is the hadronic tensor, which contains the information about the weak structure of the nucleons. A model independent definition of the hadronic tensor can be constructed using six structure functions $W_i(\nu, Q^2)$ [72]. Contracting both tensors, the following cross section is obtained:

$$\begin{aligned}\frac{d\sigma}{dE_\mu d(cos\theta_{\nu\mu})} = \frac{G_F^2}{\pi}|\vec{p_\mu}|^2 \Bigg\{ & 2W_1 sin^2\frac{\theta_{\nu\mu}}{2} + W_2 cos^2\frac{\theta_{\nu\mu}}{2} + W_3\frac{E_\nu + E_\mu}{m_t}sin^2\frac{\theta_{\nu\mu}}{2} \\ & + \frac{m_\mu^2}{m_t^2}\left[W_4 sin^2\frac{\theta_{\nu\mu}}{2} - W_5\frac{m_t}{2|\vec{p}_\mu|}\right] \\ & + \frac{E_\mu - |\vec{p_\mu}|}{|\vec{p}_\mu|}\left[W_1 + W_2\frac{1}{2} + W_3\frac{E_\nu + E_\mu}{2m_t} + W_4\frac{m_\mu^2}{2m_t^2}\right]\Bigg\}\end{aligned} \tag{4.8}$$

It is important to notice that up to Eq. 4.8 no assumptions were made on the hadronic system (parametrized with the W functions). To proceed further, the hadronic vertex must be modeled Such models are computed using exclusive processes and, they are classified using three channels: quasielastic (QE), resonant (RES) and deep inelastic scattering (DIS).

4.1.1 Quasielastic Charged Current Interaction

In charged current quasielastic interactions, the incoming nucleon changes its identity. In particular, ν_μ CCQE reaction can be expressed as follows:

$$\nu_\mu + n \rightarrow \mu^- + p \tag{4.9}$$

This is a 2-2 body interaction, so the kinematics are very constrained. In particular, p_μ and $\theta_{\nu\mu}$ are totally defined by E_ν and Q^2. Therefore, it is simpler to represent the cross section as function of Q^2 as follows (assuming negligible contribution from m_μ^2/m_n^2 terms) [74]:

$$\begin{aligned}
&\frac{d\sigma}{dQ^2} = \frac{G_F^2 m_n^2 cos\theta_c}{8\pi E_\nu}\left[A + B\frac{4m_n E_\nu - Q^2}{m_n^2} + C\left(\frac{4m_n E_\nu - Q^2}{m_n^2}\right)^2\right] \\
&A = \frac{4m_n E_\nu - Q^2}{m_n^2}[(1+\frac{Q^2}{4m_n^2})F_A^2 - (1-\frac{Q^2}{4m_n^2})F_1^2 + \frac{Q^2}{4m_n^2}(1-\frac{Q^2}{4m_n^2})F_2^2 + \frac{Q^2}{m_n^2}\tau F_1 F_2 \\
&B = \frac{Q^2}{m_n^2}F_A(F_1+F_2) \\
&C = \frac{1}{4}(F_A^2 + F_1^2 + \tau F_2^2)
\end{aligned} \tag{4.10}$$

Where F_1 and F_2 are the vector form factors and F_A is the axial form factors of the nucleon. Such form factors are related to the internal structure of the nucleon. In the case of $\bar{\nu}_\mu$ CCQE interactions, a minus sign goes next to the B factor.

On the one hand, vector form factors can be related to the form factors from electron scattering (using the conserved vector current hypothesis [75]). Consequently, they can be expressed as function of the Sachs form factors of the nucleon (G_E and G_M) [76]:

$$\begin{aligned}
F_1(Q^2) &= \frac{(G_E^p(Q^2) - G_E^n(Q^2)) + \frac{Q^2}{4m_n^2}(G_M^p(Q^2) - G_M^n(Q^2))}{1+\frac{Q^2}{4m_n^2}} \\
F_2(Q^2) &= \frac{(G_M^p(Q^2) - G_M^n(Q^2)) - (G_E^p(Q^2) - G_E^n(Q^2))}{1+\frac{Q^2}{4m_n^2}}
\end{aligned} \tag{4.11}$$

The simplest parametrization of G_E and G_M is the dipole form. However, current models already assume deviations from that shape [77].

On the other hand, the axial form factor is only accessible through weak interaction processes. Nowa-

days, most models assume a dipole shape for this factor:

$$F_A(Q^2) = \frac{g_A}{1 + Q^2/M_A^2} \tag{4.12}$$

Where g_A is the axialvector coupling constant (measured in polarized neutron β decay [78]) and M_A is the axial mass (measured using either ν scattering off protons or nuclei and π^+ electroproduction [79]).

4.1.2 Resonant Charged Current Interaction

For energies around 1 GeV and above certain threshold ($\sim$ 0.2 GeV), single π production becomes feasible through the inelastic excitation of the nucleon. Those processes are called single π resonant interactions. Three different channels contribute (mediated by W^+ boson):

$$\begin{aligned} \nu_\mu + p &\rightarrow \mu^- + p + \pi^+ \\ \nu_\mu + n &\rightarrow \mu^- + p + \pi^0 \\ \nu_\mu + n &\rightarrow \mu^- + n + \pi^+ \end{aligned} \tag{4.13}$$

The resonance state of the nucleon will have isospin 1/2 or 3/2 (higher contributions are negligible for energies around 1 GeV) and it decays with a certain branching ratio into a π-nucleon pair ($\Gamma_{R\rightarrow N\pi}/\Gamma_{total}$). Adding the contribution from different resonances, the cross section for processes in Eq.4.13 can be expressed as follows [80]:

$$\frac{d\sigma^{\nu N\rightarrow lN\pi}}{dQ^2 d\nu d\Omega_\pi^{CM}} = \frac{1}{4\pi} \sum_i^{resonances} \frac{d\sigma^{\nu N\rightarrow lR_i}}{dQ^2 d\nu} \frac{\Gamma_{R_i\rightarrow N\pi}}{\Gamma_{total}} \tag{4.14}$$

Where Ω_π^{CM} is the solid angle between the outgoing lepton and π in the center of the mass frame. The resonance production cross section can be expressed as function of $L^{\mu\nu}$ and $H^{\mu\nu}$ (see Eq. 4.7).

The structure of $H^{\mu\nu}$ depends on the features of the resonance and it is parametrized using vector and axial form factors, which depend on Q^2. In order to compute such factors, two different approaches are usually considered: extracting parameters using data from e^- and ν scattering [81] or using quark models [82].

It is worth to notice that, below 1 GeV, the main contribution comes from the Δ resonance. Above that

energy, other resonances become relevant. In this dissertation other inelastic resonant processes are not discussed (for instance single photon, kaon or multi-π productions). The prediction of the models for such processes have not been intensively tested due to the lack of experimental data. Nevertheless, at the 1 GeV energy range they are expected to be negligible.

4.1.3 Deep Inelastic Scattering Charged Current Interaction

When the energy of the ν increases, the transferred momentum can be high enough to resolve the quark content of the target nucleon. Then, the scattered quark produces an hadronic jet. This kind of interactions are studied using a particular formalism.

From Eq. 4.8, neglecting the μ mass, it is possible to define the cross section as function of x and y variables.

$$\frac{d\sigma}{dxdy} = \frac{G_F^2 m_t}{\pi} E_\nu \left[\frac{y^2}{2} 2xF_1 + \left(1 - y - \frac{m_t xy}{2E_\nu}\right) F_2 + y\left(1 - \frac{y^2}{2}\right) F_3 \right] \tag{4.15}$$

Where the F terms contain all the information about the structure of the nucleon. Such factors can be related to the W functions from Eq. 4.8, in the following way:

$$\begin{aligned} F_1(x, Q^2) &= m_t W_1(\nu, Q^2) \\ F_2(x, Q^2) &= \nu W_2(\nu, Q^2) \\ F_3(x, Q^2) &= \nu W_1(\nu, Q^2) \end{aligned} \tag{4.16}$$

If the nucleon is composed by point like structures (such as quarks), the factors F_i will only depend on x at very high values of Q^2, so called Bjorken scaling [83]. Besides, if the point like structures have 1/2 (Dirac like) spins, then the Callan-Gross relation is fulfilled $2xF_1 = F_2$ [84].

The parton model explains the scaling at high Q^2 [71]. It models the nucleon as a complex structure composed by point like structures (partons) carrying a fraction of the nucleon momentum (associated to the x variable). In this model, structure functions F_i is linked to the distribution of the partons:

$$\begin{aligned} F_1(x) &= \sum_i q_i(x) + \sum_i \bar{q}_i(x) \\ F_2(x) &= 2x \sum_i q_i(x) + 2x \sum_i \bar{q}_i(x) + 4xk \\ F_3(x) &= 2 \sum_i q_i(x) - 2 \sum_i \bar{q}_i(x) \end{aligned} \tag{4.17}$$

The q_i (k) factors, usually called parton distribution functions (PDF), gather the probability of finding Dirac (scalar) partons of type i carrying a fraction of the nucleon momentum x. The summation runs over all quark components that can be scattered in the nucleon (e.g. for CC interactions with protons it sums over d, s, and b or $\bar{u}$, $\bar{b}$ or $\bar{t}$). Currently, PDF has been studied in detail using complex parametrization applying QCD-based models (e.g. [86]).

4.2 Neutrino-Nucleus Scattering

At present, all ν experiments use targets composed by complex nuclei (carbon, iron, argon, etc.). The main reason for that is the proportionality between the number of interactions and the target density. In future, Hydrogen-based targets in a high pressure regime might be feasible, although plenty of technology development is still needed.

In order to understand this kind of interactions, models that predict the behavior of the particles within the nuclear media are needed. Those models must take into account the kinematics and dynamics of both the bound nucleons and the outgoing particles.

4.2.1 Nucleons Initial State

One of the main drawbacks of using complex nuclei is that the assumption of ν interactions with free nucleons is not valid any more. In fact, the boson mediator does not probe the nucleons as isolated objects but as moving particles in a nuclear medium.

Most models are based on the Impulse Approximation (IA) scheme, in which the nucleons are probed individually and the remaining A-1 particles act as spectators [87]. Under this assumption, the cross section with a nucleus can be described as an incoherent sum of processes only involving one nucleon. Such nucleons are not at rest, so the summation must take into account the probability distribution for the momentum and energy of the nucleons (referred as spectral function $F(\vec{p}, E)$). Besides, actually, those nucleons are bound, so certain amount of energy is needed to excite the nucleus (called binding energy E_B). In the IA this energy is assumed to be absorbed by the spectator system.

Therefore, the key ingredient to model the initial state in the nuclear media is the spectral function $F(\vec{p}, E)$. For very light nucleus (number of nucleons lower than 5) such functions can be calculated

analytically. In the case of more complex nuclei, most methods combine theoretical calculations (using approximations) with e^- scattering data [88] [89]. In the simplest approach, called Relativistic Fermi Gas (RFG), the nucleus is modeled as an ideal gas composed by fermions [90]. Using the Pauli exclusion principle, it can be derived that the momentum of the nucleons must be lower than the Fermi momentum k_F of that nucleus, giving to $F(\vec{p}, E)$ a characteristic step function shape. A more sophisticated model, called Local Fermi Gas (LFG), includes a radial dependency of k_F taking into account the nuclear density distribution [91]. Furthermore, a phenomenological approach have been developed based on scaling and superscaling phenomena observed in electron scattering [92]. Fig. 4.2 compares the spectral function obtained using different models.

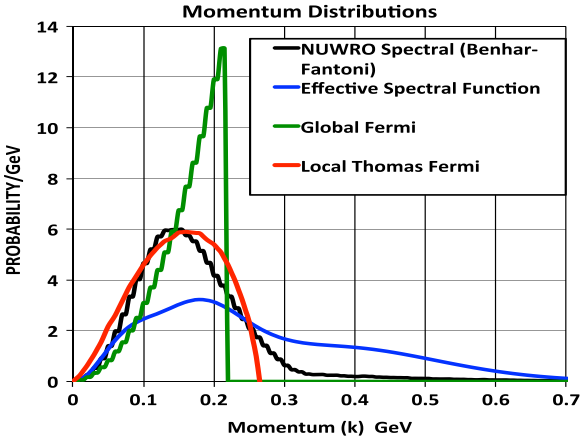

Figure 4.2: Prediction of nucleon momentum distribution in ^{12}C for different nuclear models: RFG (green), LFG (red), [88] (black) and [89] (blue).

4.2.2 Nucleons Correlations

The assumption of the interaction between the W boson with an isolated nucleon is a very simplistic approach. Possible couplings between the nucleons, which are usually named nucleons correlations, must consider.

One of the most interesting properties of this kind of system is that multinucleon interactions are likely to happen. In the simplest scenario, the interaction could happen with two nucleons, which is usually referred in literature as 2p2h reactions. Typically, this kind of interactions arise from either short range correlation of nucleon pairs (SRC) or meson exchange currents (MEC). Such processes have already shown a non negligible contribution in e^- scattering measurements (see Fig.4.3). Currently, they are being widely studied in the ν field because they have shown a clear impact on oscillation and cross

section measurements at energies around 1 GeV (it is difficult to distinguish them from QE reactions) [93] [94].

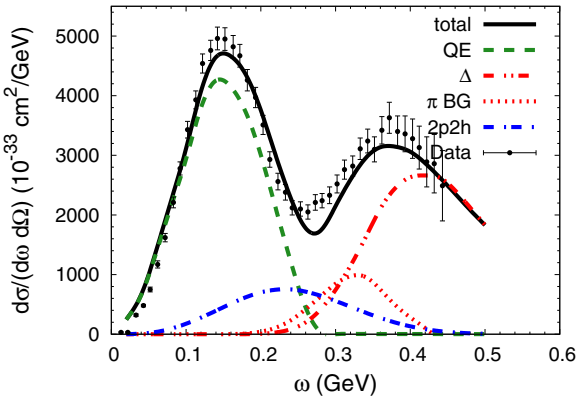

Figure 4.3: Cross section for $e^- - {}^{12}C$ scattering at 560 MeV and 60 deg. 2p2h filled the dip region in between QE and RES reactions [95].

Besides, at low Q^2 the boson mediator is not able to probe individual nucleons but a collectivity. This many-body system is usually associated to long range correlations. Currently, the response of this kind of system is modeled using an effective theory based on random phase approximation (RPA), in which the boson propagator is recomputed including extra terms (see Fig.4.4). Using this approach, large differences have been predicted with respect to the RFG model [93] [94]. On the one hand, at very low Q^2 the cross section is suppressed, while at medium Q^2 it is enhanced. On the other hand, at large Q^2 the effect disappear and it converges asymptotically into the IA prediction.

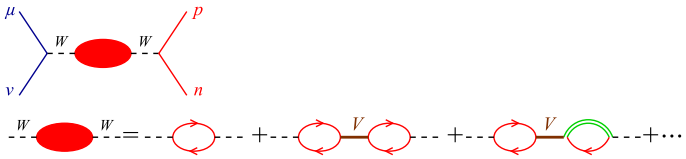

Figure 4.4: RPA computation in CCQE scattering. Solid lines pointing to the right (left) denote particle (hole) states. The double line indicates the Δ contribution. The V represents the contribution of an effective nucleon-nucleon interaction (mainly consisting in π exchange).

Another interesting property of nuclear medium is the Pauli blocking. This feature affects to the phase space of the outgoing nucleon, reducing the interaction probability. In the Fermi gas picture, all quantum states are occupied so, in order to eject one nucleon, its initial momentum plus the momentum transferred by the ν must be higher than k_F. Therefore, implementing it in the RFG is straight forward. In the case of other nuclear models with more complicated spectral functions, its implementation is not trivial, but in most cases a hard cut off on the momentum is applied (as in the RFG case).

4.2.3 Final State Interactions

Another drawback associated to interactions with a complex nuclei comes from the fact that outgoing particles (mainly hadrons) will be immersed in a highly dense nuclear medium. Consequently, they are very likely to interact. Therefore, the features of the primary outgoing particles and the particles leaving the nucleus might be completely different. This effect is commonly called final state interaction (FSI) and in the 1 GeV regime is particularly important for nucleons and π, which are not very energetic. The main idea is to understand all possible collisions that hadrons might suffer within the nuclear medium. They can be summarized as: elastic scattering, hadron production or absorption and charge exchange.

Nowadays, FSI interactions are modeled using two different approaches: the semi-classical model [81] and the intranuclear cascade [96] [97]. Both approaches assume the nucleus as a collection of isolated nucleons, in which the hadron will be likely to interact with them depending on the kind of reaction. Such probabilities are the key ingredient and in both approaches they are computed through free cross sections including medium corrections. In the semi-classical approach the computation of such cross sections is very sophisticated, making it very slow. Meanwhile, the intranuclear cascade is much faster, but more based on empirical measurements, making it less reliable for hadron energies below 200 MeV.

4.3 Neutrino Event Generators

One of the key ingredients of the ν experiments is the Monte Carlo simulation. It plays an important role in predicting the rate of ν at different energies convoluting the flux, cross section and detector response.

The ν event generator is particularly important. It creates a set of particles emerging from the scattering between ν and nuclear targets. Currently, several generators are available in the market and most of them share the same philosophy trying to predict accurate rates without being extremely heavy computationally. The work flow breaks down in the following steps: ν-nucleon interaction, hadronization and nuclear effects. The main differences between them are due to the usage of different models and parametrization. During this dissertation two of them have been used (`NEUT v5.3.2` [99] and `GENIE v2.8.0` [98]) and they will be described in the following paragraphs.

Firstly, for ν-nucleon interaction, the two generators distinguish between QE, RES and DIS interactions.

- **QE**: This interaction is modeled in the same way by both generators [74]. Both of them use the same parametrization for the vector form factors [100] and the axial form factor (dipole shape with $g_A = -1.267$). However, the default axial mass parameter is different: $M_A^{NEUT} = 1.21$ and $M_A^{GENIE} = 0.99$.

- **RES**: This interaction is modeled in the same way by both generators [82]. On the one had, `NEUT` uses this model to generate only single hadron reactions. Besides, it uses 18 resonances taking into account their interferences and the lepton mass. The default parameters for the form factor are taken from [106]. On the other hand, `GENIE` does not restrict the model to single hadron production. It incorporates 16 resonances without including interference terms and neglecting the lepton mass. The default form factors are taken from [107].

- **DIS**: This interaction is modeled using the same PDF parametrization [86] including a modification to describe scattering at low Q^2, which is slightly different in both generators: `NEUT` uses [101] and `GENIE` uses [102].

- **RES-DIS transition**: At low Q^2, DIS model predicts hadron production which has already been taken into account by the RES model. Both generators define a similar way to avoid double counting for low multiplicity inelastic reactions. On the one hand, `NEUT` generates single hadron reaction using its RES model at $W < 2$ GeV and DIS model above that value. Multiple hadron production is generated only by its DIS model. On the other hand, `GENIE` does not distinguish between single or multiple π production. Its RES model is switched off when $W > 1.7$ GeV, while its DIS model is suppressed below that value using a parameter which depends on the multiplicity of the interaction.

Secondly, for the hadronization process, both generator take as a reference the value of the W. In the case of `GENIE`, for $W < 2.3$ GeV it uses the phenomenological description based on [103] and for $W > 3$ GeV it uses [104], while in between those values the model for high invariant mass is incorporated gradually. In the case of `NEUT`, for $W < 2$ GeV it uses a custom model based on [103] and for $W > 2$ GeV it uses [105].

Finally, the nuclear effects are treated differently by both generators:

- **Nuclear model**: In the case of `GENIE`, the RFG is used (including short range correlation between nucleons [108] and Pauli blocking) for all processes. Mass density is obtained from [109]

using the Woods Saxon distribution [110]. For NEUT, a more sophisticated model is used based on spectral functions [88] (including Pauli blocking using a hard cut off on the momentum).

- **Coherent scattering**: Both generators use the Rein-Sehgal model [111] including a correction that takes into account the lepton mass [112]. However, the implementation of the model is slightly different in both generators, so different cross sections are obtained mainly for low momentum π^+.

- **2p2h**: NEUT includes multinucleon reactions using a Nieves model [94], while GENIE does not include this kind of reactions.

- **FSI**: Both generators simulate FSI using the intranuclear cascade approach, but using different prediction for the cross sections. In the case of NEUT, π interaction probabilities are computed differently depending on the momentum of the π: if $p_\pi < 500$ MeV/c it uses a density dependent model [96] and if $p_\pi > 500$ they are extracted from $\pi - N$ scattering experiments [113]. For nucleons, the same approach is used, taking into account elastic scattering (extracted from [114]) and π production by delta decay (extracted from [115]). In the case of GENIE, the default model, called INTRANUKE hA, extracts the interaction probabilities from several experiments up to 300 MeV/c, while for higher energies it is mainly based on [116].

Fig 4.5 compares the predicted cross section for both event generators for $\nu_\mu - CH$ interactions via CC.

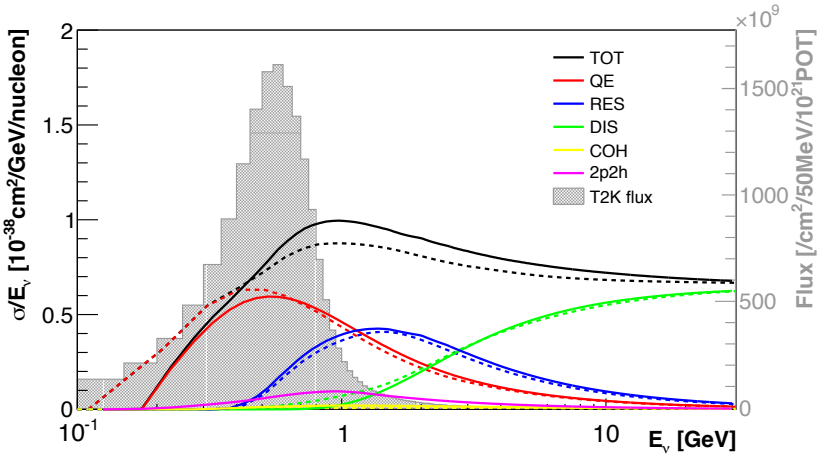

Figure 4.5: Predicted cross section for $\nu_\mu - CH$ interactions via CC (including exclusive channels) for NEUT (dashed lines) and GENIE (continuous lines) using the default parameters described in Sec. 4.3.

4.4 Experimental Status

In the 70's and 80's, ν cross sections were measured using bubble chambers filled with either Hydrogen or Deuterium. Four experiments provided measurements using reaction channels: ANL, BNL, BEBC, FNAL. Fig. 4.6 shows several measurements published by such experiments.

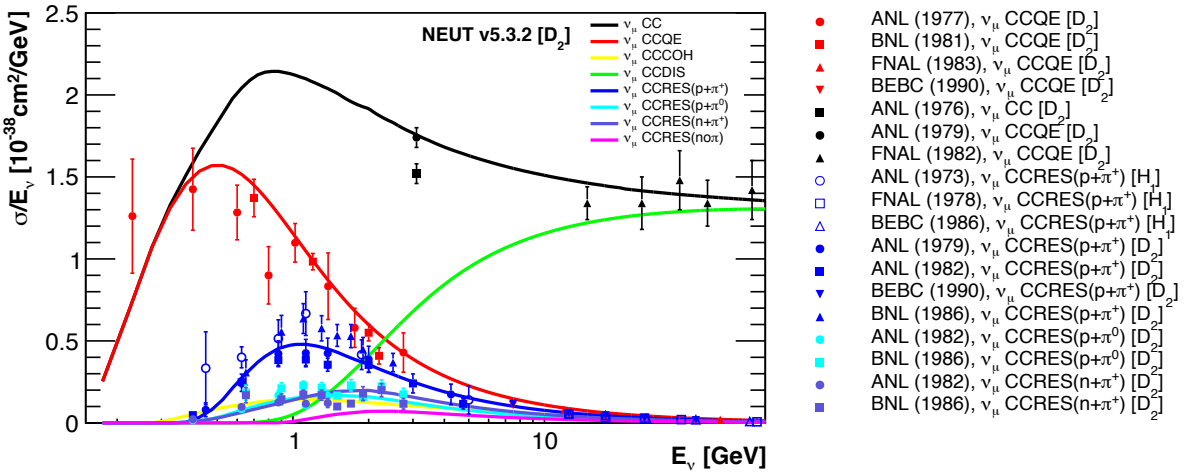

Figure 4.6: Cross section as function of the ν energy for ν_μ $-^2$ $D(^1H)$ interactions for different experiments and reaction channels. Lines represent the prediction from NEUT for Deuterium.

The main advantage of those measurements was the low Z, so they could be compared with free nucleons models. However, they were very limited by statistics and the poor knowledge of ν flux produced in accelerators, leading to measurements with large uncertainties.

In most cases, the ν flux was determined using a sample of CCQE events (assuming a value for M_A). This procedure is dangerous if the same data are used to extract the cross section. Nevertheless, those experiments provided very good constraint on the value of M_A studying the shape of such cross section versus Q^2 (independent from the flux normalization). The averaged result, $M_A = 1.026 \pm 0.021$ GeV, is in good agreement with the value extracted from π^+ electroproduction $M_A = 1.069 \pm 0.016$ GeV [79].

Since the 90's, several experiments have measured cross sections of ν interactions with complex nuclei. In current oscillation experiments, the main targets are either carbon or water. Fig 4.7 shows some of the cross section published by such experiments for ν_μ scattering using both inclusive (SciBooNE [117], NOMAD [118], T2K-ND280 [119], T2K-INGRID [120] and Minerva [121]) and QE (NOMAD [122], MiniBooNE [123], Minerva [124] and T2K-INGRID [125]) channels.

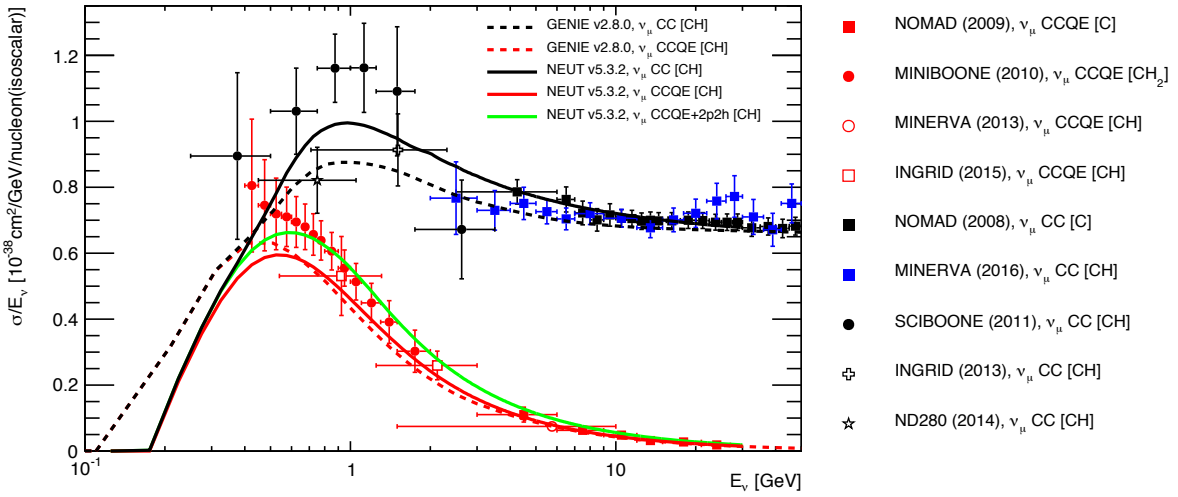

Figure 4.7: Cross section as function of the ν energy for $\nu_\mu - C$ interactions for different experiments and reaction channels. Lines represent the prediction from `GENIE` and `NEUT` for CH.

Currently, both experimentalists and theoreticians are concerned about the way to present a cross section measurement. In order to extract a cross section, the number of reconstructed events for certain signal must be corrected taking into account the acceptance of the detector. That correction factor is obtained from Monte Carlo simulation biasing the result depending on the model predictions.

In the case of CC interactions, experiments define as signal those events in which a μ is detected. Current detectors are optimized to reconstruct μ going forward, while efficiency for backward and high angle region of the phase space is very low. This fact leads to an efficiency correction which is very model dependent in those regions. Therefore, it is not advisable to compute the cross section as function of either Q^2 or E_ν because such model dependency is spread along all their range [124] [123].

The signal definition, particularly important in the QE reactions, can also lead to model dependencies. Current experiments can not distinguish between QE, 2p2h and RES with π^+ absorption reactions. This problem was visible in the case of MiniBoone results for CCQE cross section. That result pointed to a very large value of M_A for models. This could be explained by the fact that in the cross section computation only RES contribution was subtracted (not the 2p2h). However, such background extraction is dangerous when the predictions for certain processes are not reliable (i.e. 2p2h reactions or π FSI). Such problem can be also extrapolated for measurements of RES interactions.

In the case of events with high multiplicity, applying an efficiency correction as function of the kinematics of a single particle (usually μ) can also lead to model dependencies. That problem is very

difficult to overcome, because the solution would be to apply a efficiency correction that depended on the kinematics of all the particles coming out from the interaction. The main problem is that current knowledge of the kinematics of the hadrons is very poor, so model predictions are not reliable.

In conclusion, it is advisable to apply the efficiency correction as function of basic reconstructed variables (as p_μ , $\cos\theta_{\nu\mu}$, p_π or $\cos\theta_{\mu\pi}$), to assume big uncertainties and to publish such corrections. In exclusive channels, it is important to define the signal using topologies of particles leaving the nucleus. This approach has been followed by the latest cross section results from T2K [126].

Part II

T2K Experiment

Introduction

The main goal of T2K (Tokai-to-Kamioka) has been to provide competitive measurements of ν oscillation parameters. It was designed to study with high precision the ν_μ disappearance channel and to improve current knowledge of $\nu_\mu \rightarrow \nu_e$ appearance channel. This second channel is particularly important, because it has been the first experiment using an accelerator able to measure ν_e appearance.

T2K configuration has been the common one for long baseline experiments (see Fig 4.8). An accelerator is used to produce ν, which are detected in a far detector (Super-Kamiokande, which is 295 km away from the production target) and a near detector complex (both INGRID and ND280, which are 280 m away). One of the main features of T2K is that the ν beam is not straightly directed to the far detector but with an angle of 2.5° (off-axis technique). In this configuration, the ν_μ beam pointing to the far detector has a narrow band in the energy spectrum peaked at ~ 0.6 GeV (which maximes the oscillation probability at 295 km and reduces the ν_e contamination in the beam).

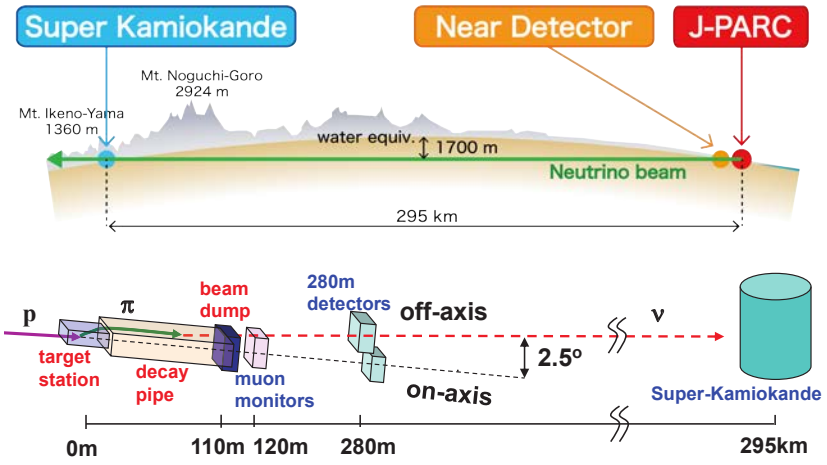

Figure 4.8: Schematic representation of the T2K configuration.

In addition to oscillation measurements, such particular energy spectrum provides an excellent opportunity to study ν interactions using the near detectors. In the following chapters the different subsystems that constitute the T2K will be reviewed. They can be broken down in: beam (including accelerator

and beamline), near detector (including ND280 and INGRID) and far detector. In particular, in this dissertation, ND280 has been used for the main analysis, so it will be described in more detail. In fact, a very detailed description of T2K experiment can be found in [127].

Chapter 5

T2K Beam

In this chapter, the ν production is explained. In order to produce them several stages are followed. Firstly, protons are accelerated in the J-PARC facilities. Secondly, they are collided with a graphite target pointing to Super-Kamiokande using a primary and secondary beamline. Finally, mainly ν and μ appear from the decay of the produced hadrons, which are focused using magnetic horns [128]. μ are used to monitored the stability of the beam (although most of them are stopped in a beam dump) [129].

The flux prediction of the ν depends mainly on in the hadron production model. In order to understand such production, data from an external experiment (NA61/SHINE) are used. A very detailed description about the T2K flux can be found in [130].

5.1 Accelarator (J-PARC)

Placed at Tokai, the accelerator is composed by three subsystems (see Fig. 5.1). Firstly, a **linear accelerator (LINAC)** accelerates H^- up to 181 MeV, which are then converted to H^+ by charge-stripping foils. Secondly, the beam is directed to a **rapid synchrotron (SRC)** and it reaches 3 GeV with a 25 Hz cycle. Finally, the proton beam is injected to the **main ring (MR)** which has a length of 1567 m. The beam is accelerated up to 30 GeV with a maximum power of 750 kW (currently $\sim$ 400 GeV is used). Each spill consists of $\sim 3.4 \times 10^{14}$ of $\sim 5\mu s$ width which is split into 8 bunches. In order to discriminate backgrounds in the different detectors, the timing information of the different spills and bunches is crucial, so that events can be synchronized with near and far detectors.

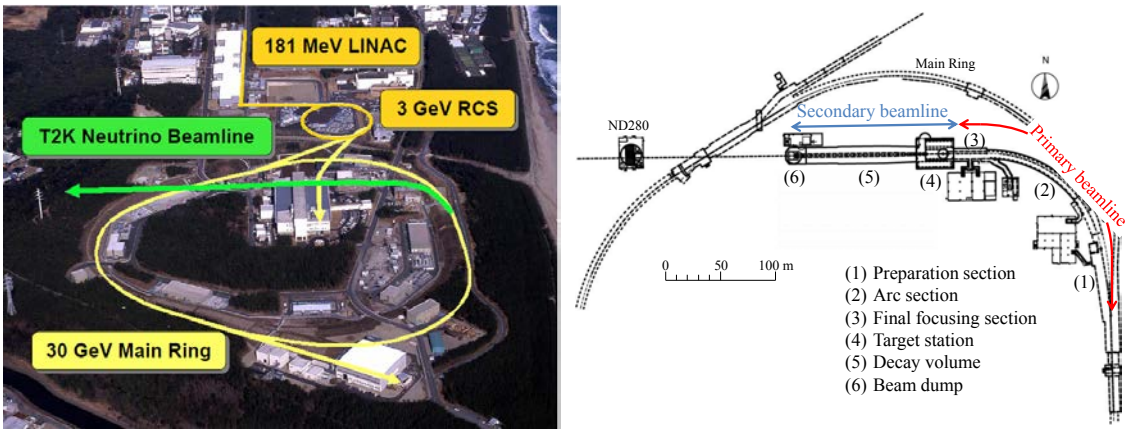

Figure 5.1: Left: Top view of the J-PARC facilities. Right: Overview of the T2K beamline.

5.2 Primary Beamline

The primary beamline is used to extract the proton beam from MR and point it towards Super-Kamiokande (see Fig. 5.1). In a **preparation section** (54 m long), the proton beam is tuned using 11 normal conducting magnets, so that the beam can be used by the arc section. In the **arc section** the beam is bent toward Super-Kamiokande using a set of 14 doublets of superconducting combined functions magnets (working at 4360 A and cooled helium at 4.5 K). The beam orbit is corrected using superconducting steering magnets (both horizontal and vertical). Finally, in the **focusing section**, the beam is focused onto a target (directing it downward by 3.6° with respect the horizontal plane) by 10 normal conducting magnets.

The stability of the proton beam is essential in order to reach high-power operation. Thus, beam features, such as intensity, position, profile and losses are monitored.

5.3 Secondary Beamline

This beamline consists of a target station, a decay volume and a beam dump (see Fig. 5.2).

The **target station** is connected with the primary beamline on one side and with the decay volume on the other side. It is located within a steel vessel (10 cm thick) filled with Helium. Firstly, a graphite baffle with a hole of 30 mm diameter is used to collimate the proton beam, whose profile is measured by an optical transition radiation monitor (OTR).

Secondly, the proton beam collides with a graphite target (denser materials would be melted) sealed inside a titanium case. It is a cylinder of 91.4 cm long and 2.6 cm diameter, which is cooled by helium gas and it is expected to reach 700° C with an intense proton beam.

From the collision *pi* are mainly generated, which are collected by a first magnetic horn. Then, two consecutive horns are used to focus the hadrons. With an operation current of 320 kA, they create a toroidal magnetic field with a maximum field of 2.1 T, which increases by a factor of ~ 16 the ν flux at Super-Kamiokande (compared to a configuration without horns). The horns dimension is (minimum diameter and length respectively) 54 mm and 1.5 m for the first, 80 mm and 1 m for the second and 140 mm and 2.5 m for the third. Positive or negative hadrons can be focused or deflected depending on the polarity of such horns. When positive hadrons are focused it is called forward horn current (FHC) and in the case of negative hadrons it is referred as reverse horn current (RHC). In this dissertation, FHC configuration is used because it provides a very pure ν_μ beam (while RHC provides a $\bar{\nu}_\mu$ beam).

The beam of hadrons enter a **decay volume** where they decay into ν and leptons (mainly μ). This volume is ~ 96 m long, 1.4 (3.0) m wide and 1.7 (5.0) m high at the upstream (downstream) end. Such a length was optimized in order to enhance the ν_μ contribution and reduce the ν_e and $\bar{\nu}_\mu$ contamination from μ decay.

Finally, at the end of the decay volume a graphite and iron **beam dump** is placed. Its dimensions are 3.17 m long (plus 2.40 m of iron thickness), 1.94 wide and 4.69 m high. Only ν and very energetic μ (above 5 GeV) will go through the dump.

5.4 Muon Monitor (MUMON)

Just after the beam dump, two detectors arrays (ionization chambers and silicon PIN photodiodes) are placed in order to monitor the ν beam intensity and direction with a precision better than 3% and 0.25 mrad respectively. They are optimized to measure the distribution profile of μ (produced along with ν in hadron decay), whose center is directly related with the ν beam direction.

For a 320 kA horn current with 3.3×10^{14} protons per spill, the intensity of μ per bunch in these detectors is 1×10^7 cm^{-2}. The profile is estimated to be Gaussian-like with a 1 m width and composed mainly by μ ($\sim 90\%$).

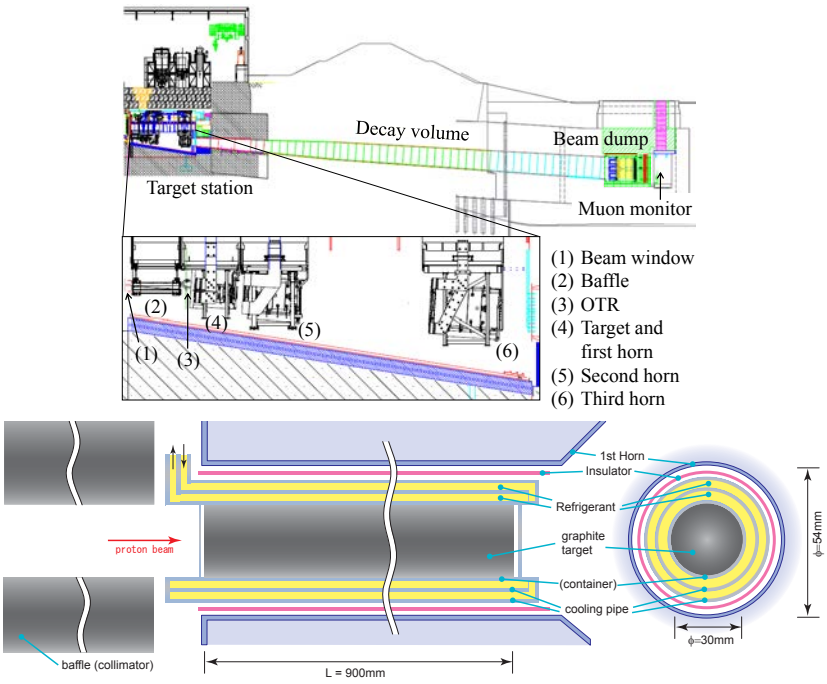

Figure 5.2: Top: Side view of the T2K secondary beamline. Bottom: Sketch of the target.

5.5 Neutrino Flux

The ν flux prediction is based on simulation. The primary proton beam collisions in the baffle and target are handled by FLUKA2011 [131]. Then, particles emitted from such collisions are propagated through the horns and decay volume using JNUBEAM program, which is based on GEANT3 [132] and modeled hadronic interactions with GCALOR [133]. The particles are tracked until they decay or their kinetic energy drops below 10 MeV. ν are generated from the decay of $\pi^{\pm}$, $K^{\pm}$, K^0_L and $\mu^{\pm}$ using the current best knowledge of branching ratios (see Table 5.1).

Particle	Decay Products	Branching Fraction (%)
π^+	$\mu^+\nu_\mu$	99.9887
	$e^+\nu_e$	0.0001
K^+	$\mu^+\nu_\mu$	63.55
	$\pi^0 e^+\nu_e$	5.07
	$\pi^0\mu^+\nu_\mu$	3.35
K^0_L	$\pi^- e^+\nu_e$	40.55
	$\pi^-\mu^+\nu_\mu$	27.04
μ^+	$e^+\nu_e$	100

Table 5.1: Decay modes with a ν in the final state considered in JNUBEAM and their branching ratio ($\bar{\nu}$ are not taken into account).

The parameters used to model hadronic interactions are tuned and constrained using external measurements of π and K yield performed by NA61/SHINE experiment [62]. Those data were obtained using a graphite target of 2 cm thickness with a proton beam similar to T2K placed at CERN. The errors associated to these interactions are the main uncertainty in the flux prediction. Other uncertainties are associated with the alignment of the proton beam and the magnetic horn, but they are smaller. It will be described in more detail in Sec. 9.1.

With all the tools described previously, the ν flux shape and normalization is predicted for each flavor in both near and far detectors (see Fig. 5.3).

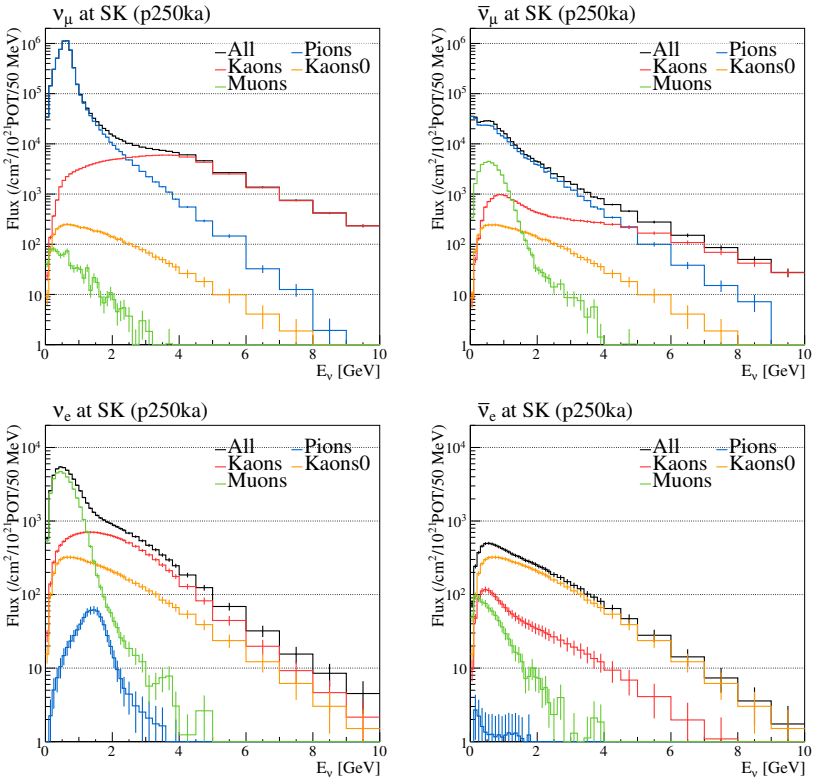

Figure 5.3: ν flux prediction at Super-Kamiokande in FHC mode (version 13a) broken down into the different flavors: ν_μ (top left), $\bar{\nu}_\mu$ (top right), ν_e (bottom left) and $\bar{\nu}_e$ (bottom right). Colors indicate the parent of the ν.

The first important aspect is related to the flux shape, which is connected with the off-axis configuration (see Fig 5.4). The dominant channel in the ν production is the π decay ($\pi^+ \rightarrow \nu_\mu + \mu^+$). As a two-body decay, the energy of the ν can be inferred from the energy of the π, the masses and the off-axis angle θ_{OFF}:

$$E_\nu = \frac{(1 - m_\mu^2/m_\pi^2)E_\pi}{1 + \frac{E_\pi^2 \tan\theta_{OFF}}{m_\pi^2}} \tag{5.1}$$

This relation shows that the dependency between E_ν and E_π becomes less linear when the off-axis angle increases. Using this technique a peaked spectrum is achieved in comparison with the on-axis configuration. It was designed to have its peak centered at the first oscillation maximum in the far detector (enhancing both ν_μ disappearance and ν_e appearance channels). Besides, the CCQE reaction, which is the channel used to reconstruct the energy of the incoming ν at the far detector, becomes dominant at this energy,.

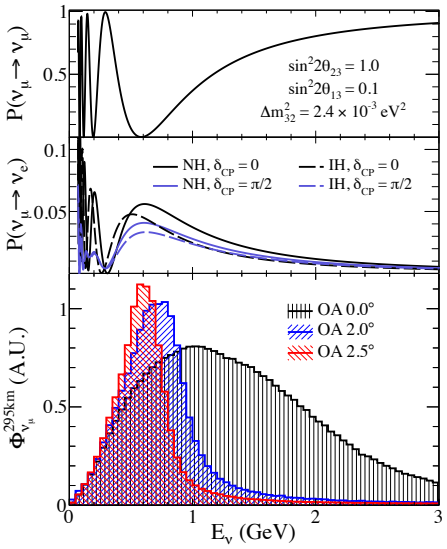

Figure 5.4: ν_μ disappearance and ν_e appearance probability at the far detector and ν flux for different off-axis angles.

The second important aspect is related with the flavor composition of the ν beam. In summary, in the FHC configuration, the ν_e ($\bar{\nu}_\mu$) contamination is around 1% (5%) below 1.5 GeV. At intermediate energies (1.5-3 GeV), the fraction of π^- that are not focused increases, enhancing ν_μ contribution. For higher energies, the contribution from K becomes significant so the impact of $\bar{\nu}_\mu$ decreases.

Chapter 6

T2K Detectors

T2K detectors have different purposes, although their main goal is to provide information for the oscillation analysis. They have been placed in two different locations: Kamioka (295 km away from the production target) and Tokai (280 m away).

On the one hand, the far detector called Super-Kamiokande is in Kamionka. Its main goal is to measure the ν rate at different energies in order to characterize the flux.

On the other hand, the near detector complex is composed by two main detectors: one on-axis (INGRID [134]) and one off-axis (ND280 [135]). INGRID is used to monitor the ν beam direction and profile. Meanwhile, ND280 has been designed to measure ν_μ flux (and ν_e contamination). Besides, it provides information about ν interactions with complex nuclei (as carbon or water). Therefore, the near detector complex has become crucial in order to reduce oscillation analysis uncertainties.

6.1 Far Detector: Super-Kamiokande (SK)

SK is an enormous cylindrical tank built underground (1 km depth) filled with pure water (roughly 50 kton). It is surrounded by ∼13,000 photomultiplier tubes (PMT) that detect Cerenkov light produced by charged particles crossing the detector. The detector was built in 1996 and it has provided competitive results in different fields of particle physics as proton lifetime or atmospheric, solar and accelerator ν oscillation. The performance of SK is very well understood thanks to its long lifetime.

It consists of two volumes separated by a stainless steel structure. The inner volume (33.8 m diameter and 36.2 m height) is surrounded by ∼11,000 PMTs with 50 cm of diameter, while the outer volume houses ∼2,000 20 cm PMTs. A pictorial view of the detector is shown in Fig. 6.1.

In the oscillation analysis within T2K, the far detector measures the flux from the T2K beam. The strategy is to measure the rate of ν_μ and ν_e CCQE interactions within the detector volume observing

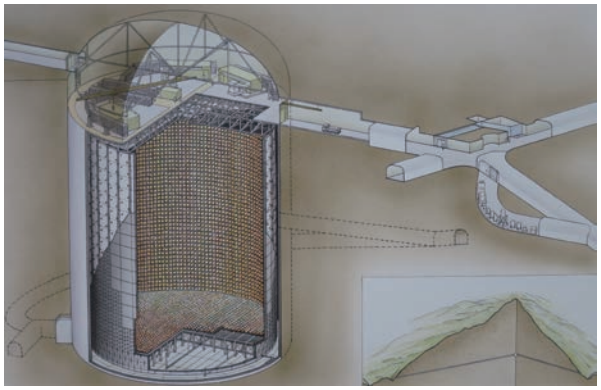

Figure 6.1: Diagram of the Super-Kamiokande Detector.

the produced leptons. Those charged particles, above certain energy threshold, produce a cone of Cerenkov light which is detected by the PMTs. The light pattern is used to extract the position of the interaction and momentum of the outgoing particles. Then, the ν energy is reconstructed using the outgoing lepton kinematics (assuming four body process). The main contamination comes from NC interactions in which a π^0 is generated and the light pattern is created by the two photons produced in the π^0 decay. When the two photons are collinear, it is likely to be misidentified as a ν_e CCQE interaction.

The outgoing lepton identity is extracted studying the features of the light pattern in the PMTs. In the case of μ, a sharp ring of hits is seen. Meanwhile, for $e^{\pm}$, electromagnetic showers are induced producing a blurry ring pattern. A comparison between them is shown in Fig. 6.2.

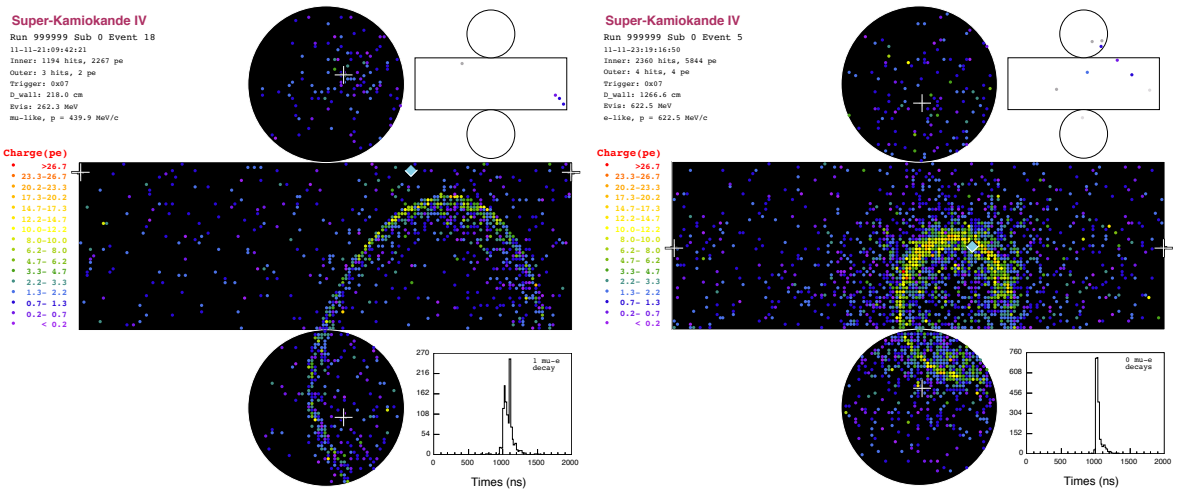

Figure 6.2: PMTs hit pattern observed in SK for mu-like (left) and e-like (right) events.

6.2 On-axis near detector: INGRID

INGRID is composed by 14 identical modules which are horizontally and vertically arranged in both groups (see Fig. 6.3). The center of the cross is placed 280 m away from the production target, at 0 degrees with respect to the proton beam direction. Its main purpose is to monitor the beam direction and intensity using a sample of ν interaction with iron (high statistics are obtained daily). Comparing the number of events observed in each module, a 0.4 mrad precision is obtained for the beam center.

The 14 modules are composed by 11 tracking scintillator planes interleaved with iron plates and surrounded by veto scintillator planes. The dimension of the iron plates is 124 cm high and wide and 6.5 cm deep, which makes a total iron mass of 7.1 tons. Each scintillator plane consists of 24 polystyrene scintillator bars ($1.0 \times 5.0 \times 120.3$ cm^3) both in horizontal and vertical directions. The veto scintillator plane is used to reject any interactions from outside the module.

An extra module, called proton module, is used to detect CCQE events, in which a μ and a proton are produced (see Fig. 6.3). The main purpose is to provide information for this interaction channel using an energy spectrum different from ND280. The main differences with respect to the other modules are that it does not contain iron planes and the scintillator bars are thinner in order to improve the tracking capabilities.

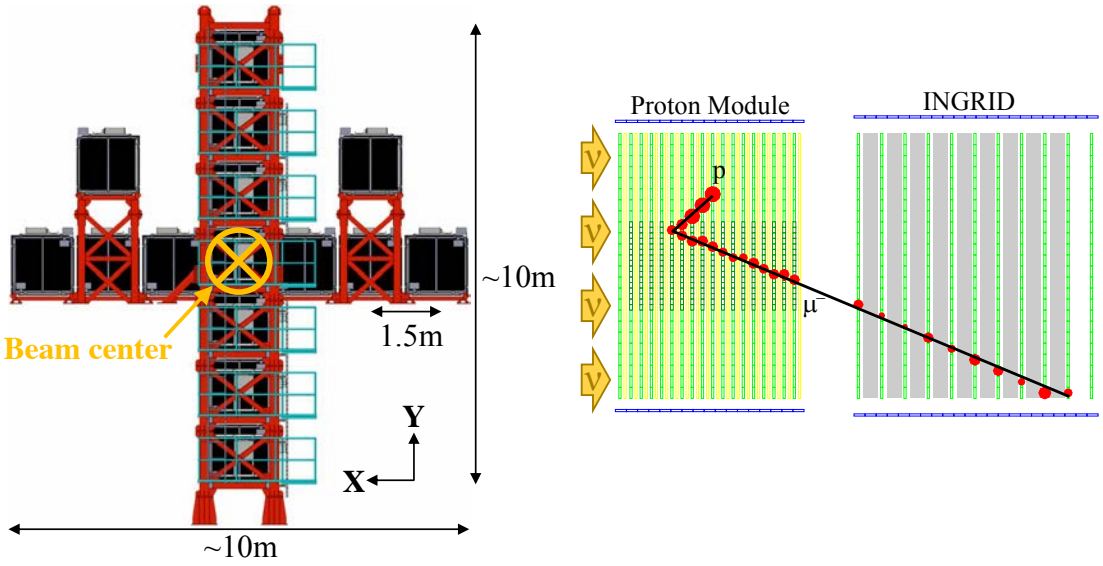

Figure 6.3: Left: Diagram of the INGRID Detector. Right: Typical ν CCQE interaction in INGRID. Green (blue) cells represents a tracking (veto) scintillator, red circles indicate the observed signal in that cell.

6.3 Off-Axis Near Detector: ND280

The near detector placed off-axis is made up of two main subdetectors: P0D and TPC/FGD sandwich (tracker). Both of them are contained in a metal basket box surrounded by electromagnetic calorimeters (ECal) and a magnet recycled from UA1 experiment (see Fig. 6.4). A very detailed description of the different subdetectors is also available: TPC [136], FGD [137], ECal [138], SMRD [139] and P0D [140].

This fine grained multi detector has been used to measure the ν flux and flavor content. Firstly, it provides a very detailed information about the ν_μ flux in the off-axis configuration. Secondly, it constrains the ν_e contamination in the flux (roughly 1%), which has an impact on the appearance analysis. Finally, it is also used to characterize relevant ν interactions in Super-Kamiokande measuring rates of either inclusive or exclusive reaction channels. CCQE and NC single π^0 are particularly important, which are the signal and main background in the oscillation analysis.

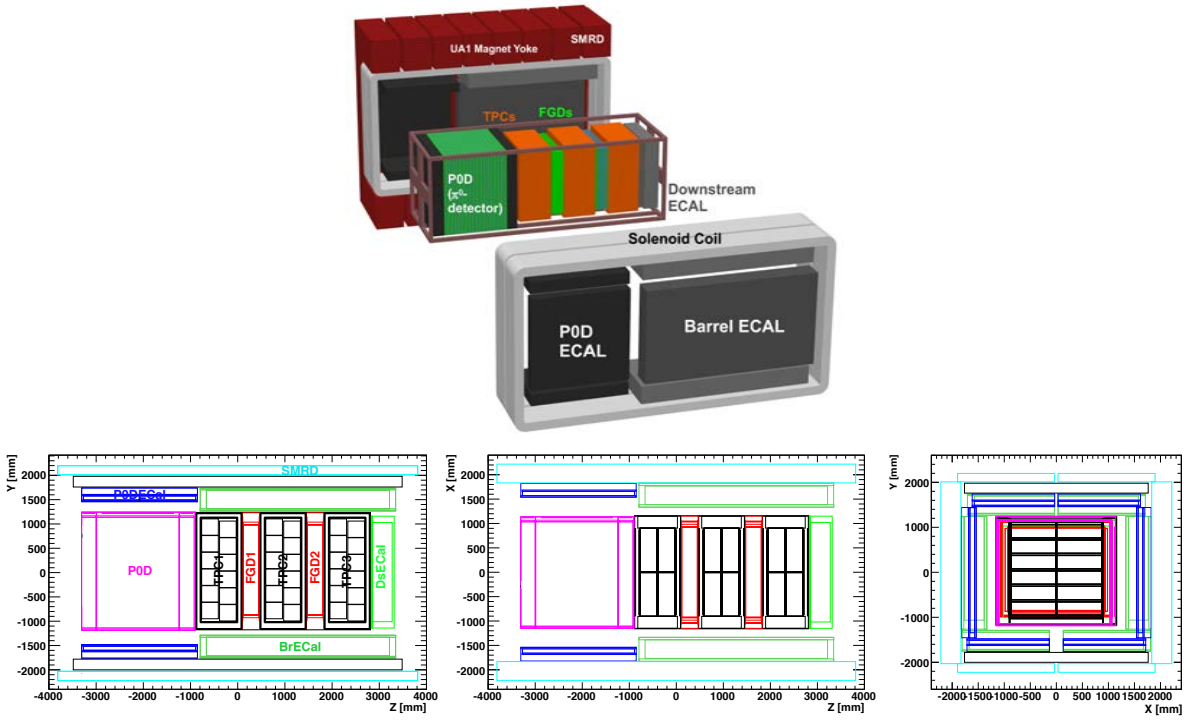

Figure 6.4: Top: Side view of the different subdetectors composing ND280. Bottom: Top view (left) and bottom view (right) of the ND280 basket and its containing subdetectors.

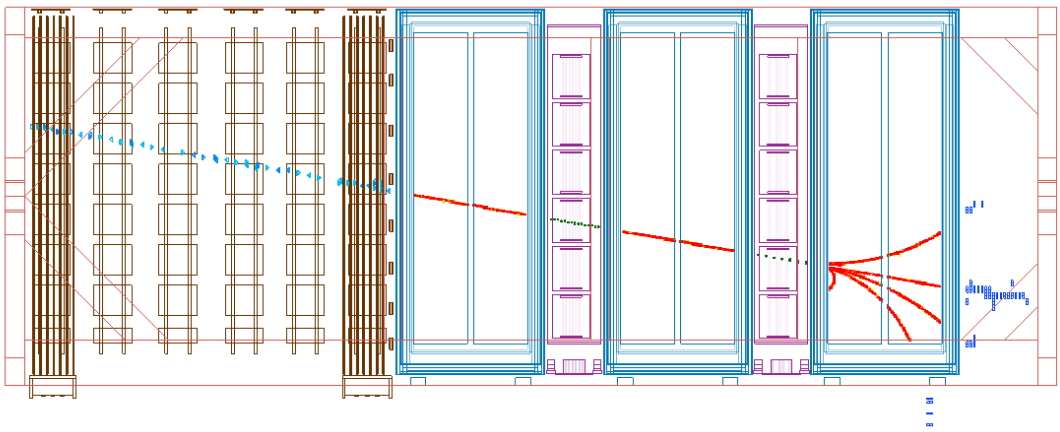

Figure 6.5: Event display of a cosmic ray crossing ND280 and interacting in the second FGD.

6.3.1 UA1 Magnet

The magnet consists of 16 C-shaped iron yokes and 4 aluminum coils, which creates a magnetic field of 0.2 T orthogonal to the beam direction when the nominal current is 2900 A. Water flows through the aluminum bars in order to cool them down. The dimensions of the inner (outer) volume are $7.0 \times 3.5 \times 3.6$ m^3 ($7.6 \times 5.6 \times 6.1$ m^3) making a total weight of 850 tons.

The charged particle motion in ND280 magnetic field (not parallel to the particle direction) will be bended. From such curvature, momentum of those particles can be obtained with a precise resolution. Besides, the charge sign can be extracted, so good rejection criteria for $\nu/\bar{\nu}$ interactions are possible.

A dedicated mapping procedure of the magnetic field was performed using a device equipped with Hall probes [141]. The measurements were performed at a magnetic field of 0.07 T, so it was necessary to scale to the nominal field (0.2 T). In the end, a final magnetic field uncertainty of 2 G was achieved for each component. Fig. 6.6 shows the magnetic field in one slice of the x plane. A very precise knowledge of the magnetic field is crucial in order to reduce the uncertainty in the momentum reconstruction.

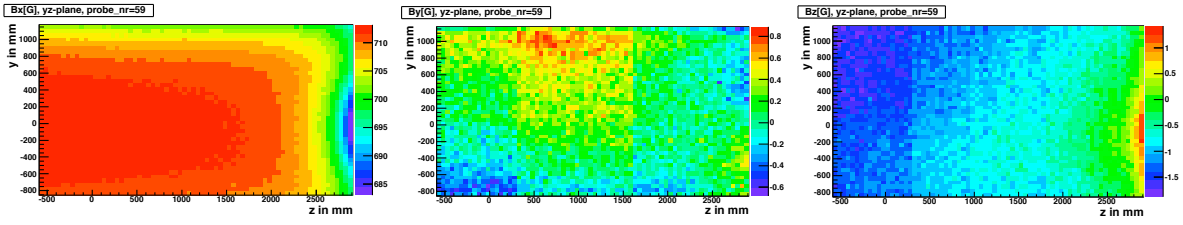

Figure 6.6: Magnet field in the x (left), y (middle) and z (right) direction for the plane $x = 0$. The colors indicate the values in G.

6.3.2 Side Muon Range Detector (SMRD)

The SMRD is made up of 440 scintillator modules placed in the 15 air gaps between the magnet yokes. Fig. 6.7 shows a single yoke structure and the slits in between iron plates. Each module is placed either horizontally or vertically and it consists of four or five scintillator counters ($875\times170\times7$ mm^3) respectively. Wave length shifter (WLS) fibers are glued into each scintillator bar and attached to Multi-Pixel Photon Counters (MPPCs), whose signal is read by the Trip-T front-end boards (TFBs).

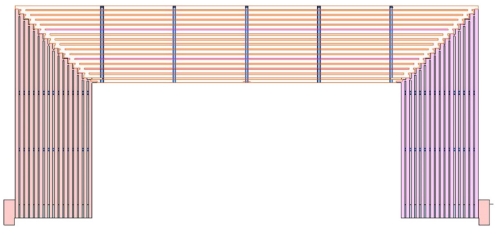

Figure 6.7: Sketch of a single yokes rotated 90 degrees. The iron plates are held by a series of bolts, which divide the device in four horizontal, four vertical and two corner sections.

This detector is used to trigger on interactions which do not happen inside the basket (as cosmic rays or interactions within the iron of the magnet). Moreover, it is also used to measure length/momentum of μ with high angles with respect to the beam direction, for which curvature can not be reconstructed.

6.3.3 Pi-Zero Detector (P0D)

The detector design can be seen in Fig. 6.8. Its central part consists of 26 scintillator modules interleaved with water layers (28 mm thick) and brass sheets (1.5 thick). Each water layer contains two fillable water bags. The upstream and downstream parts are made up of 7 scintillator modules interleaved with stainless steel clad lead sheets (4 mm thick). The detector size is 2103 mm wide, 2239 mm high and 2400 mm long, making a total mass of 16.1 and 13.3 tons with and without water respectively.

Each scintillator module has two perpendicular arrays made of triangular polystyrene scintillator bars (134 vertical and 126 horizontal). The cross section of each triangle is 33 mm base and 17 mm high. Each bar contains a wave length shifter fiber (1 mm of diameter) mirrored at one end and attached to a MPPC at the other end. The MPPCs are connected to TFB electronic boards.

The goal of this detector is to observe NC interactions on water, in which a π^0 is emitted. This reaction one of the main backgrounds in the ν_e selection at Super-Kamiokande. In this detector, electromagnetic

showers are usually contained within its volume. Besides, lead layers provide a veto region before and after the water target region. The detector segmentation is good enough to reconstruct charged particle tracks (μ and π) and electromagnetic showers ($e^{\pm}$ and photons from π^0 decay). Furthermore, the cross section for interactions happening in water can be computed using a subtraction technique.

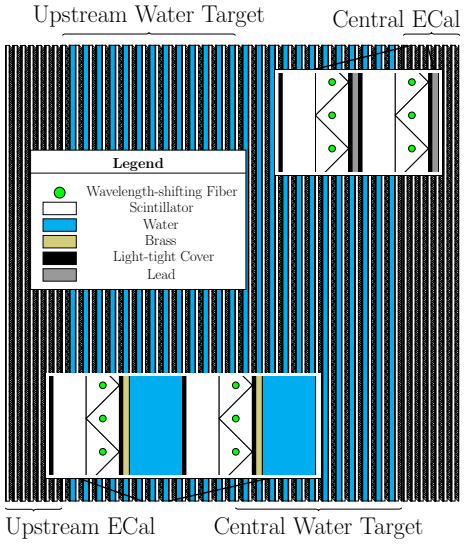

Figure 6.8: Sketch of P0D. Details of the different layers can be seen in the insets.

6.3.4 Electromagnetic Calorimeter (ECal)

The ECAL consists of 13 modules surrounding the inner detectors. The tracker module is covered by six modules in the sides (BarrelECal) and one module in the downstream exit (DsECal). Other six modules surround the sides of P0D (P0DECal). This configuration provides an almost full coverage for all particles exiting the inner detectors.

The modules are made up of polystyrene scintillator bars (4.0×1.0 cm^2 of cross section) interleaved with lead sheets (thickness of 1.75 mm for BarrelECal and DsECal or 4 mm for P0DECal). A 1 mm diameter WLS runs along a hole in the center of each bar which is attached to a MPPC in both ends. The MPPC signal is readout by TFB cards.

The BarrelECal modules contain 31 layers with scintillator bars of 1.52 m, 2.36 m and 3.84 m long running in the x, y and z direction respectively. In the case of DsECAL, there are 34 layers each one with 50 scintillator bars of 2.04 m long. For P0D-ECal, it has six layers with scintillator bars of 2.34 m long running along the z direction for all layers.

The main role of this detector is to complement the reconstruction of the inner detectors. In the case of the tracker, it is used to detect photons that do not convert into e^+e^- pairs within it, so ν interactions with π^0 production can be identified. In the case of P0D, it provides information about the escaping energy in π^0 production. Moreover, as SMRD, it is used to measure length/momentum of μ escaping with high angles with respect to the beam direction, for which curvature can not be reconstructed. Finally, in the case of particles which do not crossing the TPCs, BarrelECal and DsECal play a key role in order to identify such particles.

The clusters produced by $\mu^{\pm}$ (track-like), $e^{\pm}$ and hadrons (shower-like) have completely different features when crossing an electromagnetic calorimeter. Therefore, those features provide a particle identification tool (PID).

6.3.5 Fine Grained Detector (FGD)

The two fine grained detectors (FGDs) are sandwiched in between the TPCs. They consists of polystyrene scintillators bars ($9.61\times9.61\times1864.3$ mm^3), which are perpendicularly oriented to the beam in both x and y directions. Each bar (coated with TiO_2) has WLS fiber in its center, it is mirrored at one end and attached to a MPPC at the other end. In contrast with the other scintillator-based detectors, the MMPCs are connected to completely different front-end boards.

One FGD is made up of 15 modules, in which two layers of 192 scintillator bars are oriented alternatively in the x and y directions, which are perpendicular to the ν beam. The second FGD consists of 7 scintillator modules interleaved with layers of water (2.5 mm thick). Both of them have a total dimension of $2300\times2400\times365$ mm^3 (width×height×depth in beam direction) and they contain 1.1 ton of material.

Their main purpose is to provide a target mass for ν interactions. Their fine segmentation permits the vertex to be reconstructed with precision and it also allows the reconstruction of charged particles that go through the detector. Besides, through the comparison of the interaction rates in both FGDs, cross sections on carbon and water can be extracted.

6.3.6 Time Projection Chamber (TPC)

The three TPCs are identical. Their outer structure is a aluminum box with an approximate size of $2.3\times2.4\times1.0$ m^3. It holds an inner box (made of copper-clad) that contains the drift gas (Ar:CF_4:iC_4H_{10} [95 : 3 : 2]), which is divided into two volumes by a central cathode. This inner box is isolated through immersion in CO_2 gas and it is attached to two readout planes placed in the side walls (see Fig. 6.9).

Each readout plane is made up of 12 bulk MicroMegas modules (342×359 mm^2) arranged in two vertical columns. There is an small offset between the two columns to avoid alignment between inactive regions. Each MicroMegas module has a thin metallic micro-mesh (pitch $\sim$50 μm) that separates the volume into two regions: drift volume and amplification gap (100 μm thin). On one side of the amplification gap a Printed Circuit Board (PCB) covered by photoimageable film is segmented in 36×48 pads of 6.85×9.65 mm^2.

An uniform electric field parallel to the magnetic field is obtained in the two inner volumes by setting to -350 V and -1150 V the central cathode and the micro-mesh respectively (the pads plane is at ground potential). The inner gas mixture was chosen because of its fast speed, low diffusion and good performance with MicroMegas.

When an charged particle passes through the TPC, it ionizes the gas producing e^-. Because of the electric field, the cloud of e^- will drift towards the readout plane. Then, in the amplification gap, such e^- are accelerated and they unbound other e^- from the gas. Consequently, an avalanche process is created until the amplified cloud reaches the pad plane. Such amplification is quantified using a multiplication factor, called gain. Fig. 6.9 summarizes the TPC and MicroMegas principles.

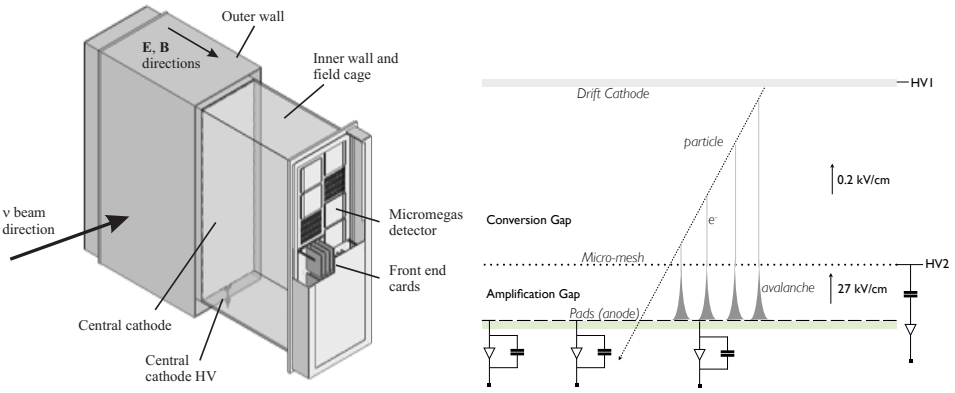

Figure 6.9: Left: Sketch of the TPC design. Right: TPC readout principle.

The TPCs perform a key role in the near detector. This technology has very precise tracking capabilities. Using the MicroMegas hit information (obtained from the pads), charged particles can be reconstructed in the YZ plane with precise resolution. In the case of the X coordinate, it is determined by comparing the timing from outer detectors (FGD, ECal or P0D), and the timing from the hits in the pads plane. Using this precise imaging feature, patterns from charged particles can be known accurately, such as multiplicity, orientation and curvature.

Moreover, it is possible to quantify the amount of ionization produced by a charged particle by measuring the deposited energy in the pads plane. This quantity is very different depending on the identity of the particle and its momentum. Therefore, it provides a powerful tool for PID.

6.3.7 Software Overview

The ND280 software is based on very well known programs. The system structure is based in CMT [142]. Meanwhile, most of the framework and storage model for data is based on ROOT [143]. Most of the simulation libraries are obtained from Geant4 [144]. In order to manage and store all the packages, CVS is used [145].

The structure of the software is split in different packages, which perform different roles in different stages (see Fig. 6.10). In the case of MC, some extra packages must be used before reaching the same stage as real data (related to the simulation of ν interaction and detector response). The different stages and packages are explained in the following sections.

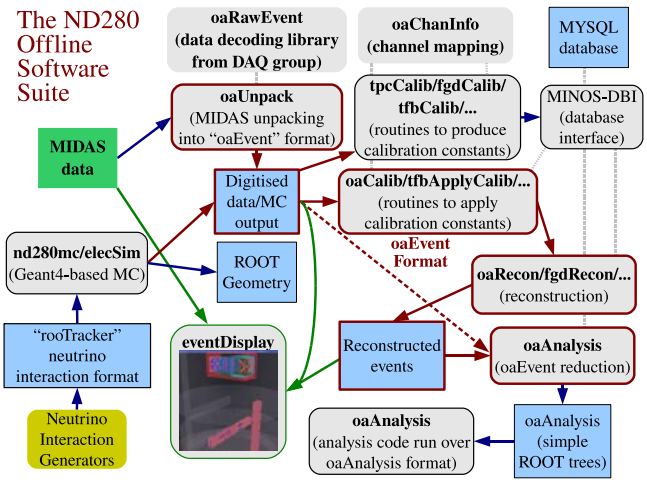

Figure 6.10: General structure of the ND280 Software including only main packages.

Simulation

The simulation of events within ND280 is done using two fundamental packages: `nd280mc` and `elecSim`.

- `nd280mc`: From the event generator (see Sec. 4.3) ν interactions can be simulated and grouped into spills. The amount of interactions per spill is computed using beam intensity predicted by the flux (see Sec. 5.5). Then, Geant4 is used to simulate the particles coming out from the ν interaction and passing through the ND280 subdetectors, whose geometry is simulated with ROOT. The information from Geant4 is saved in a list called "hits", which stores position, time and energy deposited by each particle.
- `elecSim`: This package simulates the response of the detector. In the case of the TPCs, it simulates the drift of e^- and the response of MicroMegas. For the other detectors, it simulates the light production in the scintillator bars and WLS and the response from the MPPCs. As `nd280mc`, this package is only needed for MC production. Nevertheless, its output is the same as the real data, including the charge and timing information for all electronic channels in ND280.

Calibration

The calibration of the real data is managed by a package called `oaCalib`. One of the main features of that software is the online storage of all the calibration constants using a database. Such constants are computed in parallel with data collection. The calibration is divided in four main items: charge, timing, alignment and physics calibration.

With the charge calibration, gain and pedestal constants are used to achieve the same reference system for the charge collected by each electronics channel in the same subdetector. The pedestal constants are obtained measuring the charge collected when no events are recorded ("pedestal" trigger), so it allows "dark noise" to be measured. In the case of the gain constant, they are computed by comparing the observed and expected energy deposited by a high energetic cosmic μ.

In the timing calibration, the same reference system for the time measured by each channel is pursued. The scaling and shifting constants are obtained from cosmic μ and they account for both bar-to-bar variation and delays introduced by the readout electronics. Bar-to-bar variations can be calibrated comparing the observed and expected time differences between two adjacent channels when a cosmic ray crosses them.

An important fact is that the response from scintillator (light yield) is not linear. This fact increases the complexity of the charge and timing calibration. This issue has obstructed the timing calibration for detectors based on TFB electronics (P0D, ECAL and SMRD), showing worse performance than FGDs.

Alignment calibration is needed in order to reduce the detectors geometry imperfections created during their construction or installation. Two different constants must be taken into account: shift and rotation. Currently, this calibration is applied in order to correct misalignments between different subdetector or MicroMegas modules in the TPCs. In order to extract those constants, events in which the magnetic field is zero are observed (particles pass through ND280 as straight lines).

Physics calibration tries to reduce the impact of charge attenuation or electromagnetic field distortions. Therefore, charge, time and position information from the channel are affected. This process is particularly important for the TPCs, in which electromagnetic distortions, signal attenuation or drift velocity need to be studied in more detail.

Firstly, a monitoring chamber which is connected to the same gas distribution as the TPCs is used to measure the signal attenuation and the drift velocity. Secondly, a pattern of thin aluminum strips is glued to the cathode in the TPCs. It releases e^- when it is illuminated with a 266 nm laser. Comparing the real and observed pattern, field distortions can be studied (drift velocity is also measured using this system). Thirdly, magnetic field distortions are also quantified using results from the survey with Hall probes (see Sec. 6.3.1). Finally, electric field distortions are quantified by studying the hit pattern in the pads plane for very energetic μ, whose curvature is negligible (see App. B).

Reconstruction

The reconstruction in ND280 has two phases. Firstly, each subdetector reconstructs locally hits pattern (either from MPPCs or MicroMegas) into two different objects (so called segments): tracks or showers. Secondly, a global reconstruction is performed by merging such segments. Position, orientation and momentum are extracted from global reconstruction. In the case of the PID, local information from each subdetector is used.

In this dissertation, most variables extracted from reconstruction use information from TPC, FGD and ECal. Therefore, their local reconstruction algorithms will be explained in the following paragraphs (`fgdRecon`, `tpcRecon` and `ecalRecon`).

`tpcRecon`

In the first stage, this software determines the periods in which the waveforms (charged acquired along the readout time) from each pad of the MicroMegas are above certain threshold. In the following step, clusters are created by connecting waveforms that overlap in time and are consecutive in space. Then, a pattern recognition (based on a cellular automaton) connects the cluster and forms the track segments in the YZ plane.

The x coordinate is reconstructed using the drift time of the e^- cloud. A reference time value, so called t_0, is obtained looking at the time of the closest hits from other detectors (FGD or ECAL). Then, the x coordinate is computed by, firstly, subtracting t_0 to the time of each peak in the waveforms and, secondly, multiplying it by the drift velocity.

From the reconstructed track segments in each single TPC, it is possible to extract position (x, y, z) or curvature (ρ) by applying a maximum likelihood fit with a helix parametrization (magnetic field B bends charged particles). From such reconstructed variables, momentum (p) and charge (Q) can be inferred:

$$p[\text{MeV}] = \frac{0.3B[\text{T}]}{\rho[\text{mm}]}\sqrt{\frac{1+\tan^2\theta_{XZ}+\tan^2\theta_{YZ}}{1+\tan^2\theta_{YZ}}} \qquad Q = \frac{\rho}{|\rho|} \tag{6.1}$$

Where θ_{XZ} (θ_{YZ}) is the projected angle in the XZ (YZ) plane. Transverse momentum (p_T) can be obtained assuming $\theta_{XZ} = 0$.

The ionization energy deposited along the TPC is also measured (so called, energy loss). Such a quantity, in certain momentum regime, is proportional to p/m (Bethe-Bloch formula). Therefore, it is possible to perform a PID, for a known value of the reconstructed momentum, by comparing the measured energy loss (dE/dx_{meas}) with the expected one under different particle hypothesis ($dE/dx_{exp}(\alpha)$) (see Fig. 6.11).

A discriminator factor for each particle hypothesis, so called pull (δ_α), is calculated using the following formula:

$$\delta_\alpha = \frac{dE/dx_{meas} - dE/dx_{exp}(\alpha)}{\sqrt{\sigma_{meas}(\alpha)^2 + \left(\frac{dE/dx_{exp}}{dp}\right)^2 \sigma_p^2}} \tag{6.2}$$

Where $\sigma_{meas}(\alpha)$ (σ_p) is the uncertainty on the measured ionization energy loss (momentum), which depends on the particle type hypothesis that was assumed.

`ecalRecon`

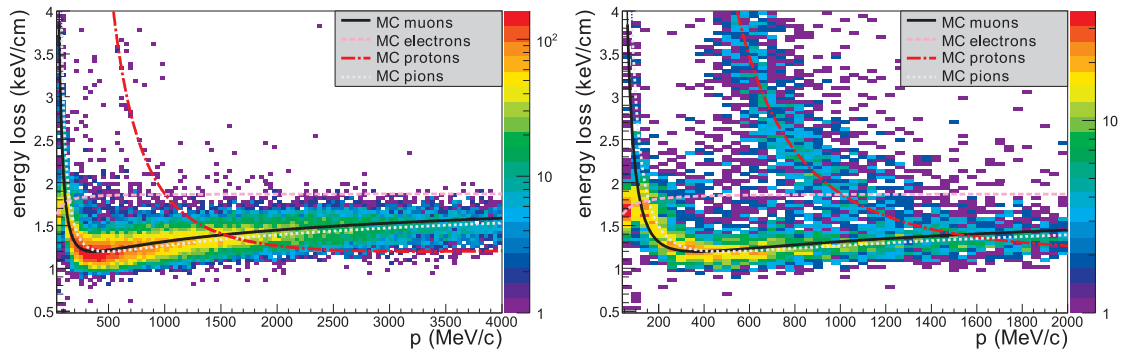

Figure 6.11: Expected (lines) and measured (colored points) energy loss distribution as a function of the momentum for negative (left) and positive (right) charged particles produced in ν interactions in ND280.

Particles entering in ECal can cause both track and shower patterns due to the lead content of this detector. Therefore, this software attempts to reconstruct those two objects. Shower patterns are more complex because hits (triggered scintillator bars) are more likely to spread across the detector.

Firstly, hits from adjacent layers are clustered in 2D using a nearest-neighbor algorithm. In the next step, 2D clusters are matched to create 3D cluster objects. The matching criteria comes from a likelihood, which is formed using features of the 2D cluster as charge and position. Then, different variables are computed in order to characterize the 3D cluster:

EM energy is the free parameter obtained from a likelihood, in which the charge distribution from a 3D cluster is compared with an electromagnetic shower.

Circularity gives a measure of how round a cluster is.

QRMS is the standard deviation of the hit charges in the cluster. This is divided by the mean of the hit charges to produce a dimensionless quantity that is independent from the overall charge scale.

Truncated Max Ratio is the ratio of charge in the highest and lowest charge layers. Before computing the charge per layer, the highest and lowest charge hits are removed to reduce the sensitivity to noise or saturated channels.

Front Back Ratio is a measure of the dE/dx along a track. It is defined as the total charge in the back quarter divided by the total charge in the front quarter. It is sensitive to the the dE/dx profile of stopping tracks.

Fig. 6.12 shows that the predicted distributions for the last four variables have certain discrimination

power for PID.

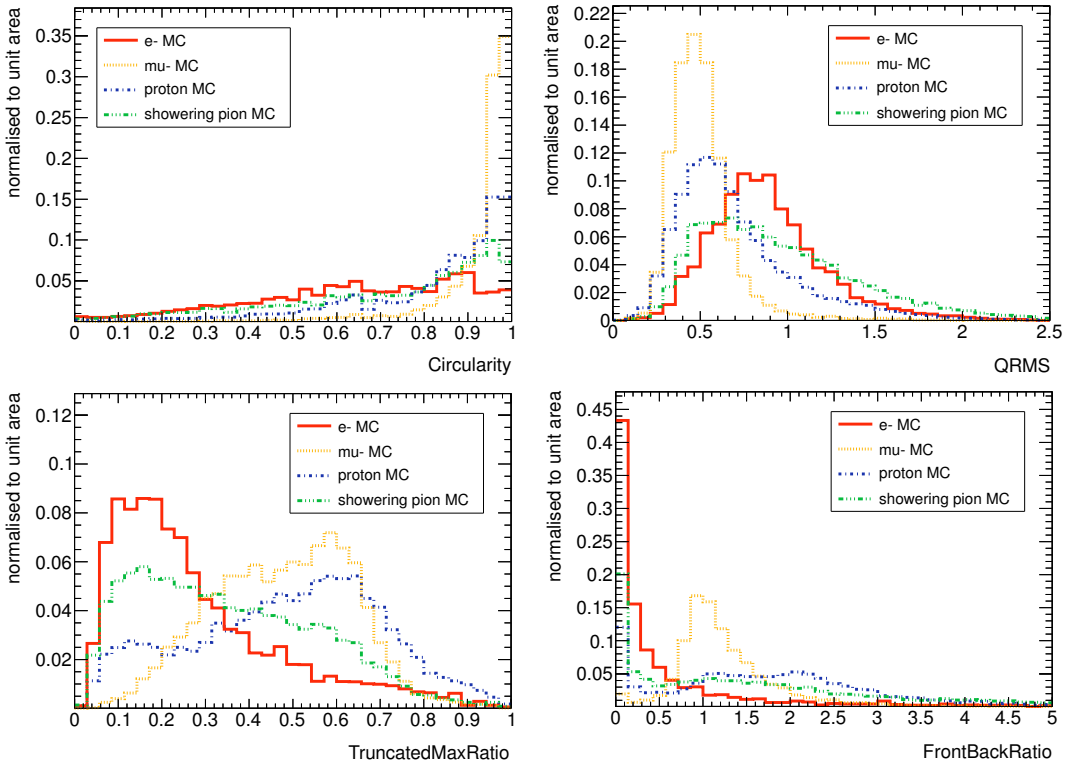

Figure 6.12: Normalized distribution for the Circularity (top left), QRMS (top right), Truncated Max Ratio (bottom left) and Front Back Ratio (bottom right). It has been generated using particle gun MC entering in DsECal.

Using such variables, several PID variables are constructed based on a log-likelihood ratio. In this dissertation one of them is used, the so called MIP-EM, to separate shower-like and track-like clusters.

`fgdRecon`

This software is used to propagate TPC tracks through the FGD and to reconstruct tracks that do not enter in the TPCs. Firstly, FGD hits are separated using timing information into set of hits (time bins). For each time bin, FGD reconstruction will be run separately after TPC reconstruction is finished.

In the following step, TPC tracks are matched to FGD hits whose time bin matches the t_0 (from TPC segment). A Kalman filter algorithm (implemented in `RecPack` [146]) is used to perform the propagation. This algorithm performs an incremental addition of hits to the track, starting from the closest layer to the TPC segment. If no hits are found in two consecutive layers, then propagation stops and no more hits are added to the track.

Finally, a stand-alone reconstruction is performed for those hits that have not been matched to a TPC segment. Furthermore, a PID scheme has been developed based on charge deposition and timing information (Michel e^-). However, such capabilities have not been exploited in this analysis.

`globalRecon`

This package produces the final output from the reconstruction process. Its main goal is to combine the results obtained in the local reconstruction. Most tools for matching and fitting objects are inherited from `RecPack`.

Firstly, each tracker object (TPC-FGD segments) is trying to match objects in the adjacent detectors (P0D, ECal). This is done by extrapolating the reconstructed state of both objects to a matching plane. Then, they are merged when position, direction and time of such states fulfill certain criteria. Once they are merged, the combined information from both objects is fitted using the Kalman filter algorithm. From this fit, more reliable information about the reconstructed kinematics is obtained. This process is repeated for every segment in the event.

This global reconstruction, together with the calibration, plays a critical role in the extraction of timing information for track crossing the different subdetectors of ND280. In this analysis, timing information is used to reconstruct the sense of the track. Global objects store two time variables associated to each local segment:

- The average time: all the hit times weighted by their deposited charge are averaged.
- The average of two time stamps: the hit time versus the hit position is fitted with a linear function. The values of this function at the first and last hit position (T_1 and T_2) are averaged ($T := \frac{T_1+T_2}{2}$).

The width and tails of the distribution associated with the average of two time stamps is smaller than the average time (see Fig. 6.13).

Therefore, it was decided to use the average of two time stamps as the time reference for each local object. Comparing the time references between two local objects in the detector X and Y, the time of flight (ToF) variable can be constructed as follows:

$$ToF_{X-Y} = T_Y - T_X \tag{6.3}$$

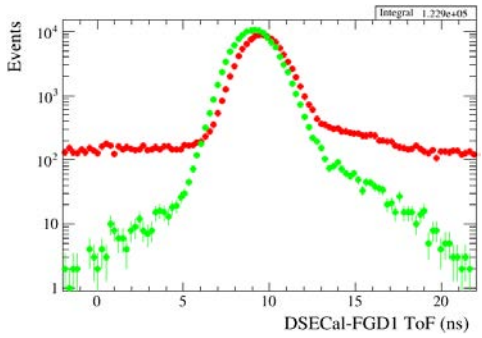

Figure 6.13: Sand μ reconstructed time of flight between DSECal and FGD1, red is calculated with the average time, green with the average of two time stamps.

The capability of the ToF variable as a sense discriminator was tested using two samples of events whose direction was already known: sand μ going through the whole detector, which contained tracks going forward, and cosmic μ, which contained downward events. Using those samples, ToF between different pairs of detectors can be studied. In this analysis, P0D-FGD1, FGD2-FGD1 and BarrelECal-FGD1 pairs are used. The following mean values are obtained for its associated ToF:

- P0D-FGD1: sand μ (forward going) mean ToF$\sim 8ns$
- FGD2-FGD1: sand μ (forward going) mean ToF$\sim 4.5ns$
- BarrelECal-FGD1: cosmic μ (backward going) mean ToF$\sim 3.5ns$

The calibration stability for the ToF variable is checked by plotting the ratio of events whose ToF sign is wrong in the sand and cosmic samples, as a function of the run number (Fig.6.14). It shows a good performance, with a very low failure rate in the case of P0D-FGD1 and FGD2-FGD1. For the BarrelECal-FGD1 case, the rate is a higher but acceptable.

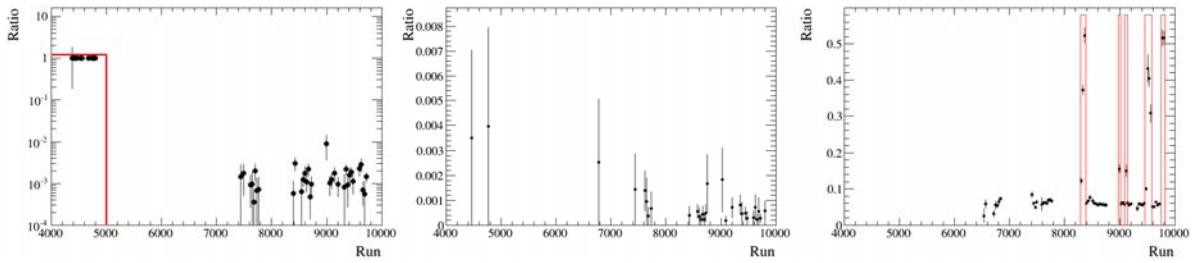

Figure 6.14: Ratio events whose ToF sign is wrong ToF as a function of the run number for P0D-FGD1 (left), FGD2-FGD1 (middle), BarrelECal-FGD1 (right). Red lines indicates runs in which ToF is not properly calibrated and they are not used for physics analysis.

Analysis Framework

The information that is obtained after the reconstruction stage is stored in a specific format, inherited from a class called `oaEvent`. This format stores all available information both from local and global reconstruction but also low-level information, such as hits. Thus, this format is suitable for calibration and reconstruction packages because it allows to rerun at any level (i.e. to check the performance of a new calibration constant) without requiring a rerun using files with raw data.

Nevertheless, the size and complexity of those files make them impractical for any statistical analysis. Therefore, two packages are used in order to simplify and process the data: `oaAnalysis` (in charge of data reduction) and `HighLAND` (in charge of event selection and detector systematics).

`oaAnalysis`: This package handles the simplification of the `oaEvent` class and it is the last stage in the ND280 software. It stores all the information in pure ROOT format and it reduces the size of the output files by a factor of four. It can be broken down in three directories that contain: beam and data quality information per spill (HearderDir); true trajectories and vertices copied directly from the generators (TruthDir); global and local reconstruction information (ReconDir).

`HighLAND`: This a a generic package that handles the selection of events stored in output files from `oaAnalysis`. Besides, it handles the corrections applied to certain reconstructed variables (due to data-MC reconstruction discrepancies) and the propagation of its associated systematic uncertainty. These tools are inherited from a second framework called `psyche`. The structure of the framework is split into different packages, which perform different roles in different stages (see Fig. 6.15).

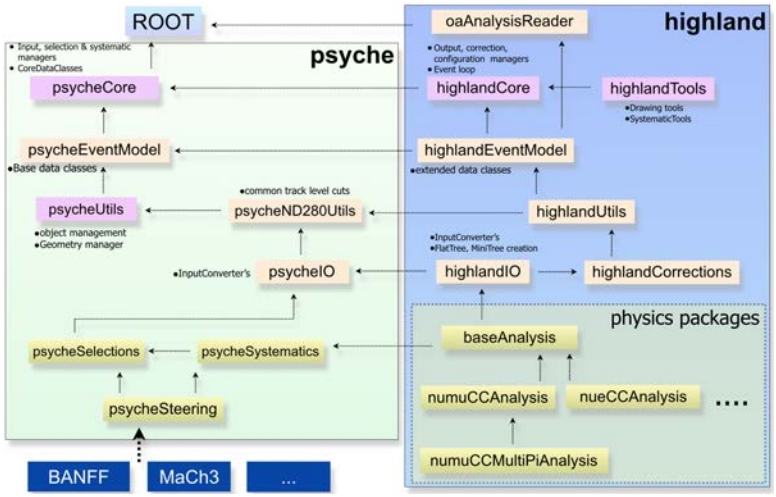

Figure 6.15: General structure of `HighLAND` including only main packages.

The workflow of this framework starts by converting the information from the input files, stored in a spill by spill, into a bunch by bunch format. Then, corrections are applied to the nominal reconstructed variables event by event. Finally, event selection is performed and detector systematic uncertainties are propagated using toy experiments, in which each event is reweighted to account for such errors.

The output from `HighLAND` has a simple TTree format, which can be easily exploited using a package called `DrawingTools`. This package can handle all data-MC comparisons and evaluation of detector systematic uncertainties.

Chapter 7

Current Results

The main objective of T2K is to measure ν oscillations. In particular, this experiment is sensitive to two oscillation channels: ν_μ disappearance ($\nu_\mu \to \nu_\mu$, mainly ruled by θ_{23} and Δm^2_{23}) and ν_e appearance ($\nu_\mu \to \nu_e$, mainly ruled by θ_{12}, octant of θ_{23}, δ_{CP} and mass hierarchy).

Currently, T2K is one of the few experiments that can provide some information about the only value that remains unknown in the PMNS matrix, δ_{CP}, thanks to its $\nu/\bar{\nu}$ production capability. If $\delta_{CP} \neq 0$, certain asymmetry should be observed between $\nu_\mu \to \nu_e$ with $\bar{\nu}_\mu \to \bar{\nu}_e$ channels. For instance, if $\sin\delta_{CP} < 0$, then ν_e (ν_e) appearance is suppressed (enhanced). Such asymmetry and the T2K flux without oscillation are shown in Fig. 7.1

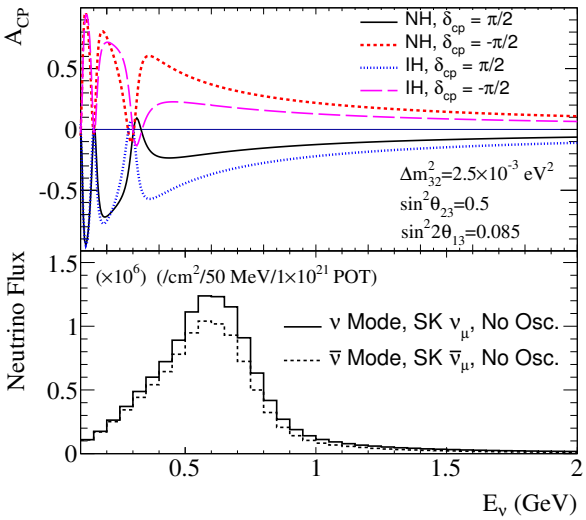

Figure 7.1: Top: CP asymmetry ($A_{CP} = [P(\nu_\mu \to \nu_e) - P(\bar{\nu}_\mu \to \bar{\nu}_e)]/[P(\nu_\mu \to \nu_e) + P(\bar{\nu}_\mu \to \bar{\nu}_e)]$) as function of ν energy under different hypothesis for δ_{CP} and mass hierarchy. Bottom: Predicted ν_μ ($\bar{\nu}_\mu$) flux at far detector assuming no oscillation in the ν ($\bar{\nu}$) mode production.

T2K experiment has been running for almost seven years, collecting in total 7.482×10^{21} protons on target (POT) in ν mode and 7.471×10^{21} in $\bar{\nu}$ mode for the far detector analysis (see Fig. 8.1). In the case of the near detector analysis, 5.82×10^{21} protons on target (POT) in ν mode and 2.84×10^{21} in

$\bar{\nu}$ mode.

In the oscillation analysis, the strategy has been to estimate the mixing parameters by comparing predicted and observed rate of ν interactions in the far detector. Previously, ND280 was used to tune the nominal flux prediction (see Sec. 5.5) and to estimate the uncertainties coming from flux and cross section modeling. The current analysis includes both ν and $\bar{\nu}$ in the fitting framework.

The interactions of ν are simulated using a modified version of `NEUT` (see Sec: 4.3). The CCQE interactions are modeled using a RFG that includes RPA (instead of the SF).

In the near detector analysis, a binned maximum likelihood fit is performed over several samples. It aims to constrain flux and cross section uncertainties, which are included in the fit as free parameters. In the ν mode, CC interactions happening either in FGD1 or FGD2 are selected by requiring a reconstructed μ^- in the TPCs. Then, such an inclusive selection is split into three subsamples depending on the number of reconstructed π and their charge (CC0π, CC1π^+, CCOthers). In the $\bar{\nu}$ mode, the inclusive selection looks for both μ^+ and μ^- (contamination of ν_μ in the flux is high). Such samples are binned according to the momentum of the reconstructed μ (p_μ) and its angle with respect to the central axis of the near detector ($\cos\theta_\mu$).

Fig 7.2 shows the impact of the near detector fit in some of the flux and cross section parameters that quantify the uncertainties. In summary, nominal flux is increased by a factor of $\sim$ 10% and M_A^{QE} is fitted to 1.12$\pm$0.03 GeV/c^2, while the nominal value was 1.20$\pm$0.07 GeV/c^2. This fit reduces the uncertainties (due to flux and interaction modeling) on the predicted rate of ν_μ (ν_e) events in ν mode at far detector from 7.6% (9.3%) to 2.9% (4.7%). Meanwhile, the effect in the total error (including both statistical and all systematic uncertainties) is also noticeable, for ν_μ (ν_e) events it is reduced from 12.0% (14.5%) to 5.1% (7.4%)

In the far detector analysis, the oscillation parameters $\sin\theta_{23}^2$, Δm_{23}^2, $\sin\theta_{12}^2$ and δ_{CP} are estimated with a maximum likelihood fit including four different samples (ν_μ, $\bar{\nu}_\mu$, ν_e and $\bar{\nu}_e$,), which are composed by CCQE interactions in the far detector. In the case of $\nu_\mu/\bar{\nu}_\mu$ samples, the selection criteria look for μ with reconstructed momentum higher than 200 MeV/c. In the case of $\nu_e/\bar{\nu}_e$ samples, $e^\pm$ with an energy lower than 1.2 GeV are traced. Finally, the energy of the incoming ν is obtained using the QE formula (based on four body kinematics).

Fig: 7.3 shows a comparison between the predicted (with and without assuming oscillations) and observed rates for the four different samples. Confidence intervals for the fitted mixing parameters

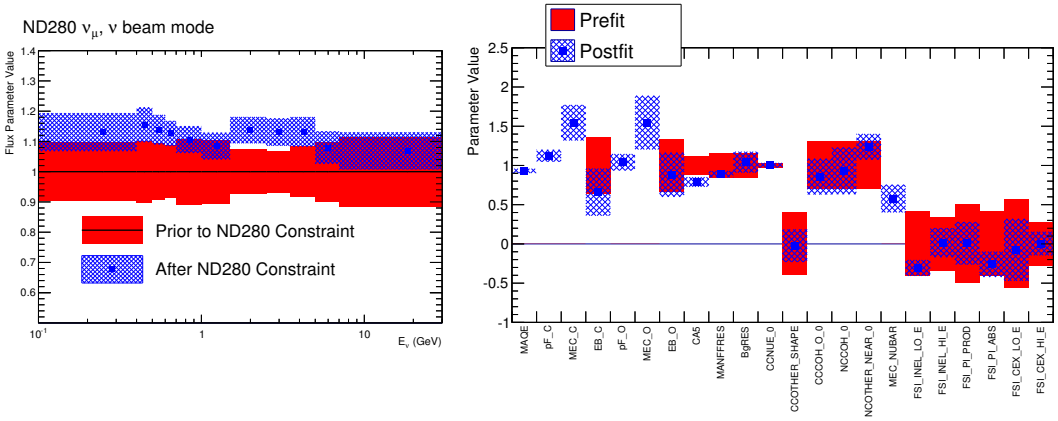

Figure 7.2: Prior and fitted values and uncertainties for the SK ν_μ flux parameters in the ν mode (left) and cross-section parameters (right) constrained by the near detector fit.

and their associated correlations are shown in Fig. 3.4 and Fig. 3.5. From those plots two important conclusions can be extracted: T2K provides the most precise measurements of θ_{23} and Δm^2_{23} and it is the first experiment that is able to exclude the CP conservation hypothesis ($\delta_{CP} = 0$) at 90% confidence level.

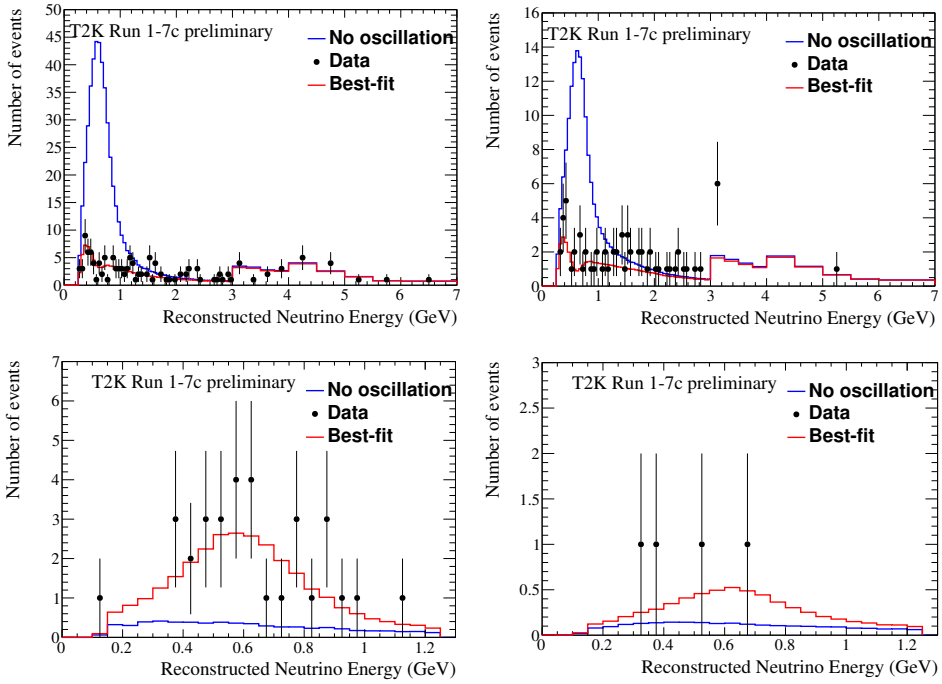

Figure 7.3: Distributions for the reconstructed energy for different samples: ν_μ (top left) and ν_e (bottom left) candidates in the ν mode; $\bar{\nu}_\mu$ (top right) and $\bar{\nu}_e$ (bottom right) candidates in the $\bar{\nu}$ mode. Best fit distributions (red lines) are compared with the predicted spectrum under no oscillation hypothesis (blue lines).

Part III

ν_μ CC Analysis

Motivation

In ND280, the main analysis selects μ^- coming out from ν_μ interactions in FGD via CC. This selection is focused on reconstructing forward going tracks starting in FGD1 (FGD2) and with segments in TPC2 (TPC3). Fig. 7.4 shows the $\cos\theta$ distribution for the events selected in the $CC0\pi$ sample of the standard ν_μ analysis in ND280 (see Sec. 7). In such analysis, the acceptance is degraded for $\cos\theta \leq 0.75$ rad, corresponding to an angle of $\sim$40 degrees in the forward direction.

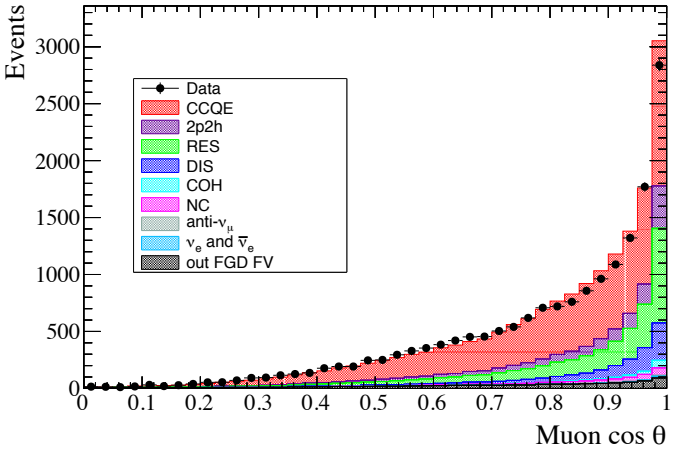

Figure 7.4: Distribution of the cosine emission angle between the reconstructed μ direction and the beam direction in the events selected at ND280.

The ND280 samples are then used to constrain the flux and cross-section parameters reducing the systematics errors in the T2K oscillation analyses (see Sec. 7). The events selected at Super-Kamiokande, given the 4π symmetry of the detector, are selected without requiring a forward going track. The expected distributions of $\cos\theta$ are shown in Fig. 7.5 for the ν_e and ν_μ samples. In the ν_e appearance analysis among all selected events, a non negligible contribution is from e^- going backward. The effect is less evident in the ν_μ analysis, since there is no cut on the reconstructed energy (backward events tends to have smaller momenta). Nevertheless, a non negligible amount of leptons are not produced forward in the ν_μ sample ($\sim$25% of events in ν_μ sample have $\cos\theta < 0.6$).

The analysis described in this dissertation provides a sample of μ^- tracks with a 4π coverage from

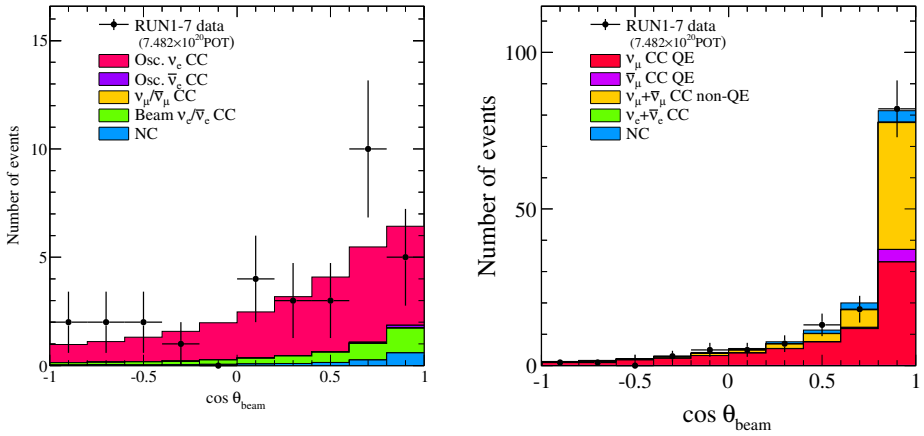

Figure 7.5: Distribution of the cosine of the angle between the reconstructed ring direction and the beam direction in the events selected for the ν_e appearance (left) and ν_μ disappearance analysis at SK.

the FGD1 thanks to the addition of the backward going sample with μ^- going from FGD1 to P0D or BarrelECals, and the high angle sample with μ^- either crossing barely or not crossing the TPCs. This new ND280 ν_μ CC 4π sample will be used in the ND280 fit to reduce and cross-check the assumptions made in the fit to constrain the expected number of events at SK.

Furthermore, this sample can be used to study ν_μ-nucleus interactions via charged current (CC). This analysis aims to reconstruct the outgoing μ^-, so an inclusive CC cross section result will be reported. This approach has several advantages.

- Very pure samples can be obtained without jeopardizing statistics. In the energy region of the T2K flux (peaked at ~700 MeV), most of the interactions are NC or CC QE, thus the multiplicity is not very high. Consequently, the misidentification of the μ^- coming out from the CC interaction is not likely to happen.

- Current models are not able to predict properties of the hadrons, so it is difficult to make a model independent hadron selection based on Monte Carlo predictions. Moreover, the efficiency correction, which is based on the model, becomes unreliable when hadrons are taken into account. Therefore, an inclusive selection, in which only the μ^- is tagged, is the least model dependent analysis that can be performed.

Furthermore, the result will be flux integrated double differential cross section as function of the μ^- momentum and direction (with respect to the ν beam). On the one hand, cross section will be flux

integrated so that no assumption is done on the ν energy. On the other hand, the simplest leptonic variables are used to avoid the inclusion of model dependencies.

This cross section measurement will give a step forward in the T2K ν oscillation measurement. The latest T2K ν oscillation measurements are claiming statistical and systematic uncertainties of the same order (see Sec. 7). Next iteration of experiments aims to increase current sensitivity on delta-CP and mass hierarchy (DUNE [147], HyperK [148], T2K-II [149]). Therefore, systematic uncertainties must be better understood. As it was shown in Sec. 3, in oscillations analysis, ν interactions with complex nucleus play a main role. Understanding such reactions is crucial because it affects both the background estimation and the ν energy reconstruction.

During last decades, ν interactions have been studied both theoretically and experimentally for a wide range of nuclei and energies (Sec. 4). Up to now, several ν cross sections with different heavy nucleus have been reported in both inclusive and exclusive channels. Nonetheless, in most cases the results have been either limited by statistics or restricted to certain kinematical regions. Therefore, models can not be fully tested.

In particular, T2K has already published an inclusive CC cross section measurement in ND280 [119]. Nonetheless, the analysis detailed in this dissertation complement the previous result in several aspects. Particularly important is the increment of statistic due to two main facts: the collected POT is roughly five times larger and the angular acceptance has been increased for high angle and backward going μ^-. Moreover, a new event selection has been carried out in order to reduce the background contamination.

Chapter 8

Event Selection in ND280

8.1 Samples

Data samples used for this analysis include ND280 runs collected during Run 2, Run 3 and Run 4 periods (see Fig. 8.1). Run 1 was excluded for several reasons. The first one was related to calibration issues during that run, the second one was the absence of the top module in BarrelECal and the third one was the small amount of statistics that it represents. Since Run 5, Run 6 and Run 7 periods have a tiny accumulated number of POT in forward horn current, i.e. ν beam mode, it was decided not to take them into account. The total sample used for this analysis contained 5.735515×10^{20} protons on target.

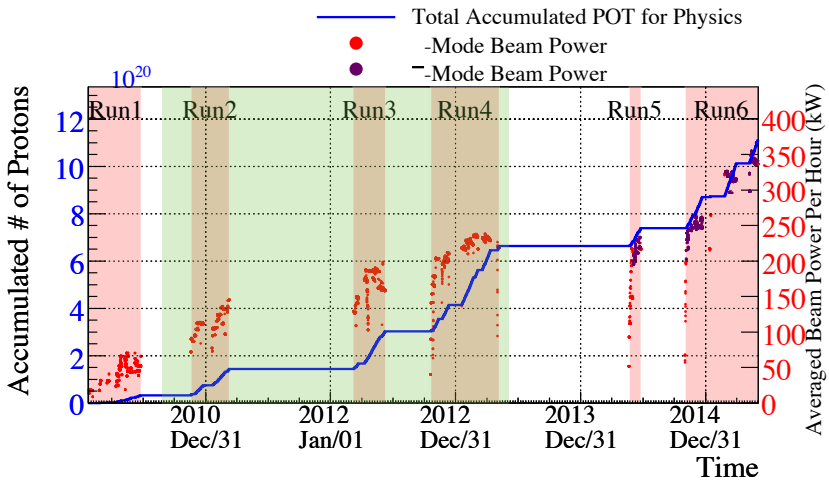

Figure 8.1: Accumulated POT and beam intensity at T2K over until Run 6 (Run 7 is not included). Shaded region indicates the runs used in this analysis.

MC interactions within ND280 subdetectors (magnet) and outside ND280 (sand) were generated separately. In the magnet case, two different MC samples were used for testing purposes (`NEUT`, `GENIE`). Both of them had the same flux, detector simulation and reconstruction. The difference between them was at the interaction level and the simulation within the nucleus. The code versions used in each of

the inputs are summarized in Table 8.1.

SAMPLE	GENERATOR	ND280 software
DATA	-	v11r31p5 (Prod 6K)
MAGNET	NEUT v5.3.2	v11r31 (Prod 6B)
	GENIE v2.8.0	v11r31 (Prod 6B)
SAND	NEUT v5.3.2	v11r31 (Prod 6B)

Table 8.1: Version of the generators used for each MC sample and its associated ND280 software.

The original MC sample was generated using a flux prediction (version 13a), which takes into account different beam conditions in each run. Detector conditions were considered by generating separate MC samples. Runs were split into periods with "water in" and "water out" in the P0D sub-detector and considering different dead channel configuration. Table 8.2 shows a summary of collected and generated POT for each run period used for this analysis.

	POT ($\times 10^{20}$)			
Sample	Data	NEUT magnet	GENIE magnet	MC sand
Run 2-Water In	0.42858	12.0343	12.8234	11.1964
Run 2-Water Out	0.355096	9.22436	9.84393	11.1964
Run 3b-Water Out	0.2146	4.47876	4.7108	11.1964
Run 3c-Water Out	1.34779	26.3072	28.0462	11.1964
Run 4-Water In	1.62699	34.9774	35.4343	11.1964
Run 4-Water Out	1.76246	34.9717	37.2547	11.1964
Total	5.735515	121.99372	128.11333	67.1784

Table 8.2: Number of POT for each data set of collected data and generated MC samples. The POT in data correspond to the collected data after good spill and data quality criteria are applied.

8.2 Fiducial Volume

The aim of this analysis is to select ν_μ CC interactions with their vertex within certain fiducial volume in FGD1. The fiducial volume is defined as:

$$\begin{aligned} |x| &< 874.51 \text{ mm} \\ |y - 55| &< 874.51 \text{ mm} \\ 115.625 \text{ mm} &< z < 447.375 \text{ mm} \end{aligned} \tag{8.1}$$

Where 55 mm offset in y direction reflects 55 mm shift of the XY modules according to the center of the ND280 coordinate system (see Fig. 8.2). Cuts in x and y direction only accept those interactions which

have their vertex 5 bars distant from the edge of the XY module of FGD1. The fiducial volume cut in z definition includes all 15 modules of FGD1 (in previous analysis the first module was not included).

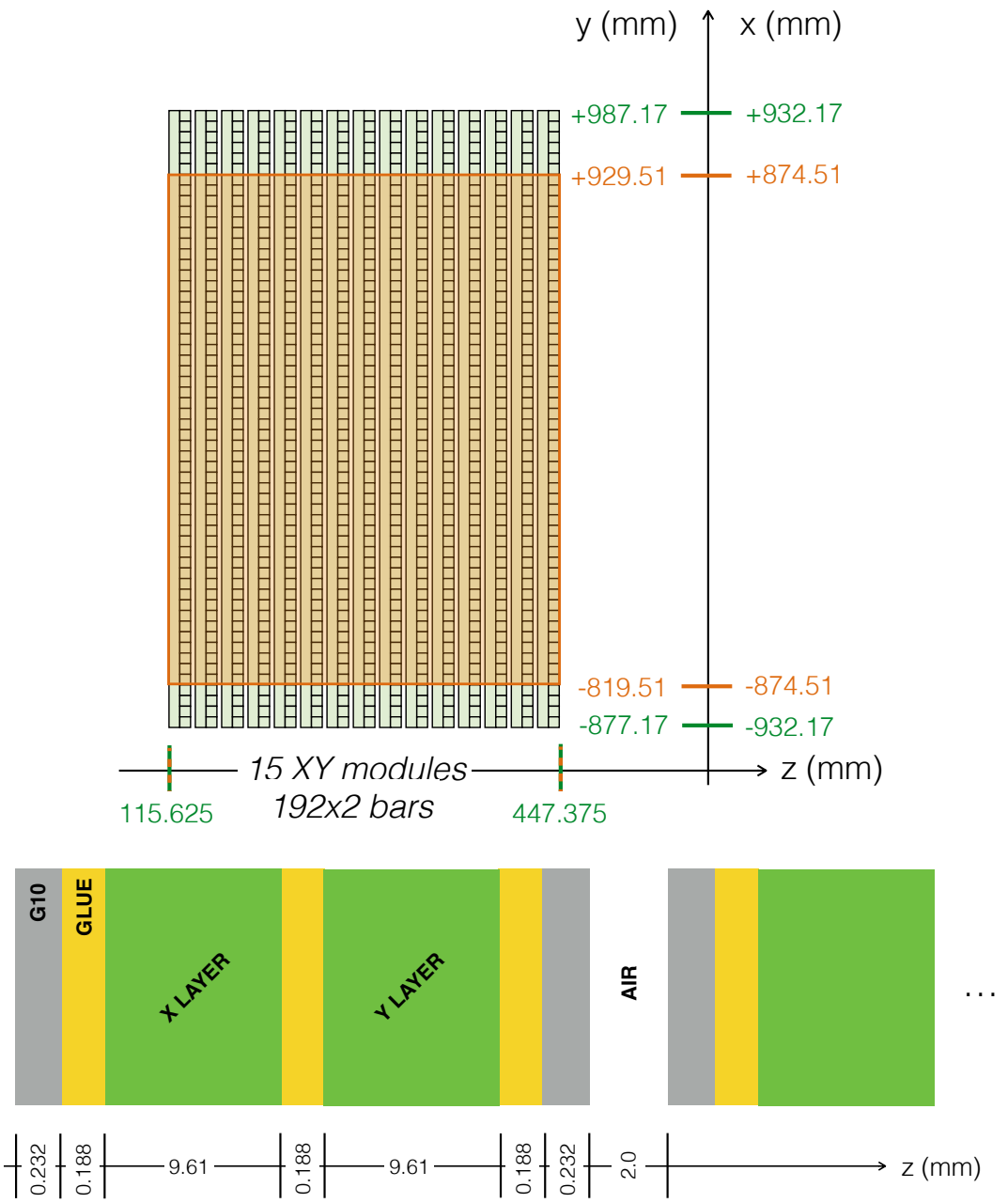

Figure 8.2: Top: FGD1 scheme. Orange lines indicate the fiducial volume. Bottom: Scintillator layers scheme.

8.3 Topology Categorization

The events are categorized depending on the topology of the interaction, using the true information provided by the MC. In this dissertation, the categorization uses the information of the true μ^- coming out from ν_μ CC interactions.

- ν_μ*CC-*μ: ν_μ CC interactions taking place within FGD1 FV, with the true μ^- selected as μ^- candidate.
- ν_μ*CC-no*μ: ν_μ CC interactions taking place within FGD1 FV, with the true μ^- not selected as μ^- candidate.
- *out FV*: Any interaction not happening within FGD1 FV but within the ND280 geometry.
- *sand* μ: Interactions happening outside the ND280 geometry. It includes interactions in the magnet.

8.4 Event Corrections

A set of corrections are applied to both data and MC samples in order to reduce the discrepancies between them. Those corrections are either based on known hardware failures or on very precise studies using control samples, as in the case of the TPC dE/dx or momentum corrections.

Those corrections must be applied prior to the selection steps, to propagate them to the reconstructed variables is produced. They will be explained in the following items:

- dE/dx_{exp}: This correction improves the estimation of dE/dx_{exp} for a given particle and momentum used in `tpcRecon`. It was observed that the average dE/dx for a given particle and momentum did not coincide with the values assumed during the reconstruction process, leading to a bias in the pulls (see Eq. 6.2). The value of dE/dx_{exp} is computed using 5 parameters:

$$\frac{dE}{dx}_{exp} = \frac{c_0}{\beta^{c_3}} \left[c_1 - \beta^{c_3} - ln\left(c_2 + (\beta\gamma)^{-c_4} \right) \right] \tag{8.2}$$

  Where c_i are the tunable parameters. Fig. 8.3 shows the dE/dx_{exp} as function of $\beta\gamma$ using two different set of parameters.

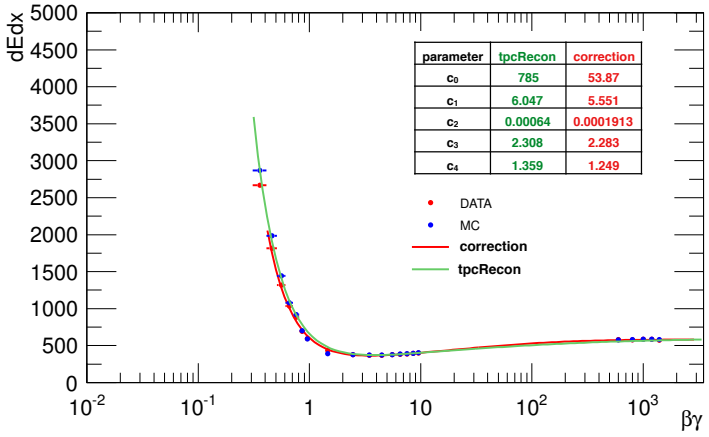

Figure 8.3: $dEdx$ distribution as a function of the $\beta\gamma$ of the particles for two different set of parameters from Eq. 8.2 (shown in the table).

- dE/dx_{meas}: The TPCs are continuously monitored during the data-taking and their performance is expected to be constant, but the performance of each subsystem can be slightly different (for instance gas density or temperature). Therefore, dE/dx_{meas} is corrected in data by scaling it with a calibration factor C_{TPC}, which is close to one (within few %) and it depends on the TPC and running period (see Table 8.3). Moreover, MC overestimates the energy loss of true e^- by 1%, so a reduction factor of 1.01 is applied to this variable for true e^- in the MC.

Period	C_{TPC1}	C_{TPC2}	C_{TPC3}
Run 2-Water In	1.015	1.011	1.025
Run 2-Water Out	1.015	1.011	1.025
Run 3b-Water Out	1.017	1.017	1.025
Run 3c-Water Out	1.018	1.014	1.023
Run 4-Water In	1.027	1.017	1.032
Run 4-Water Out	1.027	1.017	1.032

Table 8.3: Correction factor for the measured dE/dx in data for different run periods and for each TPC. The error is ±0.001.

- **TPC PID**: Another set of effective corrections is applied to dE/dx_{meas} in order to have pulls centered at 0. They are explained in Sec. 9.3.1.

- **Momentum resolution**: In MC, a smearing factor up to ∼40% is applied to the p_t^{-1} of all TPC tracks as function of the x position. This smearing is explained in Sec. 9.3.1.

- **Time of flight**: This correction smears and biases the ToF variable in the MC. It is explained in detail in Appendix C.1.

- **Sense**: This correction flips the sense of the track in data and MC if certain timing criteria is fulfilled by the reconstructed track. It is explained in detail in Appendix C.2.

Furthermore, a second set of corrections associated to the flux and detector response is applied event by event. Such set of corrections is applied reweighting each event by certain factor, so that they can not produce a variation of reconstructed variables. Therefore, they are applied after the selection has been performed. That weight corrections will be explained in Sec. 9.1 and Sec. 9.3.

All these corrections have an associated systematic uncertainties. Their evaluation and propagation is explained in Sec. 9.3.

8.5 Event Selection in FGD1

In previous analysis, the ν_μ Charged Current inclusive selection was optimized to select forward going μ^- which originates from FGD1 and leaves at least 19 reconstructed clusters in TPC2. The aim of this selection is to include the so-called "high-angle" tracks which barely cross the TPCs or completely miss them, as well as backward-going tracks leaving at least 19 reconstructed clusters in TPC1. The addition of backward-going μ^- candidates in this selection is possible only because timing information is used.

The μ^- candidate can be found in one of the three following samples: forward (FWD), backward (BWD) and high angle (HAFWD and HABWD). In the FWD/BWD selections, the μ^- candidate must have long TPCs segments, while tracks with short or no TPC segment are used in the HAFWD/HABWD (see Fig. 8.4).

The selection starts with two event-based cuts:

Event quality. It is required that the full spill has a good global ND280 data quality flag. The event must occur within the bunch time windows of the ν beam. Since events can be assigned to a particular bunch based on their timing, two ν interactions which occur within the same spill but in different bunches are treated as two independent events.

Total multiplicity cut. Only those events which have at least one reconstructed track with a FGD1 segment are considered.

The idea behind those common steps is to reduce the number of events so that they must have at least

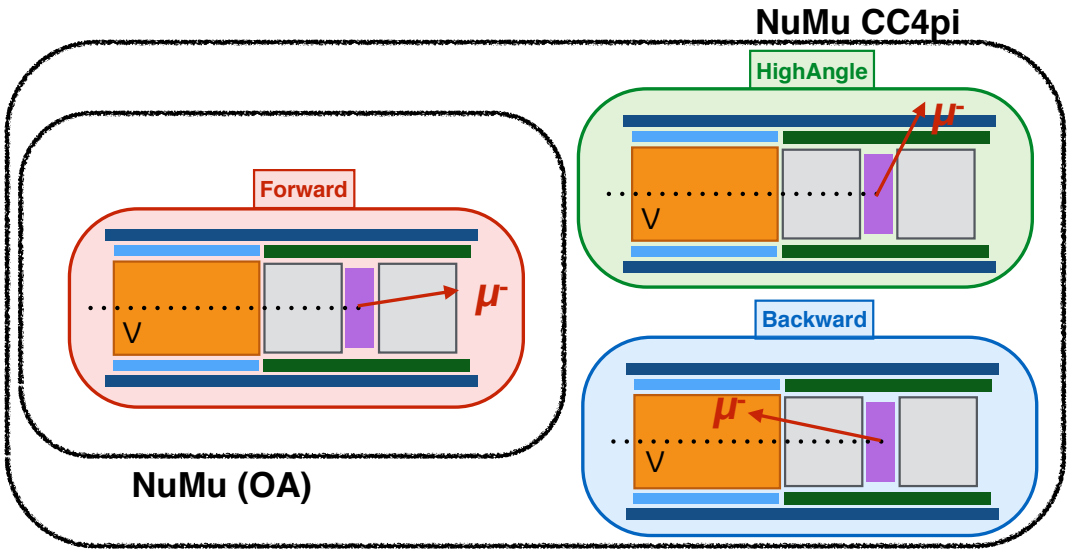

Figure 8.4: Schematic explanation of the regions of interest for each selection.

tracks at FGD1. This fact allows to reduce the computational time for the following steps. Once those steps are satisfied, the three selections FWD, BWD, and HA are processed in parallel. In order not to have more than one μ^- candidate, the presence of a μ^- candidate is tested sequentially in the three selections until a candidate is found. The ordering is FWD, BWD, and then HA. FWD and BWD have a higher priority than the HA because the μ PID from the TPCs is more accurate than in the ECals. The FWD selection has a higher priority than the BWD because forward-going μ^- happen much more often than backward-going ones.

8.5.1 Forward Selection

The forward selection (so called FWD) is very similar to previous analysis. Its aim is to select forward going μ^- coming out from ν_μ CC interactions taking place in FGD1 FV. Most of the steps are taken from the previous analysis, although some optimization has been performed. In the following sections, a detailed explanation of the selection criteria is shown.

Quality and FV

This selection only consider negative charged tracks with TPC segments containing more than 19 hits and starting in FGD1 FV. Of all those tracks, the μ^- candidate must be the one with highest momentum and forward sense. In order to reduce the OOFV contamination, tracks starting in the

most upstream layer of FGD1 are rejected (see Fig. 8.5).

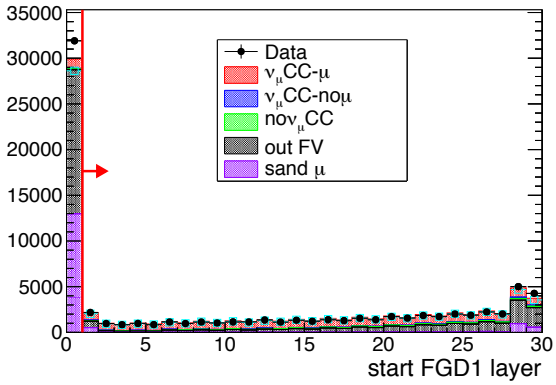

Figure 8.5: Start layer in FGD1 of the highest momentum negative charged low angle track with forward sense. Red lines mark the regions accepted by this cut. Most of the OOFV contamination is rejected with this cut. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

μ PID

μ TPC PID: For tracks which satisfy the criteria described above, particle identification (PID) procedure is applied based on dE/dx measurement in TPC module (see Sec. 6.3.7). Using Eq. 6.2, a likelihood ratio is constructed as:

$$L_i = \frac{e^{-\delta_i^2}}{\sum_i e^{-\delta_i^2}} \tag{8.3}$$

Where the summation takes into account p, e, π and μ particle hypothesis. Protons and π are rejected by requiring (Fig. 8.6):

$$L_\mu > 0.05 \tag{8.4}$$

The next cut removes e^- and it is applied only for tracks with $p < 500 MeV/c$ (Fig. 8.6):

$$L_{MIP} = \frac{L_\mu - L_\pi}{L_p} > 0.8 \tag{8.5}$$

μ FGD2 PID: FGD2 detector can be used to distinguish between μ and π (main background in this selection). The main idea is that the detector response to μ (leptons) and π (hadrons) is very different when they stop inside the FGD. High energy π are more likely to stop in FGD2 than μ. Therefore, a good rejection criteria is that the μ^- candidate ends within certain FGD2 FV with a momentum above certain threshold.

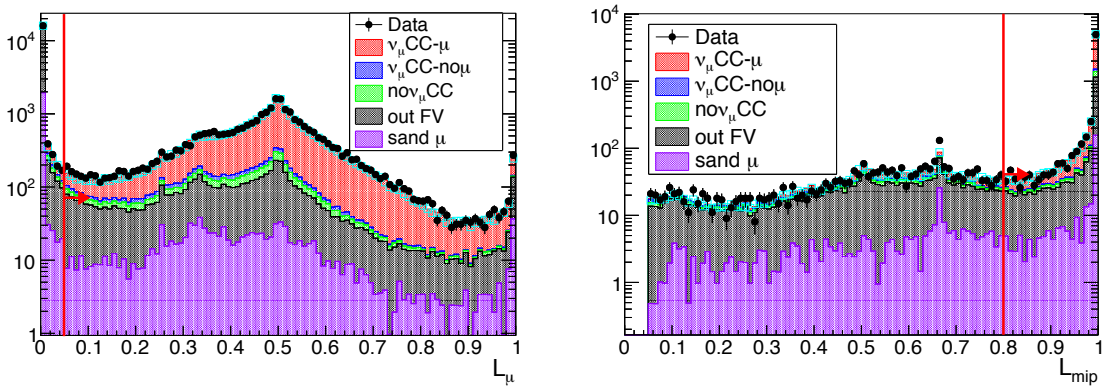

Figure 8.6: L_μ value for μ^- candidate in FWD selection (right). L_{mip} value for μ^- candidate when momentum is lower that 500 MeV/c and L_μ higher that 0.05 (left). Red lines mark the accepted regions. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

The first condition must be that the μ^- candidate ends in FGD2 volume. Besides, checking the momentum of these stopping tracks, a good rejection of π^- can be obtained. An optimization was performed by varying the FGD2 volume and momentum threshold in order to increase the purity without compromising signal efficiency.

From Fig. 8.7 shows the optimized cuts for the FGD2 volume and momentum threshold. It is important to notice that using this criteria we reduce by a factor of ∼15% the π^- contamination while only a ∼0.3% reduction is predicted for signal events.

μ ECal PID: When μ^- candidate reaches an ECal module, shower/track PID capability of these detectors can be used to distinguish between track-like and shower-like particles (see Sec. 6.3.7). Two different modules can be reached by μ^- candidate in FWD selection: BarrelECal and DsECal (see Fig. 8.8).

On one hand, when the μ^- candidate reaches BarrelECal, a PID variable, called MIP-EM (see Sec. 6.3.7), can be used to reject shower-like particles, mainly π^- (see Fig. 8.9).

On the other hand, when the μ^- candidate reaches DsECal, most μ either leave the detector or stop in the first layer. Therefore, non μ contribution is enhanced by requiring that tracks end within certain volume of the Downstream ECal (see Fig. 8.10). The π contribution in tracks stopping within the DsECal volume is reduced using the MIP-EM variable (see Fig. 8.10).

Similarly to μ FGD2 PID cut, the final cuts where chosen based on an optimization. The idea was

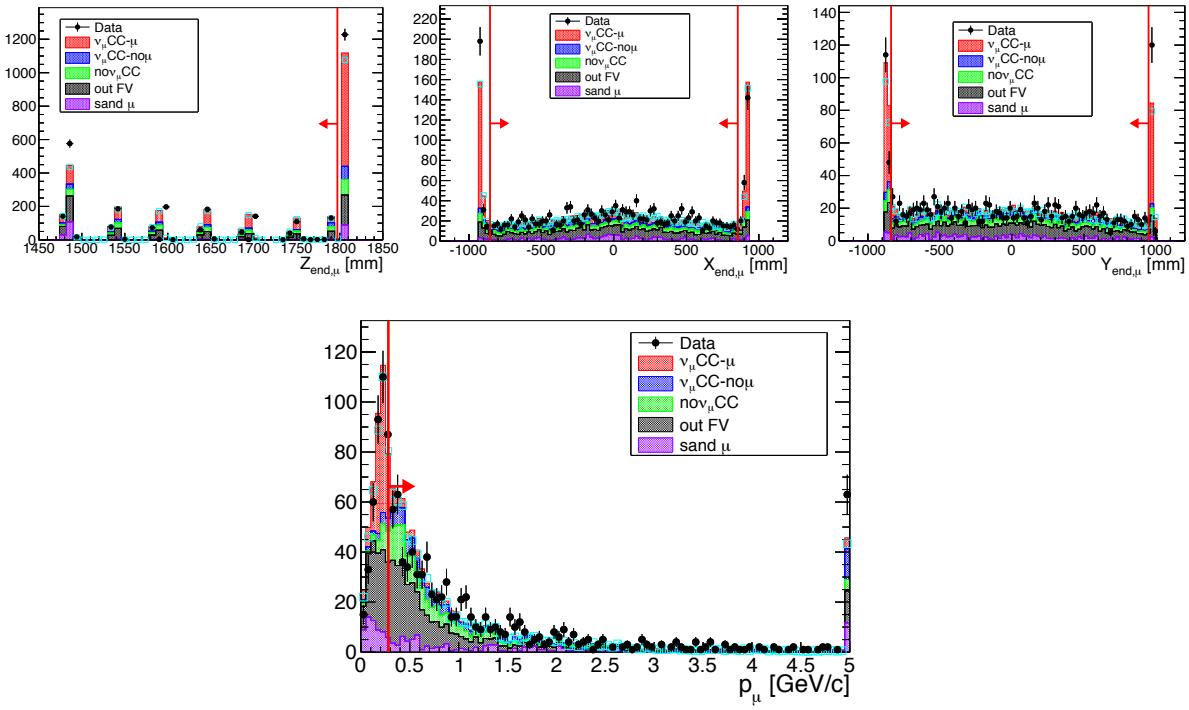

Figure 8.7: Top: End position of the μ^- candidate that fulfills TPC μ PID criteria ends in FGD2 detector. Red lines mark the defined fiducial volume in FGD2 ($|x| < 855.29$ mm, $|y - 55| < 893.73$ mm and 1473.625 mm $< z <$ 1797.25 mm). Bottom: Momentum of tracks stopping in FGD2 volume. Red line marks the region of momentum rejected by this cut ($p > 280$ MeV/c). Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

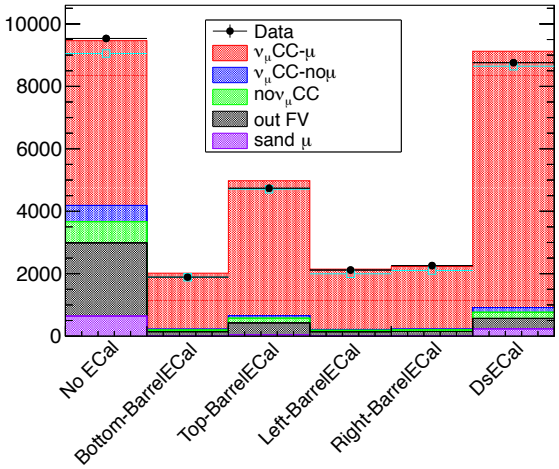

Figure 8.8: ECal subdetector that the μ^- candidate reaches. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

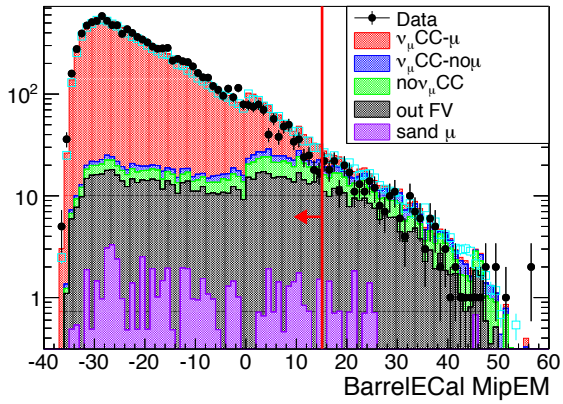

Figure 8.9: MIP-EM value of the μ^- candidate that fulfills TPC μ PID criteria and reaches one BarrelECal detector. Red line mark the regions not rejected by this cut (MIP-EM<15). Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

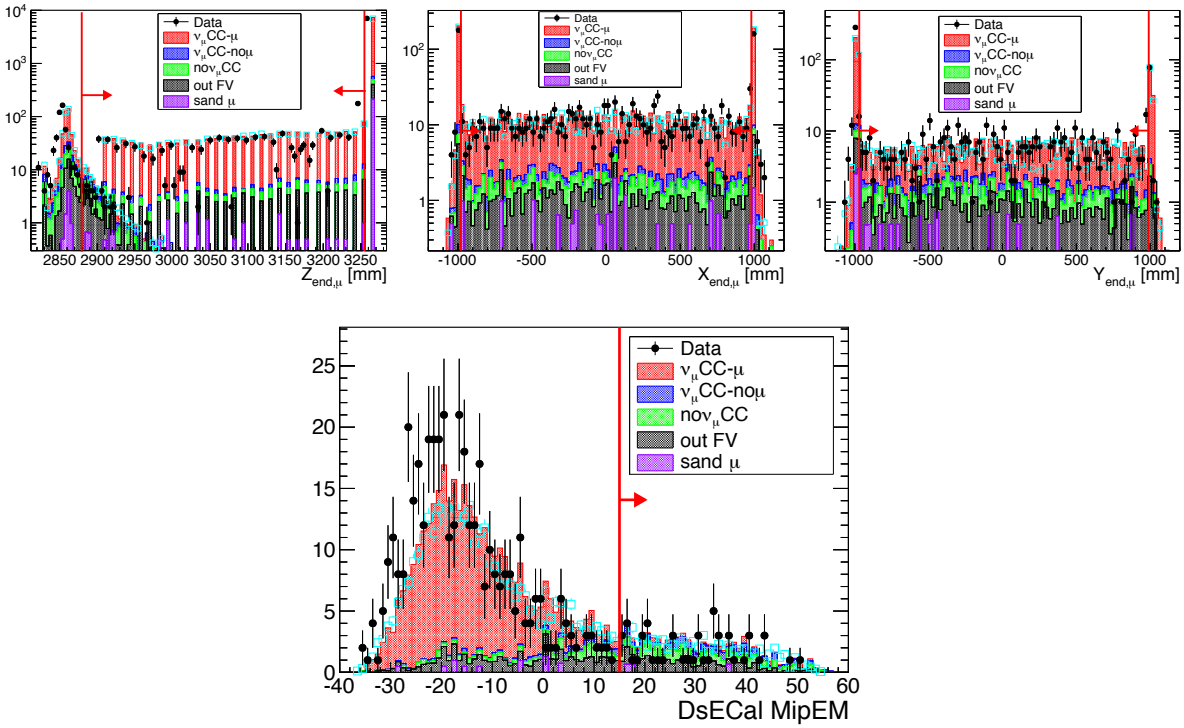

Figure 8.10: Top: End position of the μ^- candidate that fulfills TPC μ PID criteria and reaches DsECal detector. Red lines mark the defined fiducial volume in DsECal ($|x| < 979.97$ mm, $|y - 10| < 979.97$ mm and 2880.03 mm $< z <$ 3255 mm). Bottom: MIP-EM value of the μ^- candidate that fulfills TPC μ PID criteria and ends within DsECal volume. Red lines mark the regions rejected by this cut (MIP-EM>15). Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

to increase the purity without affecting the signal efficiency. In fact, using ECAL PID cuts the π^- contamination is reduced by a ∼7% while signal is reduced ∼0.3%.

Veto

One of the main backgrounds in this analysis are interactions happening out of fiducial volume. This contamination is reduced with the FV cut described above. However, it can be reduced further rejecting certain events with activity in subdetectors upstream FGD1, so called upstream background veto. Moreover, the out of fiducial volume contamination increases when the vertex is reconstructed in the most downstream layer of the FGD1 (see Fig. 8.5). This is due to mis-reconstruction of certain tracks and it is possible to reject them, so called broken track cut.

Upstream background veto: This cut is applied to reject events in which the μ^- candidate is coming from out of fiducial volume. Due to reconstruction failures and multiple scattering, a reconstructed track can be broken into two unmatched segments. One of those can have its start position in the fiducial volume, faking the vertex position of an interaction. In the previous analysis, such events were rejected if the second highest momentum track started 150mm upstream of the μ^- candidate. This cut was found to be too restrictive (it rejected events in which the μ^- was going forward and a second particle was going backward, see Fig. 8.11).

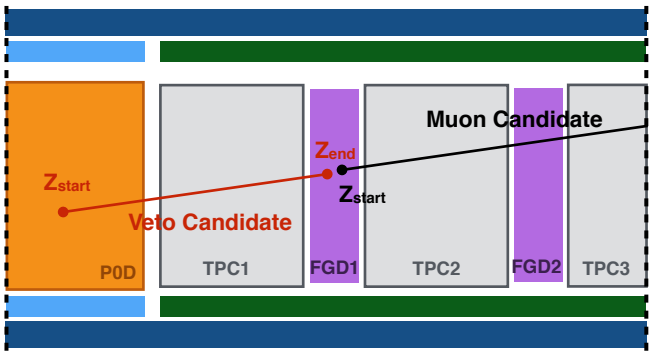

Figure 8.11: Schematic view of the tracks that will be rejected by the veto cut.

To reduce the rejection of those events, the ratio between the momentum of the μ^- candidate and the veto track is used. Ideally, if the μ^- candidate is a broken track, this ratio should be bigger than one (the first segment of the track has higher momentum than the second segment). Therefore, the distance between both tracks and their momentum ratio are used to reduce the restrictiveness of the previous cut. The values that provide the best purity and efficiency are shown in Fig. 8.12.

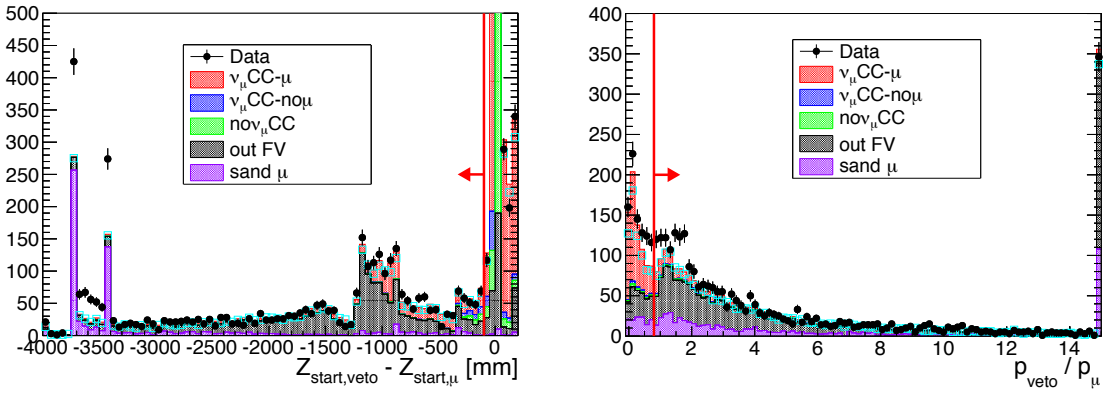

Figure 8.12: Distance between the start position of the veto track and the μ^- candidate that fulfills μ PID criteria in FWD selection (left). Ratio between the momentum of the veto track and the μ^- candidate when the previous distance is < -100mm (right). Red lines mark the regions rejected by these cuts. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

Broken track cut: This cut was applied to reject events with mis-reconstructed tracks, where instead of one μ^- candidate track originating in FGD1 FV the reconstruction procedure broke this track into two components: one track consisting of only one FGD1 segment, followed by a second track which starts in the last layers of FGD1 and goes through the TPC module. Therefore the second track is considered as μ^- candidate. For those events, the start position of μ^- candidate track is within the two most upstream layers of FGD1. To reject them, it is required to have at least one reconstructed track with a FGD1 segment but no TPC segments in the same event (see Fig. 8.13).

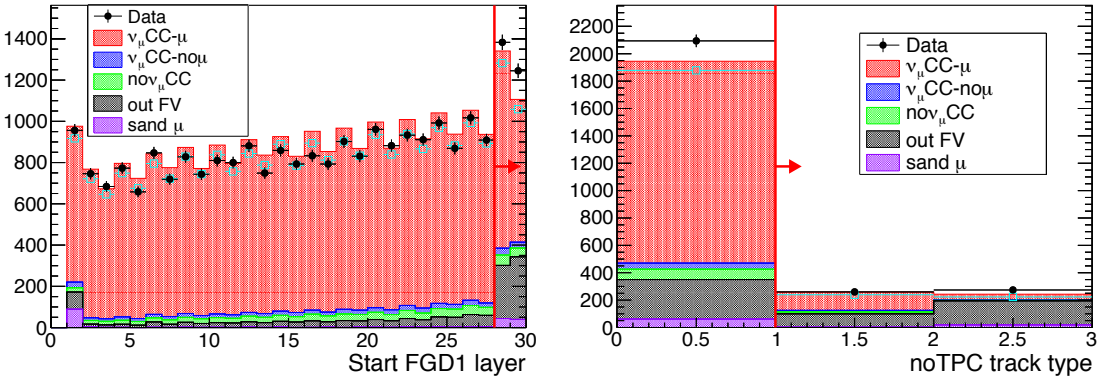

Figure 8.13: Start position of μ^- candidate that fulfills PID and upstream veto criteria (left). Tracks with FGD1 and no TPC segments types found per event when μ^- candidate starts in one of the two most upstream layers of FGD1 and fulfills PID and upstream veto criteria: 0 = FGD1+noTPC tracks are not found; 1 = FGD1+noTPC tracks are found and all of them start within FV; 2 = FGD1+noTPC tracks are found and at least one of them starts out of FV (right). Red lines mark the regions rejected by these cuts. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

Data-MC Comparisons

Data and Monte Carlo results are compared for basic distributions. Fig. 8.14 shows the reconstructed momentum and angle distributions when the μ^- candidate fulfilled the FWD selection criteria.

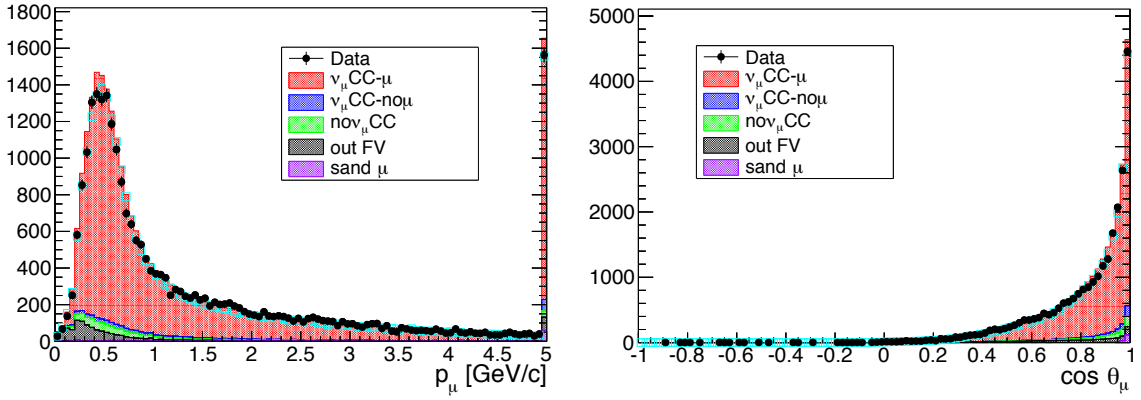

Figure 8.14: Reconstructed momentum (left) and cosine of emission angle (right) for the μ^- candidate when all criteria are fulfilled in the FWD selection. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

8.5.2 Backward Selection

The aim is to select backward going μ^- coming out from ν_μ CC interactions in FGD1 FV. The selection criteria is described in detail in the following section.

Quality and FV

This selection only consider negative charged tracks with TPC segments containing more than 19 hits and starting in FGD1 FV. Of all those tracks, the μ^- candidate must be the one with highest momentum and backward sense. In order to reduce the OOFV contamination, tracks starting in the most upstream of FGD are rejected (see Fig. 8.5). It was decided not to reject events from other layers (such as downstream layers), so statistics was not further reduced. Moreover, this sample is affected by vertex migration, which is detailed in Appendix D). Such study concluded that the position of the reconstructed vertex might be biased between data and MC.

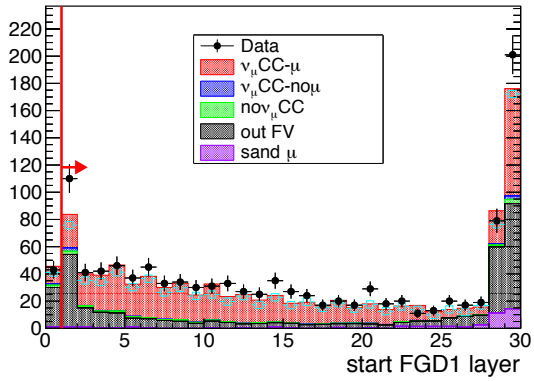

Figure 8.15: Start layer in FGD1 of the highest momentum negative charged low angle track with backward sense. Red lines mark the regions accepted by this cut. Most of the OOFV contamination is rejected with this cut. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

μ PID

μ TPC PID: When the μ^- candidate is backward going, the only PID criteria that is used is based on the L_μ (see Fig. 8.16). The value of the applied cut is the same as the one used for the FWD selection. In this angular region, the contamination of e^- is very low, so the L_{MIP} cut is not used in order not to loose statistics. In fact, using the same criteria as in the FWD selection the background (signal) would be reduced by a ∼15% (∼5%).

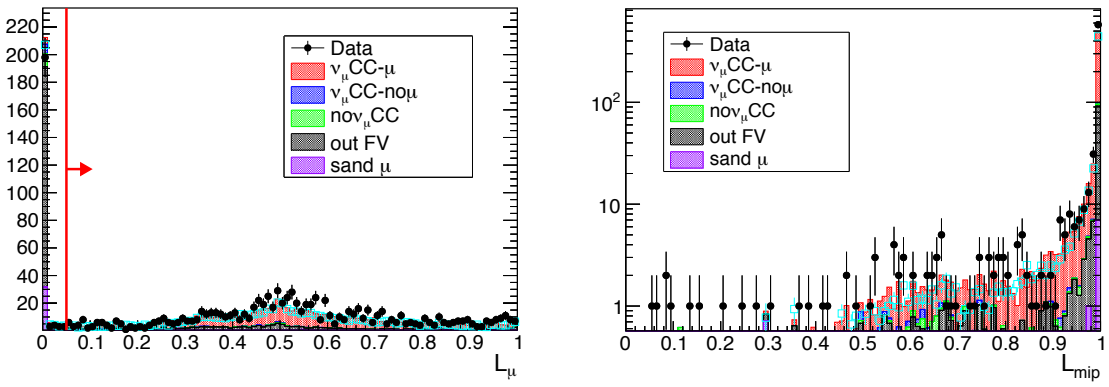

Figure 8.16: L_μ value for μ^- candidate in BWD selection (right). L_{mip} value for μ^- candidate when momentum is lower that 500 MeV/c and L_μ higher that 0.05 (left). Red line marks the accepted regions. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

Data-MC Comparisons

Data and Monte Carlo results are compared for basic distributions. Fig. 8.17 shows the reconstructed momentum and angle distributions when the μ^- candidate fulfilled the BWD selection criteria.

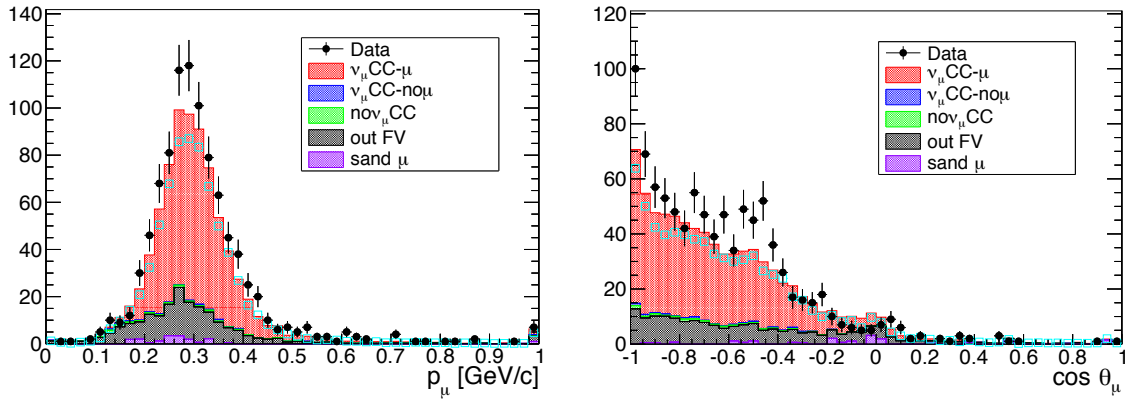

Figure 8.17: Reconstructed momentum (left) and cosine of emission angle (right) for the μ^- candidate when all criteria are fulfilled in the BWD selection. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

8.5.3 High Angle Selection

The aim is to select high angle going μ^- coming out from ν_μ CC interactions in FGD1 FV. In general, these μ^- produce a track with no TPC segments (or very short). Consequently, they are usually contained in FGD, ECal and SMRD subdetectors. It is important to notice that within this section, forward (HAFWD) and backward (HABWD) going tracks are treated independently. Indeed, the veto cut is optimized taking into account this feature. In the following section, a detailed explanation of the selection criteria is shown.

Quality and FV

High angle tracks starting in FGD1 FV and stopping either in SMRD or BarrelECal FV (see Appendix A) are used in this sample. The stopping requirement is needed to compute the momentum of the tracks (by range). In order to reduce the OOFV contamination, if the high angle track starts in the most upstream or downstream layer of FGD, it is rejected (see Fig. 8.18).

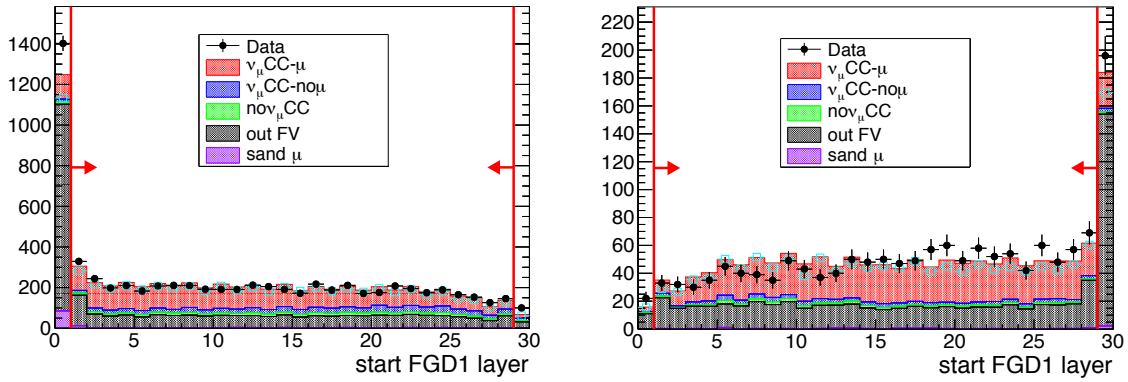

Figure 8.18: Start layer in FGD1 of high angle tracks with forward (left) and backward (right) sense. Red lines mark the regions accepted by this cut. Most of the OOFV contamination is rejected with this cut. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

μ PID

μ Stopping SMRD PID: For high angle tracks there is not reliable likelihood variable from TPCs. Therefore, SMRD or BarrelECal PID capabilities are needed. Requiring that the μ^- candidate reaches the SMRD is a good μ PID criteria. However, in the HABWD case, most of these tracks are coming from out of fiducial volume. Thus, tracks reaching SMRD will be rejected in the HABWD selection (see Fig. 8.19).

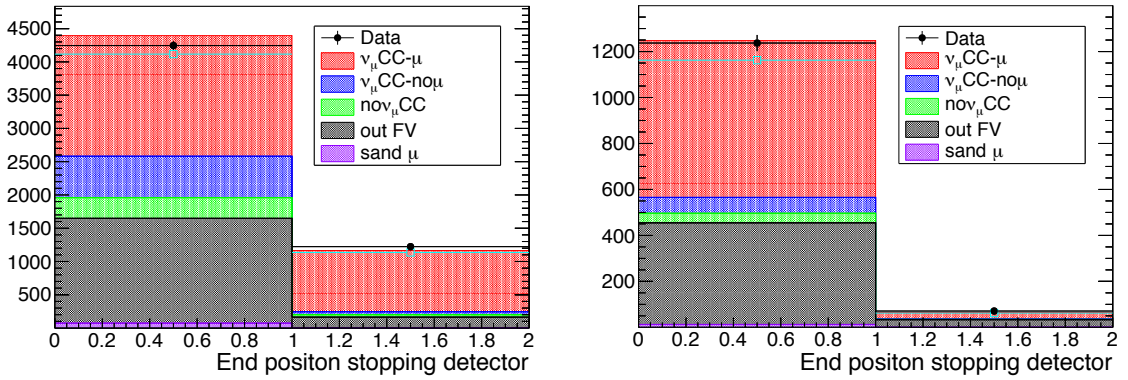

Figure 8.19: Subdetector in which μ^- candidate stops in HAFWD (left) and HABWD (right) selections (0=BarrelECal and 1=SMRD). Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

μ BarrelECal PID: For tracks not reaching SMRD (stopping in BarrelECal), some extra conditions are added. The variables used to select the μ^- candidate are the length of the track in BarrelECal,

the released electromagnetic energy and the so called MIP-EM variable. First, a cut in the MIP-EM variable is applied that allows to reject a big part of the non μ contamination (see Fig. 8.20).

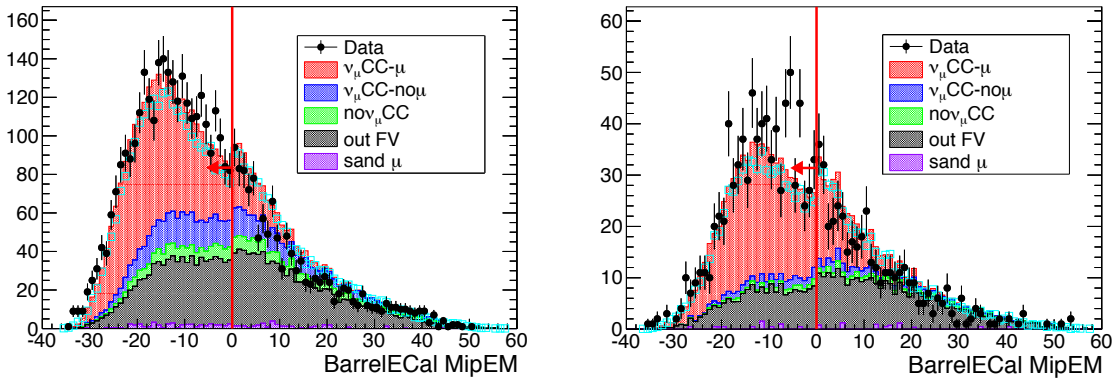

Figure 8.20: MIP-EM distribution of the μ^- candidate in HAFWD (left) and HABWD (right) selections. Red lines mark the regions accepted by this cut. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

The relation between the length and the released energy is different for μ and other particles. Therefore, the ratio between the track length in BarrelECal and the released electromagnetic energy is also used to reject non μ contamination. Overall, such distribution do not agree in data and MC (see Fig. 8.21). However, the region where background concentrates is well represented by the MC.

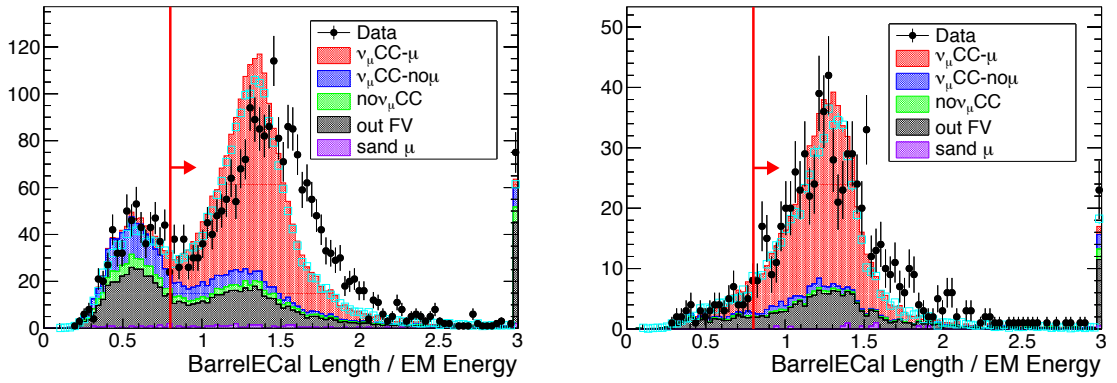

Figure 8.21: Length over EM energy distribution of the μ^- candidate when MIP-EM<0 in HAFWD (left) and HABWD (right) selections. Red lines mark the regions accepted by this cut. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

Veto

Upstream background veto: The concept of the veto cut is the same as the one introduced in the FWD selection. In this sample it is possible to distinguish between the forward and backward going tracks. Therefore, the distance and momentum ratio relation is optimized for each sense independently (see Fig. 8.22 and Fig. 8.23).

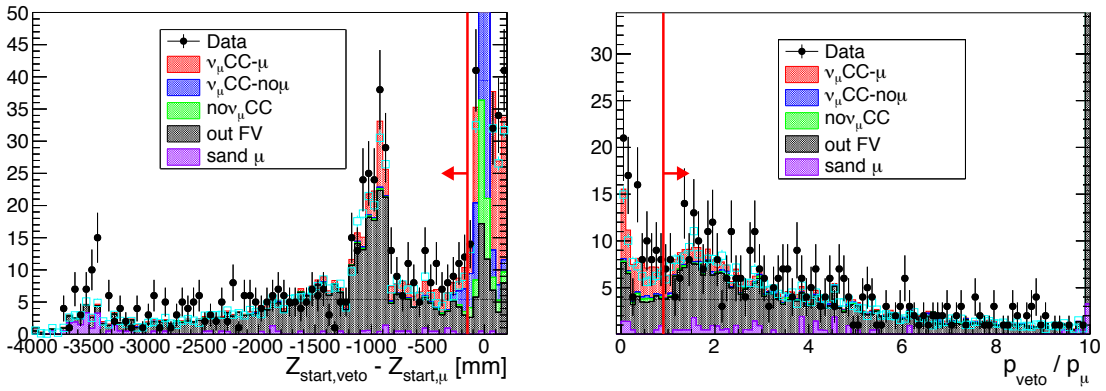

Figure 8.22: Distance between the start position of the veto track and the μ^- candidate that fulfills μ PID criteria in HAFWD selection (left). Ratio between the momentum of the veto track and the μ^- candidate when their distance is < -150mm (right). Red lines mark the regions rejected by this cut. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

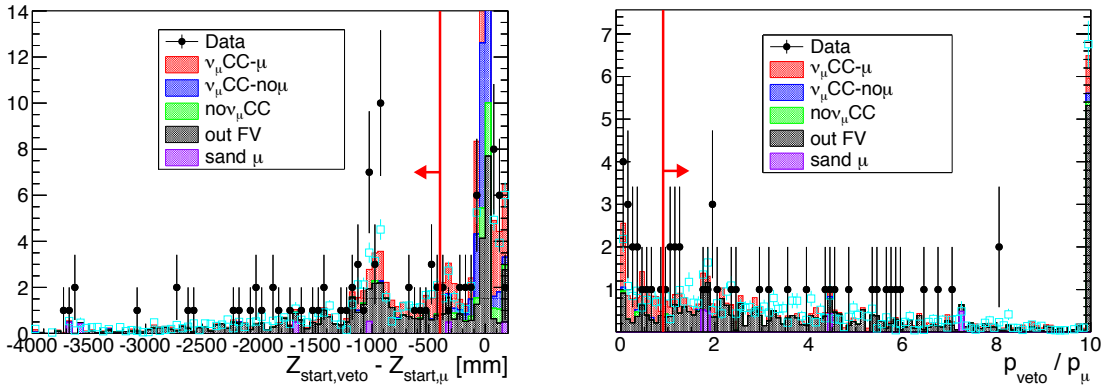

Figure 8.23: Distance between the start position of the veto track and the μ^- candidate that fulfills μ PID criteria in HABWD selection (left). Ratio between the momentum of the veto track and the μ^- candidate when their distance is higher than < -400mm (right). Red lines mark the regions rejected by this cut. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

Data-MC comparisons

Data and Monte Carlo results are compared for basic distributions. Fig. 8.24 shows the reconstructed momentum and angle distributions when the μ^- candidate fulfilled the HAFWD selection criteria. Fig. 8.25 shows the reconstructed momentum and angle distributions when the μ^- candidate fulfilled the HABWD selection criteria.

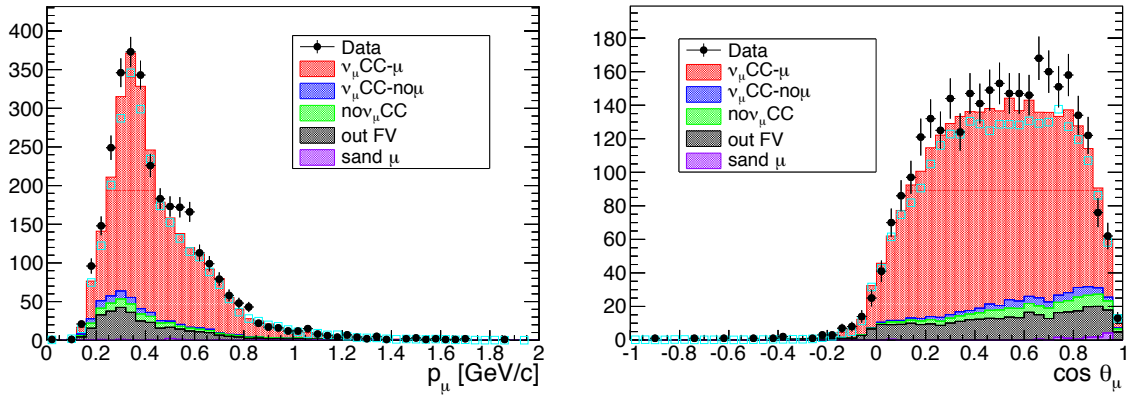

Figure 8.24: Reconstructed momentum (left) and cosine of emission angle (right) for the μ^- candidate when all criteria are fulfilled in the HAFWD selection. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

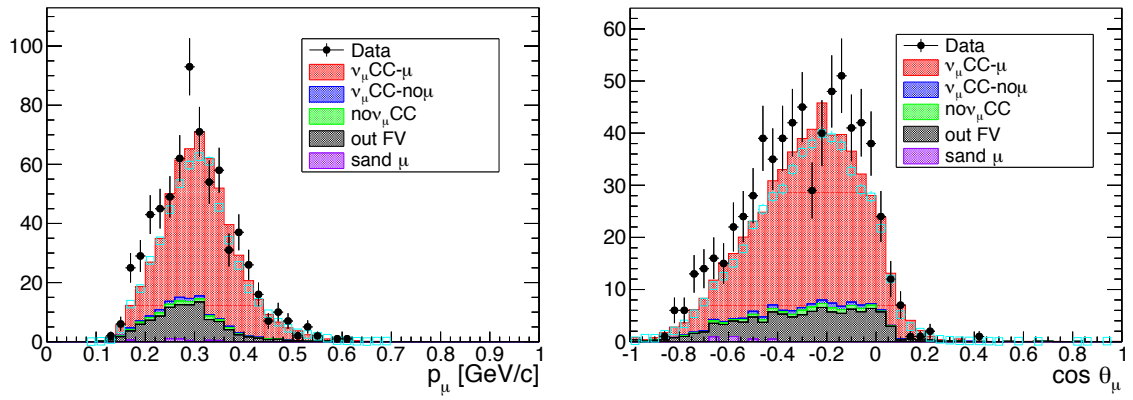

Figure 8.25: Reconstructed momentum (left) and cosine of emission angle (right) for the μ^- candidate when all criteria are fulfilled in the HABWD selection. Colors indicate different topologies of interaction. Light blue markers show the same distribution using `GENIE` as generator.

8.6 Selection Composition

Within this section a detailed information about the composition of the selected events in each sample is provided. This analysis only aims to tag μ^- coming out from the ν_μ charge current interactions taking place in FGD1 FV.

Table 8.4 summarizes how the selection steps affects the number of events in each sample comparing data and MC. Table 8.5 summarizes how such criteria affect the purity of ν_μCC-μ events. As expected, both PID and veto cuts play a main role in all the samples as they highly increase the purity.

	FWD			BWD			HAFWD			HABWD		
Cut	DATA	NEUT	GENIE	DATA	NEUT	GENIE	DATA	NEUT	GENIE	DATA	NEUT	GENIE
Quality	82155	81222	77596	11861	1050	962	7225	7121	6678	1582	1566	1461
FV	50519	51648	49285	1165	1025	939	5669	5764	5446	1356	1360	1270
μ PID	29140	29750	28169	940	799	716	3712	3487	3306	779	684	631
Veto	25669	26656	25215	940	799	716	3270	3107	2936	730	645	588
Hierarchy	25669	26656	25215	940	799	716	3082	2857	2681	682	591	542

Table 8.4: Selected number of events after each reduction step for data and MC (normalized to the data from both event generators).

	FWD		BWD		HAFWD		HABWD	
Cut	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE
Quality	32.3	31.7	58.5	56.6	41.8	41.9	48.9	47.4
FV	48.7	48.0	58.8	57.1	49.2	49.1	54.1	52.6
μ PID	81.6	81.0	73.5	72.5	71.7	71.3	72.7	72.0
Veto	89.3	88.8	73.5	72.5	79.2	79.1	75.9	76.0
Hierarchy	89.3	88.8	73.5	72.5	81.9	82.1	79.1	79.2

Table 8.5: Percentage of ν_μCC-μ events in each sample depending on the selection step with respect to the previous one for both event generators.

The composition of the selected events according to the true particle produced in the interaction is summarized in Table 8.6. The μ^- purity is rather uniform for all samples. In low angle selections, the dominant background is associated with π^- which are mis-identified as μ^- in TPC. In the high angle selections, in which there are no TPC segments, π^+ are the dominant background because the charge (curvature) of the track is not reconstructed.

In Table 8.7 the breakdown is done using the categorization defined in Sec. 8.3. Moreover, each category is split depending on its true interaction type. The main background contribution is coming from NC interactions or CCDIS interactions in which the $\mu-$ candidate is not the true $\mu-$ coming out from the interaction.

	FWD		BWD		HAFWD		HABWD	
True particle	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE
μ^-	90.8	90.2	86.8	86.3	88.4	88.0	90.8	90.8
μ^+	0.3	0.3	0.2	0.4	1.9	1.9	0.9	0.8
e^-	0.6	0.6	3.7	3.3	1.4	1.3	1.8	1.9
e^+	0.2	0.2	0.5	0.5	1.1	1.0	1.4	1.3
π^-	5.8	6.1	5.6	6.2	1.3	1.3	0.8	1.0
π^+	1.0	1.2	0.6	0.5	3.6	3.8	3.3	2.9
proton	0.3	0.3	0.1	0.0	1.9	2.0	0.6	0.6
other	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.1
notruth	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1
Sand μ	0.8	0.9	2.4	2.7	0.4	0.4	0.5	0.5

Table 8.6: μ^- candidate composition according to particle type using both event generators.

The OOFV category is split depending on the detector where the interaction took place. One of the main contributions comes from interactions taking place in the borders of FGD1, in which the closest hits to the interactions are not reconstructed. In the case of interactions in BarrelECal, the problem is that FGD and BarrelECal reconstructed segments are not matched. Finally, the contribution from interactions happening in P0D is mainly composed by neutral particles that scatter within FGD1.

	FWD		BWD		HAFWD		HABWD	
True topology	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE
ν_μCC-μ	**89.3**	**88.8**	**73.5**	**72.5**	**81.9**	**82.1**	**79.1**	**79.2**
QE	44.6	48.8	82.0	86.7	67.3	72.9	83.2	87.4
2p2h	7.5	0.0	5.5	0.0	7.2	0.0	5.3	0.0
RES	25.4	33.0	8.7	11.8	17.6	22.1	8.0	10.8
DIS	20.0	17.2	3.8	1.5	7.2	4.8	3.4	1.8
COH	2.5	1.0	0.0	0.0	0.7	0.2	0.0	0.0
ν_μCC-noμ	**2.3**	**2.0**	**1.0**	**0.5**	**2.6**	**2.4**	**1.7**	**1.0**
QE	1.9	1.8	4.6	0.0	6.3	5.3	3.4	0.9
2p2h	0.3	0.0	0.0	0.0	1.6	0.0	0.6	0.0
RES	6.4	8.6	25.1	47.4	59.1	67.0	59.5	71.7
DIS	91.1	89.4	69.7	52.6	31.7	27.0	36.5	27.5
COH	0.3	0.2	0.6	0.0	1.3	0.6	0.0	0.0
noν_μCC	**3.3**	**3.8**	**1.4**	**1.6**	**3.7**	**4.0**	**1.9**	**1.9**
NC	75.3	76.9	68.2	72.7	51.5	55.0	69.3	69.5
$\bar{\nu}_\mu$	15.9	15.3	14.3	18.0	39.3	36.2	15.4	19.8
$\nu_e, \bar{\nu}_e$	8.8	7.8	17.5	9.3	9.1	8.7	15.3	10.7
Out of FGD1 FV	**4.4**	**4.4**	**21.6**	**22.5**	**11.3**	**10.9**	**16.8**	**17.2**
TPC1	5.2	5.7	12.6	11.3	8.7	9.4	19.6	16.1
TPC2	9.0	7.4	12.0	12.5	9.0	7.5	4.9	4.6
TPC3	0.1	0.1	0.5	0.4	0.0	0.0	0.1	0.2
FGD1	13.4	12.1	16.9	13.6	35.6	34.0	36.2	29.6
FGD2	1.3	1.3	3.4	2.6	0.2	0.1	1.2	0.7
DsECAL	0.1	0.1	0.3	0.2	0.0	0.0	0.0	0.0
BrECAL	20.5	24.6	15.6	19.3	13.8	17.4	24.4	32.2
P0DECAL	6.6	7.4	5.3	6.0	5.9	6.0	3.0	3.4
P0D	27.2	24.9	19.1	19.8	14.5	13.2	2.8	3.0
SMRD	8.6	9.2	9.1	8.2	6.5	7.1	4.3	6.0
other	7.8	7.3	5.2	6.1	5.8	5.2	3.4	4.0
Sand μ	**0.8**	**0.9**	**2.4**	**2.7**	**0.4**	**0.4**	**0.5**	**0.5**

Table 8.7: μ^- candidate composition using the topologies defined in Sec. 8.3 from both event generators in bold. True reaction composition for each topology is shown as plain text. The Out of FGD1 FV category is split depending on the detector where the interaction took place).

Chapter 9

Evaluation of Systematic Uncertainties

The prediction of the rates provided by the MC presents both statistical and systematic uncertainties. Systematic uncertainties are related to the current knowledge of different factors: flux, ν interactions (so called modeling) and detector response. Flux and modeling are evaluated and parametrized by different groups within the T2K experiment, so they will be summarized in the following sections. The evaluation of the detector response will be described in more detail.

9.1 Flux

The nominal flux used during generation of the MC (version 13a) has been tuned using a set of correction factors following the recommendation from the beam group. Such tuning (v1.1) has mainly been based on latest 2009 thin target hadron production data from NA61/SHINE [62]. It is applied in an event by event basis, weighting by a factor that depends on flavor and the true energy of the ν. Fig 9.1 shows the value of the correction factors.

The uncertainties associated to the flux are parametrized as a function of the ν energy and flavor. There are several sources affecting the total error that can be summarized as follows:

- **Proton profile**: For the proton beam, the dominant sources of uncertainty are the Y alignment uncertainty of the OTR monitor (4.13 mm) and the relative alignment uncertainty in the Y direction between the primary and secondary beam lines (0.539 mrad).
- **Off-axis angle (beam direction)**: Its uncertainty was calculated using INGRID data from Run 1 to 2. The current uncertainty in the beam direction is 0.38 mrad.
- **Absolute Horn current**: The horn current is estimated to be 250 ± 5 kA. In order to study the effect of this uncertainty on the ν flux prediction, JNUBEAM simulations have been conducted by changing by $\pm 1\sigma$ the 3 horn current values with respect to their nominal settings (± 250 kA).

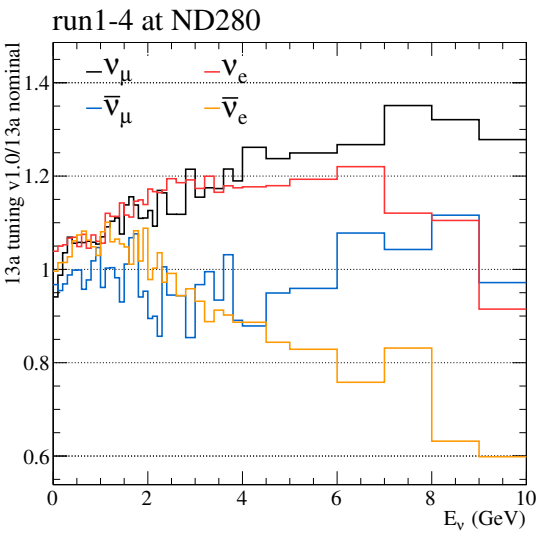

Figure 9.1: Flux tuning weights for Runs 14 at ND280. Weights are generated as a function of ν flavor and energy.

- **Horn field**: The presence of asymmetric fields can be detected by measuring the magnetic field along the horn axis. Azimuthal field measurements are within 1% of the expected theoretical value for the average peak current. For on-axis measurements, a field with unusual time-dependence was observed. Such perturbation is taken into account in the error propagation.

- **Horns alignment**: The uncertainties on the translational alignment are estimated to be 1 mm in all three directions. The uncertainty on the flux prediction arises mainly from the horns vertical displacement. The uncertainties on the angular alignment are mainly coming from Horn 1. The estimated angular alignment uncertainty is 2 mrad in both the horizontal and vertical directions. The uncertainties were evaluated by shifting the horns by their alignment uncertainties and re-running the flux simulation.

- **Target alignment**: The uncertainty in the target alignment is estimated to be 1 mrad in both directions (horizontal and vertical) based on the measured alignment of the target on the downstream end. The uncertainty due to the target misalignment is estimated by generating flux predictions with the target displaced by +1mrad in X and Y separately, and compared to the nominal flux predictions.

- **Material modeling**: Some materials have not been considered so far in the secondary beam line geometry simulation performed in JNUBEAM. Therefore, their effects are not yet included in the flux prediction, but those effects are taken into account as additional uncertainties. The first

new source is the cooling water inside the horns. The second type of material considered is the striplines, which carry the current to the horns. The flux prediction uncertainties are estimated by taking the ratio of the flux with cooling water or stripline geometry included to the nominal flux that does not include those materials.

- **Proton number**: An absolute normalization error on the flux is calculated based on the uncertainty of the current transformer (CT) beam current, and therefore POT measurement. The CT measurement uncertainty comes from the uncertainty on the absolute calibration of the most downstream CT.

- **Multiplicity**: The meson multiplicity is studied using NA61 2009 thin target data. Systematic uncertainties are evaluated by propagating the errors from that data set. For regions of the phase space not covered by NA61, the difference between FLUKA model and extrapolations based on NA61 data is set as an uncertainty.

- π **rescattering**: The interaction rate of π in the target and other beam line materials was tuned and uncertainties were evaluated. The tuning of those interactions was done using HARP measurements of $\pi^{\pm}$ production in interactions with π incident on nuclei [151]. The difference between the tuned flux and the flux without this tuning will be used as a systematic uncertainty.

- **Nucleons**: The production of secondary baryons in proton interactions is studied using 2009 NA61 proton multiplicity data. NA61 has provided a covariance matrix, which describes the uncertainties on their data points, that will be used as systematic uncertainty.

- **Interaction length**: NA61 group estimates production cross section of 31 GeV/c protons on Carbon and its associated uncertainty. For higher momentum ($p > 20$ GeV/c) protons, assigned error size is $+3.0\%$ / -1.9%. For interactions outside the target (iron or aluminum), the uncertainties are obtained using external production cross-section data [150].

While sources of uncertainty can be assumed as uncorrelated, it is expected to observe a correlation between different energy bins for each source. Consequently, covariance matrices are computed for each source independently and they are added in quadrature to obtain the overall uncertainty. Fig 9.2 shows the final covariance matrix taking into account all source of uncertainties and their relative errors.

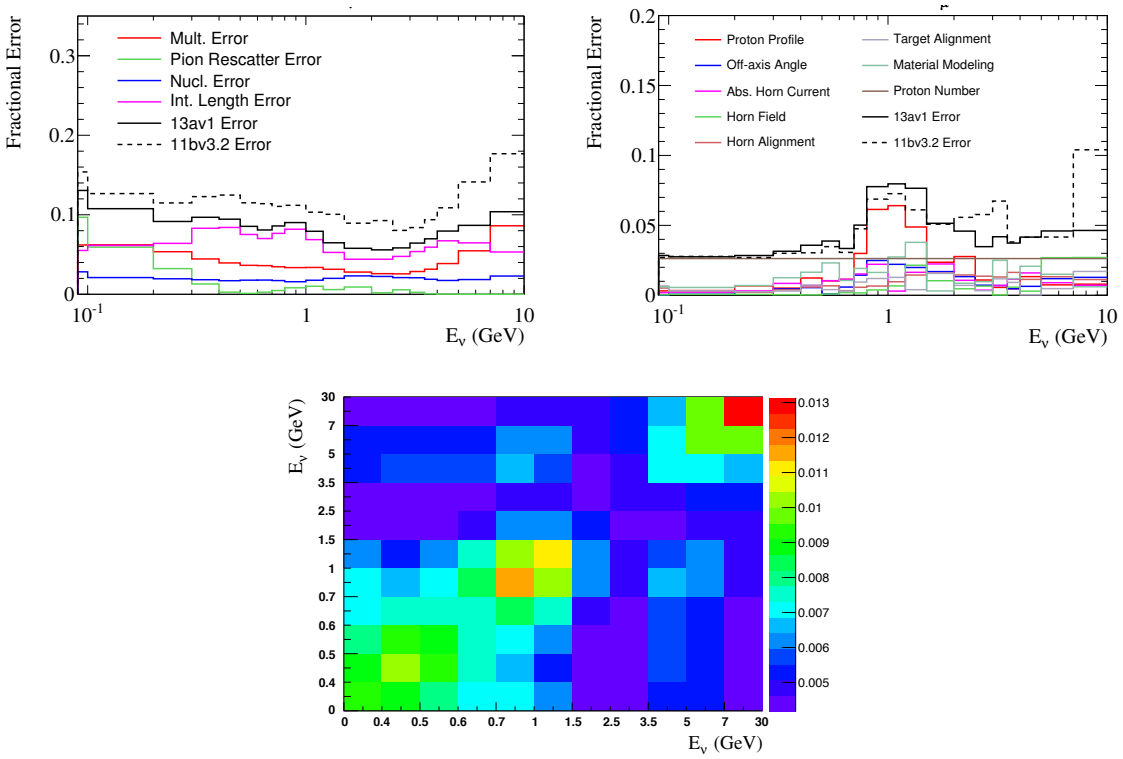

Figure 9.2: Flux uncertainties of the ν_μ prediction in FHC as function of the ν energy associated to hadron interaction model (top left) and others (top right). Dashed lines show the uncertainty used in previous analysis for comparison. Bottom: Flux covariance matrix in bins of ν energy.

9.2 Cross Section Modeling

The uncertainties associated to the cross section models used by the MC, usually called "modeling", are explained in this section. In T2K experiment, a group called NIWG is in charge of understanding the impact of the modeling mainly for the oscillation analysis.

In this analysis, ν_μ CC cross section measurement, the impact of the model used to generate the MC should not be dominant. The main reason is that it is an inclusive selection (it only aims to reconstruct the outgoing μ^-), so the purity is very high (uncertainties associated to background modeling are low). Besides, the reconstruction efficiency for μ is well understood in ND280.

In the following subsections, uncertainties associated to each interaction channel relevant for the ν_μ CC analysis will be discussed.

9.2.1 Charged Current Quasielastic Interactions

This particular reaction plays a fundamental role in T2K because it is the main channel in the oscillation analysis. Therefore, it has been deeply studied by different groups within the experiment. In `NEUT` (see Sec. 4.3), this interaction is modeled using three main parameters:

- `MAQE`: Axial mass (see Sec. 4.1.1).
- `pF`: Fermi momentum (see Sec. 4.2.1), whose value is different depending on the target material.
- `EB`: Binding energy (see Sec. 4.2.1), whose value is different depending on the target material.

NIWG performed an study in which two different nuclear models (SF and RFG+RPA from sec. 4.2.1) were fitted to external data [160] (see Fig. 9.3). The aim was to constrain the uncertainty in the parameters described above. However, in this analysis the values and uncertainties obtained in such study have not been used. The main reason is that when comparing distributions from best-fit MC and data some tensions were found between different experiments.

Therefore, it was decided to loosen the uncertainties from such study. This strategy has been recommended for cross section analyses. The purpose has been to use so large errors that they cover the current knowledge for the physical parameters and nuclear models. Table 9.1 summarizes the result.

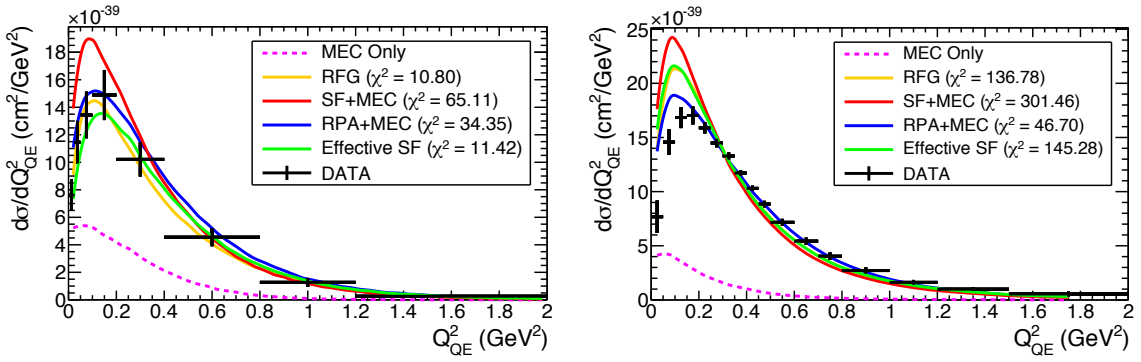

Figure 9.3: Nominal model predictions for the MINERVA with $\theta_\mu < 20$ (left) and MiniBooNE datasets with `MAQE`= 1.01 GeV, and all other model parameters at their default values.

Dial	Nominal	Error
`MAQE`	1.2	0.3
`pF_C`	217	30
`pF_O`	225	30
`EB_C`	25	30
`EB_O`	27	30

Table 9.1: Nominal value and error of the dials in the CCQE channel for `NEUT`.

9.2.2 2p2h interactions

`NEUT` includes multinucleon reactions using a Nieves model (see Sec. 4.2.2 and Sec. 4.3). This process produces an enhancement of the cross-section resulting from ν interacting with more than one nucleon.

The lack of data for this reaction channel makes it very difficult to constrain current models. Therefore, it was decided to include a 100% normalization uncertainty to this reaction independently from the target type or ν type.

9.2.3 Single Resonant π interactions

ν interactions on nucleons in which a single π is produced are generated by `NEUT` using the so called RES channel (see Sec. 4.1.2 and Sec. 4.3). The model used by this generator has three main parameters:

- `MANFFRES`: The resonant axial mass. It controls both the overall cross section and the shape of the cross section as function of Q^2.
- `CA5`: The normalization of the axial form factor. It controls the overall normalization of the cross

section.

- `BgRES`: The scale of the $I = 1/2$ non-resonant background.

The nominal values and their associated uncertainties for the three parameters were obtained by comparison with available neutrino-nucleon scattering data (mostly on deuterium). For `MANFFRES` and `CA5`, $\nu_\mu + p \to \mu^- + p + \pi^+$ was used (pure $I = 3/2$ process), while for `BgRES` both $\nu_\mu + n \to \mu^- + p + \pi^0$ and $\nu_\mu + n \to \mu^- + n + \pi^+$ were used (once `MANFFRES` and `CA5` were fixed). Table 9.2 summarizes the result.

Dial	Nominal	Error
`CA5`	1.01	0.12
`MANFFRES`	0.95	0.15
`BgRES`	1.3	0.2

Table 9.2: Nominal value and error of the dials in the RES channel for `NEUT`.

A comparisons between data (see Sec. 4.4) and nominal MC with $\pm 1\sigma$ bands can be seen in Fig. 9.4.

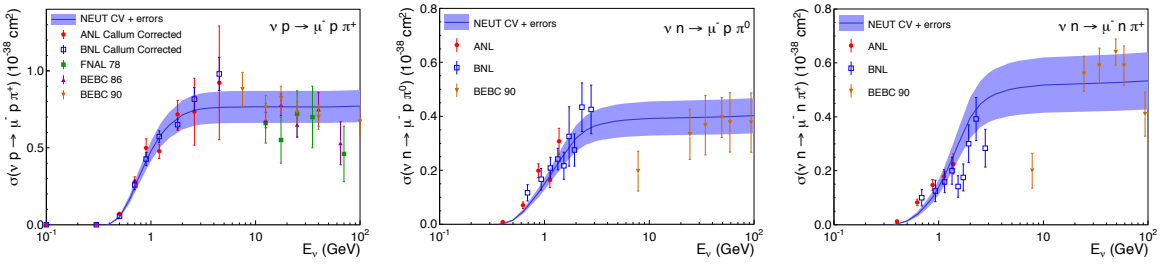

Figure 9.4: The effect on the total cross section changing each of the parameters by $\pm 1\sigma$ for the three CC single π channels.

9.2.4 Others

In this section a summary of the remaining uncertainties associated to other interaction channels which are less dominant in the T2K energy range are shown.

Coherent π production: Recently, MINERVA experiment has reported results which are consistent with coherent π production at ν energies around 1 GeV [161]. From that result, it was decided to include a 30% normalization uncertainty to coherent interaction via CC. In the case of NC, comparison between data [163] and NEUT coherent π^0 production has been performed. Such comparison motivated a 30% normalization uncertainty assigned to the NC coherent π^0 cross section.

Other CC Interactions: The interactions contributing to this category are mainly the ones generated by the DIS channel in `NEUT` (see Sec.4.3). From external data sets [164], the uncertainty is known to be on the order of 10% at 4 GeV. Therefore, an energy dependent uncertainty is defined as follows:

$$\delta\sigma_{CC-Other} = \frac{0.4}{E_\nu} \tag{9.1}$$

NC Interactions: This category groups all NC interaction channels. These channels are assigned a 30% normalization uncertainty, motivated by the poor constraints from external data.

ν_e **interactions**: Possible differences between ν_e and ν_μ ν QE cross sections have been investigated [165]. Differences can arise from many factors in the theoretical cross-section because of the difference in final state lepton mass. Available experimental data are used to limit the possible size of the effect. Currently, a normalization uncertainty of 3% is used for ν_e CC interactions. An anticorrelation term is added to take into account the fact that the ν_e/ν_μ ratio and $\bar{\nu}_e/\bar{\nu}_\mu$ cross section ratios may be slightly different.

9.2.5 π Final State Interactions

Final state interactions affect hadrons before they leave the nuclear medium, changing their true identity and kinematics (see Sec. 4.2.3). In this analysis, the main hadronic contribution comes from π^- wrongly tagged as μ^- candidate (proton contribution is negligible). Therefore, only π FSI contribution will be taken into account.

In `NEUT`, the cascade model is parametrized with several dials:

- `FSI_PI_ABS`: It scales the microscopic cascade interaction mechanism probabilities for π absorption.
- `FSI_PI_PROD`: It scales the microscopic cascade interaction mechanism probabilities for π production.
- `FSI_INEL_LO_E`: It scales QE scattering when $p_\pi < 500$ MeV/c.
- `FSI_INEL_HI_E`: It scales QE scattering when $p_\pi > 500$ MeV/c.

- `FSI_CEX_LO_E`: It scales single charge exchange branching fraction when $p_\pi < 500$ MeV/c.
- `FSI_CEX_HI_E`: It scales single charge exchange branching fraction when $p_\pi > 500$ MeV/c.

The uncertainties on π FSI are currently estimated from fits to π-nuclear scattering data on Carbon using the parameters described above ($\vec{a}_{fsi}$). The following χ^2 was minimized:

$$\chi^2 = \sum_{i=0}^{\#exp} \frac{1}{n_i} \sum_{j}^{n_i} \frac{(\sigma_j^{MC}(\vec{a}_{fsi}) - \sigma_j^{data})}{(\Delta\sigma_j^{data})^2} \tag{9.2}$$

Where i (j) runs over the number of experiments (data points) and σ represents the measured and predicted cross sections. Fig 9.5 shows the covariance matrix obtained from the fit for $\vec{a}_{fsi}$.

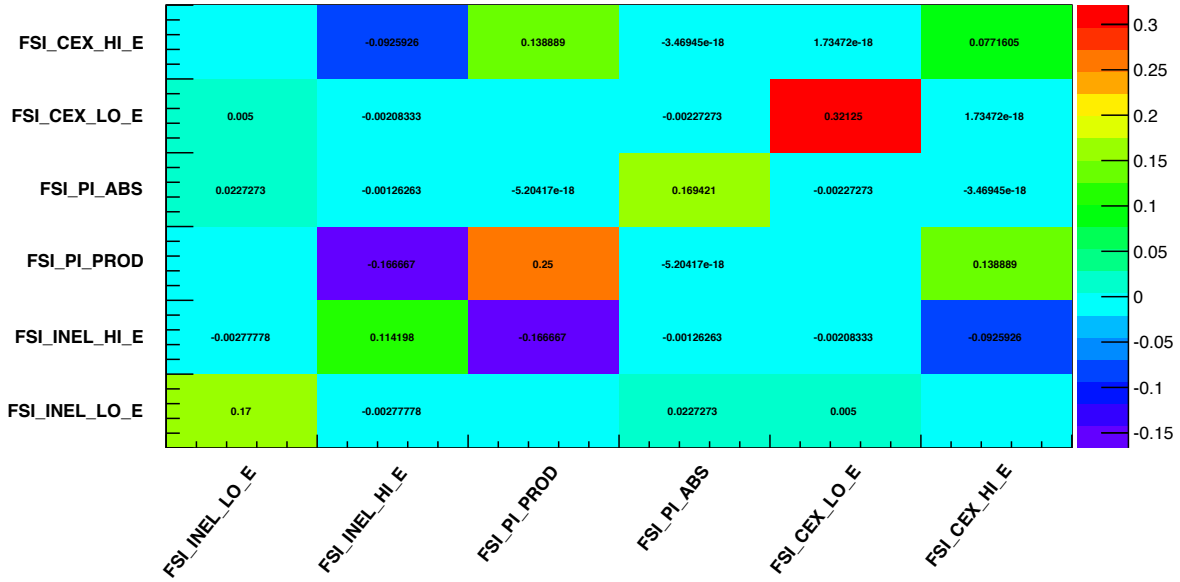

Figure 9.5: Covariance matrix for all π FSI dials.

A comparisons between data and MC best-fit with $\pm 1\sigma$ bands can be seen in Fig. 9.6.

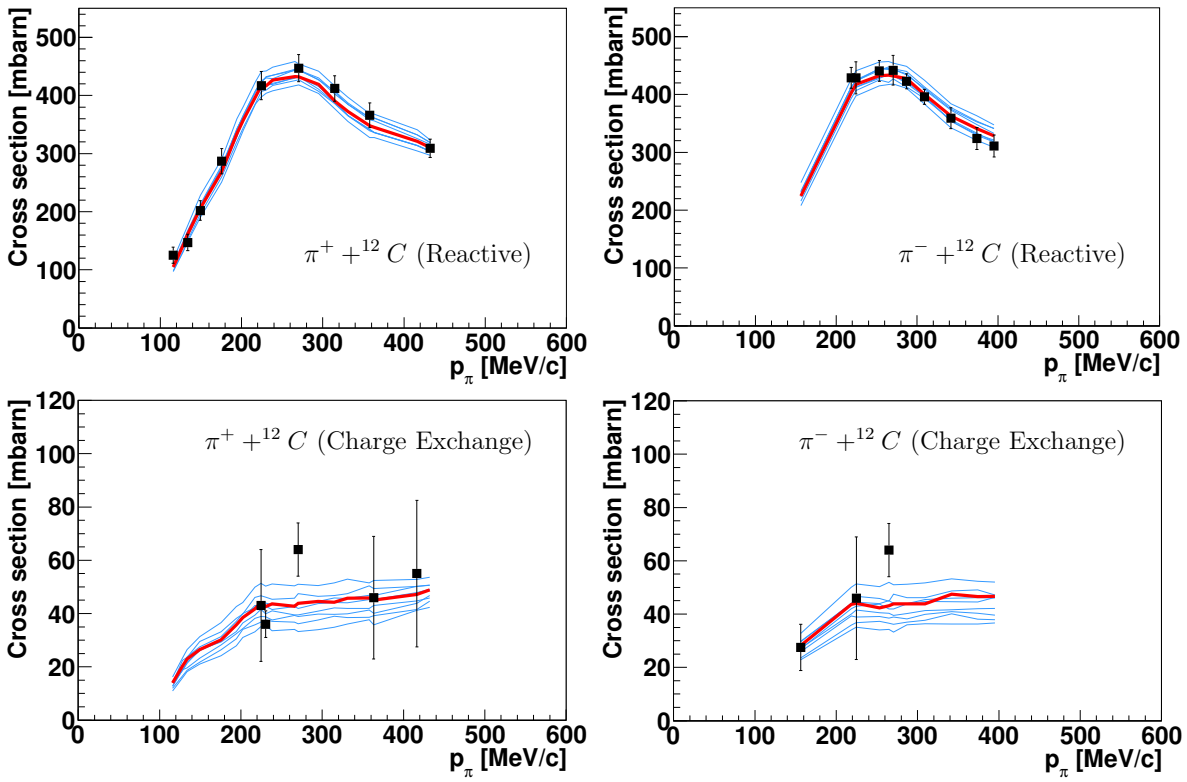

Figure 9.6: π-C cross section comparisons between data ([152][153][154][155][156][157][158][159]) and NEUT best fit (red) with $\pm 1\sigma$ errors for different interaction channels.

9.3 Detector

9.3.1 Evaluation

In this section, different sources of systematic uncertainty associated to the detector response are explained. Most sources have been evaluated using dedicated control samples that mimic the properties of the events in this selection [1]. This work has been developed by several groups in the T2K experiment.

In the following sections, the different sources affecting the ν_μ CC selection will be explained in detail.

Track charge sign identification

The charge sign reconstruction relies on the sign of the curvature measured inside the TPC. The uncertainty of misidentify the sign of the charge is parametrized using an observable that reflects the real dependencies of the misidentification. There is strong a correlation between the curvature (or inverse transverse momentum) error and the probability to misidentify the charge:

$$P = 1 - \frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty \exp^{\frac{(p_T - p_T^{true})^2}{2\sigma^2}}\, dp_T \tag{9.3}$$

where p_T is the signed transverse momentum. The curvature error measured by the trajectory of a particle in a magnetic field is given by:

$$\Delta\rho \propto \frac{\epsilon}{L^2}\sqrt{\frac{720}{N+4}} \tag{9.4}$$

where L is the length of the particle trajectory, N is the number points measured along the track and ϵ the spacial resolution at each measured point. The trajectory fit done by the reconstruction provides an estimation of this error. Using this error, it is possible to look for dependencies on the charge confusion probability.

The global and the local charge determination are correlated depending on the number of TPC segments in the track and on whether they predict the same charge. Two errors must be taken into account:

- Probability of swapping the local TPC charge identification. The ratio of data over MC for this variable is 1.00074 ± 0.00023.

[1] For instance, cosmic μ stopping in a subdetector of ND280 or through going μ crossing the TPCs

- Probability of the global tracking to swap the sign of the charge obtained from the best fit among the local segments. In this analysis, two scenarios are possible:
  - 1 TPC segment: The value of local and global charges always coincide. The ratio of data over MC for this variable is 0.99852 ± 0.00022.
  - 2 TPC segments: Local charge values from the two segments can be the same or opposite. In the first case, the ratio of data over MC for this variable is 1.00016 ± 0.000076. In the second case, it is important to quantify which of both segments has more impact on the global charge determination, which is the one with the smallest error. The ratio of data over MC for this variable is 1.0201 ± 0.0055.

TPC cluster efficiency

When a charged particle crosses the TPC, it produces a long trace of ionized gas. This efficiency quantifies the probability of finding a cluster of TPC pad hits corresponding to such trace. Those hits are clustered either vertically (Y direction) or horizontally (Z direction) depending on the absolute value of the track local angle with respect to the Z direction in the ZY plane.

In the selection, a TPC quality criteria is used in order to distinguish between low (FWD/BWD) and high angle (HAFWD/HABWD) tracks, checking the number of clusters in TPC1 or TPC2. Therefore, it is important to quantify the TPC cluster efficiency in data and MC.

Those efficiencies are computed comparing the distributions of the number of clusters in the reconstructed μ track for data and MC. For vertical cluster, μ originated from ν_μ CC interactions in FGD1 are selected. For horizontal cluster, cosmic μ tracks crossing TPC2 vertically are selected.

The ratio between data and MC for the cluster reconstruction efficiency was found to be 0.9993 ± 0.0001 (0.9989 ± 0.0002) for horizontal (vertical) clusters.

TPC track reconstruction efficiency

This efficiency evaluates how likely a track is reconstructed within a TPC when a charged particle crosses it. Thus, it evaluates TPC pattern recognition algorithm and the likelihood fit. If a segment in TPC1 or TPC2 track is not reconstructed, it has an effect in the ν_μ CC analysis because the rate of event is affected.

From different studies, it has been concluded that the TPC track reconstruction efficiency is very high and does not depend on track momentum, angle, length or the presence of a second track in the same TPC.

A control sample of through-going μ from the beam and cosmic events is used to evaluate the track reconstruction efficiency. The upstream and downstream detectors (except for the FGDs) around the tested TPC are used to select events in which a single μ crossed the tested TPC. For TPC1 reconstruction efficiency, events with segments in P0D and TPC2 are used as a reference. If no segment is found in the tested TPC, it is classified as inefficient.

Table 9.3 shows the results of the efficiency for data and MC.

	ϵ_{data}	ϵ_{MC}
TPC1	$0.999^{+0.001}_{-0.001}$	$0.996^{+0.002}_{-0.003}$
TPC2	$0.997^{+0.002}_{-0.007}$	$0.995^{+0.003}_{-0.004}$

Table 9.3: TPC track reconstruction efficiencies for both TPCs.

ECAL reconstruction efficiency

In this analysis all tracks (with and without TPC segments) are used. The high angle selection criteria (see Sec. 8.5.3) includes tracks without TPC segments but with BarrelECal segments. Therefore, an evaluation of how likely a track is reconstructed within BarrelECal when a charged particle crosses it is needed. Such efficiency is compared for data and MC, distinguishing between the two objects that ECal reconstruction algorithm includes: showers and tracks.

In order to do so, a control sample of tracks crossing a TPC an pointing to BarrelECal is used. To select samples that should produce either a track or a shower in BarrelECal, it was decided to take an $e^{\pm}$ enhanced sample ($\sim 80\%$ pure) for the shower-like and a μ^{-} enhanced sample ($\sim 90\%$ pure) for the track-like using the TPC PID capabilities. Then, the shower/track criteria for BarrelECal is defined using the value of the MIPEM variable.

The track (shower) efficiency is defined as the ratio between the number of events in μ^{-} ($e^{\pm}$) the enhanced sample with a track(shower)-like reconstructed object in BarrelECal and the number of events in μ^{-} ($e^{\pm}$) the enhanced sample. Fig. 9.7 shows the results of the efficiency for data and MC. This efficiency depends on the particle type and its momentum.

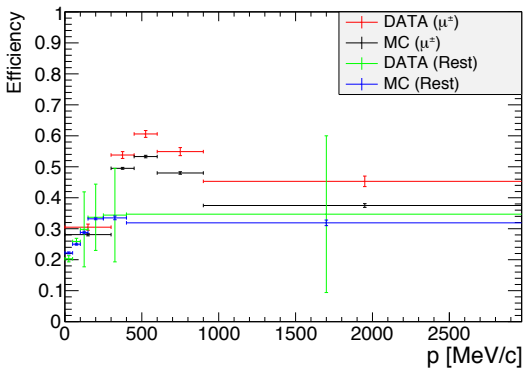

Figure 9.7: BarrelECal reconstruction efficiencies as a function of momentum for different particle types in data and MC.

FGD-TPC matching efficiency

The rate of events in ν_μ CC analysis is affected by how well the reconstruction matches FGD and TPC segments. This efficiency evaluates how likely a TPC segment is matched to a hit in the upstream FGD.

A control sample of through-going μ crossing two consecutive TPCs is used. The matching efficiency is computed from the rate of events that have also a segment in the FGD in between. This efficiency is 100% in data and MC. However, this control sample only spans a limited phase-space.

Another control sample of cosmic μ is used. Here, tracks with TPC2 segments pointing to FGD1 are selected. The matching efficiency of the ratio of the number of events with a reconstructed FGD1-TPC2 track over the number of events with hits in the FGD1. It was found to be 100%.

This matching efficiency depends on the FGD hit efficiency. Therefore, it was decided to evaluate the FGD hit efficiency in the two downstream layers of FGD1 and FGD2. Table 9.4 presents the results of the efficiency for data and MC. Its difference is used as uncertainty and it is propagated only for tracks that only contain one or two reconstructed hits in the FGD segment.

	ϵ_{data}	ϵ_{MC}
FGD1	0.969±0.8	0.976±0.45
FGD2	0.965±0.85	0.971±0.50

Table 9.4: Average hit efficiency for both FGDs.

TPC-ECAL matching efficiency

As in the previous section, the rate of event in the analysis is affected by the matching efficiency between TPC and ECal detectors (mainly in FWD and BWD selections). Possible data and MC discrepancies for the matching efficiency must be taken into account, so systematic uncertainties are calculated for tracks entering the BarrelECal and the DsECal from a TPC.

The same control sample as in Sec 9.3.1 is used but the efficiency is defined in a different way:

$$\varepsilon = \frac{\#\text{tracks TPC segment and ECal segment}}{\#\text{tracks TPC segment}} \tag{9.5}$$

The TPC PID is used to distinguish between different particle types. Fig. 9.8 shows the results of the efficiency for data and MC. This efficiency depends on the particle type, momentum and ECal subdetector.

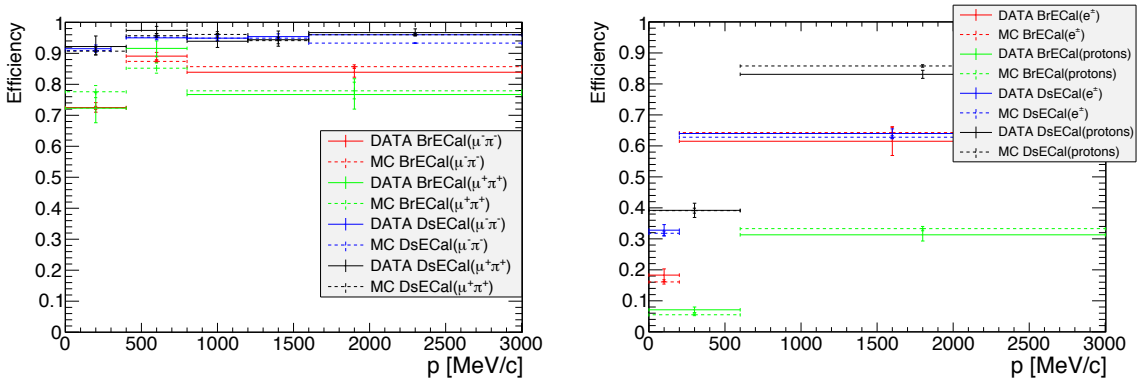

Figure 9.8: Matching efficiencies between TPC and BarrelECal or DsECal as function of the momentum for different particles types (μ and π are grouped together) in data and MC.

TPC-P0D matching efficiency

In the BWD selection, many of the tracks are reconstructed by looking at the time of flight information between FGD1 and P0D. Therefore, the rate of events in that sample is affected by how well the reconstruction matches P0D and TPC segments. Then, this efficiency evaluates how likely a TPC segment is matched to a P0D segment.

A control sample made of cosmic μ is used and events are selected by requiring a reconstructed segment in TPC1 starting close to P0D. The efficiency is defined as the ratio between the number of events with

a matched TPC1-P0D segment and the total number of events in the control samples.

This efficiency has been computed for data and MC in different momentum regions. Fig. 9.9 shows the values reported as function of the momentum for each TPC and particle hypothesis.

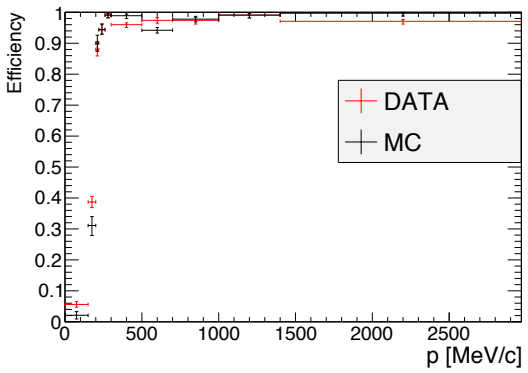

Figure 9.9: Matching efficiencies between TPC and P0D as function of the momentum in data and MC.

FGD-ECAL(SMRD) matching efficiency

In the HA selection, the μ^- candidate is most times contained within FGD, BarrelECal and SMRD subdetectors. Therefore, this sample is affected by how well the reconstruction matches FGD, BarrelECal and SMRD segments. This efficiency evaluates how likely a FGD segment is matched to a reconstructed object in BarrelECal and SMRD.

A control sample of through-going μ with an BarrelECal (SMRD) segment that points to FGD is used. In order to mimic the kinematics of the μ^- candidate, it is required that the μ stops within the FGD (Michel e^- is found). The matching efficiency is computed from the ratio between the number of events with a matched FGD-BarrelECal (or FGD-BarrelECal-SMRD) segment and the total number of events in the control samples.

This matching efficiency has been computed for data and MC in different angular regions and different SMRD modules. Fig. 9.10 shows the values reported as function of the angle. Such efficiency is much lower for the Top/Bottom SMRD modules. This is because the coils that surround the ND280 subdetectors (see Sec. 6.3.1) are in between those modules and BarrelECal modules (see Fig. 6.4).

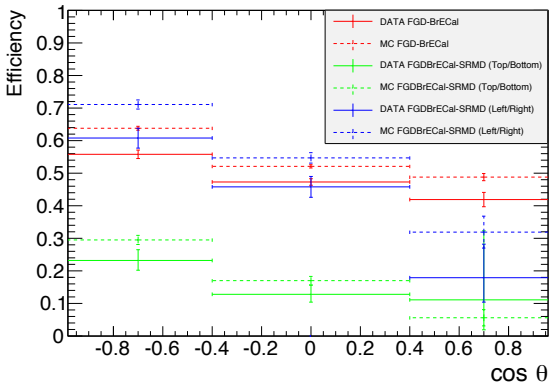

Figure 9.10: Matching efficiencies between FGD and BrEcal or SMRD as function of the direction in data and MC.

ECAL EM resolution and scale

The EM Energy variable is a quantity assigned to each reconstructed cluster object in the ECal, which is strongly correlated with the sum of the hit charges (see Sec. 6.3.7). This variable is used to reduce the proton contamination in the high angle selections as described in Sec. 8.5.3.

The confidence in this variable can be obtained by comparing it with an independent energy measurement. The TPC provides a precise momentum measurement of charged particles, which for the purpose of this comparison can be treated as the the true energy. Discrepancies between data and MC in the distribution of (EM Energy-p_{TPC}) are used to bias and smear the variable in the MC and to set a systematic uncertainty.

In order to make the comparison work, the $e^{\pm}$ control sample from Sec 9.3.1 is used requiring that the ECal segment is reconstructed. Similarly to TPC momentum, this variable has two different uncertainties associated: scale (due to MC bias) and resolution (due to MC smear). Currently, a 3.9% uncertainty is associated to the resolution and a 4.7% to the scale.

TPC PID

The TPC PID plays a key role in the ν_μ CC analysis, because μ^- can be misidentified when selecting the lepton candidate. It is based on the measurement of the energy loss by a particle when crossing the TPC gas (see Sec. 6.3.7).

This systematic is evaluated for $\mu^{\pm}$, $e^{\pm}$ and protons using control samples directly extracted from beam

events. The uncertainties are extracted from the differences observed between data and MC for the pull distributions. Firstly, the pulls distributions in MC are biased and smeared using the mean and width differences between data and MC. Then, the estimated error for the width is used as systematic uncertainty.

Fig. 9.11 shows the values of the pull mean and width as function of the momentum in TPC2 (which is the main TPC in this analysis because is downstream FGD1) for $\mu^{\pm}$, $e^{\pm}$ and protons.

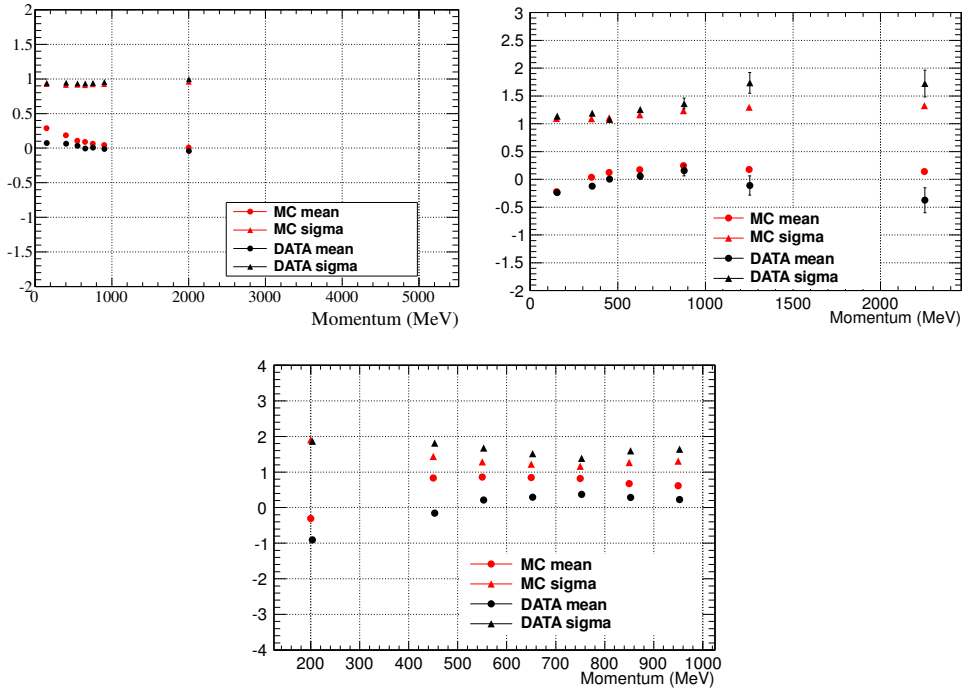

Figure 9.11: Pull mean (points) and sigma (triangles) values for $\mu^{\pm}$ (left), $e^{\pm}$ (middle) and protons (right) in TPC 2 as a function of the momentum. Data points are shown in black while MC is shown in red.

ECAL PID efficiency

Both FWD and HA selections include an PID cut based on ECal information. In particular, they use a variable called MIP-EM, which aims to distinguish between track-like and shower-like particles (see Sec. 6.3.7). The systematic uncertainty in the ECal PID is estimated by comparing selection efficiencies in data and MC for high purity $e^{\pm}$, $\mu^{\pm}$ and proton control samples.

Such control samples are constructed as in Sec 9.3.1, distinguishing between different particles using the TPC PID criteria. For each control sample a cut was placed on MIPEM and the efficiency for each

particle type was calculated. This efficiency is calculated for both real data and MC samples. Any difference between data and MC is interpreted as a systematic error in the modeling of the ECal PID for that particle type. Fig. 9.12 presents the results of the efficiency for data and MC depending on the particle type and ECAL subdetector.

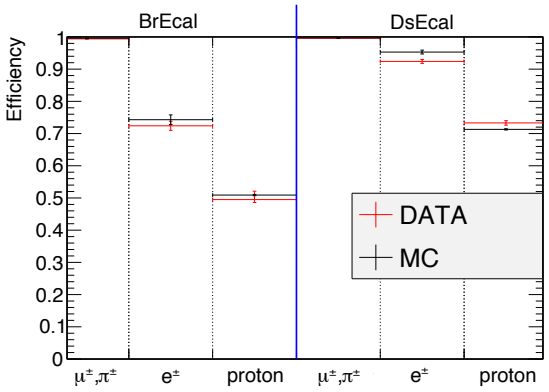

Figure 9.12: ECAL PID efficiencies using BarrelECal and DsECal for different particle types in data and MC. For $\mu^{\pm}\pi^{\pm}$ and protons ($e^{\pm}$) efficiency is computed using MIP-EM<0 (MIP-EM>0).

Magnetic field intensity

The trajectory of charged particles will be deflected when crossing a magnetic field. From that curvature r, transverse momentum p_T of the track can be computed using the following relation:

$$p_T[\text{MeV/c}] = 0.3B[\text{T}]r[\text{mm}] \tag{9.6}$$

Where B is the mean value of the magnetic field. The field intensity has an uncertainty associated to it. From a series of measurements made with magnetic field probes an error of 0.57% was estimated for the mean magnetic field.

TPC field distortions

Currently, the MC simulation assumes an homogeneous magnetic field $\vec{B}$ in ND280. Besides, it assumes an uniform electric field $\vec{E}$ in the drift volume of the TPCs. However, this ideal case is not fulfilled in the real TPCs, so discrepancies for some observables which depend on such fields (as momentum) are expected between data and MC.

On the one hand, $\vec{B}$ has been measured inside the ND280 basket using a Hall probes and distortions of the field were found (see Sec 6.3.1). Therefore, a correction is applied to the data at the reconstruction level based on the $\vec{B}$ map. This is done through the evaluation of the change in Y and Z position, when each cluster in the TPC is drifted back from the pad plane to the ionization point using the distorted map. Applying this correction to tracks crossing the TPCs, the discrepancies between data and MC for the reconstructed momentum were reduced.

Moreover, $\vec{B}$ has been measured using a laser system within the TPCs, which illuminates aluminum dots on the cathode. For each dot the expected and measured position are compared, when the magnet is on (including previous correction) and off. However, this method measures the distortions integrated over the whole drift distance from the cathode to the anode. Therefore, this map is not included in the correction at the reconstruction level, but it is used as a estimation of the systematic uncertainty due to magnetic distortions.

On the other hand, an empirical parametrization of $\vec{E}$ electric field within the TPCs has been done, showing clear non uniformities in the field distribution (see App. B).

TPC momentum resolution

A method has been developed to compare the global momentum resolutions of data and MC. If differences are found, MC is corrected applying an smearing factor and an uncertainty is associated to it.

A control sample of μ crossing multiple TPCs is selected in data and MC. Then, an observable is computed comparing the inverse of the transverse momentum of two consecutive TPCs (subtracting the energy loss in the intermediate FGD). The distribution of this difference is Gaussian and centered at zero. The width of such distribution is related to the momentum resolution.

That distribution has been studied as function of different observables (momentum angle, drift distance), showing clear discrepancies between data and MC. Preliminary studies have shown that the main cause for this discrepancy could be misalignment between the MicroMegas planes of the TPCs.

Therefore, it was decided to apply a smearing factor (as function of the drift distance) to the reconstructed momentum in the MC (see Fig. 9.13). To account for dependencies not considered and for the deficiencies of the method, the error for each bin was enlarged to 0.1 (10% systematic uncertainty on

the knowledge of the momentum resolution).

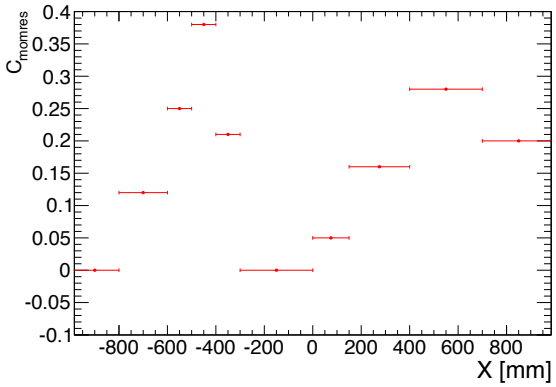

Figure 9.13: Smear factor (correction) applied in MC to the inverse transverse momentum for different x ranges.

Momentum by range resolution

The μ^- candidate in the HA sample do not include a TPC segment. Therefore, its momentum must be computed from the range of the track (it must stop within the ND280 detector). Possibles discrepancies between the reconstructed momentum by range in data and MC are studied. Then, the momentum by range in MC will be smeared using such differences.

A control sample of μ crossing one TPC and with both ends stopping within in FGD1 and BarrelECal (or SMRD) are selected in data and MC. Then, the distribution of the difference between momentum provided by the TPC segment and the momentum by range are compared in data and MC.

The mean of such distribution is centered at zero in data and MC, so a correction factor is not needed. However, some differences were found in the width of the distribution. That difference is used to set the uncertainty. Fig. 9.14 shows the propagated smear factor as function of the angle and for tracks ending in either BarrelECal or SMRD. In the case of SMRD, the systematic uncertainty was found to be significantly bigger than for BarrelECal ($\sim$50% and $\sim$20% for backward and forward going tracks respectively).

Time of flight

The timing information is mandatory to reconstruct the sense of the tracks. As it was shown in App. C.1, the time reconstruction appears to be different in data and MC. Therefore, an smearing

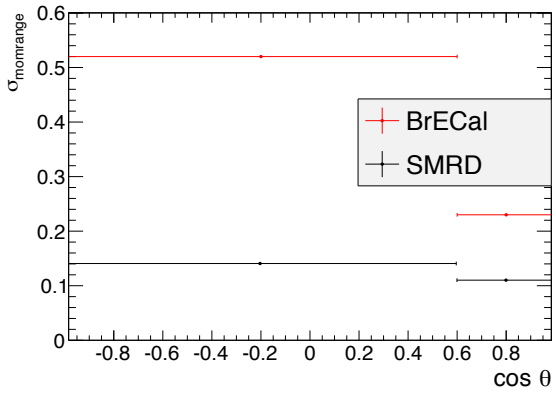

Figure 9.14: Smear factor (systematic) applied to the momentum by range for different direction ranges and detectors.

factor is applied to the ToF variable in MC.

This variable plays a fundamental role in this analysis, so it was decided to associate an uncertainty to it. That error is set equal to the maximum bias or resolution correction. With this criteria, an underestimation of its error is very unlikely.

Table 9.5 shows the smearing values depending on the properties of the track (see App. C.1) and the detector used to compute the ToF.

ToF	σ_{ToF} [ns]
LA-Start	
FGD1-FGD2	1.484
FGD1-BrECal$_{Track}$	1.574
FGD1-BrECal$_{Shower}$	1.437
LA-End	
FGD1-P0D$_{Track}$	1.470
FGD1-BrECal$_{Track}$	1.670
FGD1-BrECal$_{Shower}$	2.107
HA-Start	
FGD1-BrECal$_{Track}$	1.209
FGD1-BrECal$_{Shower}$	1.378
HA-End	
FGD1-BrECal$_{Track}$	1.589
FGD1-BrECal$_{Shower}$	1.875

Table 9.5: Smear factor (systematic) applied to the time of flight depending on the topology of the track (see App. C.1).

Sand events background

ν interactions happening outside the ND280 detector might produce particles that cross the detectors. Such events, so called sand μ, can be misclassified as ν interactions in the FGD1, affecting the ν_μ CC selection.

Sand μ are generated using a dedicated MC simulation. The rate of sand μ has a uncertainty associated, which is computed comparing the rates of tracks entering the upstream wall of the P0D in two different simulations and in data.

The data/MC discrepancy is below 10% for one of the simulations and ∼10% for the other. Therefore, it was decided to apply a 10% systematic uncertainty to the predicted numbers of sand events.

Event pile up

In FWD, HAFWD and HABWD selections, events with activity in TPC1 can be rejected due to the veto criteria. Tracks with TPC1 segments are usually coming from interactions upstream the detector (sand μ) or P0D. However, sand μ are not included in the standard MC simulation (they are generated independently), so a scale is done to take into account events that are rejected due to coincidence with a sand μ. Therefore, the number of selected events in MC is reduced using the following formula:

$$w_c = 1 - C_{pileup} \tag{9.7}$$

This correction is evaluated for each running period separately. The method to extract that number is to count the number of events with a segment in TPC1 per bunch in sand μ simulations.

This correction has a systematic uncertainty associated because sand μ are simulated with a 10% error. Besides, the actual and simulated beam intensity and side-band materials might be different in data and MC. The error is computed comparing the number of events with a segment in TPC1 per bunch in data and MC (including sand μ).

Table 9.6 shows the final values for the correction and systematic uncertainty.

	C_{pileup}	σ_{pileup}
Run 2 air	0.0097	0.00097
Run 2 water	0.0083	0.00083
Run 3b air	0.0096	0.00096
Run 3c air	0.0105	0.0024
Run 4 air	0.0118	0.0040
Run 4 water	0.0134	0.0028

Table 9.6: TPC1 Pile-up correction and uncertainty for each data set.

OOFV background

Out of Fiducial Volume (OOFVV) category tags interactions happening outside the FGD1 FV but they are reconstructed as being originated inside it. It is the main background in the four samples of the ν_μ CC analysis. The rate of OOFV events has an associated uncertainty coming from two factors: rate and reconstruction.

Rate uncertainty: For interactions happening in P0D, ECal or SMRD (composed by heavy nuclei), an uncertainty in the rate of events is applied. Such uncertainty is evaluated comparing the rate of events with interactions in each subdetector for data and MC. In the case of interactions happening in the magnet or electronics, the uncertainty is set to be different between the `NEUT` and `GENIE` generators. For OOFV events originating inside the FGD or TPC but outside the FV region (either scintillator or aluminum) there is not direct rate measurement from the neutrino data that is independent of the ν_μ analysis itself. Nevertheless, the rate uncertainty associated to these kind of events should be already taken into account with the cross section modeling systematics (see Sec. 9.2).

Table 9.7 summarizes the rate uncertainty for each subdetector.

	FWD		BWD		HAFWD		HABWD		Rate
	`NEUT`	`GENIE`	`NEUT`	`GENIE`	`NEUT`	`GENIE`	`NEUT`	`GENIE`	uncertainty
FGD or TPC	29.0	26.6	42.0	40.4	53.5	51.0	62.0	51.2	0
ECAL	27.2	32.1	20.9	25.5	19.7	23.4	27.4	35.6	0.116
P0D	27.2	24.9	19.1	19.8	14.5	13.2	2.8	3.0	0.051
SMRD	8.6	9.2	9.1	8.2	6.5	7.1	4.3	6.0	0.049
Other	7.8	7.3	5.2	6.1	5.8	5.2	3.4	4.0	0.136

Table 9.7: The last column shows rate uncertainties on OOFV events depending on their origin of the interaction. Other columns show the OOFV composition in each selection from both event generators.

Reconstruction uncertainty: The OOFV background is split into 9 categories using the true information from MC. Such categories aim to distinguish different reconstruction failures and a different

uncertainty is associated to each one of them.

- *OOFV inside the FGD*: Interactions in the XY edges of the FGD (borders of the FV) normally produce hits that match to the TPC track and the event be properly identified as OOFV. However, a hit may miss when an event happens in dead material in the XY module, and the first hit will appear inside the FV. A study has shown that the FGD1 hit efficiency in those layers is the same in data and MC.

- *In tracker downstream (upstream)*: Backward (forward) going hadrons originated in the upstream (downstream) walls of the TPC2(1) can produce hits in FGD1. Those hits can be matched to the forward (backward) going lepton in the ν_μ CC interaction which will be misidentified as an event happening within the FV. The probability of that mismatching depends on the matching criterion used to associate FGD hits to extrapolated TPC tracks. Data-MC comparisons show that the distribution for data and MC differ in width by a 1.3 factor. Such discrepancy is relate to the subdetectors alignment.

- *From neutral parent*: Neutral particles can be produced in ν interactions outside the FGD1. Then, they can enter the FGD and create secondary charged particles. The reconstruction will not reject the event. However, the reconstruction for those events is expected to be the same in data and MC.

- *Backward (forward) event*: Backward (forward) going π^+ produced downstream (upstream) that stop in the FGD can be misidentified as a lepton coming out from the ν_μ CC interaction. Those events are partially rejected when the timing information is available. Therefore, the uncertainty associated to that timing must be used for those events (see Sec. 9.3.1).

- *High angle event*: Tracks entering FGD1 from BarrelECal and going through TPC1 or TPC2 can be misclassified as ν_μ CC interaction in the FV, when the reconstruction is not able to match the hits on the edges of the FGD with the other hits of the FGD and TPC segments. A sample of high angle cosmic μ triggered by the FGD is used to measure the uncertainty associated to this type of failure. The efficiency is computed counting the rate of events in which the hits from the border of the FGD are matched to the reconstructed track. The discrepancy between data and MC is used as uncertainty.

- *Last module failure*: It happens when the TPC segment has its drift velocity wrongly reconstructed, so the X coordinate is not reliable. When matching those segments with the FGD hits,

only YZ layers in the YZ projection are used. In this case, the FGD-TPC matching algorithm will match hits up to the last FGD module where the matching gives up. A control sample of μ crossing TPC1-FGD1-TPC2 is used to estimate the difference in the failure rate between data and MC. For such μ with $p < 0.5$ GeV/c, the distribution of the most upstream module in the FGD segments was used as a reference for comparison.

- *Double skipped layer*: The TPC-FGD algorithm is likely to skip one single layer if no hits are found. However, if there is a lack of FGD hits in two consecutive layers in the same projection, the algorithm stops the matching. For very forward tracks, they can pass through the dead coating material between two scintillator bars without creating hits. Therefore, a broken track pattern will appear and the selection can misidentify one of the two tracks as an interaction within the FV. The failure uncertainty is estimated similarly to the previous case, but comparing the rate in the whole FGD1 FV (instead of the last module).

- *Hard scattering*: When a μ hard scatter in the FGD, the hit pattern is difficult to match to the TPC segments. For those events, it is very likely that the algorithm picks a wrong hit to start the extrapolation, so a through going track can be misidentified as stopping inside the FGD FV. The method to assign the uncertainty is similar to the one used in the double skipped layer category, but using μ with $p > 1$ GeV/c.

Table 9.8 shows reconstructed uncertainty, which depends on the track features. It is important to notice the treatment is symmetric between forward and backward going tracks.

	FWD		BWD		HAFWD		HABWD		Reco
	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE	NEUT	GENIE	uncertainty
OOFV inside the FGD	7.5	8.0	23.2	19.5	25.9	25.8	38.4	30.7	0.0
In tracker down(up)stream	11.0	8.5	10.0	9.8	9.1	7.3	6.6	5.8	0.05
From neutral parent	27.3	25.3	15.6	15.1	21.1	20.8	11.1	12.9	0.0
Backward (forward) event	20.0	22.9	18.4	16.9	5.4	7.6	24.9	28.3	0.0
High angle event	7.8	8.4	12.4	13.5	13.8	14.1	15.8	19.3	0.33
Double skipped layer	6.9	7.1	3.0	4.3	4.0	3.8	0.2	0.1	0.55
Last module failure	0.4	0.4	0.1	0.4	0.3	0.3	0.0	0.0	0.35
Hard scattering	19.1	19.4	17.3	21.0	20.4	20.1	3.0	2.8	0.32

Table 9.8: The last column shows uncertainties on OOFV events depending on the reconstructed category. Other columns show the OOFV composition in each selection from both event generators. Notice that some categories must be flipped to be consistent for backward selections (shown in parenthesis).

Hadron Secondary Interactions

When a hadron is produced in an interaction and it leaves the nucleus, it is likely that it interacts within the detector in a so called secondary interaction. Three different types of interactions will be taken into account:

- *Absorption*: The incident hadron is absorbed by a nucleus, and no hadrons are produced in the interaction.
- *Charge Exchange*: A π^0 is produced when the incident π interacts with a nucleus.
- *Quasi-Elastic scatter*: The momentum or direction of a hadron is changed when it suffers an inelastic interaction with a nucleus.

The probability for this type of interactions is included in the Geant4 model and it has been found to be significantly different from external data results. In order to account for such a difference, a correction is applied event by event to the nominal MC. The uncertainty in the external data is used as to set the systematic error in the correction.

The phase space from the available external data does not cover the full momentum range of π in ND280. Therefore, an extrapolation technique has been used to cover those regions and the uncertainties in such regions has been enlarged. It is assume that the farther away one gets from the last data point in momentum, the less close to reality the corrected Monte Carlo is.

ν direction

In this analysis, the reference for the angular information of the outgoing μ^-, $\cos\theta_\mu$, is the ν direction. In order to determine the ν direction, two positions are used: the vertex of the interaction in FGD1 and the parent decay point of the ν in the decay tunnel.

On the one hand, the vertex position is well defined at ND280. On the other hand, the mean position in which ν parents decay in the decay tunnel is provided by the bean group (see Fig 9.15).

However, if the neutrino direction is extracted in this way, then a systematic must be associated with it (using the uncertainty in the mean decay point). This is taking into account by varying the mean decay point based on the ν parents decay position distribution in the MC (see Fig 9.15).

Consequently, the variation of the ν direction will have an impact in the $\cos\theta_\mu$ reconstructed variable.

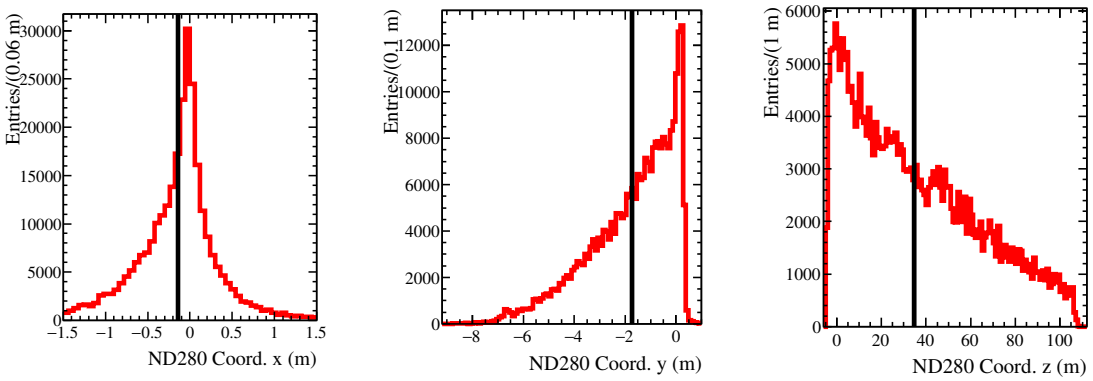

Figure 9.15: ν parent decay position distributions produced. Black lines indicate the mean positions of each distribution.

Vertex migration

In this analysis, the definition of the vertex is done using the reconstructed position of the μ^- candidate. In the FGD, this position is defined from the hit pattern in the different layers. When the multiplicity of particles increases such procedure becomes non trivial (mainly for topologies in which there are back to back particles), so a vertex might be migrated. This effect has shown a non negligible impact on the BWD and HABWD samples.

In App. D a dedicated study has been performed for this kind of events using the truth information from the MC. The idea has been to try to characterize the vertex migration and try to quantify its impact in this analysis. In summary, the main systematic uncertainty is associated to the reconstructed momentum of the μ^- candidate.

9.3.2 Propagation

The systematics uncertainties described in previous sections are propagated in order to check their impact in the rate of reconstructed events for each sample. `Psyche` framework integrates several tools that perform such propagation, which is done using toy experiments. In this framework, the systematics uncertainties are classified in three different groups:

- **Observable variation**: It concerns to a systematic that associates an uncertainty to reconstructed variables (for instance, momentum or time of flight). In this case, the propagation is performed running the selection with altered values of the reconstructed variables, so it will affect the number of selected events.

- **Efficiency-like**: It concerns to a systematic that associates an uncertainty to a reconstruction/detection probability. In this case, a weight is applied event by event, based on such probability once the selection is finished.

- **Normalization-like**: It concerns to a systematic that associates an uncertainty to the rate of events in certain subsample, scaling them up or down. This kind of systematic is propagated as the efficiency like systematics.

As described in Sec. 8.4, prior to the propagation of certain efficiency and normalization like systematic uncertainties, the nominal value of reconstructed events is corrected by applying a weight event by event based on the data and MC discrepancies described in previous sections. The systematic uncertainties are propagated after the corrections are applied.

When generating the toy experiments, the systematic uncertainties must be thrown assuming a Probability Density Function (PDF). For this analysis all PDFs are assumed to be Gaussian except for the field distortion (uniform distribution is assumed). The covariance matrix associated to the reconstructed number of events is given by:

$$V_{ij} = \frac{1}{N_{toys}} \sum_{x}^{N_{toys}} (N_i^x - \mu_i)(N_j^x - \mu_j) \tag{9.8}$$

Where the weighted number of selected events in bin i per toy x is $N_i^x = \sum_e^{N_{events}} w_{e,i}^x$ and μ_i its average value. Results were obtained after throwing 500 toys.

Table 9.9 shows the mean number of selected events and their relative error in each sample, splitting the contribution from each source of systematic uncertainty. On the one hand, the predicted rate of events is clearly affected by the corrections associated to detector response, where pile up correction dominates. On the one other hand, the correction associated to the flux is increasing the amount of selected tracks in each sample.

Fig. 9.16 shows the effect of the systematics as function of the kinematics of the μ^- candidate in each

sample. The correlation among different samples and each bin of reconstructed μ^- momentum and angle (see Sec. 10.3) is shown in Fig. 9.17.

Currently, the uncertainty associated to the matching among FGD, ECal and SMRD subdetectors dominates in both HAFWD and HABWD selections. The reason is that the misalignment between both subdetectors has not been properly corrected in data, leading to discrepancies in the matching efficiency for segments contained in those subdetectors.

In the case of the BWD sample, as expected, the matching between TPC and P0D subdetectors and the ToF resolution dominates. Meanwhile, in the FWD selection the uncertainty associated to the particle identification in BarrelECal and DsECal dominates. That uncertainty is applied in a conservative way to μ^- candidates crossing such subdetectors.

	FWD	BWD	HAFWD	HABWD
DATA	25669	940	3082	682
MC w/o corrections	23454.4	700.3	2739.5	560.7
ALL	26655.7 (1.807)	798.7 (5.852)	2857.2 (14.262)	591.5 (14.314)
Efficiency-like				
Charge ID	23445.8 (0.081)	701.4 (0.184)	2739.5 (0.000)	560.7 (0.000)
TPC Cluster	23454.5 (0.0005)	700.3 (0.0002)	2739.6 (0.0023)	560.7 (0.0008)
TPC track	23505.7 (0.798)	702.3 (0.421)	2739.3 (0.048)	560.7 (0.020)
ECAL track	23523.8 (0.162)	702.6 (0.210)	2998.7 (4.116)	606.4 (4.888)
TPC-FGD	23440.1 (0.108)	670.0 (0.080)	2739.4 (0.004)	560.7 (0.005)
TPC-ECAL	23539.5 (0.465)	701.8 (0.828)	2742.3 (0.161)	560.9 (0.079)
TPC-P0D	23454.4 (0.000)	733.4 (3.929)	2739.5 (0.000)	560.7 (0.000)
FGD-ECAL	23454.4 (0.000)	700.3 (0.000)	2590.8 (4.665)	516.1 (6.462)
FGDECAL-SMRD	23454.4 (0.000)	700.3 (0.000)	2645.4 (11.594)	560.7 (0.000)
ECAL PID	23678.0 (1.294)	700.3 (0.000)	2742.6 (0.500)	560.6 (0.349)
Normalization-like				
Pile up	23182.6 (0.232)	700.3 (0.000)	2707.7 (0.234)	554.2 (0.235)
OOFV	23454.4 (0.497)	700.3 (1.884)	2739.5 (1.021)	560.7 (1.978)
Sand mu	23454.4 (0.056)	700.3 (0.113)	2739.5 (0.023)	560.7 (0.032)
SI π	23529.6 (0.344)	700.8 (0.255)	2739.5 (0.394)	561.7 (0.253)
SI proton	23454.4 (0.014)	700.3 (0.001)	2739.5 (0.164)	560.7 (0.010)
Observable variation				
B field	23448.5 (0.007)	700.2 (0.008)	2739.4 (0.001)	560.7 (0.004)
Momentum (resolution)	23418.3 (0.023)	700.2 (0.015)	2738.9 (0.006)	560.7 (0.005)
Momentum (scale)	23453.6 (0.007)	700.3 (0.000)	2739.6 (0.004)	560.7 (0.006)
Momentum (bias)	23453.6 (0.003)	700.3 (0.000)	2739.6 (0.006)	560.7 (0.010)
TPC PID	23296.8 (0.238)	690.7 (0.392)	2739.7 (0.024)	560.7 (0.020)
Momentum by Range	23454.4 (0.000)	700.3 (0.000)	2739.1 (0.085)	560.5 (0.061)
ECAL EM (resolution)	23454.4 (0.000)	700.3 (0.000)	2720.4 (0.080)	552.8 (0.190)
ECAL EM (scale)	23454.4 (0.000)	700.3 (0.000)	2737.4 (0.809)	559.2 (1.531)
ToF	23433.6 (0.080)	673.2 (2.621)	2688.5 (2.367)	446.2 (7.307)
ν direction	23454.4 (0.000)	700.3 (0.000)	2739.5 (0.000)	560.7 (0.000)

Table 9.9: The first two rows show the total number of events for data and nominal MC without applying corrections associated to flux and efficiency-like detector systematics. The third row indicates the mean number of events and its relative error in parenthesis when propagating all detector systematic uncertainties including the flux correction. The remaining rows indicate the mean number of events and their percentage relative error in parenthesis when propagating each detector systematic uncertainty independently not including the flux correction.

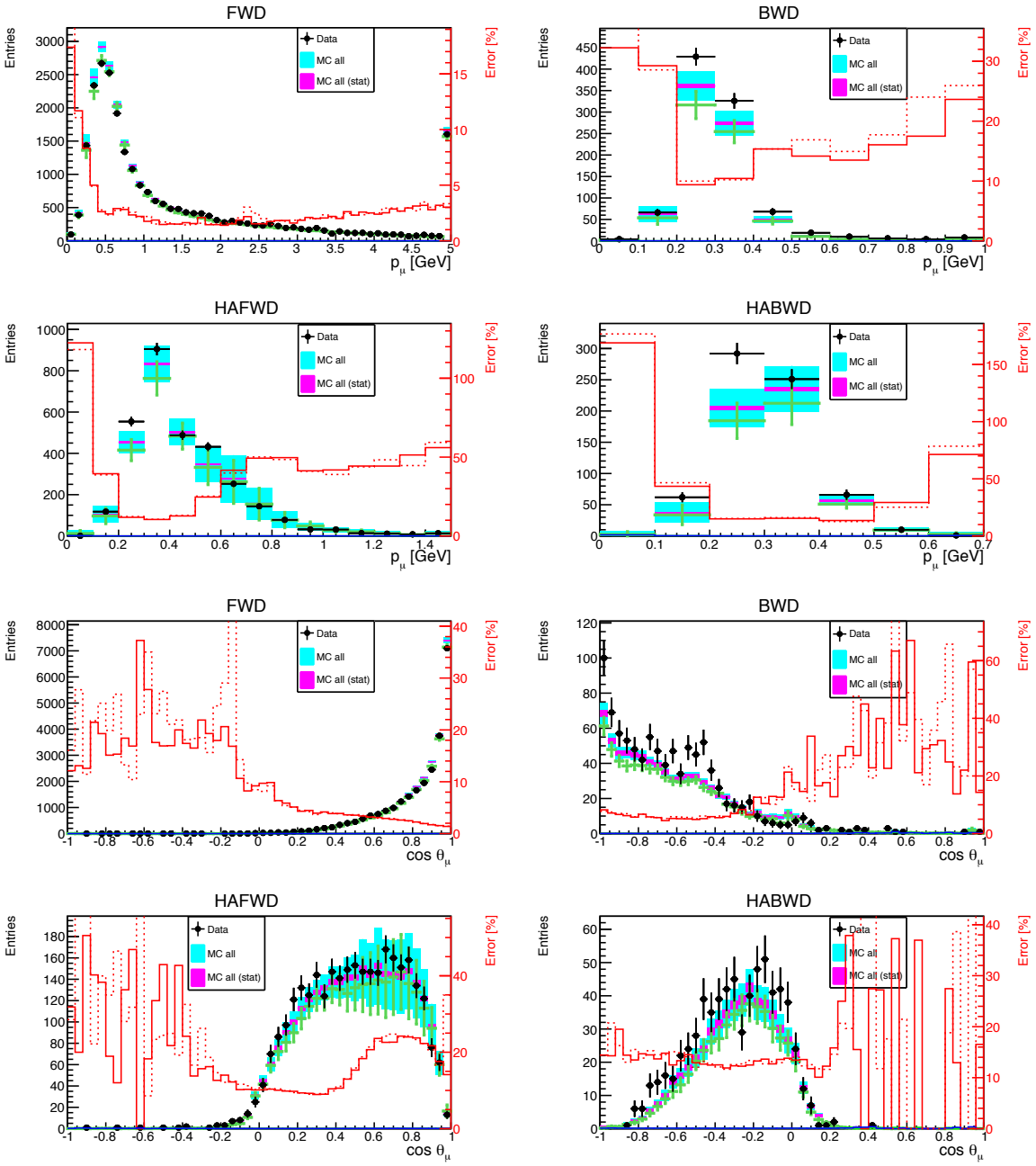

Figure 9.16: Distributions of μ^- momentum (top) and cosine of emission angle with respect ν direction (bottom) when all systematics are thrown together. Error bars represent the statistical and detector systematic uncertainties (rectangles for `NEUT` and crosses for `GENIE`). Red lines indicate the relative errors per each bin (straight line is `NEUT` and dashed line is `GENIE`).

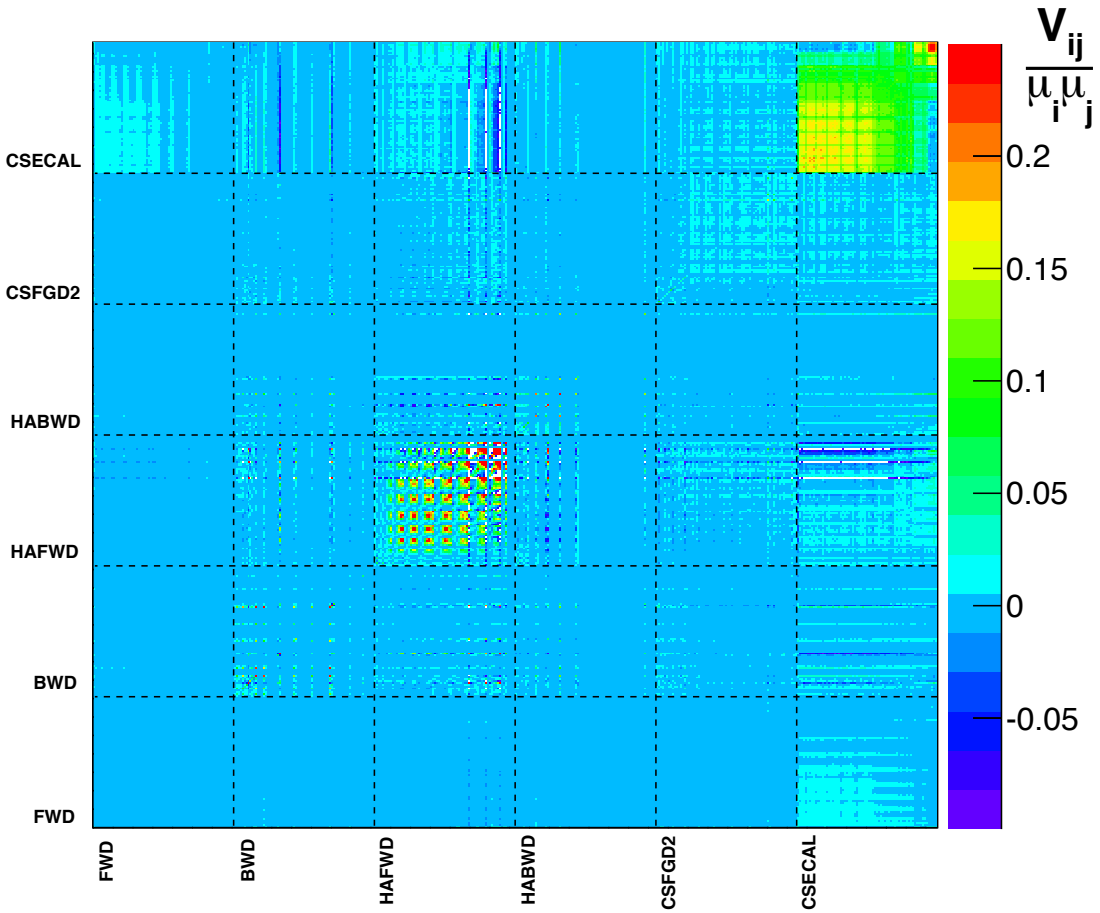

Figure 9.17: Normalized detector covariance matrix. Each bin represents the relative uncertainty in the number of selected events for certain μ^- momentum and angle. Maximum values are truncated at 0.25 for display purposes. Labels indicate where bins for a given sample start. The bins within each sample are ordered in increasing momentum intervals, each on containing all angular bins from backward to forward going.

Chapter 10

Cross Section Measurement

10.1 Cross Section Definition

This analysis presents a flux integrated double differential cross section. Assuming that the reconstruction efficiency is well modeled by the MC, Eq. 4.1 can be rewritten,

$$\frac{d\sigma}{dXdY} = \frac{N_{ij}^{Signal}}{\epsilon_{ij}^{Signal,MC}\Phi T \Delta X_i \Delta Y_j} \tag{10.1}$$

Where N_{ij}^{Signal} is the number of selected signal events with an state i and j (obtained from unfolded data, see Sec. 10.1.3). $\epsilon_{ij}^{Signal,MC}$, ΔX and ΔY are the reconstruction efficiency and bin widths for signal events predicted by the MC given the state i and j. Finally, Φ is the total integrated flux and T is the number of target nucleons.

The cross section is flux integrated in order to reduce model dependency as much as possible. Using this methodology, the cross section is not corrected by different flux contributions which depend on the relation between μ kinematics and ν energy. Consequently, the result becomes experiment dependent (no correction for the flux bin-by-bin is applied), but model independent (no assumption needs to be made on the particular ν energy distribution in each μ kinematic bin). Afterwards, cross section models can be convoluted with the T2K flux and compared with the result. A normalization uncertainty will be associated to the ν flux Φ.

10.1.1 Signal Definition

This dissertation aims to measure a double differential cross section for ν_μ CC interactions in plastic using FGD1 as a target. The results are reported using the momentum and angle (with respect to the average direction of the incoming ν) of the outgoing μ^-. However, the reconstructed variables are not an exact representation of the true initial μ^- kinematics. Therefore, an unfolding method must be used

to remove the detector dependencies away from the μ^- kinematics measurement. Several unfolding methods are available and all of them rely on MC calculations.

Reporting the result as function of variables that are directly observable (such as momentum and direction of particles) is advisable. In a first approximation, the association between their true and reconstructed values is only affected by the detector properties instead of the underlying cross section model. Nonetheless, for other variables, such as the ν energy, the association between true and reconstructed values has an strong dependency on the generators.

In this analysis, an event is defined as signal if the μ^- candidate is a true μ^- coming out from an ν_μ CC interaction happening inside the FGD1 FV (so called ν_μCC-μ, see Sec. 8.3). By requiring the μ^- candidate to be the true μ^-, it is ensured that unfolding is only removing detector dependencies. On the contrary, for ν_μCC-noμ events (see Sec. 8.3), the migration between reconstructed and true μ^- kinematics do not rely on detector dependencies but on MC theoretical model. Besides, kinematics of outgoing hadrons is not well modeled by MC generators. Consequently, it is more consistent not to unfold events in which a hadron is selected as μ^- candidate even if the event is a ν_μ CC interaction.

Meanwhile, it is important to notice that the signal efficiency will be reduced and the background will be increased by not including ν_μCC-noμ events. Therefore, some model dependencies are introduced in the analysis.

Using this signal definition, double differential cross section is written as follows:

$$\frac{d\sigma}{dp_\mu dcos\theta_\mu} = \frac{N_{ij}^{CC-\mu}}{\epsilon_{ij}^{CC-\mu,MC}\Phi N_{nucleons}^{FV}\Delta p_{\mu,i}\Delta cos\theta_{\mu,j}} \tag{10.2}$$

Where $N_{ij}^{CC-\mu}$ is the number of selected ν_μCC-μ events with momentum and angle in bin i and j respectively. In order to report a ν_μ CC inclusive cross section, reconstruction efficiency is defined as follows.

$$\epsilon_{ij}^{CC-\mu,MC} = \frac{N_{ij}^{CC-\mu,MC}}{N_{ij}^{genCC,MC}} \tag{10.3}$$

Where $N_{ij}^{CC-\mu,MC}$ is the number of selected ν_μCC-μ events in MC with momentum and angle in bin i and j respectively. While $N_{ij}^{genCC,MC}$ is the number of generated ν_μ CC events in MC.

For simplicity, p_μ and θ_μ bins i and j will be merged into one single bin in the following sections in order to simplify nomenclature in equations.

When unfolding, a relation between reconstructed and true variables for selected events is used. This is the so called transfer matrix, $M_{rt,e}^{k(s),MC}$ (which is obtained from MC). Where r and t are the reconstructed and true bins for the μ^- variables respectively. e is the true energy bin of the incoming ν, k is the true reaction type of the interaction and s is the sample in which the event is selected (see Sec. 8.5). Then efficiency can be stated as follows:

$$\epsilon_t^{CC-\mu,MC} = \frac{\sum_s^{samples} \sum_e^{E_\nu bins} \sum_r^{recobins} M_{rt,e}^{CC-\mu(s),MC}}{\sum_e^{E_\nu bins} N_{t,e}^{genCC,MC}} \tag{10.4}$$

10.1.2 Run by Run POT reweight

From run to run, detector and beam conditions change, so transfer matrix, efficiency and flux might differ. To take into account those fluctuations, a run by run POT reweight is applied as follows:

$$\hat{M}_{rt,e}^{k(s),MC} = \sum_{run} M_{rt,e,run}^{k(s),MC} POT_{run}^{DATA}/POT_{run}^{MC} \tag{10.5}$$

$$\hat{\epsilon}_t^{CC-\mu,MC} = \frac{\sum_s^{samples} \sum_e^{E_\nu bins} \sum_r^{recobins} \hat{M}_{rt,e}^{CC-\mu(s),MC}}{\sum_{run} \sum_e^{E_\nu bins} N_{t,e,run}^{genCC,MC} POT_{run}^{DATA}/POT_{run}^{MC}} \tag{10.6}$$

$$\hat{\Phi} = \sum_{run} \sum_e^{E_\nu bins} \Phi_{e,run} POT_{run}^{DATA} \tag{10.7}$$

In the following sections, when run by run POT reweight is applied to a variable, it will have the hat on top of them. Finally, using this normalized variables, eq. 10.1 is rewritten as follows:

$$\frac{d\sigma}{dX} = \frac{\hat{N}_t^{CC-\mu}}{\hat{\epsilon}_t^{CC-\mu,MC} \hat{\Phi} N_{nucleons}^{FV} \Delta X_t} \tag{10.8}$$

10.1.3 Unfolding: Maximum Likelihood Fit or Forward Fit

In order to report an experiment independent result, reconstructed kinematics from data must be unfolded. Maximum likelihood fit has shown good performance in the past (see [126]) and it will be the unfolding procedure for this analysis.

The basic idea is to vary the true spectrum of the MC and propagate its effect to the rate of events in each reconstructed bin. Then, such a rate is compared with the values from data. The variation of the true spectrum is performed scaling up or down (with a correction factor) the rate of signal events in

each true bin, as follows:

$$\hat{N}_t^{CC-\mu} = c_t \sum_s^{samples} \sum_e^{E_\nu bins} \sum_r^{recobins} \hat{M}_{rt,e}^{CC-\mu(s),MC} \tag{10.9}$$

Where t runs over the true bins of the μ^- kinematics. Those factors are obtained fitting their associated reconstructed distribution to data. The following expression is minimized:

$$\chi^2 = \sum_s^{samples} \sum_r^{recobins} 2(F_r^{(s)}(\vec{c}) - N_r^{(s),DATA} + N_r^{(s),DATA} \ln \frac{N_r^{(s),DATA}}{F_r^{(s)}(\vec{c})}) \tag{10.10}$$

On the one hand, $N_r^{(s),DATA}$ is the number of selected events in data for each reconstructed bin r and sample s. On the other hand, $F_r^{(s)}(\vec{c})$ is obtained reweighting the nominal MC prediction:

$$F_r^{(s)}(\vec{c}) = \sum_e^{E_\nu bins} \sum_t^{truebins} \left(c_t \hat{M}_{rt,e}^{CC-\mu(s),MC} + \sum_k^{Bckg} \hat{M}_{rt,e}^{k(s),MC} \right) \tag{10.11}$$

Where c_t are the correction factors, which are the free parameters of the fit with no prior constraint.

Systematic uncertainties will change the distributions of nominal MC. Therefore, efficiency, background and flux will be affected by those uncertainties. Some of them can be constraint including nuisance parameters in the fitter. Consequently, penalty terms must be included in eq. 10.10 as discussed in Sec. 10.7.

Minimization will be carried out by `MINIUT` algorithm. `MIGRAD` strategy will be used to find the minimum. However, when the minimization is very complex (many parameters), errors computed by `MIGRAD` are not absolutely reliable. Therefore, once the fit converges, `HESSE` strategy is used to compute the error matrix.

On the one hand, the main advantage of this method with respect to other unfolding techniques is that the unsmearing process is unbiased. On the other hand, this minimization process might lead to an ill-posed problem, since different sets of parameters could provide good results. In general, this would lead to a result with strong anticorrelation between nearby bins.

10.2 Sidebands

Sidebands are samples in which the background is enhanced. They are used to estimate the background contribution in the signal samples, so the background subtraction is less dependent on the MC prediction.

This analysis strongly rely on the TPC PID, which is very similar for $\mu^{\pm}$ and $\pi^{\pm}$ along the momentum spectra (see Fig. 6.11). Therefore, the main contamination comes from NC interactions in which a π^{-} is emitted or CCDIS interactions in which a π^{-} is more energetic than the outgoing μ^{-} (see Table 8.6).

Events rejected by either FGD2 or ECal PID criteria (see Sec. 8.5.1) provide two control samples that can be candidate for sidebands: CSFGD2 and CSECAL. Their main contribution are π^{-} (70%) coming out from either NC or CCDIS interactions (see Table 10.1). Moreover, the percentage of signal events is not dominant.

	CSFGD2		CSECAL	
True topology	NEUT	GENIE	NEUT	GENIE
ν_μCC-μ	**12.0**	**11.2**	**20.7**	**18.2**
CCQE	28.9	31.3	31.7	32.5
2p2h	8.6	0.0	6.5	0.0
RES	41.4	50.2	26.5	37.4
DIS	16.5	17.6	32.4	28.7
COH	4.5	0.9	2.8	1.4
ν_μCC-noμ	**24.4**	**21.4**	**20.2**	**17.0**
CCQE	0.8	1.0	1.3	1.3
2p2h	0.2	0.0	0.0	0.0
RES	5.1	5.8	4.5	6.9
DIS	93.6	92.9	94.2	91.8
COH	0.4	0.3	0.0	0.0
noν_μCC	**35.6**	**40.3**	**37.5**	**43.9**
NC	77.4	78.7	76.4	76.5
$\bar{\nu}_\mu$	16.0	14.6	14.6	15.1
$\nu_e, \bar{\nu}_e$	6.6	6.7	9.1	8.5
Out of FGD1 FV	**22.9**	**21.8**	**20.0**	**19.3**
DIS	48.4	43.2	53.1	45.7
NC	26.4	25.8	31.7	29.8
Other	25.2	31.0	15.2	24.5
Sand μ	**5.1**	**5.3**	**1.6**	**1.7**

Table 10.1: μ^{-} candidate composition in both sidebands using the topologies defined in Sec. 8.3 from both event generators in bold. True reaction composition for each topology is shown as plain text.

Fig. 10.1 shows the reconstructed kinematics for the μ^{-} candidate in each of the sidebands. Fig. 10.2 shows that the phase space covered by the sidebands is similar to the background prediction in the

selection, although the acceptance of the sidebands in the backward region is negligible.

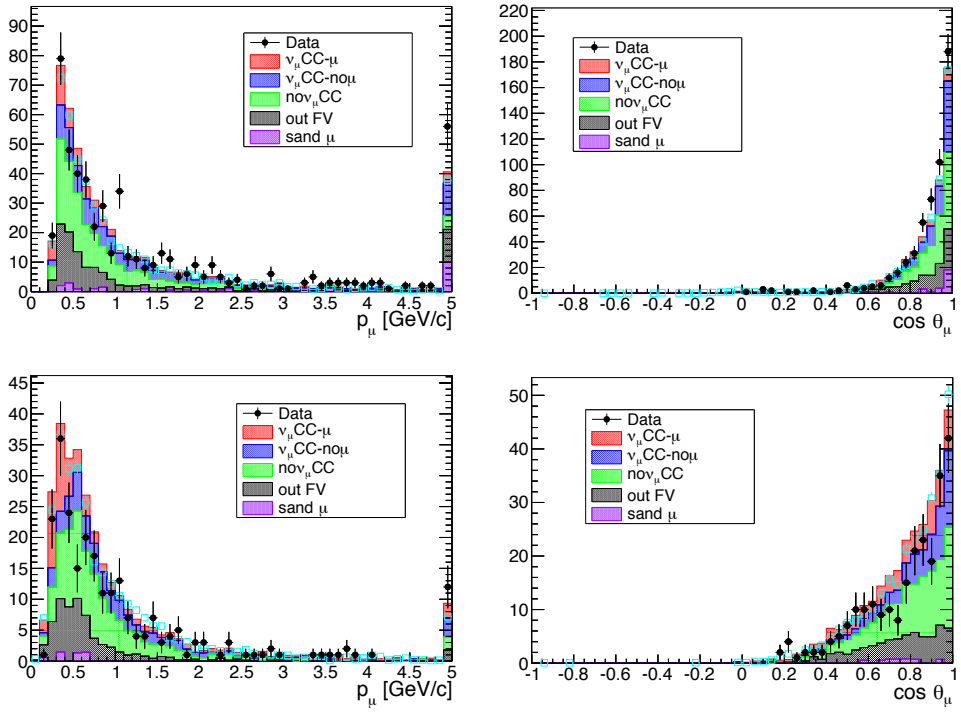

Figure 10.1: Reconstructed momentum (left column) and cosine of emission angle (right column) for the μ^- candidate when selection criteria are fulfilled in the CSFGD2 (first row) and CSECAL (second row) sideband selection. Colors indicate different topologies of interaction using `NEUT` as generator. Light blue markers show the same distribution using `GENIE` as generator.

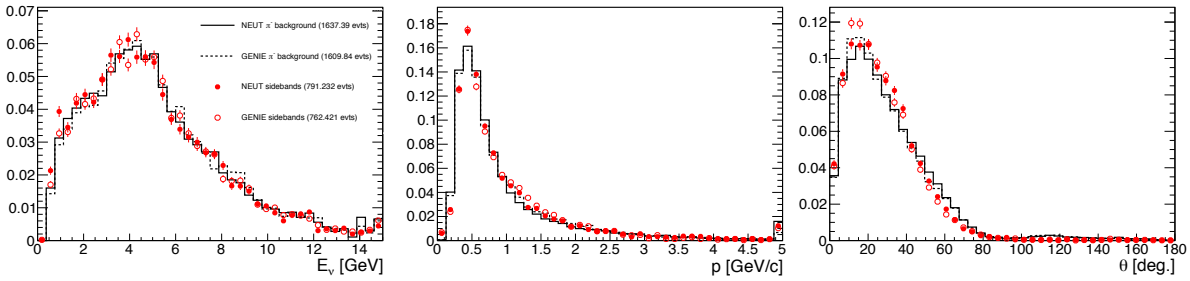

Figure 10.2: True ν energy (left), momentum (middle) and emission angle (right) for the μ^- candidate in CSECAL+CSFGD2 (dots) and for the π^- background in FWD+BWD+HAFWD+HABWD (lines) predicted by both `NEUT` and `GENIE`. The histograms are area normalized to one (number of events are shown in the legend).

In principle, such sidebands could constrain certain model parameters that affect the background prediction, basically, π FSI and NC interactions. In order to do so, the acceptance of the sideband has to match the relevant kinematical regions. It is not enough to match the phase space between the background prediction of the selection and the sidebands. Therefore, when using those sidebands, it is

important to check in which regions of the phase space π FSI and NC interactions are constrained.

10.3 Binning selection

Cross section will be reported as function of p_μ and $\cos\theta_\mu$. The 4π angular acceptance is the main motivation for the binning choice. Besides, in order to avoid a strong anticorrelation between nearby bins, the binning choice becomes extremely important. The aim is to balance the number of selected events per bin and the resolution of the detector.

In order to avoid big fluctuations in the statistical uncertainties, more than 200 selected events per bin are required, trying to keep smooth distributions.

As it was shown in Sec. 8.5, the different samples are well separated in the angular phase space. In fact, the detector response is different for the selected events in each sample. Thus, the angular binning (i.e. $\cos\theta_\mu$) has been chosen to separate as much as possible such samples.

On the one hand, statistics is limited for backward and high angle regions ($cos\theta_\mu < 0.71$), so bins are wider. Besides, BWD selection is affected by vertex migration (see App. D), which produces a bias when reconstructing both momentum and angle. The bias is produced by the outgoing hadrons coming in the opposite direction of the μ^-. Current models are not precise at predicting the kinematics of the outgoing hadrons, so the bias might be very different between data and simulators. Therefore, one single bin in both momentum and angle will be used for those events, so that vertex migration impact is reduced.

On the other hand, forward regions are very populated so narrower bins are used ($cos\theta_\mu > 0.71$). This region is mainly populated by FWD events, for which the angular resolution is optimal (see Fig. 10.3). In addition, big fluctuations between neighbor bins are avoided in order to reduce strong anticorrelation between them.

After angular division is performed, momentum bins are chosen. Momentum resolution is not as good as angular resolution, thus a very narrow binning is not recommended (lower than 100 MeV). Furthermore, the momentum resolution is very different for low and high angle cases (extracted from curvature and range respectively) and it has a clear dependency on the momentum. Therefore, bin width greater than momentum resolution has been the main binning criteria (see Fig. 10.4).

Finally, Fig. 10.5 shows the number of reconstructed events in each bin for the true and reconstructed kinematic of the μ^- candidate. As it acn be observed, the goal of 200 selected events per bin is achieved.

In summary, the contributions from each sample can disentangled with this binning. Besides, special care has been taken in order to avoid large migrations from bin to bin in the momentum phase space. Fig. 10.6 shows the transfer matrices obtained from the nominal MC using `NEUT` as generator.

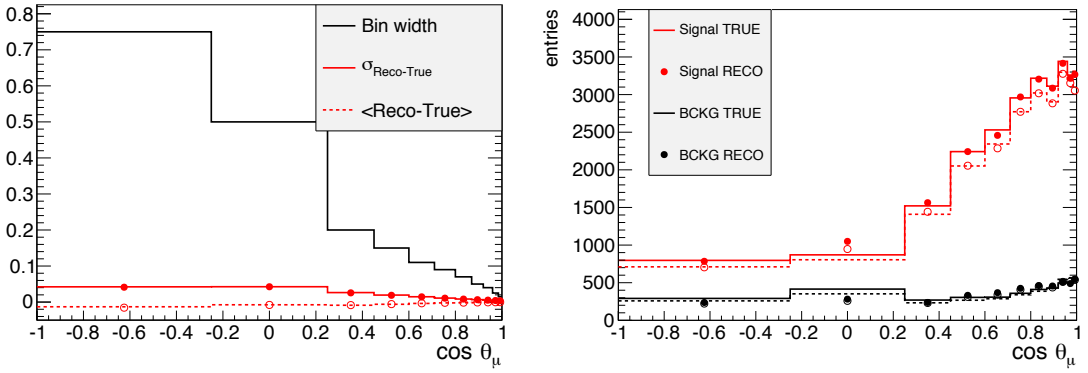

Figure 10.3: Left: Angular resolution and bias per each angular bin. Lines (dots) show the distributions using `NEUT` (`GENIE`) as generator. The width of each bin is also shown (black lines). Right: Reconstructed (points) and true (lines) cosine of the angle for muon candidate and true muon respectively for signal and background events using the optimized binning with `NEUT` as generator. Empty dots and dashed lines show the same distributions using `GENIE` as generator.

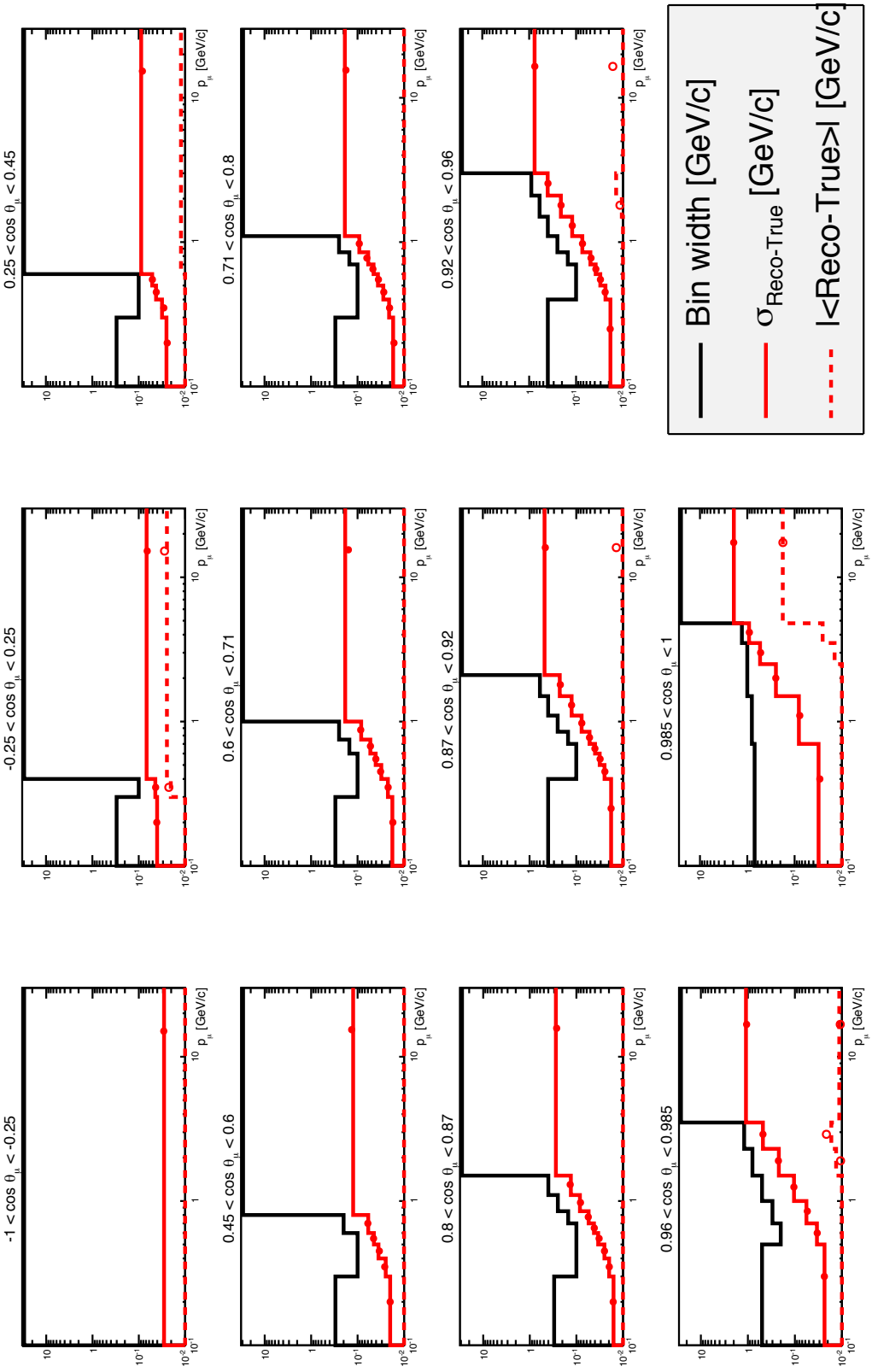

Figure 10.4: Momentum resolution and bias per each angular bin using NEUT (red lines) and GENIE (red dots) as generator. The width of each bin is also shown (black lines).

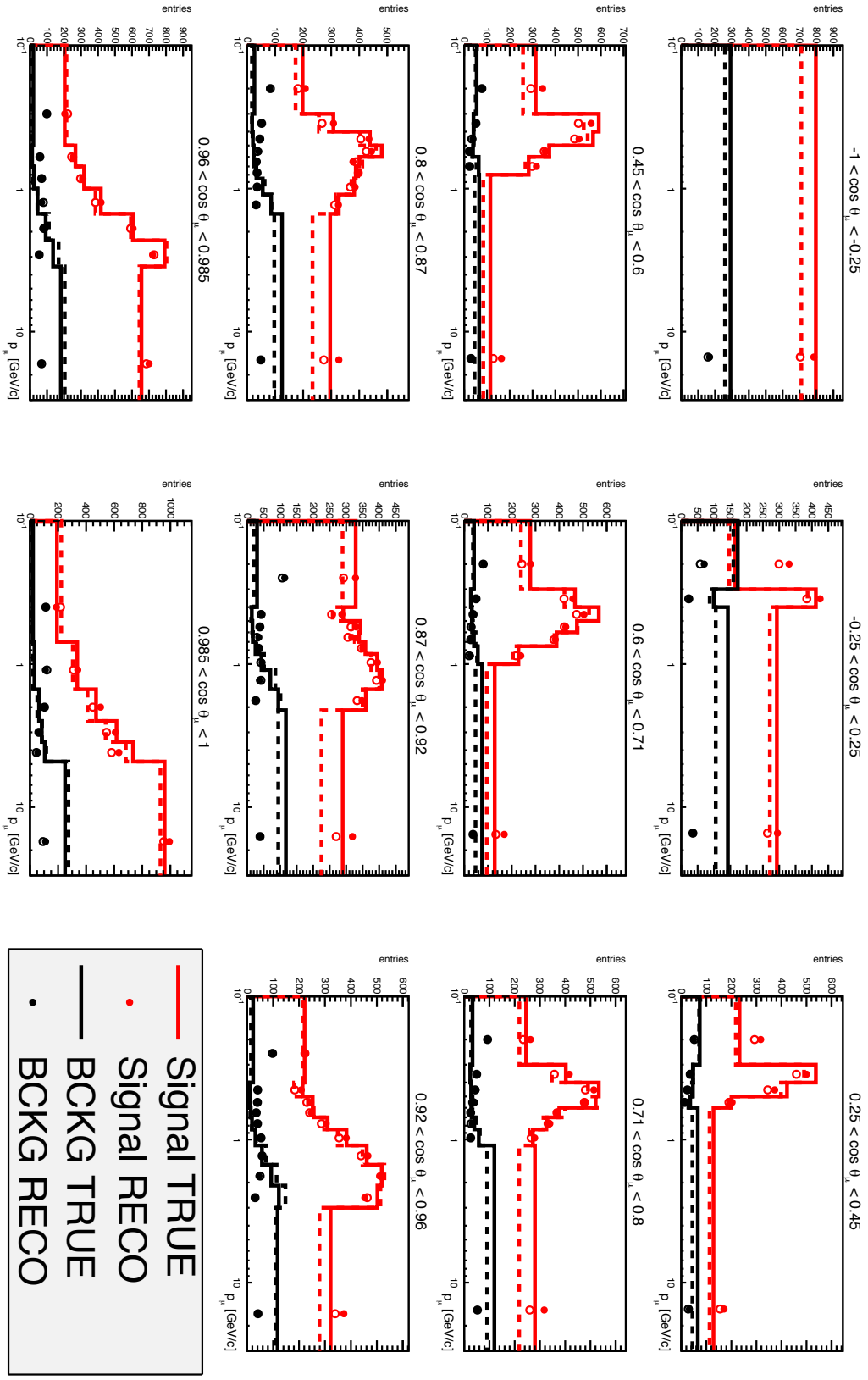

Figure 10.5: Reconstructed (points) and true (stacked) momentum of the μ^- candidate and true μ^- respectively using the optimized binning with NEUT as generator. Empty dots and dashed lines show the same distributions using GENIE as generator.

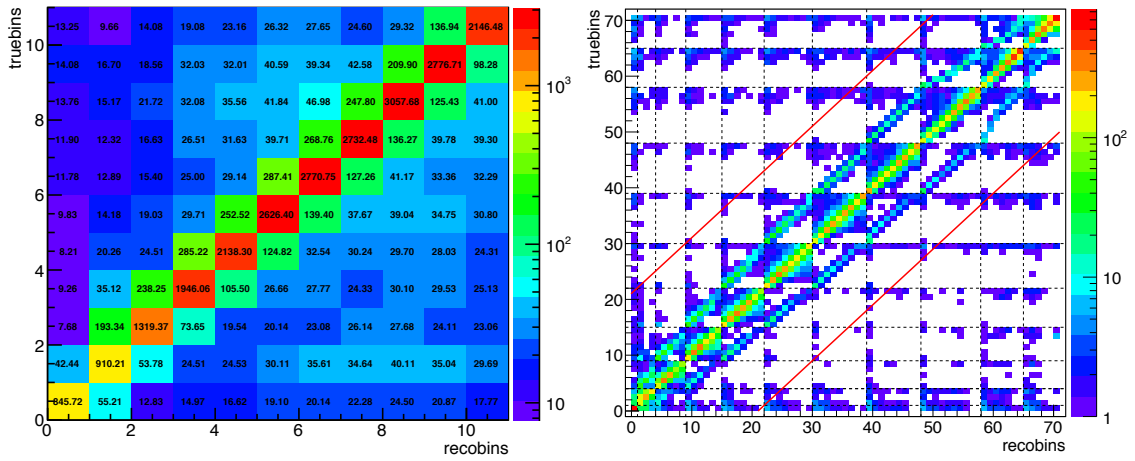

Figure 10.6: Left: Transfer matrix for angular bins obtained from nominal MC summing the contribution from all the samples for signal events. Each row/column corresponds to an angular bin going from $0 = [-1, -0.25]$ to $10 = [0.985, 1]$. Right: Transfer matrix (left) obtained from nominal MC using `NEUT` as generator summing the contribution from all the samples for signal events. Red lines define the region for which response functions are computed (see Sec. 10.7.5). Dashed lines indicate the angular boundaries. Each row/column represents a bin going from $0 = [-1 < \cos\theta_\mu < -0.25, 0\text{GeV/c} < p_\mu < 30\text{GeV/c}]$ to $70 = [0.985 < \cos\theta_\mu < 1, 4.8\text{GeV/c} < p_\mu < 30\text{GeV/c}]$.

10.4 Efficiency

As it was explained in Sec. 10.1, reconstruction efficiency plays a key role in cross section extraction. When the efficiency correction is performed as a function of a single variable, for instance the momentum, the limited acceptance of the detector becomes critical. The main problem is that a single bin of momentum contains contributions from events with different directions, specifically very different efficiency corrections. Then, the acceptance in a single bin depends on the relative number of events with a μ^- going forward or backward. Nonetheless, the efficiency corrections as function of the momentum and direction of the outgoing μ^- are mostly model-independent, since these are the actual variables measured in the detector.

The total reconstruction efficiency is shown in Fig. 10.7. Some remarks can be extracted:

- For low momentum μ^- (below $\sim$500 MeV/c) efficiency drops drastically because it is very likely that they do not exit the FGD, so they are not reconstructed.

- For high angle μ^- the main limitation arises from the stopping requirement in the event selection (see Sec. 8.5.3). Such a criteria is used to compute the momentum using the track length. Consequently, μ^- not stopping either in BarrelECal or SMRD are not reconstructed.

- The timing requirement is used to tag μ^- going backward (see Sec. C.2). This means that the μ^- must reach P0D or BarrelECal detectors and produce a long enough pattern of hits, so a segment is reconstructed. However, the energy of the μ^- going backward is rather low, so it is very likely that they stop in the detector passive edges.

Fig. 10.8 shows reconstruction efficiency for signal events using the binning presented in Sec. 10.3. In Appendix E, the efficiency is split into different reaction types. From this efficiencies, it is possible to extract information from particular regions in the momentum plus angle phase space:

- The fact of not including ν_μCC-noμ as signal reduces the efficiency for CCDIS events (see Fig. E.4). In those events it is very likely to have a π^- coming out from the interaction with similar or higher momentum than the μ^-.

- It is important to emphasize the discrepancies between both generators in the forward region at low momentum. Those differences are observed using both signal definitions. The reason is that CCDIS and CCRES modeling in the 2 GeV region (so called multi pi channel) is different between both generators.

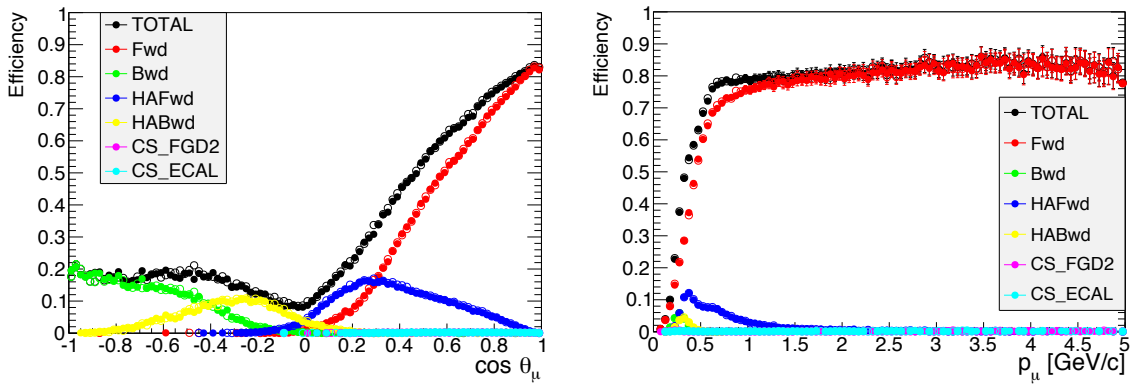

Figure 10.7: Cosine of emission angle (left) and momentum (right) efficiency for ν_μCC-μ using `NEUT` (dots) and `GENIE` (empty dots) as generator. Colors indicate contribution from different selections.

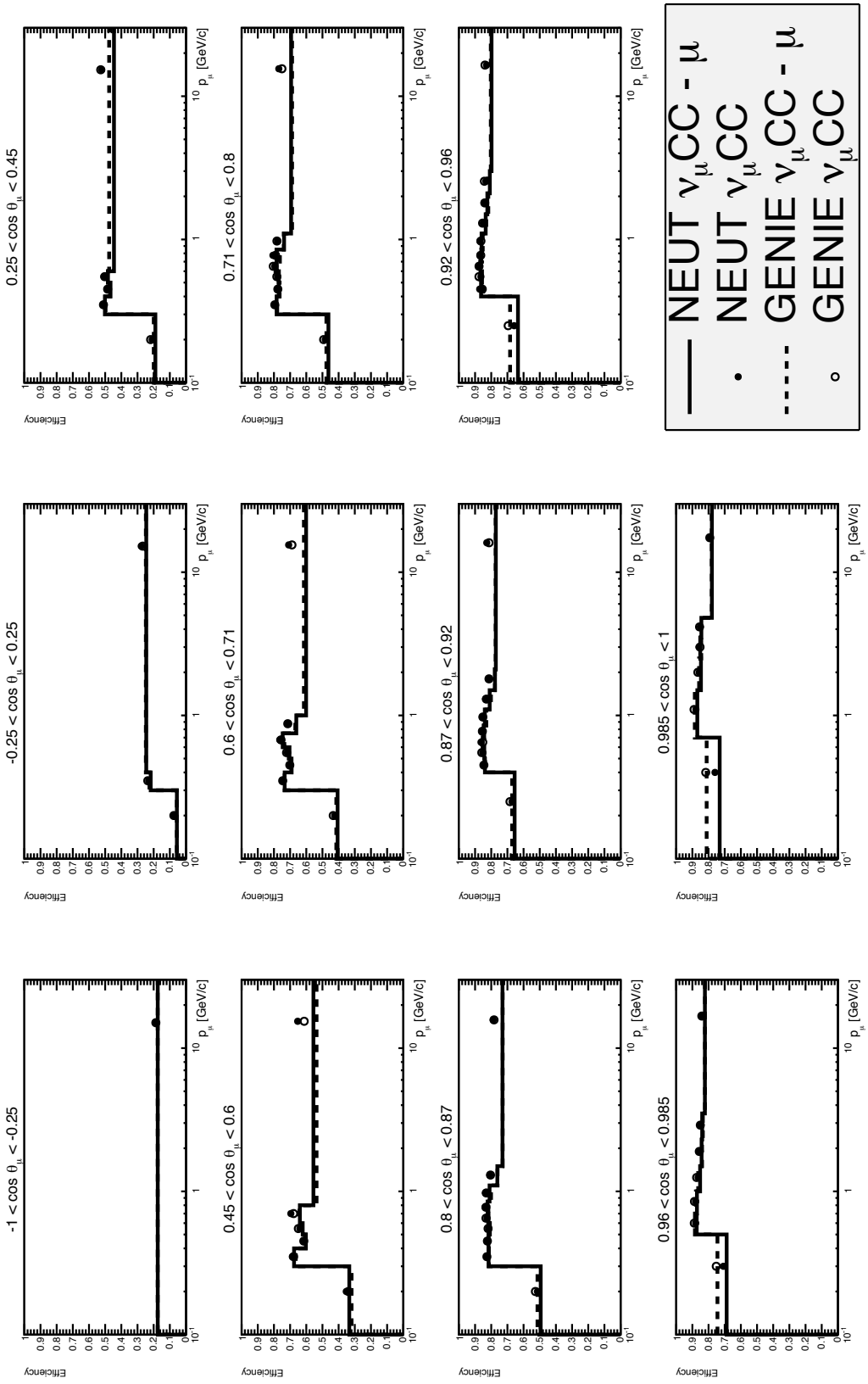

Figure 10.8: Reconstruction signal efficiency as function of the momentum and angle of the true μ^-.

10.5 Integrated Flux

This analysis reports on a flux-integrated cross section. Beam group provides the predicted flux used per each run at near detector (see Sec. 5.5) integrating over a 150×150 cm^2 centered in the basket plane (see Fig. 10.9).

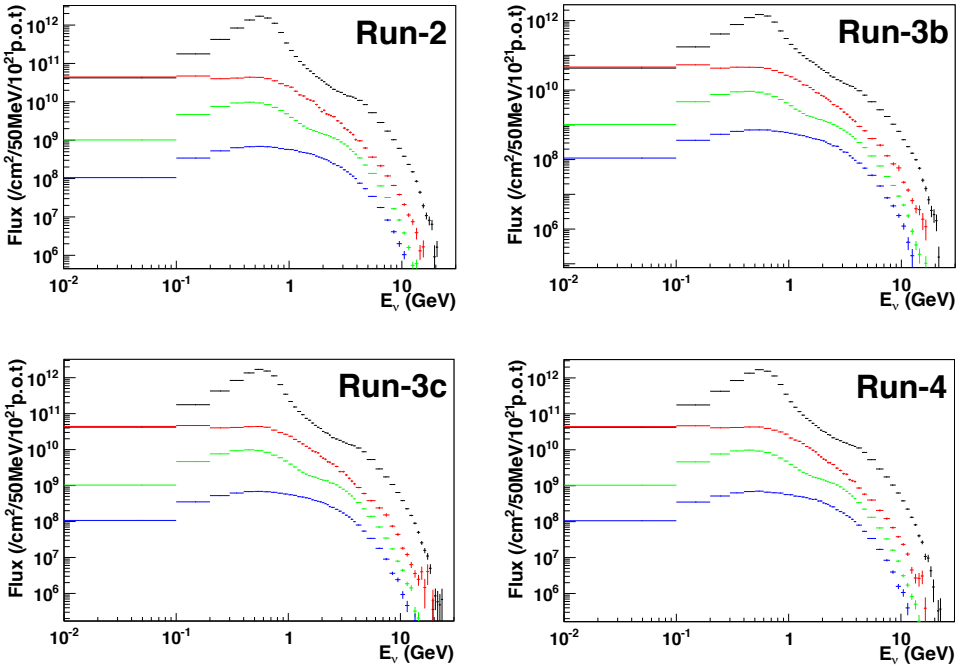

Figure 10.9: Run by run fine binned flux at the near detector for ν_μ (black), $\bar{\nu_\mu}$ (red), ν_e (green), $\bar{\nu_e}$ (blue) using flux 13a v1.1.

The flux prediction distinguishes between different ν flavors, but it main contribution comes from ν_μ. The integrated flux in each run is summarizes in Table 10.2. It is computed using Eq. 10.7 from Sec. 10.1, taking into account only the ν_μ fraction.

Run	POT ($\times 10^{21}$)	Φ ($\times 10^{13}$/cm^2/10^{21}POT)	$\hat{\Phi}$ ($\times 10^{13}$/cm^{-2})
2-Water In	0.042858	1.926	0.0825
2-Water Out	0.0355096	1.926	0.0684
3b-Water Out	0.02146	1.738	0.0373
3c-Water Out	0.134779	1.935	0.2608
4-Water In	0.162699	1.940	0.3156
4-Water Out	0.176246	1.940	0.3419
Total	0.5735515	-	1.1066

Table 10.2: Number of POT (first column), integrated ν_μ flux (second column) and run by run POT reweighted flux from Eq. 10.7 (third column) for each data set of collected data.

10.6 Number of Targets

For ν_μ CC interactions, both protons and neutrons are assumed to be the targets [1]. To extract the numbers of nucleons within a certain volume the following formula can be used:

$$T = N_A \Delta X_{FV} \Delta Y_{FV} \sum_{a}^{elements} \rho_a \Delta Z_a \frac{A_a}{M_a} \tag{10.12}$$

Where $N_A = 6.022 \times 10^{23}$ mol^{-1} is the Avogadro number, $\Delta X(Y)_{FV}$ is the total length in X(Y) axis of the fiducial volume, A_a and M_a are the average number of nucleons and atomic mass per element a respectively. Finally, the product $\rho_a \Delta Z_a$ is the areal density per element a in the fiducial volume.

In this analysis FGD1 is used as target, which is mainly composed of Carbon. Table 10.3 explains in detail the composition of the target showing its main elements.

Element	A	Natural abundance	A_a	FGD1 Fraction	$\rho_a \Delta Z_a$ [g/cm^2]
C	12	98.9	12.011	86.13	1.849
	13	1.1			
H	1	99.985	1.00015	7.36	0.1579
	2	0.015			
O	16	99.762	16.0044	3.7	0.0794
	17	0.038			
	18	0.2			
Ti	46	8.	48.024	1.65	0.0355
	47	7.5			
	48	73.8			
	49	5.5			
	50	5.4			
Si	28	92.22	28.1058	1.01	0.0218
	29	4.68			
	30	3.09			
N	14	99.634	14.00366	0.14	0.0031
	14	0.366			

Table 10.3: For each element of FGD1 fiducial volume, information used to compute the total number of nucleons.

From Sec. 8.2, the fiducial volume takes into account all 15 XY modules of FGD1 ($\Delta Z_{FV} = 15 \times 2.025 + 14 \times 0.2 = 33.175$ cm). In the X and Y dimension, a distance equivalent to five bars on each side is removed ($\Delta X_{FV} = \Delta Y_{FV} = 186.434 - 2 \times 5.766 = 174.902$ cm).

Using all the information provided above, the number of nucleon targets is 5.9×10^{29}.

[1] For inclusive CC interactions the ν can only interact with any nucleon

10.7 Error Propagation

Analytical computation for most uncertainties is not possible, so toy experiments are used to study their impact. The followed methodology reweight the nominal MC for each toy experiment. Then, the reconstructed distribution from the data is unfolded (forward fitted) using the transfer matrix from such reweighted MC. Besides, efficiency, flux and number of targets will also be computed using the reweighted MC.

The cross section per bin t for each toy experiment x is defined as follows:

$$\left(\frac{d\sigma}{dX_t}\right)^{FIT}_{(x)} = \frac{\hat{N}^{CC-\mu,FIT}_{t(x)}}{\hat{\epsilon}^{CC-\mu,FIT/PRIOR}_{t(x)}\hat{\Phi}^{PRIOR}_{(x)}N^{FV,PRIOR}_{nucleons(x)}\Delta X_t} \tag{10.13}$$

Where all factors with the "prior" label are computed using the reweighted MC for each toy experiment. In the case of the efficiency, it will depends on whether the fit does include nuisance parameters ($\epsilon^{CC-\mu,FIT}_{t(x)}$) or it does not ($\epsilon^{CC-\mu,PRIOR}_{t(x)}$).

The distribution of the cross section defined in Eq. 10.13 for many toy experiments is expected to be Gaussian. The mean and width of such distribution are used to define the average cross section ($\left\langle \frac{d\sigma}{dX_t} \right\rangle$) and its associated uncertainty ($\Delta\frac{d\sigma}{dX_t}$) per each bin t respectively. In this case, the correlation across bins is computed using the following formula:

$$\Sigma_{ij} = \frac{1}{\Delta\frac{d\sigma}{dX_i}\Delta\frac{d\sigma}{dX_j}}\frac{1}{N_{toys}}\sum_{x=0}^{N_{toys}}\left(\left(\frac{d\sigma}{dX_i}\right)^{FIT}_{(x)} - \left\langle\left(\frac{d\sigma}{dX_i}\right)\right\rangle\right)\left(\left(\frac{d\sigma}{dX_j}\right)^{FIT}_{(x)} - \left\langle\left(\frac{d\sigma}{dX_j}\right)\right\rangle\right) \tag{10.14}$$

As it was explained previously, in this methodology nominal MC must be reweighted to obtain the prior distributions. The reweighting procedure varies depending on the source of uncertainty. Besides, certain systematics can be constrained adding nuisance parameters into the fitter.

In the following sections, the reweighting procedure for each source of uncertainty will be described.

10.7.1 Data Statistics

To compute the uncertainty due to data statistics, a set of toys is produced applying a Poissonian fluctuation to the number of reconstructed events in data for each bin and sample independently.

Then, the reconstructed distribution per each toy x can be expressed as follows:

$$N_{r(x)}^{(s),FAKE} = \text{POISSON}\left[N_r^{(s),DATA}\right] \tag{10.15}$$

This fake data distribution is unfolded using minimization from Sec. 10.1.3, in which the only free parameters are the correction factors c_t. The transfer matrix used in the unfolding is extracted from the nominal MC. Thus, the fitted distribution is computed with the following formula:

$$\hat{N}_{t(x)}^{CC-\mu,FIT} = c_{t(x)} \sum_{s}^{samples} \sum_{k}^{reactions} \sum_{e}^{E_\nu bins} \sum_{r}^{recobins} \hat{M}_{rt,e}^{k(s),MC} \tag{10.16}$$

For this particular case, true distributions (transfer matrix and efficiency) remain unchanged for all the toy experiments (i.e. as they are extracted from the nominal MC). Therefore, this uncertainty will only affect the number of signal events.

10.7.2 MC Statistics

To compute the uncertainty due to MC statistics, a set of toys is produced applying a Poissonian fluctuation to the generated events by nominal MC (without normalizing to the collected POT in data). Therefore, this statistical fluctuations will affect the transfer matrix and efficiency predicted by nominal MC. Consequently, per each toy experiment x, those quantities are recomputed:

$$\hat{M}_{rt,e(x)}^{k(s),PRIOR} = \text{POISSON}_{\text{NOPOTNORM}}\left[\hat{M}_{rt,e}^{k(s),MC}\right] \tag{10.17}$$

$$\hat{\epsilon}_{t(x)}^{CC-\mu,PRIOR} = \frac{\sum_s^{samples} \sum_k^{reactions} \sum_e^{E_\nu bins} \sum_r^{recobins} \hat{M}_{rt,e(x)}^{k(s),PRIOR}}{\text{POISSON}_{\text{NOPOTNORM}}\left[\sum_e^{E_\nu bins} \hat{N}_{t,e}^{genCC,MC}\right]} \tag{10.18}$$

Then, the reconstructed distribution from data is unfolded using minimization from Sec. 10.1, in which the only free parameters are the correction factor c_t. Such unfolding is performed using the transfer matrix extracted from the reweighted MC. Therefore, the fitted distribution is obtained from Eq. 10.16. Consequently, the uncertainty associated with the MC statistics affects both the number of signal events and the efficiency.

10.7.3 Flux

For this analysis, the number of no-ν_μ selected events is below 1%, thus fluctuations of no-ν_μ flux can be neglected. Beam group provides a covariance matrix that summarizes all source of uncertainty coming from the flux knowledge in the bins of the incoming ν beam (see Fig 9.2).

The workflow per each toy experiment x is as follows. Firstly, the covariance matrix is decomposed using the Cholesky method (function `TDecompChol::Decompose()` can be found in ROOT [143]), so random variables are generated taking into account the correlations between them. Secondly, using the decomposed matrix, a set of random weights ($w^{flux}_{e(x)}$) is generated throwing them using a Gaussian PDF. Then, per each toy experiment, transfer matrix, efficiency and integrated flux must be reweighted using the thrown weights:

$$\hat{M}^{k(s),PRIOR}_{rt,e(x)} = w^{flux}_{e(x)} \hat{M}^{k(s),MC}_{rt,e} \tag{10.19}$$

$$\hat{\epsilon}^{CC-\mu,PRIOR}_{t(x)} = \frac{\sum_s^{samples} \sum_e^{E_\nu bins} \sum_r^{recobins} w^{flux}_{e(x)} \hat{M}^{CC-\mu(s),MC}_{rt,e}}{\sum_e^{E_\nu bins} w^{flux}_{e(x)} \hat{N}^{genCC,MC}_{t,e}} \tag{10.20}$$

$$\hat{\Phi}^{PRIOR}_{(x)} = \sum_e^{E_\nu bins} w^{flux}_{e(x)} \hat{\Phi}^{MC}_e \tag{10.21}$$

Then, the reconstructed distribution from data is unfolded using minimization from Sec. 10.1 with the transfer matrix extracted from the reweighted MC. In this case, the only free parameters are the correction factor c_t. Finally, the fitted distribution is obtained from Eq. 10.16 using the reweighted transfer matrix.

Consequently, the flux fluctuations will affect the number of selected events, efficiency and integrated flux. In fact, a relatively flat distribution will be observed for the relative uncertainty among bins. The reason is that such error is mainly affected by integrated flux factor Φ (around 8.7%).

10.7.4 Mass

The number of nucleons within the fiducial volume is known (see Sec. 10.6) and its associated error is computed in the section. The source of this uncertainty is related to the areal density measurement of the elements composing the FGD. The uncertainty of the areal density of each element and their correlations are obtained from Fig. 10.10.

The workflow per each toy experiment x is the following: Firstly, the covariance matrix is decom-

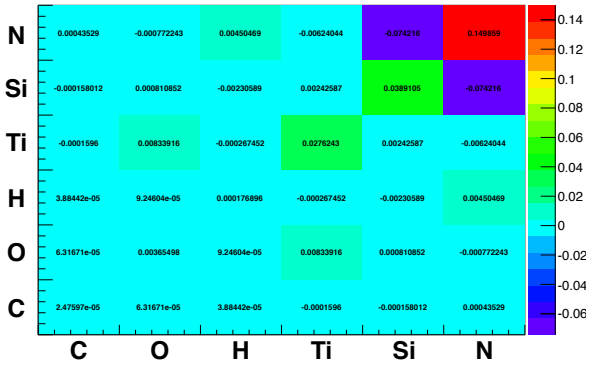

Figure 10.10: Mass covariance matrix.

posed using the Cholesky method. Secondly, using the decomposed matrix, a set of random weights ($w^{mass}_{element(x)}$) is generated throwing them using a Gaussian PDF. Then, per each toy experiment, the areal density of each element is reweighted using the thrown weights. Consequently, the number of nucleons in the FGD FV will vary for each toy experiment.

A relative uncertainty of 0.67% is obtained, which is flat among the bins.

10.7.5 Modeling

Current theoretical models and their associated parameters have some limitations reproducing data (see Sec. 9.2). Therefore, efficiency and background predictions from nominal MC must have an associated uncertainty. `T2KReWeight` package allows to estimate the impact of the theoretical parameters variation (so called dials).

Prior to the toy experiments, response functions must be generated in order to have a direct dial-kinematics relation. Different response functions are generated:

- $w(a)^{k(s)}_{rt}$ ratio between the number of selected events by varying dial a with respect to the nominal MC for each sample s, reaction type k, reconstructed bin r and true bin t.
- $w(a)^{(s)}_{t}$ ratio between the number of selected signal events by varying dial a with respect to the nominal MC for each sample s and true bin t.
- $w(a)_t$ ratio between the number of generated signal events by varying a with respect to the nominal MC for each true bin t.

Once those response functions are generated, toy experiments are produced and the following workflow is followed for each toy x. Firstly, the covariance matrix is decomposed using the Cholesky method. Secondly, using the decomposed matrix, the set of dials are thrown $a_{(x)}$ using a Gaussian PDF. This procedure is repeated for each toy. Finally, transfer matrix and efficiency are reweighted using the response functions evaluated with thrown dial values:

$$\hat{M}^{k(s),PRIOR}_{rt,e(x)} = \prod_{a}^{dials} w(a_{(x)})^{k(s)}_{rt} \hat{M}^{k(s),MC}_{rt,e} \tag{10.22}$$

$$\hat{\epsilon}^{CC-\mu,PRIOR}_{t(x)} = \frac{\sum_{s}^{samples}\sum_{e}^{E_\nu bins}\sum_{r}^{recobins}\prod_{a}^{dials} w(a_{(x)})^{(s)}_{t}\hat{M}^{CC-\mu(s),MC}_{rt,e}}{\sum_{e}^{E_\nu bins}\prod_{a}^{dials} w(a_{(x)})_t \hat{N}^{genCC,MC}_{t,e}} \tag{10.23}$$

Then, the reconstructed distribution from data is unfolded using minimization from Sec. 10.1.3 with the transfer matrix extracted from the reweighted MC. However, certain systematics might be constrained using sidebands (see Sec. 10.2). This is done including nuisance parameters in the fit and their correspondence penalty term in the minimization formula:

$$\begin{aligned}\chi^2_{(x)} = \sum_{s}^{samples}\sum_{r}^{recobins} & 2(F^{(s)}_r(\vec{c}_{fit(x)},\vec{a}_{fit(x)}) - N^{(s),DATA}_r + N^{(s),DATA}_r \ln\frac{N^{(s),DATA}_r}{F^{(s)}_r(\vec{c}_{(x)},\vec{a}_{fit(x)})}) + \\ & + (\vec{a}_{fit(x)} - \vec{a}_{(x)}) V^{-1}_{cov}(\vec{a}_{fit(x)} - \vec{a}_{(x)})\end{aligned} \tag{10.24}$$

Where, $\vec{a}_{fit(x)}$ are the free nuisance parameters, $\vec{a}_{(x)}$ their prior value per each toy and V_{cov} the corresponding covariance matrix. Besides, $F^{(s)}_r(\vec{c_{(x)}},\vec{a}_{fit(x)})$ will be affected by these parameters.

$$\begin{aligned}F^{(s)}_r(\vec{c_{(x)}},\vec{a}_{fit(x)}) = \sum_{e}^{E_\nu bins}\sum_{t}^{truebins}\Bigg(c_{t(x)}\prod_{p}^{nuisances} w(a_{fit(x)})^{CC-\mu(s)}_{rt}\hat{M}^{CC-\mu(s),PRIOR}_{rt,e(x)} + \\ + \sum_{k}^{Bckg}\prod_{p}^{nuisances} w(a_{fit(x)})^{k(s)}_{rt}\hat{M}^{k(s),PRIOR}_{rt,e(x)}\Bigg)\end{aligned} \tag{10.25}$$

Consequently, after minimization is found, transfer matrix and efficiency must be recomputed using the free parameters from the fit.

$$\hat{M}^{k(s),FIT}_{rt(x),e} = \prod_{p}^{nuisances} w(a_{fit(x)})^{k(s)}_{rt}\hat{M}^{k(s),PRIOR}_{rt,e(x)} \tag{10.26}$$

$$\hat{\epsilon}^{CC-\mu,FIT}_{t(x)} = \frac{\sum_{s}^{samples}\sum_{e}^{E_\nu bins}\sum_{r}^{recobins}\prod_{p}^{nuisances} w(a_{fit(x)})^{k(s)}_{t}\hat{M}^{CC-\mu(s),PRIOR}_{rt,e(x)}}{\sum_{e}^{E_\nu bins}\prod_{p}^{nuisances} w(a_{fit(x)})_t \hat{N}^{genCC,PRIOR}_{t,e(x)}} \tag{10.27}$$

Furthermore, unfolded selected signal events include correction factors.

$$\hat{N}^{CC-\mu,FIT}_{t(x)} = c_{t(x)} \sum_{s}^{samples} \sum_{k}^{reactions} \sum_{e}^{E_\nu bins} \sum_{r}^{recobins} \hat{M}^{k(s),FIT}_{rt,e(x)} \tag{10.28}$$

It is important to realize which uncertainties associated to the modeling can be constrained by adding nuisance parameters. The nuisance parameters associated with uncertainties that affect the distribution of signal events will be very correlated with the correction factors c_t. Therefore, only the uncertainties that affect the rate of background events rather than the signal will be constrained in order to avoid model dependencies.

As it was shown in Sec 9.2, most uncertainties are associated to CC interactions, so they cannot be constrained. Consequently, only π FSI and NC interactions will be constrained using nuisance parameters.

The sensitivity of such nuisance parameters to the variation of their associated systematic sources was tested. 500 fake data distributions were generated reweighting the nominal MC by a set of thrown weights associated to π FSI and NC interactions. Such fake data distributions were unfolded using the nominal MC including the nuisance parameters. If the nuisance parameter is sensitive, a correlation between the thrown and fitted value must be observed. Fig 10.11 shows the test result. π FSI nuisance parameters are correlated to the thrown values, although that correlation is not very large. Meanwhile, the correlation between the fitted and thrown value for the parameter associated to NC interactions is higher.

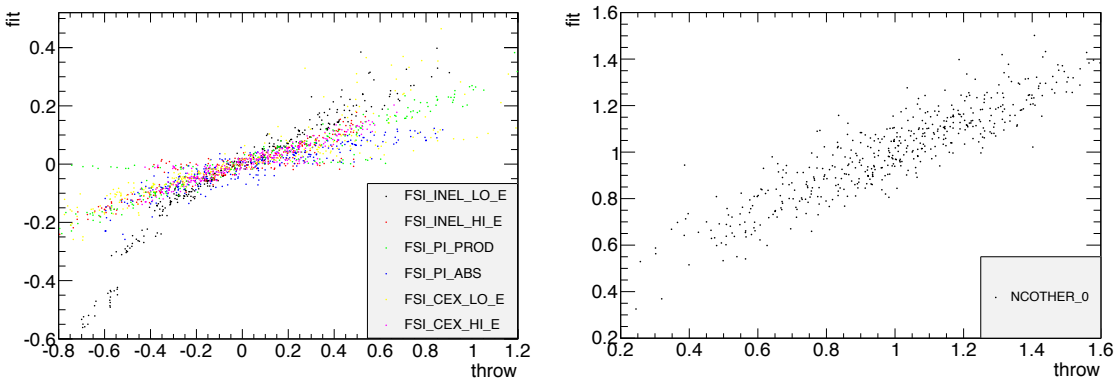

Figure 10.11: Fitted value versus thrown value for nuisance parameters associated to π FSI (left) and NC interactions (right) uncertainties for 500 toy experiments generated by applying fluctuations due to π FSI and NC systematic uncertainties respectively.

10.7.6 Total Uncertainties

In this section a summary of all sources of uncertainty is presented. 500 toys are generated applying fluctuations associated with each source of uncertainty. The relative uncertainties in each bin can be obtained dividing the width and mean of the distribution defined in Eq. 10.13.

Fig 10.12 shows relative uncertainties comparing them with the error independently extracted from each source. The dominant systematic uncertainty is the flux (except in the backward and high angle regions were cross section modeling and detector dominates respectively). Flux uncertainty is almost flat among bins because it is mainly affected by Φ factor in Eq. 10.1 (around 8.7%). Its impact in the cross section shape is much smaller. The uncertainty associated with statistics becomes dominant in the high momentum region where the number of reconstructed events is lower (except in the very forward region). It is also interesting to notice that theoretical systematic uncertainties play a secondary role in an inclusive cross section. Further studies on each source of uncertainty are described in App. F.

Besides, the correlation across bins were independently computed for each source of uncertainty using Eq. 10.14. Fig 10.13 shows that the correlations behave as expected. Firstly, anticorrelation between nearby momentum and angular bins for statistical uncertainties. Secondly, errors associated to the flux are almost fully correlated across the whole phase space. Thirdly, correlations of detector systematic uncertainties in nearby bins.

It is also interesting to analyze the nuisance and sidebands impact on the total uncertainty. Three different scenarios have been compared: using sidebands with and without nuisance parameters and not using sidebands without nuisance parameters. Fig 10.14 shows relative uncertainties in the three scenarios. It can be concluded that nuisance parameters slightly reduce the overall uncertainty, but they do not play a major role in total uncertainty.

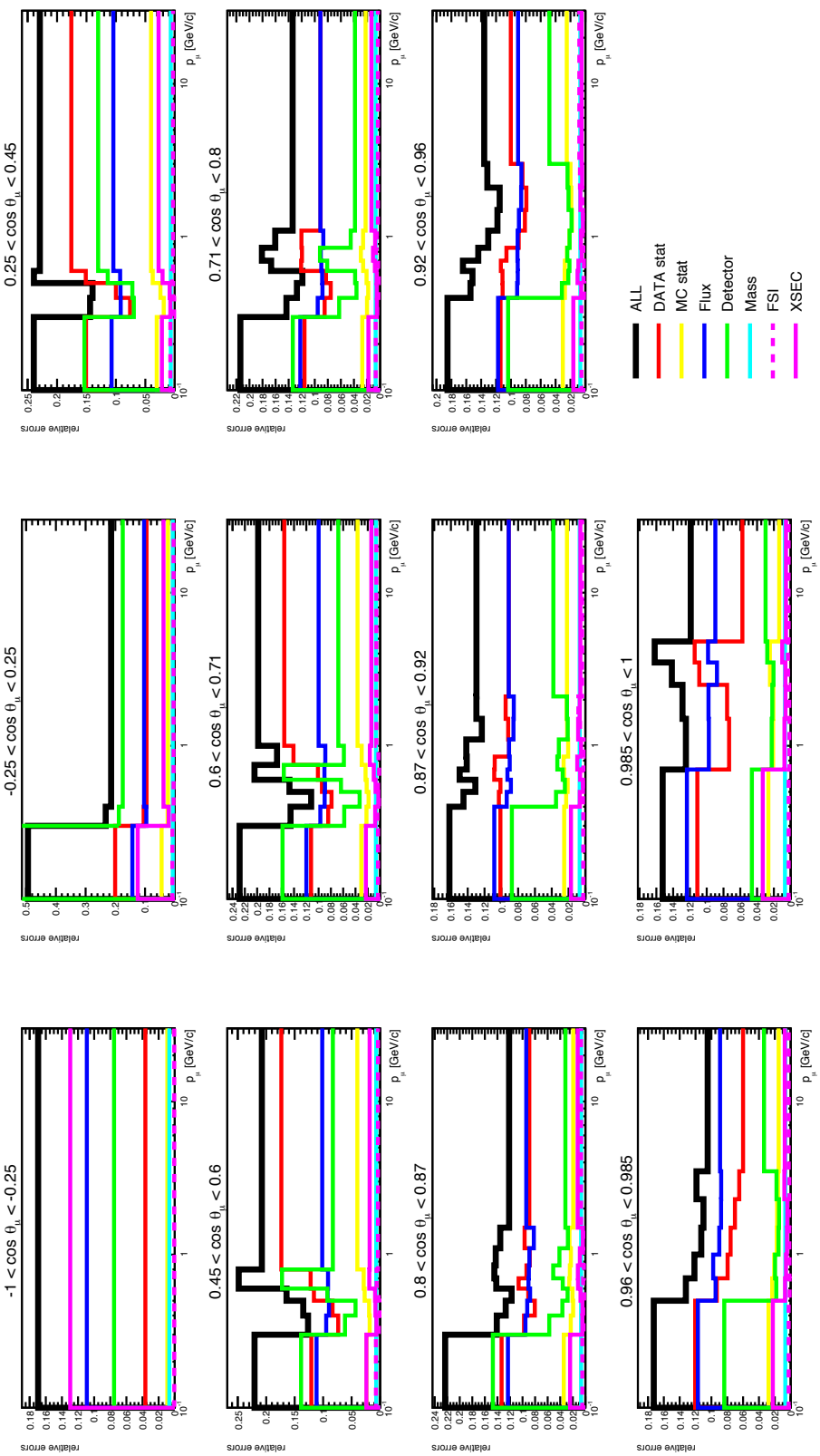

Figure 10.12: Relative total uncertainty for cross section when nuisance parameters are included for π FSI and NC interactions.

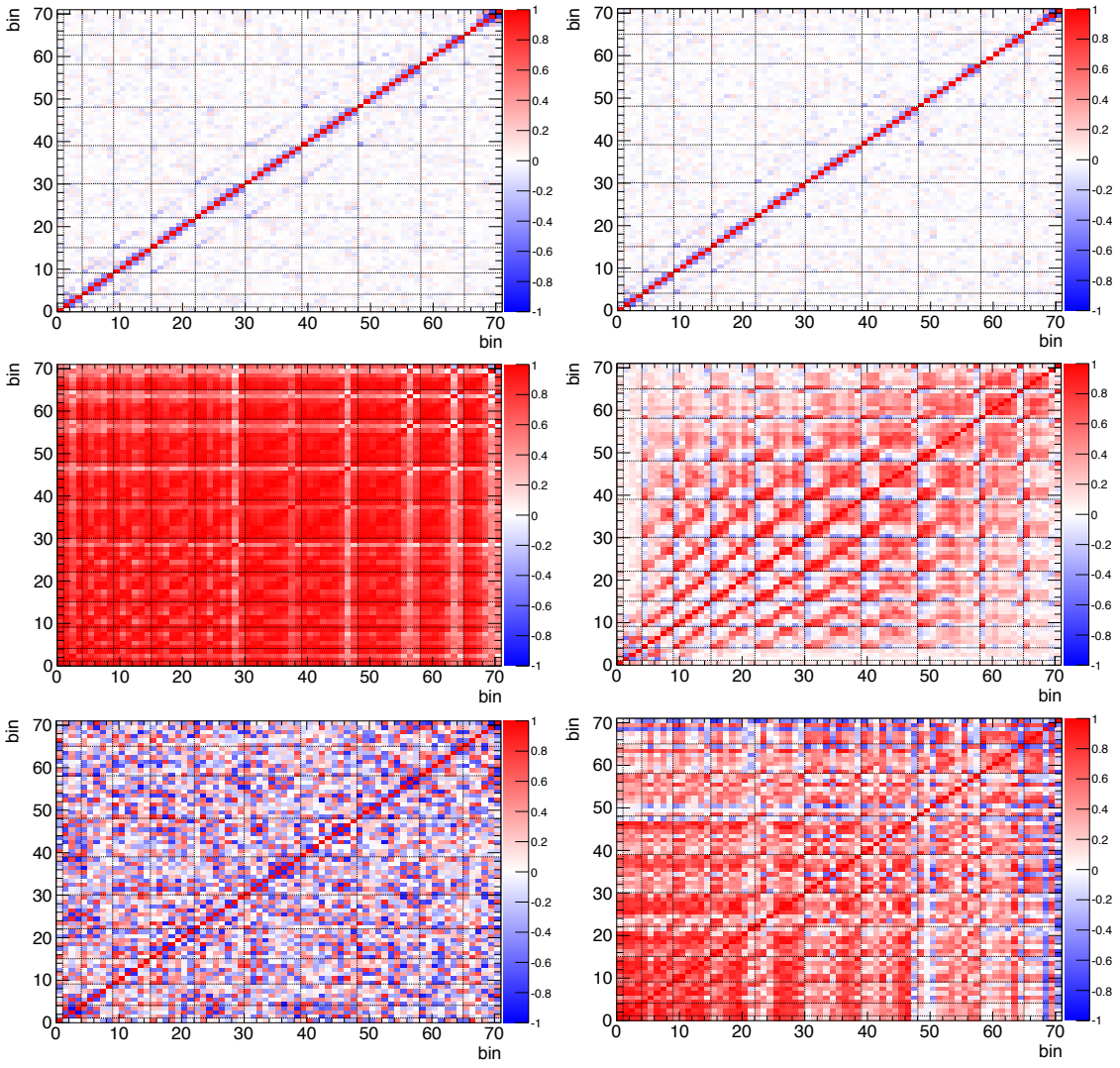

Figure 10.13: Correlation matrix obtained from Eq. 10.14 for each source of uncertainty: data statistics (top left); MC statistics (top right); flux without including the integrated flux factor Φ (middle left); detector response (middle right); π FSI (bottom left) and cross section modeling (bottom right). Dashed lines separate different angular bins.

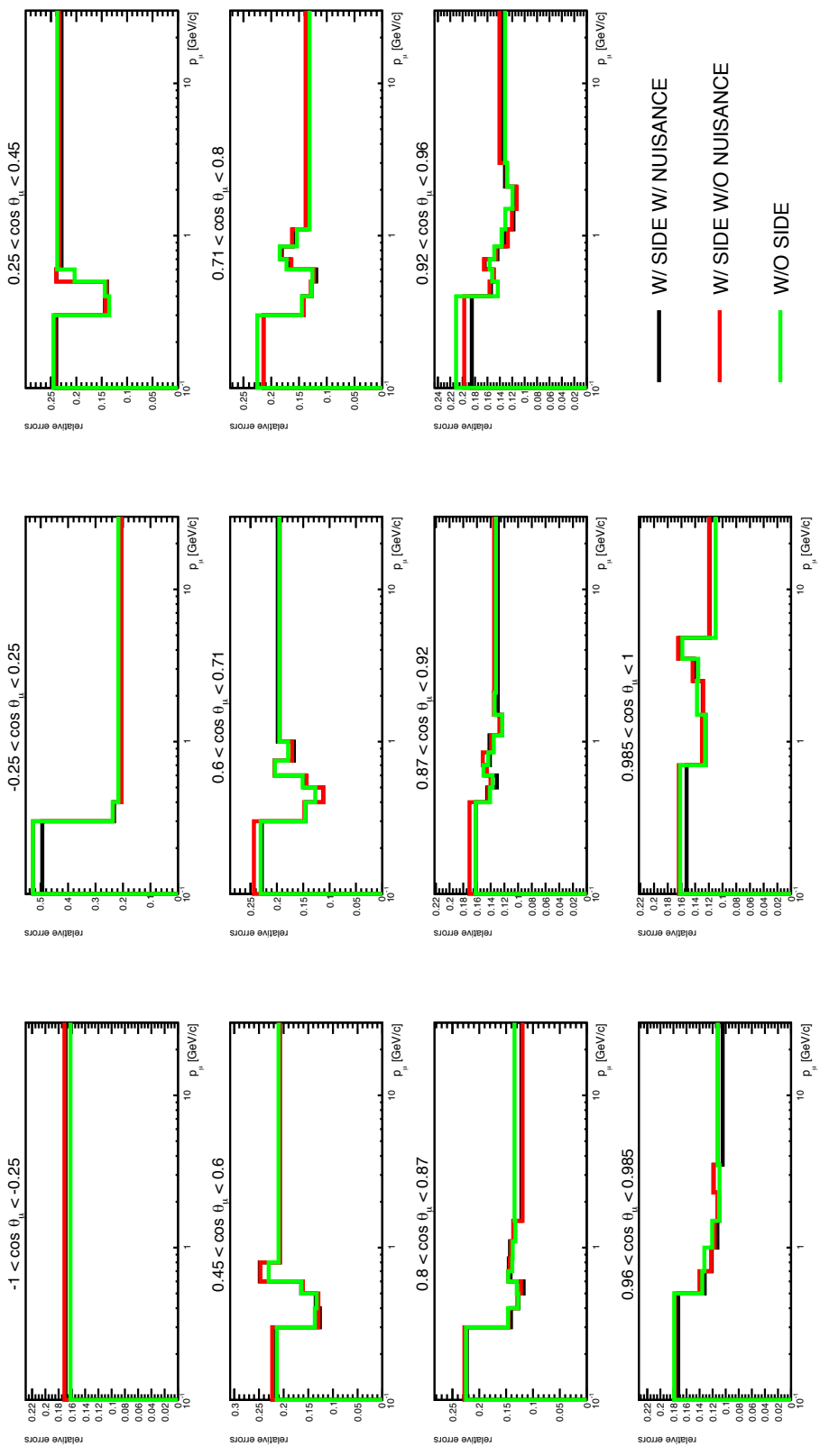

Figure 10.14: Relative total uncertainty for cross section when sidebands are not included (green); when sidebands are included without nuisance parameters (red) and with nuisance parameters for π FSI and NC interactions (black).

10.8 Double Differential Cross Section

In this section, the final cross section extracted from the data is studied in detail. In App. G, several tests were performed to validate the unfolding method. Overall, good performance was obtained except for an small bias in the low momentum and very forward region. This bias was produced by the efficiency correction, which was found to be slightly different between `NEUT` and `GENIE` (see Fig. 10.8).

The reconstructed distribution from data is unfolded using both `NEUT` and `GENIE` as prior, including nuisance parameters for π FSI and NC interactions. The statistical and systematic uncertainties and their correlations can be computed using toy experiments (see Sec 10.7).

It is not possible to compare the extracted cross section with the truth distribution from the data. However, the number of reconstructed events for the fitted distribution after 500 toys and the data can be compared (see Fig. 10.15). It can be concluded that the agreement is remarkable in all the phase space. Besides, the result obtained using either `NEUT` or `GENIE` as prior is equivalent. This test shows that this unfolding method does not depend on `NEUT` and `GENIE` models.

The $(d\sigma/dX_t)^{FIT}$ distribution in each bin (see Eq. 10.13) for 500 toys is shown in Fig. 10.16. The mean and width of such distributions are used to set the average cross section and its associated uncertainty per each bin respectively (see Fig. 10.17). As it was observed in App. G.3, small disagreement is observed in the low momentum and very forward region when using different event generators as prior. This bias is understood and it is due to the different efficiency predictions in that region of the phase for `NEUT` and `GENIE` (see Sec. 10.4). In fact, it has been checked that current uncertainties cover that bias.

The correlation across bins which is computed using Eq. 10.14 is shown in Fig. 10.18, when `NEUT` is used as prior.

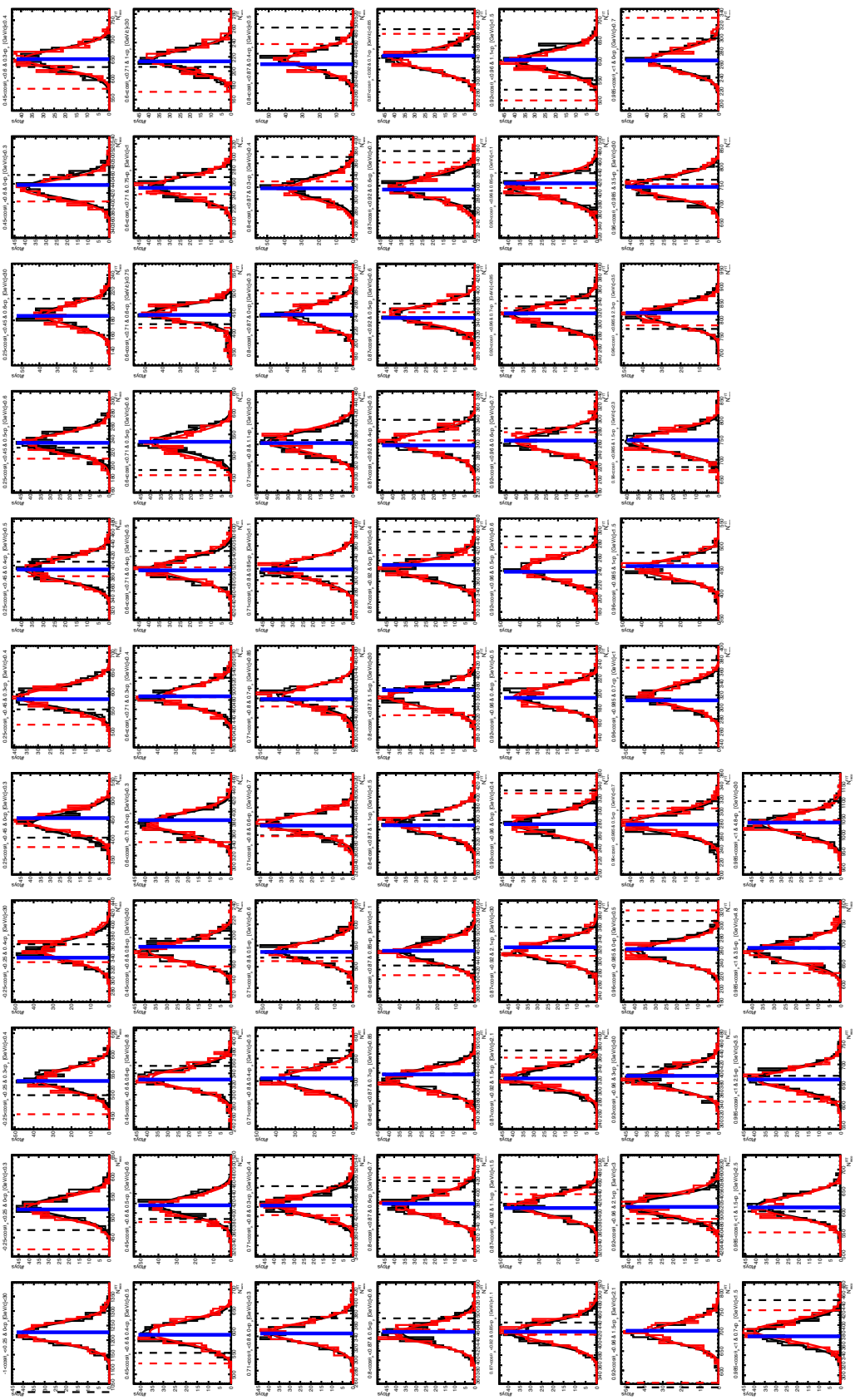

Figure 10.15: Distributions for the fitted number of reconstructed events (including signal and background) for 500 toys experiments using data and as prior NEUT (black) or GENIE (red). Plot legend indicates its associated bin increasing horizontally from $[-1 < \cos\theta_\mu < -0.25, 0\text{GeV/c} < p_\mu < 30\text{GeV/c}]$ (top left) to $[0.985 < \cos\theta_\mu < 1, 4.8\text{GeV/c} < p_\mu < 30\text{GeV/c}]$ (bottom right). Blue line indicates the number of reconstructed events in data. Dashed lines are centered at NEUT (black) and GENIE (red).

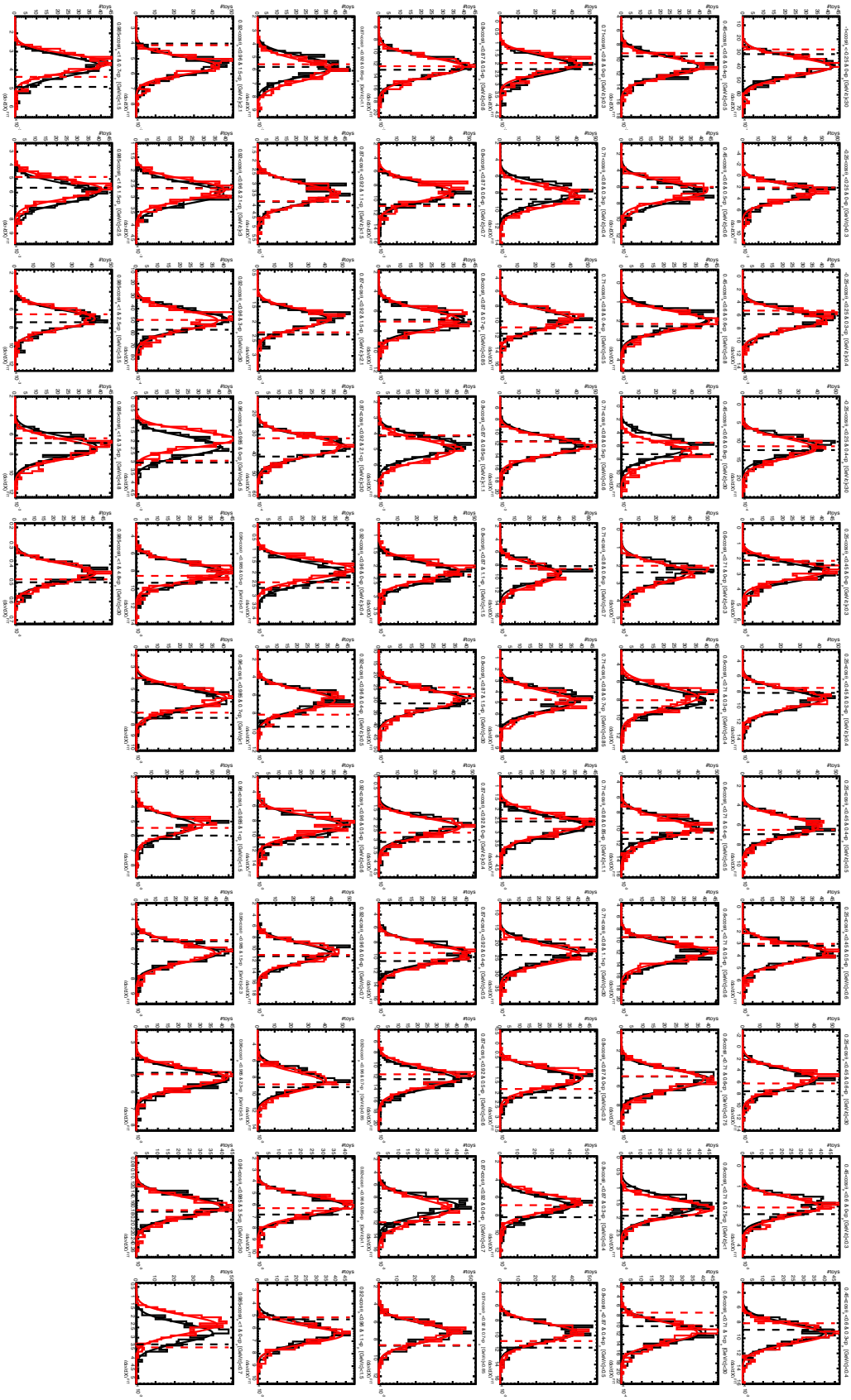

Figure 10.16: $(d\sigma/dX_t)^{FIT}$ distribution (see Eq. 10.13) for 500 toy experiments using real data and as prior NEUT (black) or GENIE (red). Plot legend indicates its associated bin increasing horizontally from $[-1 < \cos\theta_\mu < -0.25, 0\text{GeV/c} < p_\mu < 30\text{GeV/c}]$ (top left) to $[0.985 < \cos\theta_\mu < 1, 4.8\text{GeV/c} < p_\mu < 30\text{GeV/c}]$ (bottom right). Dashed lines are centered at the NEUT (black) and GENIE (red).

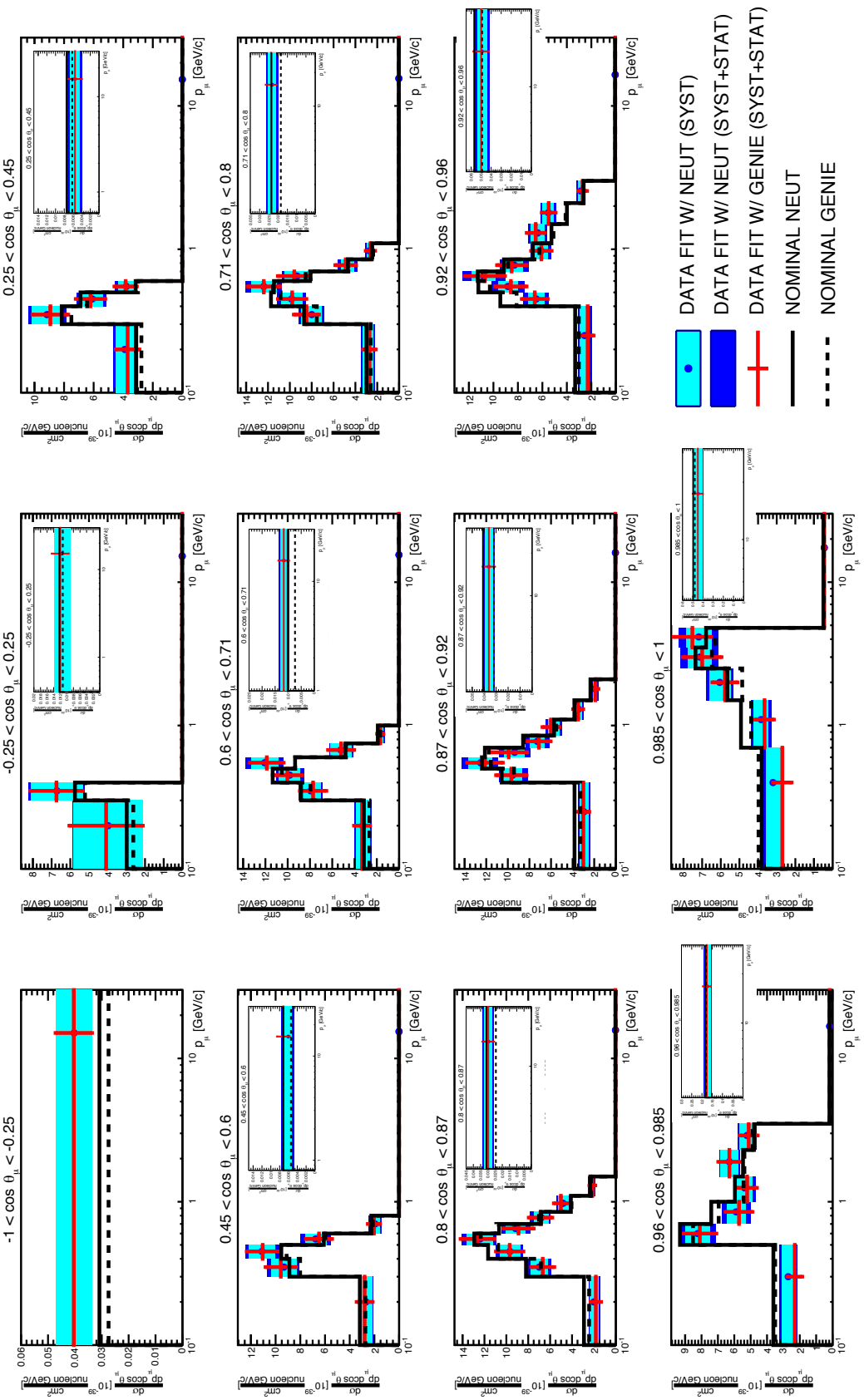

Figure 10.17: Flux-integrated double differential cross section per nucleon for NEUT (black line), for GENIE (black dashed line) and average result after fitting 500 toys with either NEUT or GENIE.

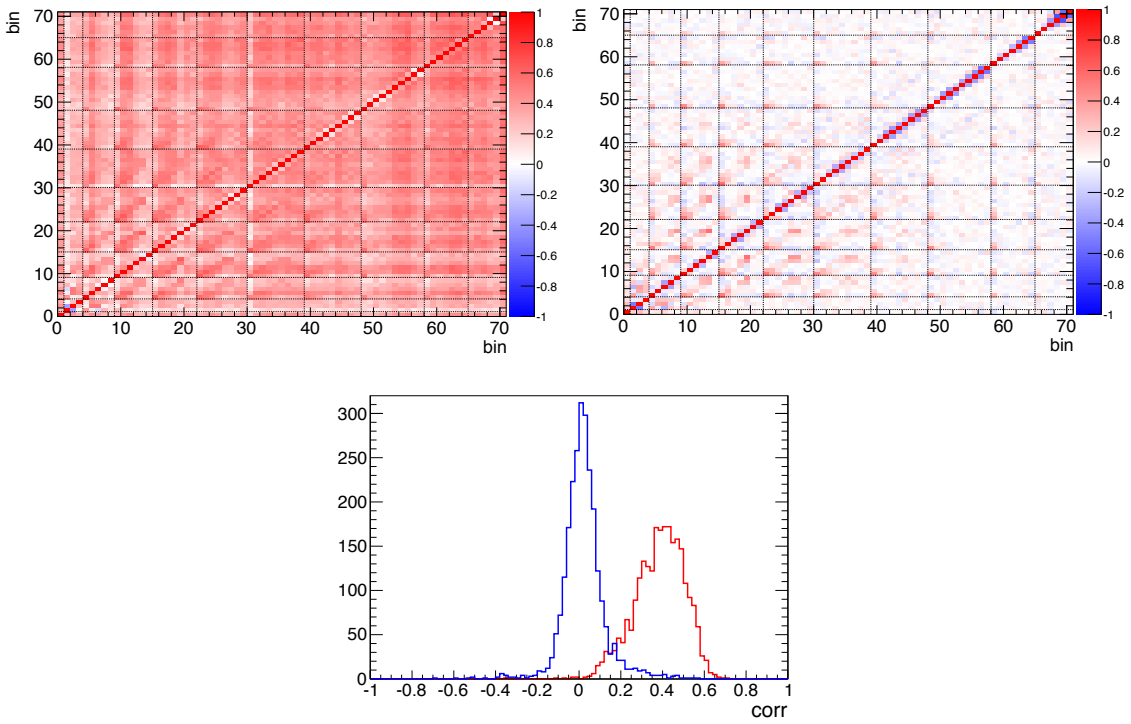

Figure 10.18: Top: Correlation matrix obtained from Eq. 10.14, with (left) and without (right) including the integrated flux factor Φ, for real data unfolded using as prior `NEUT`. Dashed lines separates different angular bins. Bottom: Distribution of the correlation in each bin with (red) and without (blue) including the integrated flux factor Φ.

10.9 Model comparisons

In this section, the results from Sec. 10.8 are compared with the prediction of different models. Previously, the predictions from nominal `NEUT` and `GENIE` have been used. On the one hand, the nominal `NEUT` models can be tuned using `T2KReweight`. On the other hand, another neutrino generator is included, called `NUWRO` [166].

Table 10.4 summarizes the parameters used by each model for the CCQE and 2p2h channels. In the case of RES channel, the model used by `NUWRO` is significantly different than the one used by `NEUT` and `GENIE`. Finally, the DIS channel is modeled in a similar way by the three generators. However, the transition between RES and DIS channels is treated in a different way by each generator.

In order to give an estimation of the agreement between average result for the fitted cross section and the true cross section from the different models a χ^2 is computed. This factor is computed using the following expression:

$$\chi^2 = \sum_{i=0}^{N_{bins}} \sum_{j=0}^{N_{bins}} \frac{1}{\Delta\frac{d\sigma}{dX_i}\Delta\frac{d\sigma}{dX_j}} \left(\left\langle \frac{d\sigma}{dX_i} \right\rangle - \frac{d\sigma^{MODEL}}{dX_i} \right) \Sigma_{ij}^{-1} \left(\left\langle \frac{d\sigma}{dX_j} \right\rangle - \frac{d\sigma^{MODEL}}{dX_j} \right) \tag{10.29}$$

Where $\left\langle \frac{d\sigma}{dX_i} \right\rangle$ ($\Delta\frac{d\sigma}{dX_i}$) and $\frac{d\sigma^{MODEL}}{dX_i}$ are the average result (uncertainty) for the fitted cross section and the true cross section for a particular model in the bin i. Finally, Σ_{ij} represent the correlation across bins (see Fig. 10.18).

	`NEUT`-SF	`NEUT`-RFG	`NEUT`-RFG+RPA	`NEUT`-NIWG	`GENIE`	`NUWRO`
Nuclear model	SF	RFG	RFG	RFG	RFG	LFG
M_A^{QE} [GeV/c^2]	1.21	1.21	1.21	1.15	0.99	1.2
p_F [MeV/c]	217	217	217	223	221	-
E_B [MeV]	25	25	25	25	25	25
MEC-Nieves %	100	100	100	27	0	100
RPA	No	No	Yes	Yes	No	No
χ^2 (DATA fit w/ `NEUT`)	211.9	284.8	173.4	160.9	191.3	471.4
χ^2 (DATA fit w/ `GENIE`)	227.4	300.6	192.6	175.3	190.5	465.7

Table 10.4: Comparison between CCQE/2p2h dials for different models.

Fig. 10.19 shows the cross section result from Sec. 10.8 compare to the predictions from nominal `NEUT` and `GENIE`. Fig. 10.20 shows the cross section result for different CCQE models based on `NEUT` generator. The main differences appears in the peak region where CCQE channel dominates. The SF

predicts a lower cross section than the RFG nuclear model. Meanwhile, the RPA correction significantly suppressed the cross section in the T2K energy range. Besides, the NIWG model (used as nominal MC in T2K oscillation analysis) predicts less rate of CC interactions along all the muon phase space.

A more sophisticated nuclear model than the RFG is included in `NUWRO` generator, the Local Fermi Gas (LFG). This model includes a radial dependency on the initial momentum of the nucleons (see Sec. 4.2). Fig. 10.20 shows the cross section result for `NUWRO` and `NEUT`-RFG. None of them includes the RPA correction. The prediction for the QE channel is very similar in both cases (except in the very forward region). The main discrepancies appear in the RES channel.

Even though an inclusive cross section is not the best channel to extract information about nuclear effects, this result points to an overestimation of simplistic nuclear models as RFG. Such disagreement can be reduced using nuclear models that take into account nucleon correlations, such as RPA or SF.

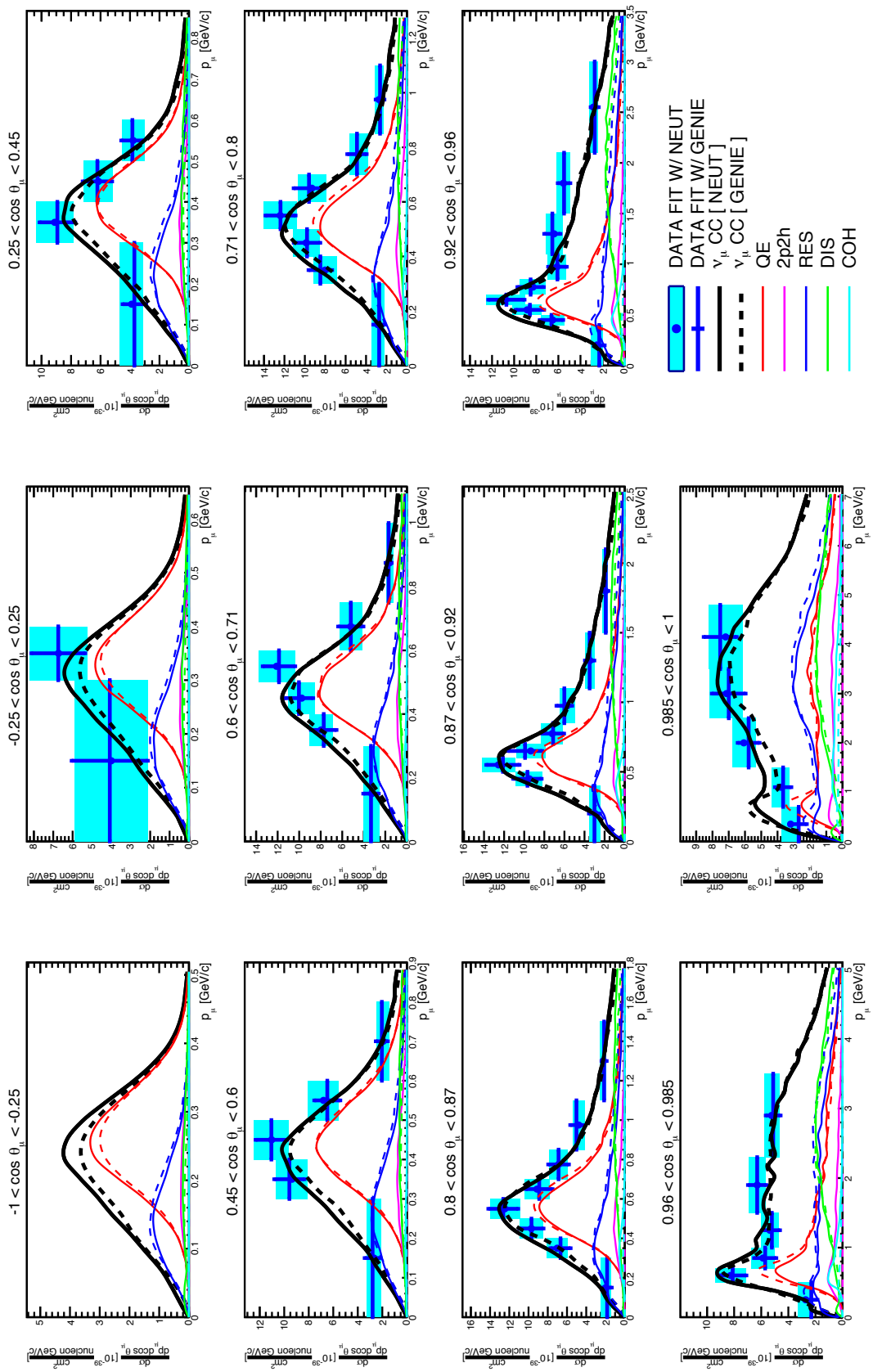

Figure 10.19: Flux-integrated double differential cross section per nucleon using an smother binning for the momentum for NEUT (black line), for GENIE (black dashed line) and average result after fitting 500 toys with either NEUT or GENIE.

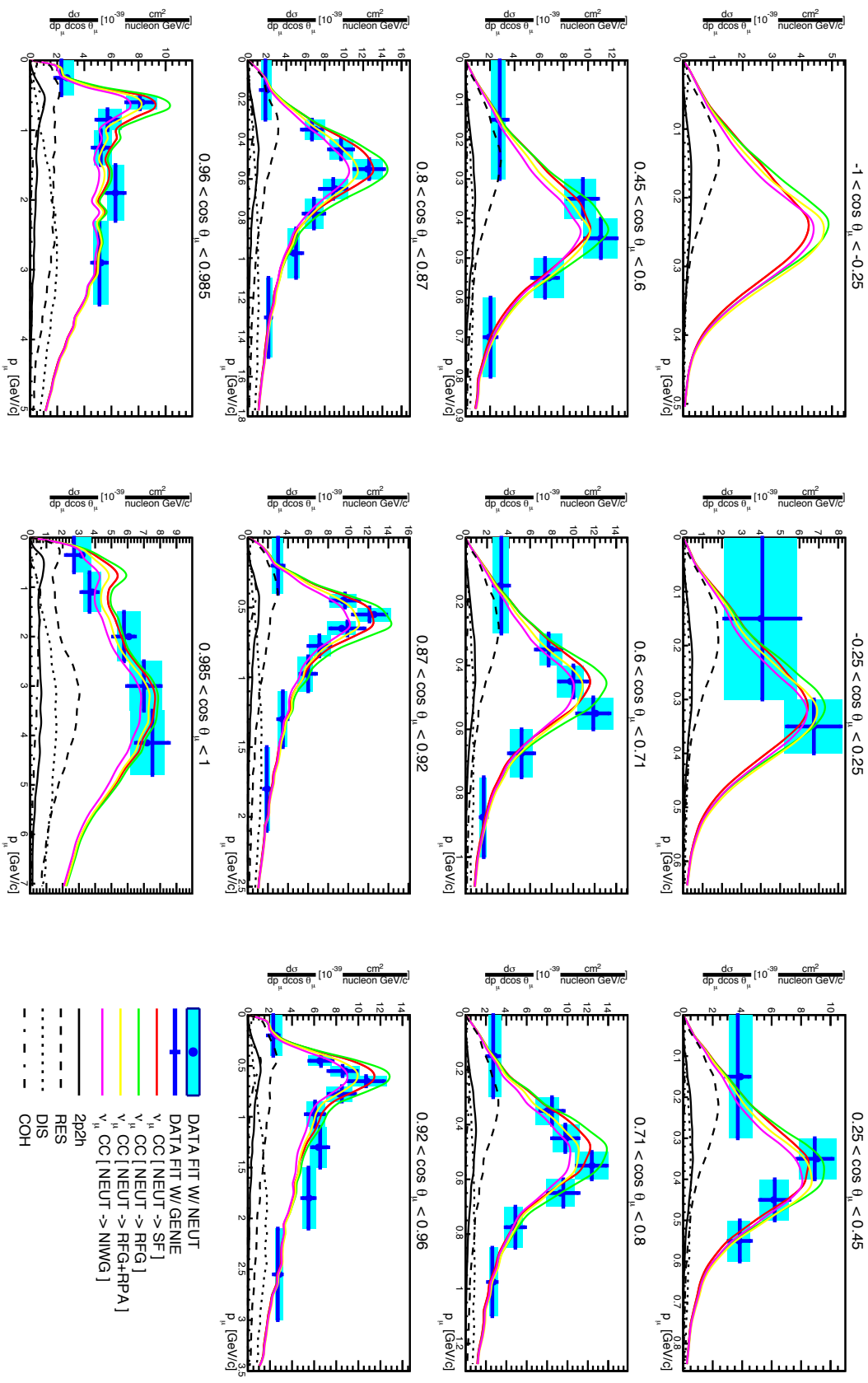

Figure 10.20: Flux-integrated double differential cross section per nucleon for different models based on NEUT (lines) and average result after fitting 500 toys using nominal NEUT (blue rectangles) or nominal GENIE (blue crosses) as prior.

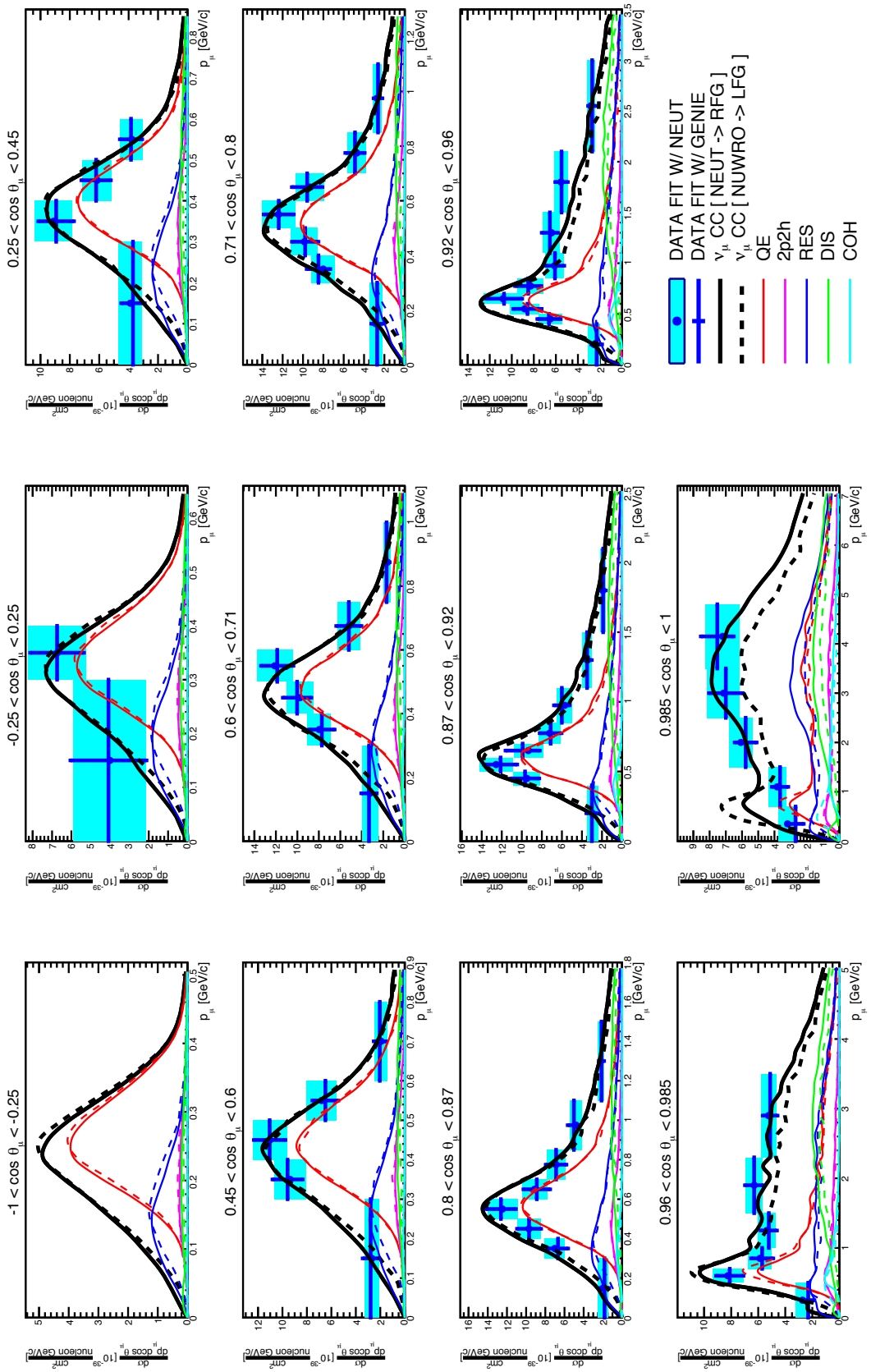

Figure 10.21: Flux-integrated double differential cross section per nucleon for different models based on `NEUT` and `NUWRO` (lines) and average result after fitting 500 toys using nominal `NEUT` (blue rectangles) or nominal `GENIE` (blue crosses) as prior.

10.10 Integrated Cross Section

Using the result from previous section, it is possible to compute the integrated cross section using the following formula:

$$\sigma = \sum_i \sum_j \frac{d\sigma}{dp_{\mu,i} d\cos\theta_{\mu,j}} \Delta p_{\mu,i} d\cos\Delta\theta_{\mu,j} \tag{10.30}$$

This value integrates both the number of signal events and the efficiency over the μ^- phase space. As it was shown in Sec. 10.4, it is known that the detector performance is quite different depending on the angle of the outgoing μ^-. Therefore, the extracted value using the integrated cross section must be interpreted cautiously.

The integrated cross section for both `NEUT` and `GENIE` has been extracted:

$$\begin{aligned} \sigma_{\text{NEUT}} &= 7.108 \times 10^{-39} \frac{\text{cm}^2}{\text{nucleon}} \\ \sigma_{\text{GENIE}} &= 6.564 \times 10^{-39} \frac{\text{cm}^2}{\text{nucleon}} \end{aligned} \tag{10.31}$$

The integrated cross section is extracted for real data. This is done through the methodology described in the previous section. Then, the double differential result is integrated over the full phase space of the μ^-. Fig. 10.22 shows the distribution for 500 toys, from which the average result can be extracted:

$$\begin{aligned} \sigma_{\text{DATA (NEUT)}} &= (7.00 \pm 0.64) \times 10^{-39} \frac{\text{cm}^2}{\text{nucleon}} \\ \sigma_{\text{DATA (GENIE)}} &= (6.90 \pm 0.64) \times 10^{-39} \frac{\text{cm}^2}{\text{nucleon}} \end{aligned} \tag{10.32}$$

Discrepancies are found when using `NEUT` or `GENIE` as prior. As shown previously, this bias is produced by the different efficiency prediction of both event generators. This result entirely agrees with the one obtained in [119]. As in the previous analysis, the current result is dominated by the normalization uncertainty associated to the flux ($\sim 11\%$ in [119] and $\sim 9\%$ in this analysis).

The result can be compared with other experiments, dividing the integrated cross section by the mean energy of the T2K ν flux. This mean energy is 850 MeV and it is slightly shifted with respect to the peak position of the flux (due to the high energy tail). The errors associated to the mean energy were obtained from [119].

Fig 10.23 shows the integrated cross section divided by the mean energy computed in this dissertation and other measurements using Carbon as main target. In the same energy range, SciBoone reported

higher cross section results. However, data agree with MC predictions in this energy range.

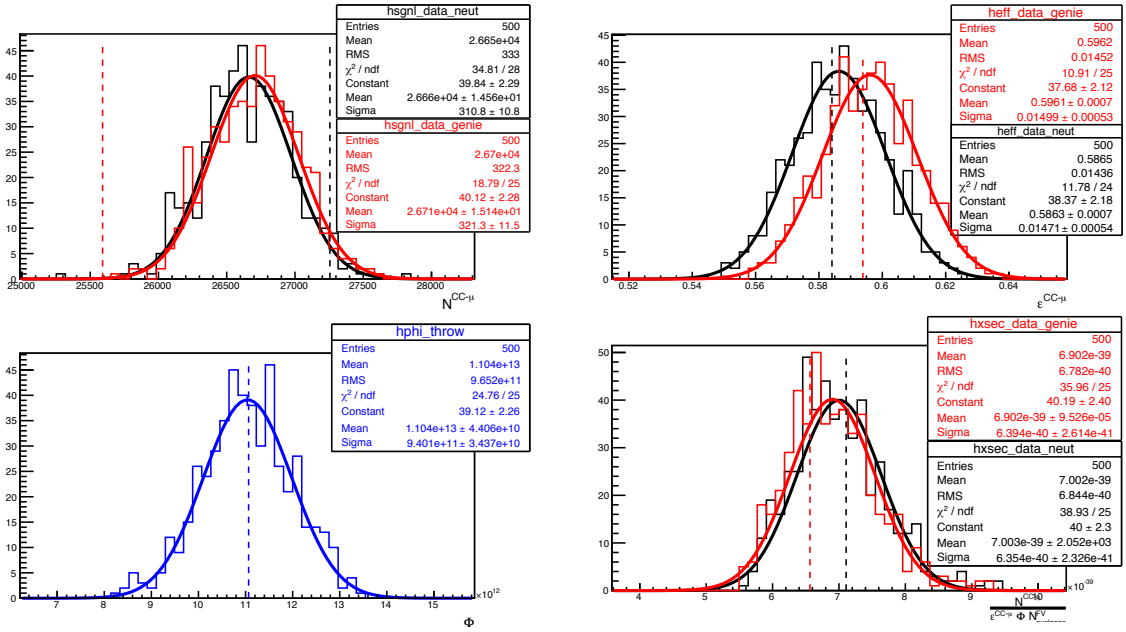

Figure 10.22: Distributions for the integrated number of signal events (top left), efficiency (top right), flux (bottom left) and cross section (bottom right) obtained from the fit using as prior `NEUT` (black) and `GENIE` (red). Dashed lines are centered at nominal value for `NEUT` (black) and `GENIE` (red).

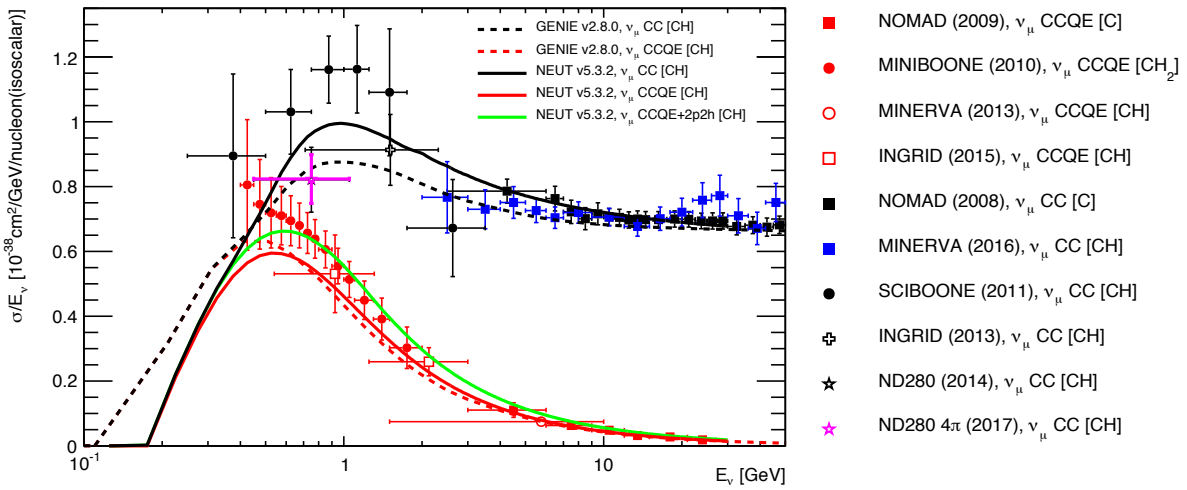

Figure 10.23: Cross section as function of the ν energy for ν_μ−C interactions for different experiments and reaction channels. Lines represent the prediction from GENIE and NEUT for CH.

Chapter 11

Conclusions

Neutrino oscillation has become one of the most widely studied topics in the last decades within High Energy Physics. Many experiments have been constructed in order to improve our understanding of such phenomena.

In particular, T2K collaboration has proved that long baseline experiments (using an accelerator as ν source) play a fundamental role in this research. The peaked ν_μ flux allows to study the ν_μ disappearance channel in detail for both ν and $\bar{\nu}$. Currently, the most precise measurement of θ_{23} and $\Delta m_2 3^2$ is provided by T2K [167]. Moreover, while studying the ν_e appearance channel, certain sensitivity on δ_{CP} has been observed [167].

Up to now, those results were dominated by statistical uncertainties. However, for next iteration of measurements, systematic errors will become of the same order. Therefore, one of the main objectives is to narrow down such uncertainties. For that purpose, placing a near detector is mandatory. ND280 was designed to reduce systematic uncertainties, due to our current knowledge of the T2K flux and cross sections. Such uncertainties are reduced by a factor of three by studying ν_μ CC interactions in the FGDs of ND280.

The extrapolation of the systematic uncertainties between near and far detector must be treated cautiously as they are completely different detectors. Currently, one of the main concerns is that ND280 can only observe ν_μ CC interactions in which the μ^- coming out from the interactions goes forward (starting in the FGD and crossing the downstream TPC), while SK does not have this limitation. Therefore, the uncertainties associated to a limited region of the phase space are extrapolated to different regions of the phase space.

In order to avoid that tension, a new analysis has been developed in ND280. This dissertation describes such analysis, in which ν_μ CC interactions happening within the FGD are selected increasing the angular acceptance of previous analysis. Using timing information between different subdetectors and PID capabilities from BarrelECal, μ^- coming out from the interaction can be reconstructed when

going backward or with high angle. Fig. 11.1 shows the reconstruction efficiency in this new analysis as function of the emission angle of the outgoing μ^-. Those promising results will be used in the coming iteration of T2K oscillation measurement.

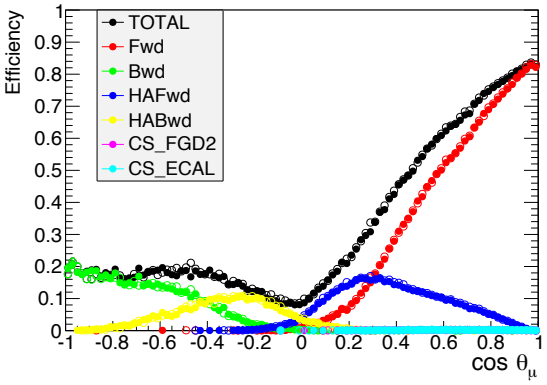

Figure 11.1: Cosine of emission angle efficiency for $\nu_\mu CC - \mu$ using `NEUT` (dots) and `GENIE` (empty dots) as generator. Colors indicate contribution from different selections. Previous analysis has similar performance than red dots.

Furthermore, apart from reasonable efficiency, this sample of ν_μ CC interactions has high purity. Therefore, it provides a perfect condition to perform a cross section measurement. It is important to notice that our current knowledge of ν_μ CC interactions with complex nucleus is very limited. Thus, this kind of measurement is extremely useful for theoreticians, because their models can be tested.

The cross section measurement presented in this dissertation is pioneer because it reports results in regions of the phase space that had not previously been studied. Moreover, the procedure followed to extract the cross section is aimed to become a standard within the field. Several tools and warnings that have been described might help future analysis make a rigorous measurement.

Both double differential and integrated cross section for ν_μ CC interactions on plastic have been reported. The double differential measurement shows some disagreement with respect to the model predictions from `NEUT` and `GENIE`. It is particularly important the bias observed in peak structure in different angular regions. This peak is dominated by the CCQE channel, so those discrepancies point to a mismodeling of such a channel (possibly due to nuclear effects). Another interesting result is that both `NEUT` and `GENIE` predictions underestimate the cross section in the backward region.

Finally, as closure test, the integrated cross section presented in this dissertation is consistent with previous measurements. However, it should be underlined the completely different methodology followed in this measurements.

In the future, this analysis can be further improved reducing the statistics and flux uncertainties. Moreover, the angular acceptance for backward going μ^- is likely to be increased. In this analysis, we have realized that most of the outgoing μ^- goings backward have very low momentum $p \sim 200MeV$. Therefore, they are very likely to stop very close to the walls of the P0D or BarrelECal producing few hits in such subdetectors. That fact makes the reconstruction of those hits as isolated segments challenging. If those hits are not reconstructed, the timing information is missed and those μ^- can not be tagged.

Currently, a new reconstruction algorithm is been developing, which aims to recover those hits. Preliminary studies are very promising and the efficiency for backward going μ^- could be increased from $\sim 20\%$ to $\sim 50\%$.