
Global Analysis of Muon-Neutrino Charged-Current Data for the Characterization and Tuning of Cross-Section and Hadronization Models in the GENIE Neutrino Event Generator



DEPARTMENT OF PHYSICS

Thesis submitted in accordance with the requirements of the University
of Liverpool for the degree of Doctor in Philosophy by

Júlia Tena Vidal

March 2022

Abstract

The GENIE neutrino Monte Carlo event generator seeks to model neutrino interactions for all probes and targets across the entire kinematic phase space relevant for neutrino experiments. The lack of a complete theory to model all neutrino scattering processes forces neutrino event generators to rely on empirical models to simulate exclusive final states for neutrino scattering off nucleon and nuclei. These models need to be tuned to data in order to provide a better description and uncertainties for systematic analyses. This thesis focuses on the development of GENIE global analysis of neutrino scattering data, the tuning of simulations and the data-driven characterization of simulation uncertainties.

The first analysis aims to improve the description of the shallow-inelastic scattering region at the free nucleon level using all the relevant neutrino-nucleon inclusive and exclusive pion-production cross-section data on hydrogen and deuterium. Tensions between inclusive and exclusive datasets are resolved by improving the systematic handles of historical data. The bare-nucleon global tune improves the agreement with both datasets, solving the longstanding discrepancies.

Neutrino-induced hadronization is a crucial modeling element for the description of inelastic processes. GENIE uses a hadronization model, known as AGKY, which includes an empirical model for low hadronic invariant mass events and incorporates the PYTHIA model at high invariant masses. Only the empirical low-mass model parameters were extracted from neutrino hadron production data, but previous efforts lacked a rigorous analysis procedure and hadronization uncertainties were missing. In addition, comparisons of the GENIE model against neutrino-induced hadron shower data exposed significant disagreements at $W > 3 \text{ GeV}/c^2$. This thesis presents the first tune of the AGKY hadronization model against charged averaged multiplicity data on hydrogen and deuterium targets.

The last analysis exploits the GENIE global analysis framework to analyse Mini-BooNE, T2K and MINERvA cross-section data on muon-neutrino charged-current interactions on carbon with no pions in the final state. This event topology is the simplest one, and the most dominant topology at $E_\nu < 1 \text{ GeV}$, but it is still not well understood due to the nuclear environment. The analysis incorporates new parametrization of modeling uncertainties in GENIE as well as data-driven constraints from the GENIE bare-nucleon tune on hydrogen and deuterium. A partial tune is performed for each of the experiments, providing a common ground for the discussion of tensions. This analysis is the first step towards a global tune that improves our understanding of neutrino-nuclei interactions.

Acknowledgements

I would like to start by thanking my supervisor, Costas. You have been an excellent guide in my academic career and I will always be grateful that you encouraged me to push myself to my limits. I remember that ice-cream in the Chicago mall after a long week of collaboration meetings, you always know how to convince me that working hard is worth it. I am looking forward working with you and the GENIE team in the near future.

There are many things great I could say about Marco. Marco is an excellent teacher and he will always be my inspiration. Thank you for teaching me so much about neutrino cross sections, coding and having so much patience with all my questions. I cannot wait to meet you at Fermilab again and beat you at table tennis! (Of course I cannot say here that you also win sometimes).

I would also like to thank every member of the GENIE Collaboration, as they have always been really supportive. I will always remember the amazing time I had in Boston working with Hugh on the bare-nucleon tune. It was also exciting to work with Steve, from who I have learnt a lot about final-state interactions.

When I arrived at Liverpool I had the warmest welcome. In particular, I would like to thank all the neutrino group members. Rhiannon, you have been a wonderful colleague and friend. Thank you for everything I have learnt from you, all the trips we had together and for sharing the enthusiasm about making NuGirl a real thing. I was also really lucky to be part of LIV.DAT, as I met so many wonderful people through it. Thank you Fraser for always making me laugh. It is still a mystery to me why you know how to say "sacapuntas". Thank you Alex, Zhen and James for all the amazing climbing and hiking trips around the UK. And thank you Phil for always being ready for last-minute plans. I also had the best after-work walks with Ulserik, who always gives me the warmest hugs.

Me mudé a Inglaterra y me encontré con Maria, que la conocía de mi pueblo dels Ports! Y gracias a ella conocí a Alessia, Elena, Zoe y Amy. Habéis sido todas las mejores compañeras de piso que podría haber tenido, espero veros pronto y vamos de fiesta, que sin vosotras no es lo mismo. Zoe, fuiste la mejor compañera de lockdown. Muchísimas gracias por todos los días de yoga y telas en el parque. Thank you very much Lauren for being such a good friend, and the cake, that cake was good! I also had amazing flatmates in Geneva: thank you Karolina, Kevin, Alex, Bruno and James for welcoming me when I unexpectedly moved in. I promise I make another round of croquetas. I would also like to thank all the climbing friends from Geneva, where I had a lot of fun and unforgettable experiences.

M'agradaria fer una menció especial als meus pares. No només m'heu donat els millors consells durant aquests quatre anys, sinó que també heu sigut la meva inspiració durant tota la meva vida. Sou els millors professors de física i química que hauria pogut tenir. I per descomptat, moltes gràcies germanets, sempre que torno a casa em feu sentir com si no hagués marxat mai.

I would like to finish by thanking the person who has been by my side through most of my PhD: James, not only you have been the best friend and office mate for the last three years, but you have been my adventure partner. Thank you for all the amazing climbing trips around Europe and the dream month in Interlaken. Those waterfalls are truly amazing.

Contents

1	Introduction	1
1.1	Thesis structure	2
2	Modeling of neutrino interactions with matter	3
2.1	Electroweak interactions in the Standard Model	3
2.2	Neutrino interactions with nucleons	6
2.2.1	Modeling of neutrino charged-current quasi-elastic interactions	9
2.2.2	Modeling of charged-current resonant interactions	12
2.2.3	Modeling of charged-current deep inelastic scattering interactions	15
2.2.4	Modeling of the shallow-inelastic scattering region	16
2.3	Neutrino interactions with nuclei	18
2.3.1	The nuclear model	19
2.3.2	Multi-nucleon production processes	20
2.3.3	Final state interactions	23
2.4	Neutrino cross-section measurements	25
2.4.1	The 70's: Bubble chamber experiments	27
2.4.2	The early modern era	29
2.4.3	The modern era	30
3	Neutrino-nucleon cross-section model tuning in GENIE v3	33
3.1	Introduction	33
3.2	Bare nucleon cross-section modeling in GENIE	36
3.3	Comprehensive model configurations in GENIE v3	40
3.3.1	CMCs available in GENIE v3	40
3.3.2	Free nucleons and CMCs	42
3.4	Data and model uncertainties review	43
3.4.1	Datasets included in the fit and their systematics	43
3.4.2	Model uncertainties	44

3.5	Bare-nucleon cross-section tuning procedure	50
3.5.1	Likelihood construction	50
3.5.2	Treatment of systematic uncertainties	52
3.5.3	Discussion of priors	53
3.5.4	Final form of the χ^2	54
3.6	Tuning results	55
3.6.1	Partial fits	55
3.6.2	Global fit	57
3.6.3	Parameter error estimation	63
3.7	Conclusions	70
4	Hadronization model tuning in GENIE v3	75
4.1	Introduction	75
4.2	The AGKY model	78
4.2.1	Effective low- W AGKY hadronization model	78
4.2.2	PYTHIA in GENIE	80
4.3	Neutrino-induced hadronization data review	83
4.3.1	Hydrogen data	84
4.3.2	Deuterium data	87
4.3.3	Sources of systematic uncertainties in the FNAL 15 ft and BEBC experiments	90
4.4	Review of previous tunes to hadronization data	92
4.4.1	Fits to bubble chamber data	92
4.4.2	The HERMES tune	94
4.4.3	Requirements for including a dataset in the AGKY multiplicity tune	95
4.5	Parametrization of model uncertainties	95
4.6	Construction of the GENIE predictions and evaluation of the likelihood	98
4.6.1	Professor interpolation cutoff condition	99
4.6.2	Parameter priors	101
4.6.3	Final form of the χ^2	102
4.7	AGKY tune results	102
4.7.1	The 2021 GENIE AGKY global tune	109
4.7.2	The 2021 GENIE AGKY ${}^2\text{H}$ tune	113
4.7.3	AGKY global and deuterium only tunes impact on other neutrino-induced hadronization observables	114
4.7.4	2021 GENIE AGKY global tune impact at the SIS region	115
4.8	Conclusions	117

5 Neutrino-nuclei CC0π cross section tuning in GENIE v3	121
5.1 Introduction	121
5.2 Review of neutrino nucleus CC0 π measurements	124
5.2.1 Kinematic quantities of interest for CC0 π measurements	128
5.2.2 MiniBooNE CC0 π cross-section measurement	131
5.2.3 T2K ND280 CC0 π cross-section measurements	134
5.2.4 MINERvA CC0 π cross-section measurements	140
5.2.5 MicroBooNE CCNp0 π cross-section measurement	144
5.2.6 Dataset overview and initial considerations	146
5.3 Discussion of CC0 π model uncertainties in GENIE	148
5.3.1 Charged-current quasi-elastic uncertainty parametrizations	148
5.3.2 Charged-current shallow-inelastic uncertainty parametrizations	152
5.3.3 Charged-current multi-nucleon uncertainty parameteterizations	153
5.3.4 Nuclear model uncertainty parametrizations	156
5.3.5 Final state interaction uncertainty parametrizations	157
5.4 Tuning procedure	159
5.4.1 Construction of the GENIE prediction	159
5.4.2 Avoiding the Peele's pertinent puzzle	160
5.4.3 Professor parametrization	161
5.4.4 Final choice of parameters for the CC0 π tune	161
5.4.5 Discussion of data driven priors	163
5.4.6 Evaluation of the χ^2	163
5.5 Tuning Results	164
5.5.1 Discussion of partial CC0 π tune results	165
5.5.2 Tension between CC0 π partial tunes	167
5.6 Tensions between ν_μ CC0 π and ν_μ CCNp0 π datasets	171
5.6.1 Investigation of tensions between CC0 π and CCNp0 π datasets	175
5.7 Conclusions	180
6 Concluding remarks	183
Appendices	189
Appendix A Neutrino-nucleon cross-section model tuning in GENIE v3	191
A.1 Comprehensive model configuration naming convention	191
A.2 Tunes for z-expansion CMCs	192

CONTENTS

Appendix B Hadronization model tuning in GENIE v3	193
B.1 Dataset compatibility study	193
Bibliography	193

Per als meus pares

CONTENTS

Chapter 1

Introduction

Neutrinos play an important role in our understanding of the universe. Yet, many questions remain about their nature. Measurements of potential CP-violation effects in the lepton sector may be the clue to understand the matter-antimatter asymmetry of the universe. Beyond that, other fundamental ingredients of the neutrino puzzle are the determination of the neutrino mass hierarchy and tests of the three flavour paradigm [1]. These will be some of the major directions in science during the next decade and beyond.

The interaction of neutrinos with nuclei at a few GeV is one of the dominant sources of uncertainties for the precise determination of neutrino properties. To meet the requirements of future oscillation experiments, we must control the systematic uncertainties associated with neutrino interactions. Neutrino experiments such as DUNE [2] need to reduce the cross-section uncertainties to the 1-2% level in order to measure the CP-violation in the lepton sector [3].

In spite of the recent experimental and theoretical progress, a comprehensive description of neutrino interactions is still missing. Recent measurements exposed the lack of detailed understanding at the few-GeV energy range and substantial disagreement between measurements from different experiments [4]. The understanding of several aspects of neutrino interactions on nucleons is still driven by small samples of just a few hundred events recorded in the bubble chamber era in the 70's and 80's. These samples have outstanding tensions that have not been resolved yet. Discrepancies with recent data on neutrino interactions with heavier targets are also present. It is not yet clear what the origin of these tensions is and these must be incorporated in systematic uncertainties studies.

Neutrino Monte Carlo (MC) event generators are a powerful tool for the study of neutrino interactions. These aim to predict the kinematics of every particle produced after a neutrino interaction for every target from low to high energies. Experiments

rely on the MC generators to reconstruct the neutrino energy, constrain backgrounds and estimate systematic uncertainties. Unfortunately, neutrino MC generators are not exact: there is no complete theory to model neutrino interactions from a few MeV to the PeV region [5]. Consequently, event generators implement various solutions to simulate neutrino interactions. These solutions, based on empirical models, rely on tuning to be able to describe neutrino interactions in regions where our model is not complete yet.

This thesis focuses on the development of GENIE global analysis of neutrino scattering data, the deployment of improved simulations and the data-driven characterization of simulation uncertainties. It demonstrates a powerful method to tune neutrino MC event generators against available data and characterize tensions between datasets. This work is based on the GENIE MC event generator but the same strategy can be implemented in other MC event generators.

1.1 Thesis structure

This thesis is organized as follows. The concepts necessary to understand the status of neutrino interaction modelling in neutrino event generators are explained in Chapter. 2. In Chapters. 3-5, the potentials of the tuning machinery are exploited for analysis of different aspects of neutrino interactions within GENIE. Chapter. 3 focuses on the tuning of the free-nucleon cross section in GENIE and it also provides with most of the concepts behind the global tuning machinery. Chapter. 4 presents the first tune of the GENIE hadronization model using neutrino scattering data. Chapter. 5 offers an insight to neutrino-nuclei tuning using data from the latest neutrino experiments on carbon and argon. The last chapter summarizes the thesis conclusions.

This thesis is a compilation of two papers, for which I am the main author. Both Chapter. 3 and Chapter. 4 have been published in Phys.Rev.D. We aim to publish Chapter. 5 after the submission of this thesis. Chapter. 5 is also structured as a paper. The two publications are included in accordance to the code of practice for PGR students at the University of Liverpool. These are not modified with respect to the published versions.

Chapter 2

Modeling of neutrino interactions with matter

This chapter offers a review of neutrino interactions with matter at the energies of interest for neutrino oscillation experiments. Emphasis is placed on the modeling aspects that require empirical solutions in order to be implemented into neutrino Monte Carlo (MC) event generators.

The analyses presented in this thesis focus on tuning some of the modeling regions described in this chapter: Sec. 2.1 is an introduction to neutrino interactions with elementary particles. The modeling of free-nucleon interactions is described in Sec. 2.2. Interactions on nuclei are covered in Sec. 2.3. Finally, Sec. 2.4 provides an overview of the available neutrino data to be used to tune these models.

2.1 Electroweak interactions in the Standard Model

The Standard model (SM) is a quantum field theory that describes strong, electromagnetic and weak interactions of elementary particles. The unified theory of weak and electromagnetic interactions was formulated independently by Weinberg [6] and Salam [7] for leptons and it was later extended by Glashow, Iliopoulos and Maiani to the quark sector [8], giving rise to the SM of electroweak interactions in 1967.

In the SM, elementary particles are classified into quarks and leptons. Quarks are organized into three families or doublets,

$$\text{Quarks: } \begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix} \quad (2.1)$$

being the t quark the heaviest out of the six. The quarks u, c, t have electric charge

$+2/3|e|$ whilst the charge for the d, s, b quarks is $-1/3|e|$. e stands for the elementary electric charge. The corresponding antiparticles, denoted as \bar{q} , have the opposite charge. Particles in a doublet have different values of the weak isospin projection I_3 , and these are common for every doublet. In particular, $I_3^{u,c,t} = 1/2$ and $I_3^{d,s,b} = -1/2$. Antiquarks have the opposite sign for the isospin projection. Similarly, leptons are classified into three flavor families:

$$\textbf{Leptons: } \begin{pmatrix} v_e \\ e^- \end{pmatrix}, \begin{pmatrix} v_\mu \\ \mu^- \end{pmatrix}, \begin{pmatrix} v_\tau \\ \tau^- \end{pmatrix} \quad (2.2)$$

There is a neutrino for each flavour family. Similarly to the quark doublets, $I_3^{v_e, v_\mu, v_\tau} = +1/2$ and $I_3^{e, \mu, \tau} = -1/2$. Whilst e^-, μ^- and τ^- have charge $-|e|$ and mass, in the SM the neutrinos do not have charge and from the interaction point of view they are treated as massless. We now know this is a simplified picture of the reality: the observation of neutrino oscillations in 1998 by Super-Kamiokande [9] was the first hint of beyond SM physics as they indicate that neutrinos have mass. The correct path towards the extension of the SM is not clear yet. Even though there are theoretical models that can describe the SM and neutrino oscillations, they bring up new open questions, such as the nature of the mechanism that generates neutrino masses. These questions need to be resolved experimentally. Until then, the SM guides our understanding of weak interactions.

The SM theory successfully postulates the existence of three massive bosons responsible of weak interactions, W^\pm and Z^0 , one massless boson for electromagnetic interactions, γ , and eight massless gluons which are the mediators of strong interactions. In addition, it requires a Higgs boson for the construction of a Lagrangian with mass terms that respect the gauge symmetries of the model. The SM model predicts all the couplings between the elementary particles and the W^\pm and Z^0 bosons. Neutrinos interact only via the weak force¹ and couple with either the W^\pm boson, charged-current (CC) interaction, or the Z^0 boson, in neutral-current (NC) interactions. Quarks and charged leptons also experience the electromagnetic force. In this review, we focus on the weak interaction for its relevance in this thesis.

The SM Lagrangian terms containing the weak couplings are:

$$\mathcal{L}_{\text{Weak}} = \mathcal{L}_{\text{CC}} + \mathcal{L}_{\text{NC}} \quad (2.3)$$

where

$$\mathcal{L}^{\text{CC}} = -\frac{g}{2\sqrt{2}} [J_W^\mu W_\mu + H.c.] \quad \text{and} \quad \mathcal{L}^{\text{NC}} = -\frac{g}{2\cos\theta_W} J_Z^\mu Z_\mu \quad (2.4)$$

¹Neutrinos also experience gravity but this force is not incorporated into the SM yet.

being g the weak coupling constant and θ_W the Weinberg angle. J_W^μ and J_Z^μ represent the charged current and neutral currents respectively. W_μ and Z_μ are the W^\pm and Z^0 bosons. The coupling and the Weinberg angle are related to the elementary electric charge as $g \sin \theta_W = e$ [10]. For historical reasons, in neutrino physics the weak coupling is expressed in terms of the Fermi constant, G_F . The Fermi constant is related with the weak coupling and the boson masses as:

$$\frac{g}{2\sqrt{2}} = \left(\frac{G_F M_W^2}{\sqrt{2}} \right)^{1/2}, \quad \frac{g}{4 \cos \theta_W} = \frac{1}{\sqrt{2}} \left(\frac{G_F M_Z^2}{\sqrt{2}} \right)^{1/2} \quad (2.5)$$

The charged current is defined as:

$$J_W^\mu = \sum_{f=e,\mu,\tau} [\bar{v}_f \gamma^\mu (1 - \gamma_5) f] + \sum_{i=u,c,t} \sum_{j=d,s,b} (\bar{q}_i \gamma^\mu (1 - \gamma_5) V_{ij}^{CKM} q_j) \quad (2.6)$$

where the summation goes over the lepton flavors, f , for the lepton term, and over the quark flavors, q_i , for the quark term. The spinors associated to neutrinos and charged leptons are represented by v_f and f respectively. γ^μ are the Dirac matrices [1]. The first term of Eq. 2.6 describes the CC coupling of a neutrino and a charged lepton. An important observation is that CC terms conserve lepton flavor. The second term describes the CC couplings of the quark sector. Quark mixing is introduced using the Cabibbo-Kobayashi-Maskawa (CKM) matrix [11]. In opposition to the lepton component, quark CC terms involve quarks from different doublets. The CKM is a 3×3 unitary matrix which is parametrized in terms of three Euler angles, θ_{12} , θ_{23} and θ_{13} , and one phase, δ . The phase δ is the source of CP violation in the quark sector [12]. The values of all mixing angles and phase for the CKM matrix must be extracted from data. Their values can be found in Ref. [13].

The neutral current terms are defined as:

$$J_Z^\mu = \sum_{f=e,\mu,\tau} [\bar{v}_f \gamma^\mu (1 - \gamma_5) v_f + \bar{f} \gamma^\mu (g_V^f - g_A^f \gamma_5) f] + 2 \sum_q \bar{q} (g_V^q - g_A^q \gamma_5) q, \quad (2.7)$$

where $g_{V,A}^f$ and $g_{V,A}^q$ are the vector and axial coupling for charged leptons and quarks respectively. The vector and axial couplings for fermions depend on I_3^p , the elementary particle charge, Q_p , and the Weinberg angle as $g_V^p = I_3^p - 2Q_p \sin^2 \theta_W$ and $g_A^p = I_3^p$.

In a lepton weak decay or scattering, two main types of processes can be identified: CC and NC. A CC interaction is defined by one first-order diagram in which a single W^\pm boson is exchanged, whilst in a NC interaction, a single Z^0 boson is exchanged. For instance, the reaction $\nu_\mu e^- \rightarrow \mu^- \nu_e$ is CC, whilst $\nu_\mu e^- \rightarrow \nu_\mu e^-$ is NC. In fact, it is possible to have an interaction that is defined by more than one first-order Feynman di-

agram from CC and NC. An example of such interaction is the scattering of a neutrino with the corresponding charged lepton from the same family, i.e. $\nu_f f^- \rightarrow \nu_f f^-$. CC interactions play a major role in oscillation experiments as the phenomena of neutrino oscillation is only visible through the study of CC interactions. For this reason, this thesis focuses on the study of CC interactions and NC interactions are only discussed briefly.

The discovery of the building blocks of the SM was not an easy journey. Neutrinos, proposed in 1930 by W. Pauli as a possible solution to explain the continuous β decay spectrum, were thought to be undetectable until 1956, when Reines and Cowan detected the first antineutrinos from a nuclear reactor [14]. It was not until 1973 that CERN and Fermilab observed the first NC interaction [15, 16]. In fact, it took another 10 years to directly prove the existence of the W^\pm and Z bosons [17–20]. The remaining piece left was the Higgs boson, discovered in 2012 at CERN [21]. Even though all the pieces that build the SM have been found, we now know that the SM is an effective theory.

It is important to highlight that even though the SM is a really successful theory, all parameters must still be determined from fits to experimental data. For instance, the interaction couplings, Weinberg angle, all the masses as well as the three mixing angles and phase from the CKM matrix must come from data. This fact is an excellent example of the close relationship between theory and experiments and it emphasizes the necessity of tuning in physics.

2.2 Neutrino interactions with nucleons

The most simple targets besides fundamental particles are nucleons. When neutrinos interact with nucleons, the nucleon internal structure must be taken into account. The information on the charge distribution and magnetization density is encapsulated within form factors. Experimentally², form factors can be extracted by comparing neutrino-hadron interactions to the alternative point-like interaction. Hence, the form factor informs on how different the hadron is from a point-like particle. The kinematic variables used to describe neutrino interactions on a free nucleon are defined in Fig. 2.1.

Neutrino CC interactions with a nucleon can be classified into three different regions. A graphical representation of the differential cross section as a function of the invariant mass is shown in Fig. 2.2. At low W , neutrinos interact via a quasielastic

²Mathematically, form factors are interpreted as the Fourier transform of a density distribution.

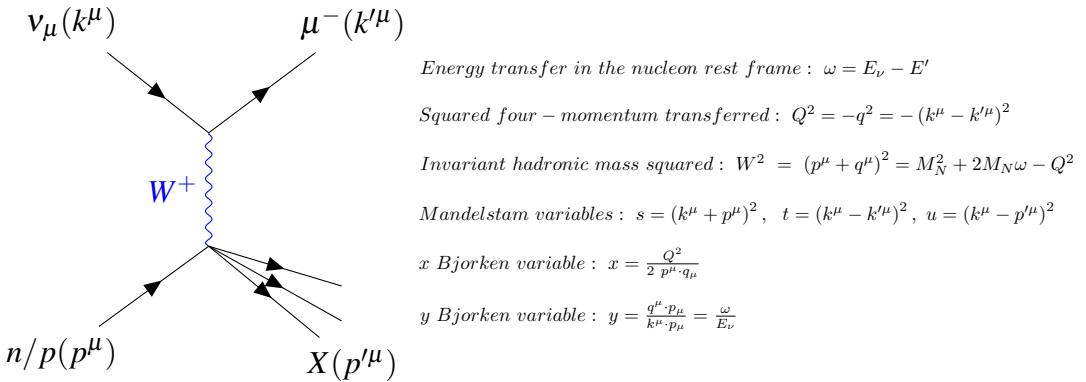


Figure 2.1: Kinematic variables for a ν_μ CC interaction on a nucleon. The four-momentum of each particle is specified in parenthesis. E_ν and E' stand for neutrino and muon energy respectively. M_N is the mass of the initial nucleon. The hadronic final state is specified with X . Next to the diagram, some relevant Lorentz invariant quantities for the description of neutrino interactions on free nucleon are presented. These variables can also be used to define any lepton-nucleon scattering. The kinematics of the final state lepton can be constrained given a pair of Lorentz invariant variables, such as Q^2 and W^2 and the information on the initial state. The energy transfer in the nucleon rest frame can be referred to as ω or v , depending on the literature source.

(QEL) interaction³. In the particular case of a QEL interaction with a nucleon of mass M_N , the invariant mass coincides with the mass of the outgoing nucleon, $W = M_{N'}$. This is observed as a peak in the differential cross section, see Fig. 2.2. At higher W , it is possible to excite the final state into a resonant state. This is known as resonant (RES) interaction. The lightest resonance, the $\Delta(1232)$, is the main contribution to the RES cross section. Resonances usually decay into pions, however, other decay modes leading to kaon, hyperon or eta production are also possible at a more reduced rate. Besides RES processes, pions can also be produced via nonresonant background (NRB) interactions in the same W region as RES. These processes define the NRB background. Finally, if we keep increasing W , we enter the deep-inelastic scattering (DIS) region, where it is possible to resolve individual quarks. The transition region between Δ peak and DIS defines the shallow-inelastic scattering (SIS) region. Even though is not clear where the boundary between SIS and DIS interactions is, it is considered that the pure DIS region starts at $W > 2 \text{ GeV}/c^2$. The corresponding Feynman diagrams for muon-neutrino on nucleon CC QEL, RES and DIS interactions are shown in Fig. 2.2 (a-c) respectively.

This section offers an overview of the different cross-section models available for each scattering region with special emphasis on their implementation in the GENIE

³This division is common between CC and NC interactions, with the exception of NC interactions at low W , for which neutrinos interact via Elastic scattering.

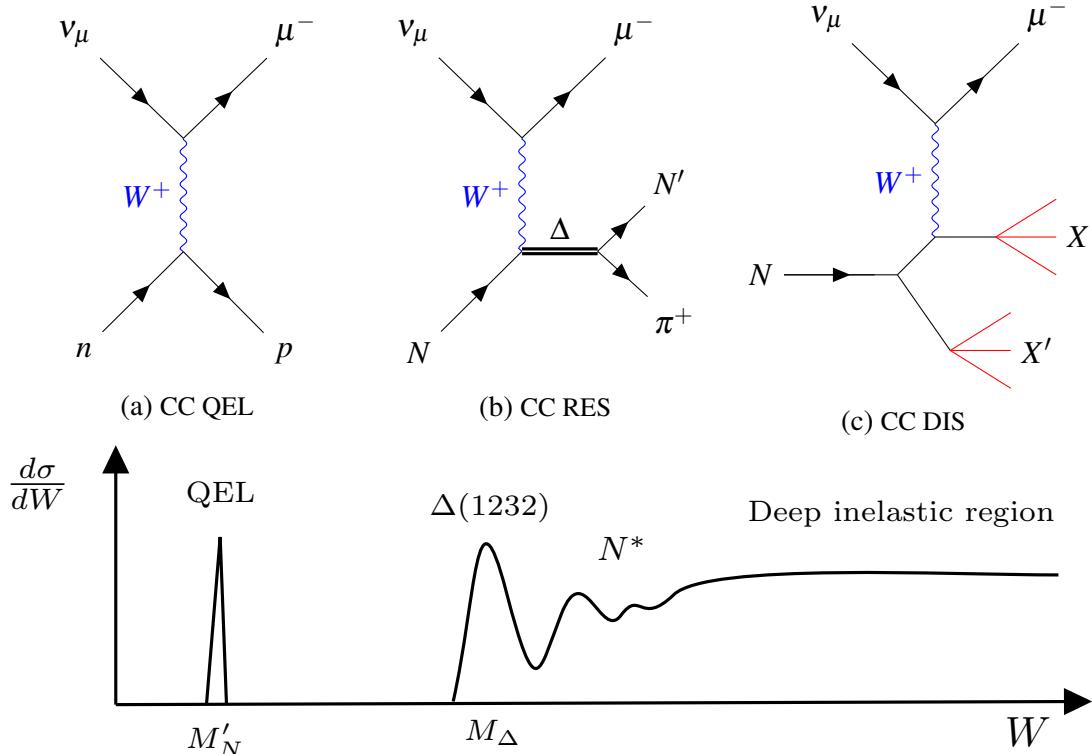


Figure 2.2: Examples of Feynman diagrams for possible charged current muon neutrino weak interaction on free nucleon (a-c). N stands for nucleon. X are showers of hadrons produced as a consequence of the hadronization (red lines). (d) Shows a graphical representation of the differential neutrino-nucleon cross section as a function of the invariant hadronic mass, W . The different regions of interest are highlighted. The differential cross section for NC interactions with free nucleons is similar to the CC one, with changes due to the different particles involved in the interaction. For instance, the elastic peak would be at $W = M_N = M_{N'}$.

event generator.

2.2.1 Modeling of neutrino charged-current quasi-elastic interactions

A QEL interaction is characterized by the production of a charged lepton and a (semi-)stable baryon in its ground state. In a (anti)neutrino CCQEL interaction on free nucleons, the dominant reactions are:

$$\nu_f + n \rightarrow f^- + p \quad (2.8)$$

$$\bar{\nu}_f + p \rightarrow f^+ + n \quad (2.9)$$

In the particular case of QEL interactions, $W = M_{N'}$, being $M_{N'}$ the mass of the final state nucleon. In terms of the energy transfer, $\omega = (M_N - M_{N'} - q^2)/2M_N \sim Q^2/2M_N$.

In the laboratory frame, where the nucleon is at rest, the CCQEL cross section is given by:

$$\frac{d\sigma^{CCQEL}}{dq^2} = \frac{G_F^2 M_N^2 |V_{ud}^{CKM}|^2}{8\pi E_\nu^2} \left[A(q^2) \mp B(q^2) \frac{(s-u)}{M_N^2} + C(q^2) \frac{(s-u)^2}{M_N^4} \right] \quad (2.10)$$

s and u are the Mandelstam variables defined in Fig. 2.1. In this case, they satisfy $s - u = 4M_N E_\nu + q^2$. V_{ud}^{CKM} is the CKM matrix element for the $u \rightarrow d$ transition. For (anti)neutrino CC QEL interactions, Eq. 2.10 has a $-$ ($+$) sign.

The $A(q^2)$, $B(q^2)$ and $C(q^2)$ are related to the nucleon form factors as:

$$\begin{aligned} A(q^2) &= \frac{m_f^2 - q^2}{4M_N^2} \left[\left(4 + \frac{q^2}{M_N^2} \right) g_A^2(q^2) - \left(4 + \frac{q^2}{M_N^2} \right) f_1^2(q^2) \right. \\ &\quad - \frac{q^2}{M_N^2} \left(1 + \frac{q^2}{4M_N^2} \right) f_2^2(q^2) - \frac{4q^2}{M_N^2} f_1(q^2) f_2(q^2) \\ &\quad - \frac{m_f^2}{M_N^2} \left((f_1(q^2) + f_2(q^2))^2 + (g_1(q^2) + 2g_3(q^2))^2 \right. \\ &\quad \left. \left. + \left(\frac{q^2}{M_N^2} - 4 \right) g_P^2(q^2) \right) \right] \end{aligned} \quad (2.11)$$

$$B(q^2) = -\frac{q^2}{M_N^2} g_A^2(q^2) (f_1(q^2) + f_2(q^2)), \quad (2.12)$$

$$C(q^2) = \frac{1}{4} \left(g_A^2(q^2) + f_1^2(q^2) - \frac{q^2}{4M_N^2} f_2^2(q^2) \right) \quad (2.13)$$

where m_f is the final lepton mass. In total, four nucleon form factors must be deter-

mined: the vector and weak magnetic form factors, $f_1(q^2)$ and $f_2(q^2)$, and the axial and pseudoscalar form factors, $g_A(q^2)$ and $g_P(q^2)$ respectively.

The vector and weak magnetic form factors are related to the magnetic, $G_M(q^2)$, and electric, $G_E(q^2)$, Sachs form factors, as:

$$f_{1,2}(q^2) = F_{1,2}^p(q^2) - F_{1,2}^n(q^2) \quad (2.14)$$

$$F_1^{p,n}(q^2) = \frac{G_E^{p,n}(q^2) + \tau G_M^{p,n}(q^2)}{1 + \tau} \quad (2.15)$$

$$F_2^{p,n}(q^2) = \frac{G_M^{p,n}(q^2) - \tau G_E^{p,n}(q^2)}{1 + \tau} \quad (2.16)$$

being $\tau = -q^2/2M_N$. In the Berit frame⁴, the charge, $G_E(q^2)$, and magnetization, $G_M(q^2)$, form factors are interpreted as the Fourier transforms of the charge density and the magnetization density. Therefore, in the limit of $q^2 \rightarrow 0$, $G_E^{p,n}(0) = Q_{p,n}$, which is 1 for protons and 0 for neutrons, and $G_M^{p,n}(0) = \mu_{p,n}$, being $\mu_p = 2.79278$ and $\mu_n = -1.91315$ the magnetic moment of the neutron and proton respectively [22]. The electric and magnetic Sachs form factors dependence on q^2 can be parametrized using different functions. One possible approach is the dipole approximation [23]:

$$G_E(q^2) = G_E^p(q^2) = (1 - q^2/M_V^2)^{-2} \quad (2.17)$$

$$G_E^n = 0 \quad (2.18)$$

$$G_M^{p,n}(q^2) = \mu_{p,n} G_E(q^2) \quad (2.19)$$

This parametrization depends on one free parameter, M_V , known as vector mass. The same parametrization is used to describe electromagnetic scattering of electrons with nucleons [24]. In fact, the vector mass is obtained from fits to electron-proton scattering and electron scattering on light nuclear targets. The best fit value for M_V from electro-scattering data is $M_V = 0.84 \pm 0.04$ GeV/c² [1]. However, as more data became available, deviations from the simplistic dipole shape approximation were observed. A better description of electron scattering data is provided by the BBBA-05 [25], the BBBA-07 [26], Kelly [27] or Punjabi [28] parametrizations. GENIE implemented the BBBA-07 parametrization to describe the vector form factor.

The the pseudo-scalar form factor can be expressed as a function of the axial form factor as:

$$g_P(q^2) = \frac{2M_N g_A(q^2)}{m_\pi^2 - q^2} \quad (2.20)$$

where m_π is the pion mass. This is known as the Goldberger-Treiman relation. The

⁴In the Breit frame, the mass of the target is much greater than the one from the projectile.

remaining form factor, the axial form factor, must be parametrized. The standard approach is to use the dipole approximation:

$$g_A(q^2) = \frac{g_A(0)}{\left(1 - \frac{q^2}{M_A^2}\right)^2} \quad (2.21)$$

$g_A(0)$ is measured from neutron β decay data, being $g_A(0) = -1.2695 \pm 0.002$ [29]. The axial mass can be obtained from fits to CCQEL neutrino and antineutrino scattering data on hydrogen and deuterium from old bubble chamber experiments. The world's fit value is $M_A = 1.014 \pm 0.014 \text{ GeV}/c^2$ [26].

Even though this parametrization is still the most used nowadays, new models are developed to improve the description of the axial form factor. An alternative is to parametrize the form factors with the z-expansion formalism [30], which describes the form factor shape using a convergent Taylor expansion, defined as $F(q^2) = \sum_{k=0}^{\infty} a_k [z(q^2)]^k$ where, $z(q^2) = (\sqrt{t_{cut} - q^2} - \sqrt{t_{cut} - t_0}) / (\sqrt{t_{cut} - q^2} + \sqrt{t_{cut} - t_0})$. Here, a_k are dimensionless coefficients that encode the nucleon structure, t_{cut} is the mass of the lightest state to be produced, and t_0 is a free parameter set as convenient. This parametrization is also available for use in GENIE.

Hence, Eq. 2.10 altogether with the parametrization of the vector and axial form factors, describes our current understanding of CCQEL interactions on free nucleons. This theory is referred to as Llewellyn-Smith model (LS) [31] and it is implemented in all neutrino MC generators. Provided that most of the free parameters are well described by electron-scattering data, neutron scattering and β decay experiments, the main uncertainty is due to the determination of the axial form factor.

The same methodology can be followed to derive a model for neutrino NC elastic interactions. Now, the parametrization of the neutral-current vector and axial form factors also involves the strangeness vector form factors [1]. In GENIE, NC elastic interactions are simulated using the Ahrens model [32].

Neutrino CCQEL interactions can lead to hyperon production (Δ^0 , Σ^- and Σ^0). This process is covered in GENIE with the Pais model [33].

2.2.2 Modeling of charged-current resonant interactions

A resonance production process in a (anti)neutrino CC interaction with a free nucleon is defined by one of the following processes:

$$\nu_f + n \rightarrow f^- + \mathcal{R}^+ \quad (2.22)$$

$$\nu_f + p \rightarrow f^- + \mathcal{R}^{++} \quad (2.23)$$

$$\bar{\nu}_f + n \rightarrow f^+ + \mathcal{R}^- \quad (2.24)$$

$$\bar{\nu}_f + p \rightarrow f^+ + \mathcal{R}^0 \quad (2.25)$$

where \mathcal{R} is a resonance excitation, such as the $\Delta(1232)$ or others. Since resonances are unstable, they will eventually decay.

Different models exist for the evaluation of the RES cross section. The most common model in MC event generators is the Rein-Seghal model (RS) model [34], which is the base for more elaborate approaches. In the original RS paper, their goal was to describe single-pion resonance production up to $W = 2 \text{ GeV}/c^2$ whilst taking into account the interference between nearby resonances. The model assumes that the final lepton is massless and, on top of the resonance contribution, they include a NRB background with isospin 1/2 which is added incoherently.

In GENIE, a modified version of the original RS model is implemented: the RS model is used to calculate the cross section for each resonance up to $W = 2 \text{ GeV}/c^2$. This procedure ignores interference between nearby resonances and the NRB background is added using a different approach (see Sec. 2.2.4). The last step is to decay the resonance into a stable or semi-stable final state. The main advantage of the GENIE implementation is that it covers single and multi-pion production processes after a RES interaction.

The main ingredient is the production cross section for a single resonance. In the RS model, it is given by:

$$\frac{d^2\sigma^{RS}}{dQ^2 dW^2} = \frac{G_F^2 |V_{ud}^{CKM}|^2}{8\pi^2} \left(\frac{1}{1 + \omega^2/Q^2} \right) \kappa (u^2 \sigma_L + v^2 \sigma_R + 2uv \sigma_S) \quad (2.26)$$

for neutrino interactions, whilst for antineutrinos it is:

$$\frac{d^2\sigma^{RS}}{dQ^2 dW^2} = \frac{G_F^2 |V_{ud}^{CKM}|^2}{8\pi^2} \left(\frac{1}{1 + \omega^2/Q^2} \right) \kappa (u^2 \bar{\sigma}_R + v^2 \bar{\sigma}_L + 2uv \bar{\sigma}_S) \quad (2.27)$$

where $\kappa = (W^2 - M_N^2)/2M_N$, being M_N the resonance mass. The variables u and v are defined as $u = (E_\nu + E' + Q)/2E_\nu$ and $v = (E_\nu + E' - Q)/2E_\nu$. These quantities do not

correspond to the u -Mandelstam variable and energy transfer, but they were introduced in the original paper as an abbreviation. The same nomenclature is used in this text.

The neutrino resonant cross section differs from the antineutrino one as left-handed and right-handed currents are interchanged. The cross section is given in terms of helicity cross sections that describe the absorption of the propagator with positive, negative or zero helicity, σ_L , σ_R or σ_S respectively,

$$\sigma_{L/R} = \frac{\pi M_R}{2M_N} \frac{1}{\kappa} \sum_{j_z} |\langle N, j_z \mp 1 | F_{\pm} | N^*, j_z \rangle|^2 \delta(W - M_R), \quad (2.28)$$

$$\sigma_S = \frac{\pi M_N}{2M_R} \frac{1}{\kappa} \frac{|\mathbf{q}|^2}{Q^2} \sum_{j_z} |\langle N, j_z | F_0 | N^*, j_z \rangle|^2 \delta(W - M_R). \quad (2.29)$$

\sum_{j_z} is the sum over all available resonance spin states. The terms $\langle N^* | F_{\pm,0} | N \rangle$ are known as helicity amplitudes. These contain information on the form factors, which depend on the resonance spin. The corresponding helicity amplitudes calculated within the RS model are tabulated for all resonances up to $W = 2 \text{ GeV}/c^2$ in Ref. [34].

Some of the parameters involved in the description of the amplitudes need to be extracted from data. Particularly, the helicity amplitudes are proportional to the resonant vector and axial form factors $G^V(q^2)$ and $G^A(q^2)$ form factors which are parametrized with a modified dipole approximation:

$$G^{V,A}(q^2) \propto \left(1 + \frac{Q^2}{4M_N^2}\right)^{1/2-n} \left(1 + \frac{Q^2}{M_{V,A}^2}\right)^{-2}. \quad (2.30)$$

The integer n represents the number of oscillator quanta present in the final resonance. In the original paper, the vector mass is set to be the same than the one obtained in QEL parametrization of the electric form factor: $M_V = 0.84 \text{ MeV}/c^2$. The resonance axial mass value is extracted from neutrino pion production data. The world average for CC RES, $M_A = 1.12 \pm 0.03 \text{ GeV}/c^2$ [35], is the value used in most MC event generators.

The term $\delta(W - M_R)$ describes a resonance with negligible width. This is known as *narrow-width approximation*. In order to take into account that resonances have a non-null width, Γ_R , the δ -function can be substituted by a Breit-Wigner factor,

$$\delta(W - M_R) \rightarrow \frac{1}{2\pi} \frac{\Gamma_R}{(W - M_R)^2 + \Gamma_R^2/4}. \quad (2.31)$$

The widths for resonances up to $W < 2 \text{ GeV}/c^2$ are extracted from pion-nucleon scattering data from Ref. [36]. In addition to the 18 resonances of the original publication, GENIE includes some more resonances from the PDG baryon tables [37].

The RES cross section for a given resonance is given by integrating Eq. 2.26 over SIS region defined by [34]:

$$W_{min} = M_N \leq W \leq W_{max} = \sqrt{s} - m_l = \sqrt{2M_N E_V + M_N^2} - m_l \quad (2.32)$$

$$Q_{min}^2 \leq Q^2 \leq Q_{max}^2 \quad (2.33)$$

$$Q_{min}^2 = \frac{(s - M_N^2 - m_l^2)}{2s} \left((s + m_l^2 - W^2) - \sqrt{(s + m_l^2 - W^2)^2 - 4sm_l^2} \right) - m_l^2$$

$$Q_{max}^2 = \frac{(s - M_N^2 - m_l^2)}{2s} \left((s + m_l^2 - W^2) + \sqrt{(s + m_l^2 - W^2)^2 - 4sm_l^2} \right) - m_l^2$$

The RS model implementation in GENIE corrects the total CC RES cross section for the lepton mass, which is neglected in the original RS model [37]. This correction is included in the RES limits presented in Eq. 2.32 and Eq. 2.2.2. Despite the number of approximations, this model is widely used in the neutrino community to describe RES interactions.

An extension of the RS model is the Berner-Seghal model (BS), which takes into account the final lepton mass, m_f [38]. The BS differential cross section is expressed in terms of six helicity cross sections, σ_+^λ , σ_-^λ and σ_0^λ , as:

$$\frac{d^2\sigma^{BS}}{dQ^2 dW^2} = \frac{G_F^2 |V_{ud}^{CKM}|^2}{8\pi^2 M_N} \left(\frac{1}{1 + \omega^2/Q^2} \right) \sum_{\lambda=+,-} \left[\left(c_L^\lambda \right)^2 \sigma_L^\lambda + \left(c_R^\lambda \right)^2 \sigma_R^\lambda + \left(c_S^\lambda \right)^2 \sigma_S^\lambda \right] \quad (2.34)$$

where $\lambda = +, -$. Helicity cross sections are calculated following the recipes from the RS model, defined in Eq. 2.28 and Eq. 2.29, with slightly different dynamical form factors. In the limit $m_l \rightarrow 0$, $c_L^- \rightarrow u$, $c_R^- \rightarrow v$, $c_S^- \rightarrow \sqrt{2uv}$, and $c_{L,R,S}^+ \rightarrow 0$, recovering Eq. 2.26. In the BS formalism, the cross section is suppressed at forward lepton angles. This suppression is particularly important when considering neutrino interactions that produce heavy leptons, such as τ .

Neither of these two models considers the interaction between resonances and the NRB background. This is possible with the Monireh Kabirnezhad model (MK) [39]. The MK model is the first resonance model that incorporates interference effects in current event generators and it is now available for wider use in GENIE [40].

The RES models can be used to describe NC RES interactions provided the NC helicity amplitudes, which are tabulated in Ref. [34].

2.2.3 Modeling of charged-current deep inelastic scattering interactions

At very high Q^2 , the neutrino can resolve the constituents of the nucleon, the quarks. Such interaction, known as DIS, is characteristic at the $W \geq 2 \text{ GeV}/c^2$ and $Q^2 \geq 1 \text{ GeV}^2/c^2$ region. As a result of the interaction and the colour confinement nature of the strong force, the nucleon breaks up and quarks materialize into a jet of hadrons, X . An example of a (anti)neutrino CC DIS interaction with a nucleon is given by:

$$\nu_f + N \rightarrow f^- + X \quad (2.35)$$

$$\bar{\nu}_f + N \rightarrow f^+ + X \quad (2.36)$$

In order to describe such interactions in event generators, this process is factorized into two models: a model that describes the DIS interaction with quarks from the nucleon, and a hadronization model that determines the nature and kinematics of the jet shower. In GENIE, hadronization is simulated using the Andreopoulos-Gallagher-Kehayias-Yang (AGKY) hadronization model [41]. The details on the AGKY model of interest for this work are discussed in Chapter 4. In this section, we focus on the description of the DIS interaction model.

The CC DIS double differential cross section is given by [42]:

$$\begin{aligned} \frac{d^2\sigma^{DIS}}{dxdy} = & \frac{G_F^2 s}{2\pi} \left(\frac{M_W^2}{M_W^2 + Q^2} \right)^2 \left[xy F_1^{vN}(x, Q^2) \right. \\ & \left. + \left(1 - y - \frac{M_N xy}{2E_\nu} \right) F_2^{vN}(x, Q^2) \pm xy \left(1 - \frac{y}{2} \right) F_3^{vN}(x, Q^2) \right] \end{aligned} \quad (2.37)$$

In the limit $m_f \rightarrow 0$, the DIS phase-space is defined by $0 < x < 1$ and $0 \leq y \leq 1/(1 + xM_N/2E')$.

$F_i(x, Q^2)$ are dimensionless structure functions that are related to the internal structure of nucleons. In the limit of $Q^2 \rightarrow \infty$ and $\omega \rightarrow \infty$, the structure functions are independent of Q^2 and ω and they can be expressed in terms of the Bjorken variable x only. This fact, known as Bjorken scaling⁵ [44], implies that the neutrino interacts with point-like particles, the quarks and gluons within the nucleons⁶. In this regime, the structure functions can be expressed in terms of the internal parton distribution functions (PDFs), which represent the probability of interacting with a specific quark

⁵Bjorken scaling was first observed using electro scattering on protons and neutrons in late 1960s. This phenomenon was later observed with neutrino scattering data from CERN [43].

⁶Originally, these particles were called "partons".

or antiquark carrying a fraction x of the nucleon momentum.

Deviations from this simplified picture exist. At low Q^2 , Quantum-Chromodynamics (QCD) becomes non-perturbative. Interactions between partons via gluon exchange introduce a Q^2 dependence in the nucleon structure functions [45–47]. This dependency can be ignored at high energies, but it becomes important at the energies of interest for neutrino experiments. Moreover, the structure functions must be modified to take into account next-to-leading order, next-to-next-to-leading order terms, target-mass corrections (TMC) or kinematic higher energy twists, and higher twist (HT) effects⁷ [48]. HT effects are a consequence of multiparticle correlations [49, 50].

In GENIE, CC and NC DIS interactions are described with the Bodek and Yang model (BY) model, which takes into account this effects [51]: it uses an effective leading order model with modifications of the structure functions that allow the description of the low Q^2 region. The structure functions are tuned against e/μ -N and $\nu/\bar{\nu}$ -A data using either leading-order or next-to-leading-order calculations. The theory also includes TMC and HT corrections. The BY model applies the same corrections to the axial and vector form factors. This introduces some uncertainty for neutrino interactions. Full details on the model can be found in Ref. [51].

2.2.4 Modeling of the shallow-inelastic scattering region

On top of the resonances, there is a continuum of nonresonant pion production processes, known as nonresonant background (NRB). Experimentally, it is not possible to disentangle the NRB background from RES as these belong to the same kinematic region. The region between the Δ peak to DIS is known as the SIS.

The SIS region plays an important role in neutrino oscillation experiments. Particularly, experiments such as DUNE expect around 50% of events from the SIS region [42]. However, there is no theoretical model available to simulate RES and NRB processes altogether. Hence, its implementation in neutrino event generators varies as there is no clear guidance.

Even though some RES models, such as the RS model, aim to describe RES and NRB pion production, some studies suggest that the RS approach is not sufficient to describe the NRB background [52]. Instead, the standard approach in event generators is to use a phenomenological description based on *quark-hadron duality*, also known as duality. Bloom and Gilman introduced the concept of duality by comparing the structure functions from inclusive electron-nucleus DIS scattering with resonance production in similar experiments [53]. They observed that the average over resonances is

⁷TMC and HT effects are relevant at low Q^2 and high x .

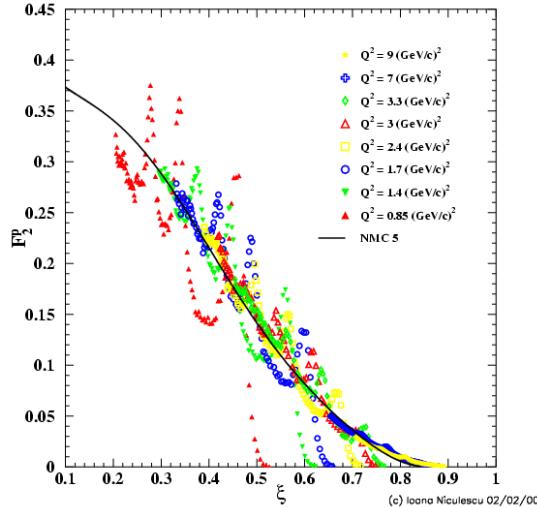


Figure 2.3: Comparison of F_2^p as a function of the Nachtmann variable ξ for a series of resonances at a given Q^2 . The data is compared to the extrapolated DIS measurement (black line) from the New Muon Collaboration (NMC) at 5 GeV [54]. Figure from Ref. [55].

approximately equal to the extrapolated results from DIS events. In terms of structure functions, this duality-condition translates as:

$$\int_{\xi_{min}}^{\xi_{max}} d\xi F_i^{RES}(\xi, Q_{RES}^2) \approx \int_{\xi_{min}}^{\xi_{max}} d\xi F_i^{DIS}(\xi, Q_{DIS}^2) \quad (2.38)$$

where ξ is the Nachtmann variable defined as $\xi = 2x/(1 + \sqrt{1 + 4x^2 M_N^2/Q^2})$. The super-scripts *RES* and *DIS* refer to kinematic values in the RES and DIS regions, whilst the sub-script *i* refers to the target type, i.e. proton. In addition, $\xi_{min} = \xi(W_{max}, Q_{RES}^2)$ and $\xi_{max} = \xi(W_{min}, Q_{RES}^2)$, being W_{min} and W_{max} the limits of the SIS region defined in Eq. 2.32 and Eq. 2.2.2. An example of duality for the F_2^p structure function⁸ is shown in Fig. 2.3. It is observed that for different values of Q^2 , the resonance production data is well described in average by the extrapolated DIS measurement.

Duality has been shown to be valid for e -N and e -A data [55]. The study of quark-hadron duality for neutrino interactions is limited to a few low-statistical bubble chamber data on hydrogen and deuterium from the 70's and 80's, which is scarce above the Δ region, restricting our understanding of duality for neutrinos. There is a growing effort to validate duality on neutrino interactions [56]. Until now, some studies show that duality may be observed for proton targets, but not for neutron targets [55]. Despite the little evidence for duality in neutrino interactions, describing the NRB

⁸The F_2^p structure function for RES interactions is related to the $F_{\pm,0}$ structure functions via polarization vectors.

background using duality is a common approach in neutrino MC generators. However, each MC generator models SIS differently. In GENIE, the NRB background is modeled by extrapolating the DIS contribution to the RES region. In order to avoid the double-counting of the RES cross section, this contribution is scaled depending on the multiplicity of the final state and it is tuned to data. Further details are given in Chapter 3.

2.3 Neutrino interactions with nuclei

Most neutrino physics experiments are using complex nuclei as targets. These experiments have the capability of collecting high statistics neutrino samples due to the target high density but incorporate the additional difficulty of understanding the nuclear medium.

The general approach to model neutrino interactions with nuclei is to simplify the problem by considering the nuclear cross section an incoherent sum of free nucleon cross sections. This approach, known as plane wave impulse approximation (PWIA) or Impulse Approximation, was first introduced in electro-scattering interactions [57–59]. The PWIA is built under the following assumptions:

1. The neutrino interacts with a single nucleon within the nucleus
2. The interaction is described by the free nucleon cross section
3. The outgoing particles can be described as noninteracting plane waves

The interaction with a nucleon within the nucleus differs from the free-nucleon interaction: the hit nucleon has a Fermi motion and it is bound. As a consequence of Fermi motion, some of the kinematic variables from Fig. 2.1 are no longer well defined. For instance, in neutrino-nucleon interactions, the energy transfer ω is calculated in the nucleon at the rest frame. In this case, as nucleons in the nuclei are bound, the nucleon at rest reference frame is no longer valid. The correct approach is to define the energy transferred in the nucleus rest frame. In order to avoid confusion, the energy transferred in the nuclei rest frame is denoted as q_0 .

In addition, the description of neutrino-nucleus interactions under the PWIA is not sufficient: because of nucleon-nucleon correlations, it is possible to transfer momentum to multiple nucleons in the same interaction (np-nh). Moreover, hadrons may suffer from final-state interactions (FSI) while leaving the nucleus. These processes, neglected in the PWIA, are essential for the correct description of neutrino interactions on heavy nuclei.

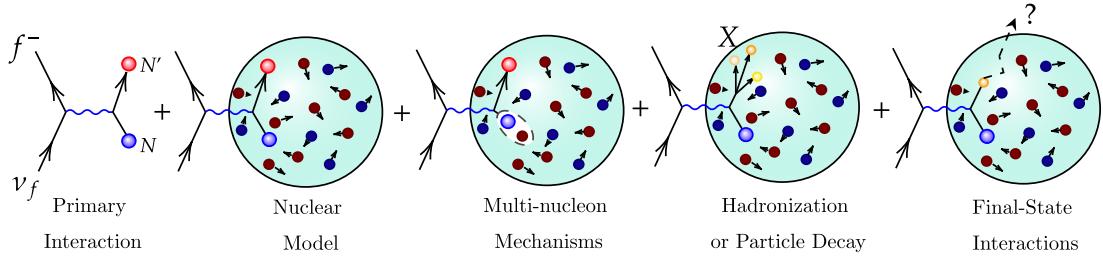


Figure 2.4: Modeling of neutrino-nucleus interaction in GENIE under the factorization approach. Each diagram represents each simulation layer.

Neutrino event generators rely on the *factorization* approach: neutrino-nucleus interactions are modeled by adding several simulation layers on the top of the neutrino-nucleon interaction. The neutrino interaction with a free nucleon, which is on-shell, is the starting point of the simulation chain. However, in the nuclear environment, the hit nucleon is off-shell. The required information to bind the hit nucleon is extracted from the nuclear model. The cross-section model, which includes multi-nucleon mechanisms, is used to determine the type of interaction and the kinematics of the final state lepton. Depending on the scattering mechanism, it may be required to decay (RES) or hadronize (DIS) the particles produced after the primary interaction. Finally, hadrons are propagated throughout the nucleus taking into account possible FSI. The different simulation layers are shown in Fig. 2.4. In this section, the details on the implementation of the layers specific to neutrino interactions on nuclei are discussed.

2.3.1 The nuclear model

The nuclear environment is complex: nucleons are neither at rest nor independent from one another. In particular, the nucleon binding energy, Fermi motion and Pauli blocking are essential to describe the nuclear ground state. This information is embedded in the nuclear model, which is a four-dimensional distribution of the hit-nucleon momentum in the laboratory frame and its binding energy. Pauli blocking for the outgoing nucleon is handled on an event-by-event basis in the cross-section model.

There are different ways to model the nuclear ground state. Within GENIE, three nuclear models are available: the Relativistic Fermi gas (RFG), the Local Fermi gas (LFG) and the Correlated Fermi gas (CFG) model. The prediction from each model for the initial nucleon momentum distribution is shown in Fig. 2.5.

The RFG is the simplest model available in GENIE. The model assumes a constant nuclear density and it describes the motion of noninteracting fermions in a constant potential well. Each of the available states is filled up to the Fermi momentum according the Pauli exclusion principle. As a consequence of short-range correlations,

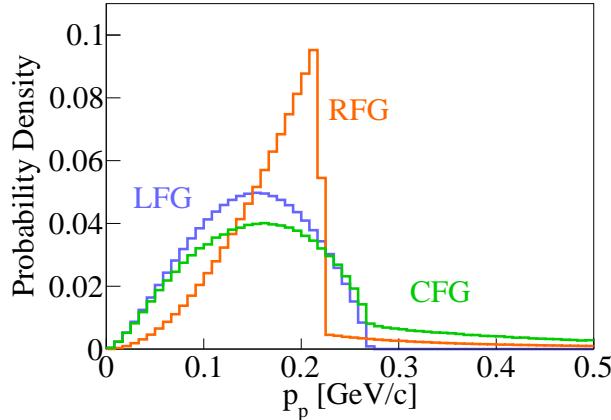


Figure 2.5: Initial nucleon momentum distribution predictions for a carbon target using GENIE’s nuclear ground models. This plot belongs to Ref. [40].

there is a finite probability of having some nucleons above the Fermi momentum in the ground state. This includes a high momentum tail in the probability distribution which is supported by electron scattering data [60]. This tail is included in the RFG implementation in GENIE with the Bodek and Ritchie model [61].

The LFG and CFG models have been included in GENIE recently [40]. The LFG is an extension of the RFG model that considers the changing density of the nucleus. As a consequence, the Fermi momentum depends on the distance from the centre of the nucleus. The implementation of the LFG model in GENIE neglects the high-momentum tail [40]. The CFG model is based on the LFG approach but takes into account the high-momentum tail due to nucleon correlations [62].

The three nuclear models available in GENIE use a constant binding energy, which depends on the target. However, nuclear models can predict a correlation between the four-momentum and the binding energy of the hit-nucleon [63]. One possibility is to use the Benhgar spectral function model [64], which is built using QEL electron-scattering data and takes into account short-range correlations. The main limitation of this model is that it is nucleus specific and difficult to scale for other nuclei. This complicates its implementation in a MC event generator which aims to describe neutrino interactions with any nuclei.

2.3.2 Multi-nucleon production processes

Nucleons in the nuclei are highly correlated as a consequence of the exchange of ρ mesons and pions. These are commonly known as short-range and long-range correlations. In many-body theory, this is handled by adding multi-particle interaction terms in the Hamiltonian. These terms take into account many possible interaction processes that would produce multi-nucleon emission. An effect of correlations is the

modification of the week-coupling strength from their free-nucleon values [65].

Because of Pauli blocking, the nucleon can only be excited to an available state, creating a *hole* in the previous occupied state. The nucleus, now in an excited state, will decay back to the ground state. This process is known as one-particle-one-hole (1p1h) interaction. This is often referred to as QEL, as the final state before FSI is the same. With respect to QEL on free nucleus, 1p1h interactions are suppressed at low- Q^2 .

In GENIE, 1p1h cross sections are computed according to the Forest prescription [66]: the cross section is calculated using the LS form factors with some modifications. In order to put the struck nucleon off-shell, the energy transfer is modified to account for the binding energy of the struck nucleon. Long-range correlations are handled with the random-phase approximation (RPA) [66, 67]. Pauli blocking as well as Coulomb distortion effects, which take into account self-energy of the charged lepton in the nuclei, are also included [68].

This scenario is even more complex when taking into account nucleon-nucleon correlations: now it is possible to transfer energy to multiple nucleons in the same interaction (np-nh). The main contribution from np-nh interactions comes from two-particles-two-holes (2p2h) interactions, which dominate the region between the QEL peak and the Δ peak, known as *dip region*.

The multi-nucleon models available in GENIE are inclusive models: these only predict the kinematics of the final state lepton. However, event generators must predict the kinematics of all final state particles. The standard procedure to simulate 2p2h events in MC event generators is the following:

1. Obtain the information on the final-state lepton kinematics from an accept-reject loop⁹.
2. Assign the four-momentum transferred to the nucleon pair, which is treated as a pseudo-particle. In GENIE, we refer to the pseudo-particle as *nucleon cluster*.
3. Infer the cluster type (pn, nn, pp) from the lepton kinematics and the 2p2h cross-section model.
4. Obtain the initial four-momentum and binding energy for each nucleon in the cluster from the nuclear model. This information is used to determine the cluster initial kinematics.
5. Decay the cluster into two nucleons. This step defines the four-momentum of the outgoing nucleons.
6. Propagate the outgoing nucleons throughout the nucleus using one of the avail-

⁹In an accept-reject loop, the available kinematic space for the lepton is randomly sampled. In each iteration, the total cross section is computed. Only events with a cross section lower than the maximum cross section are accepted.

able FSI models in GENIE, described in Sec. 2.3.3.

GENIE has three multi-nucleon production models available: Empirical 2p2h model [69], Nieves Simo Vacas or Valencia model [70, 71], and SuSAv2 [72–74]. The Valencia model is the most used nowadays in neutrino collaborations. The SuSAv2 model is relatively new and it was added in the latest GENIE release.

The Empirical 2p2h model was the first multi-nucleon model available in GENIE [69]. The model is constructed by scaling the LS QEL cross section as a function of W and Q . The dependency on W is parametrized with a Gaussian distribution in the kinematic region between the QEL and the Δ peak. The Q dependency is similar to the dipole parametrization, but it scales with Q^{-6} instead. A benefit of this model is that it can be easily tuned to data.

Alternatively, the Valencia and SuSAv2 models are theory-driven. They are based on the many-body formalism which includes one- and two-body currents. Both models are able to describe 1p1h, 2p2h as well as Δ excitations in the medium. The main difference between the two models relays on the treatment of the Δ -resonance propagator [72]. In addition, the Valencia model is limited to $q_0 \leq 1.2$ GeV. For higher-momentum transfers, the 2p2h cross section is set to zero. SuSAv2 does not have this limitation and it is validated against electron-scattering data [75–77]. Consequently, the SUSAv2 approach has a better agreement not only with electron data, but also neutrino data [74]. Another distinction between the two models arises due to the different treatment of the off-shell nucleon in the GENIE implementation [40].

Multi-nucleon models, such as Valencia and SuSAv2, are complex and a complete calculation of the cross section in run time would be inefficient. In fact, most of the computing power would go into the computation of the hadron tensors. For this reason, MC generators use pre-computed hadron tensors. These, organized in tables, provide the information on the hadronic part of the interaction as a function of the energy, q_0 , and momentum transferred, q_3 , in the nucleus reference frame. The hadron tensor tables are only available for some isoscalar nuclei, such as ^{12}C or ^{16}O , which have the same number of protons and neutrons. The 2p2h hadron tensors are estimated for nonisoscalar nuclei, such as ^{40}Ar , by scaling the hadron tensors with closest atomic number [71]. For instance, the ^{40}Ar hadron tensors are obtained from the ^{20}Ca ones. The scaling factor depends on the type of nucleon pair. When interacting with a pn pair, the scaling factor, $S(pn|_Z^AX)$, is:

$$S(pn|_Z^AX) \equiv \left(P(pn|_Z^AX) / P(pn|_Z^{2Z}X) \right)^{1/2}, \quad (2.39)$$

being $P(pn|_Z^AX)$ the probability of having a given pn pair in a nuclei X with atomic

mass A and atomic number Z . The same approach is implemented to the nn or pp hadronic tensors. This procedure is implemented in the Valencia and SuSAv2 2p2h models, but not for the Empirical 2p2h model, which does not use pre-computed hadron tensors.

Substantial differences between the different GENIE 2p2h predictions exist. These are highlighted in Fig. 2.6. The models differ substantially in the prediction of the normalization and shape of the integrated cross section. These differences are important at high energies, $E_\nu > 1$ GeV, but also at lower energies. One reason for these discrepancies is the Valencia model limit of $q_0 < 1.2$ GeV, highlighted in Fig. 2.6 (b). In addition, there are important shape differences in the q_0 - q_3 region: the Valencia model predicts two peaks, one in the QEL region, and another one at the Δ region. The second is related to 2p2h processes involving Δ excitation diagrams. Alternatively, the GENIE Empirical model and SuSAv2 predict a single peak in q_0 - q_3 .

2.3.3 Final state interactions

Particles produced in the primary interaction are propagated throughout the nuclei using an FSI model which simulates hadronic interactions in the nuclear environment. FSI are essential to describe neutrino nuclei interactions: the probability of suffering an interaction for hadrons with kinetic energies below 1 GeV is expected to be around 40-50% for protons and up to 75% for pions¹⁰ [79, 80]. In fact, around 20% of pions produced in the nucleus are absorbed in the nuclear environment [37].

Different models are available to describe FSI [79, 80]. The most common approach in neutrino event generators is to use intra-nuclear cascade (INC) models. The logic behind the implementation of INC models is simple: each hadron is propagated through the medium following a series of steps until it leaves the nucleus. In each of these steps, the probability of undergoing an interaction is evaluated. This process is repeated until all hadrons are either absorbed or leave the nucleus.

The probability of travelling a distance λ depends on the position in the nucleus r as:

$$P(\lambda, r) = e^{-\lambda/\tilde{\lambda}(E, r)} \quad (2.40)$$

being $\tilde{\lambda}(E, r)$ the mean-free path (MFP) of the hadron in the nuclear medium,

$$\tilde{\lambda}(r) = (\sigma_p \rho_p(r) + \sigma_n \rho_n(r))^{-1} \quad (2.41)$$

where σ_p (σ_n) is the total cross section for interaction on a proton (neutron). $\rho(r)$ is

¹⁰The probability of pions interacting with matter depends strongly on the pion kinetic energy.

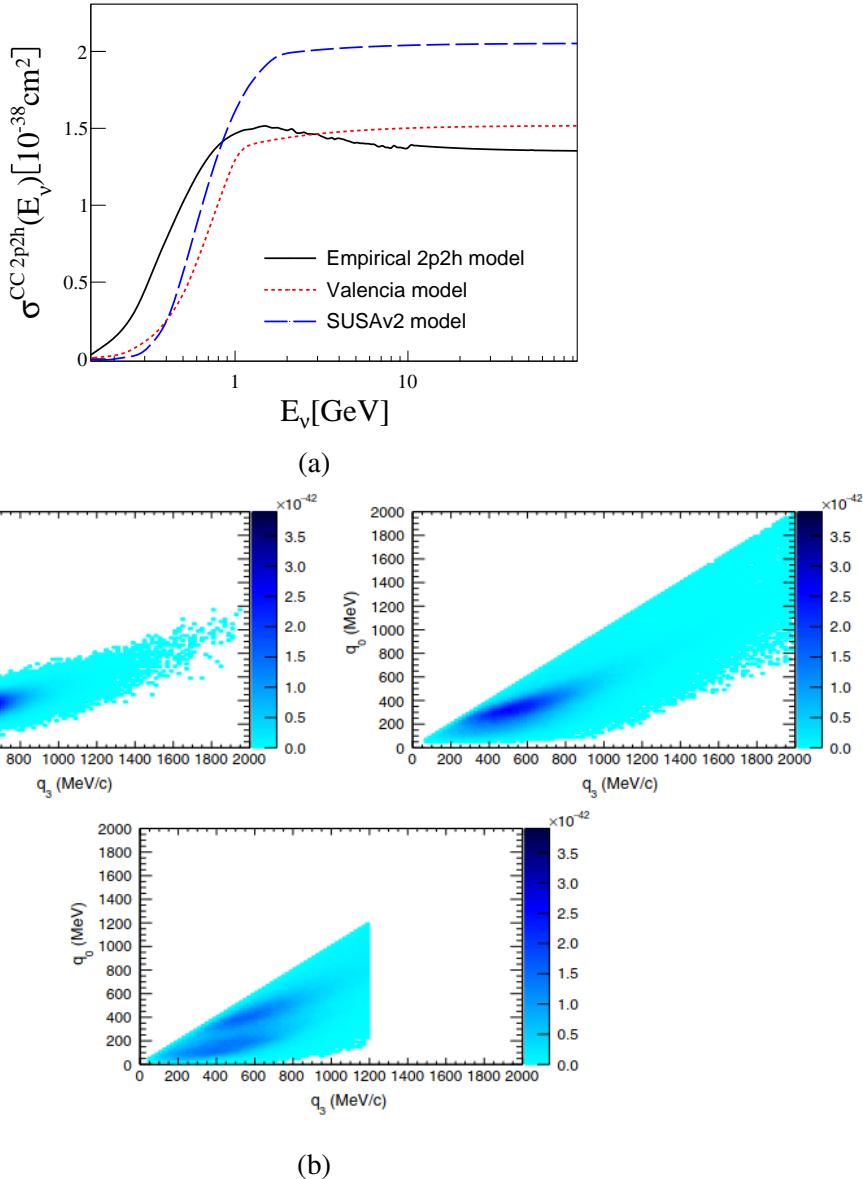


Figure 2.6: (a) Total CC 2p2h cross section on carbon for the available models in GENIE. (b) Double-differential flux-integrated 2p2h cross section as a function of the energy, q_0 , and momentum transferred, q_3 . The cross section is integrated over the T2K ND280 flux and it is given in units of cm^2/MeV^2 [78]. The plots correspond to the Empirical 2p2h (top left), SUSAv2 (top right) and Valencia (bottom) models. This plot belongs to Ref. [74].

the density of nucleons in the nucleus at a distance r from its centre. The step size, λ , must be small enough to guarantee constant nuclear density within the step path. In GENIE, $\lambda=0.05$ fm.

Hadrons produced in the nucleus do not immediately interact with the full cross section after they are created in the primary lepton scattering. Instead, while quarks hadronize in the nuclear environment, the interaction probability is dramatically reduced. This transition is known as the formation zone. In GENIE, the formation zone is modeled by adding a *free step* at the start of the cascade¹¹ in which no interactions can occur. This implementation has one free parameter which is determined from SKAT data [81].

Currently, GENIE has four FSI models based on the INC approach: *hA* [37], *hN* [40], *INCL++* [82, 83] and *GEANT4* [84]. The *INCL++* and *GEANT4* models have recently been added in GENIE [82, 83]. The details on the *hN*, *INCL++* and *GEANT4* models are not described here as these are not used in this thesis analyses.

The *hA* model is the first FSI model that was implemented in GENIE. This model is not strictly a cascade model: *hA* has a single step that is strongly based on data. It takes into account the total cross section for each process for nucleons and pions up to 1.2 GeV. Medium modifications are handled only for nucleon FSI with the Pandharipande-Pieper recipie [85]. The nuclear environment is approximated to an ensemble of isolated nucleons which have Fermi motion and binding energy, but Pauli blocking is not taken into account. The nuclear density is modeled with a Gaussian spatial nucleon density function for $A < 20$, and with the Woods-Saxon prescription for heavier nucleus [37]. Even though this model is more empirical and it is not strictly a cascade model, it successfully describes hadron-nucleus data. Its main benefit for neutrino experiments is that it treats the propagation as single step and, consequently, the model is easily reweightable.

2.4 Neutrino cross-section measurements

The study of neutrino interactions is a challenge for theorists and experimentalists alike. This section offers an overview of the evolution of neutrino experiments that study neutrino cross sections. These measurements are crucial for the validation and tuning of the theoretical models described in this chapter.

Cross-section measurements can be classified into inclusive and exclusive. This distinction depends on the event topology definition. In ν_μ CC inclusive measure-

¹¹A limitation of the INC implementation in GENIE is that the formation zone is not implemented for every step in the cascade.

ments, the event topology requires one μ^- , regardless of other final state particles. Alternatively, exclusive measurements also require a specific hadronic signature in the detector. The definition of the exclusive event topology depends on the analysis physics goals. For instance, in order to measure the ν_μ CC pion production processes, experiments search for events with a muon and pions in the final state.

Neutrino experiments measure the rate of neutrino interactions [4]. The total integrated cross section associated to a specific event topology, σ_{tot}^X , is determined from its associated event rate, N^X , and the neutrino flux, $\phi(E_\nu)$, as:

$$\sigma_{tot}^X(E_\nu) = \frac{N^X(E_\nu)}{T \cdot \phi(E_\nu)} \quad (2.42)$$

being T the number of target particles exposed to the flux. $\phi(E_\nu)$ corresponds to the neutrino flux at a given E_ν bin. This quantity is also known as flux-unfolded cross section. This approach introduces strong biases in the cross section measurement as a consequence of possible mismodeling of the neutrino flux and neutrino energy estimation.

The alternative approach is to measure flux-integrated cross sections [4]. The n-differential flux-integrated cross section for a given set of observables, Ω , is defined as:

$$\left(\frac{\partial^n \sigma^X}{\partial \Omega^n} \right)_i = \frac{N_i^X}{T \cdot \Phi \cdot \Delta \Omega_i} \quad (2.43)$$

$\Delta \Omega$ is the observable n-dimensional i-th bin volume. Φ , corresponds to integrated neutrino flux spectra over the whole beam energy range, $\Phi = \int \phi(E_\nu) dE_\nu$. The total flux-integrated cross section is obtained by integrating over the Ω phase-space. This approach does not unfold the flux from the event rate. Consequently, flux-integrated measurements obtained with different fluxes cannot be compared with each other. Instead, the flux prediction must be taken into account in the evaluation of MC predictions. The main benefit of flux-integrated measurements is that these datasets have better statistical properties and are more suited for tuning.

The different experimental efforts presented here are sub-divided into three eras: the 70's era or the bubble chamber era, the early modern era and the modern era. All the experiments discussed in this section face the same problems to measure the neutrino cross section, but these are handled in a slightly different way. The main difficulties that affect neutrino cross-section measurements are:

- Precise determination of the neutrino flux
- Neutrino energy reconstruction



Figure 2.7: Real photography of a ν_μ - e interaction candidate in Gargamelle [88].

- Treatment of statistical and systematic errors

These differences are important and must be accounted for when tuning with this data.

2.4.1 The 70's: Bubble chamber experiments

The first neutrino cross-section measurements were obtained with bubble chambers [86]. In a bubble chamber, the detector is filled with a superheated liquid. Charged particles crossing this medium leave a trace of tiny bubbles of vapour which were recorded using photographic machines. Back then, the computing capabilities were really limited and it was not possible to have automatic reconstruction software. Instead, potential neutrino events were scanned and measured by hand by a team of *bubble chamber scanners*¹² [87]. An example of a real picture of a neutrino interaction in the Gargamelle detector is shown in Fig. 2.7. The images from neutrino bubble chamber experiments are fascinating for the precision of the event topology: the spatial resolution can be of the order of 0.1 mm. In addition, it is possible to identify particles as the track bubble size is proportional to the energy loss.

Examples of neutrino bubble chambers were Gargamelle¹³, ANL 12 ft, BNL 7 ft, BEBC and FNAL. The focus of these experiments was to study a few GeV neutrino interactions, with the exception of FNAL, which used a neutrino beam whose flux peaked at 20 GeV. All these experiments measure neutrinos produced using the pion decay in flight method [89].

¹²This team was dominated by women who are usually forgotten in the history of neutrino physics. However, without such effort none of these datasets would exist, hence, they deserve a special mention in my thesis.

¹³Gargamelle data proved the existence of neutral currents in 1973.

The analyses performed by bubble chamber experiments consider different targets, with an atomic number from hydrogen to neon. In all analyses, the goal is to measure neutrino interactions on free nucleon. Nuclear effects, which are non-negligible for heavy targets, are neglected. Even though this problem is not present in hydrogen and deuterium measurements, these are extremely limited by statistics due to the low density of these targets: hydrogen and deuterium experiments observed only a few hundred events [90].

Bubble chamber experiments provided with the first flux-unfolded integrated cross-section measurements as well as differential cross-section measurements as a function of Q^2 . Even though the focus was on inclusive measurements, some analysis on exclusive processes exist: QEL, one-pion and two-pion production measurements. In addition, bubble chamber experiments also studied neutrino hadronization processes on hydrogen and deuterium.

Cross-section analyses from this era attempted to reconstruct the neutrino energy on an event-by-event basis. The reconstructed energy was then used to estimate other quantities, like Q^2 . The final results were often reported as a function of these reconstructed quantities, instead of observables that are easier to measure like the final-state-lepton energy. This approach was too optimistic on the energy reconstruction resolution and it lead to statistically unreliable datasets. Attempts were made to correct the datasets using MC simulations but these corrections were poorly implemented because of the bad simulations of the time and the lack of correct statistical treatment of systematics. In addition, exclusive datasets were reported divided by process, such as CCQEL, instead of the final state topology. This also introduced statistical issues in the released datasets.

Precise simulations of the neutrino flux were not possible at the time. The general approach was to model the neutrino flux from the CCQEL measurement in the same experiment assuming a CCQEL model. The flux prediction would then be used to extract the integrated cross section [91, 92]. This approach biased the result and, consequently, integrated cross-section measurements have a large systematic uncertainty (10-15%) due to poor knowledge of the incoming neutrino flux. This led to longstanding disagreement between ANL 12 ft and BNL 7 ft single pion production cross-section measurements on deuterium [93].

Besides the difficulties bubble chamber analyses faced, hydrogen and deuterium measurements are still crucial nowadays as these are the only available¹⁴ datasets that allow us to constrain neutrino interactions on nucleons.

¹⁴Due to safety concerns, there are no plans to perform new measurements with hydrogen and deuterium experiments.

2.4.2 The early modern era

The early modern era is characterized by the desire of increasing the statistics of the cross-section measurements in order to reduce statistical uncertainties. In order to do so, experiments focused on the study of neutrino interaction on heavy nuclear targets. Some of the experiments from this era are K2K [94], SciBooNE [95], NOMAD [96], MINOS [97] and MiniBooNE¹⁵ [98]. Successfully, these experiments provided event samples of thousands of events. For instance, MiniBooNE recorded 150,000 ν_μ CC0 π events in the 1 GeV region [99]. Fortunately, the high-statistics measurements were accompanied by the improvement in computing power: these experiments analyzed events using automatic reconstruction software.

The increase in the statistics allowed MiniBooNE to extract the first double-differential flux-integrated cross-section measurements as a function of the final lepton kinematics [99]. Integrated measurements were also released in order to facilitate the comparison with previous measurements.

An improvement with respect to the 70's era is the determination of the neutrino flux. Neutrino interaction data is not used in any way to obtain the flux prediction. Instead, dedicated simulations are performed to characterize it. This approach allowed them to estimate the systematic errors associated with the flux [100]. Flux uncertainties can include bin-to-bin correlations in the cross-section measurements, but such correlations are not always reported [99].

Besides some improvements, there were problems in the way the data was analyzed. Experiments still aimed to measure the neutrino cross section associated with a specific interaction process, such as QEL. Defining a topology that is specific for an interaction is not possible due to FSI. In fact, what experiments measured was the CCQEL-Like event rate. The definition of the CCQEL-Like topology depends on the experiment. The general requirement is to select events with a muon and no pions. This is also referred to as CC0 π topology. This definition includes non-QEL events, such as pion production events in which the pion is absorbed. In MiniBooNE, a subset of CC1 π^+ events was used to estimate the non-QEL contribution to the CC0 π sample by tuning their simulation to the CC1 π^+ data. The estimated non-QEL contribution was removed from the final sample used for the evaluation of the QEL cross section. This approach introduces a strong model dependency in the reported cross section. MiniBooNE also released the CC0 π cross-section measurement [99].

The importance of multi-nucleon interactions in neutrino measurements was not well understood. In particular, MiniBooNE, K2K, SciBooNE, and MINOS reported a

¹⁵This review focuses on the MiniBooNE analysis for its relevance in this thesis.

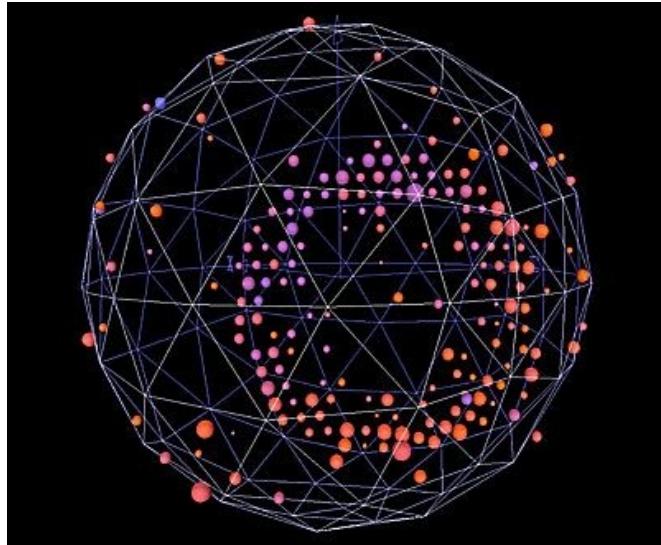


Figure 2.8: Example of a Cherenkov ring in the MiniBooNE detector.

higher QEL cross section than the one predicted by theoretical models [99, 101–103]. This observation suggested an increase of the QEL axial mass of 10%–30%. In the case of MiniBooNE, $M_A^{QEL, eff} = 1.34 \pm 0.17$ GeV/c² [99], which is in disagreement with hydrogen and deuterium bubble chamber and NOMAD data [104]. This tension is known as QEL or MiniBooNE puzzle. Nowadays, this puzzle is understood as a consequence of multi-nucleon processes, which were ignored at the time.

These experiments had additional limitations with respect to the 70’s era. In the particular case of MiniBooNE, the fine spatial resolution provided by bubble chamber experiments was lost, as highlighted in Fig. 2.8. In addition, the efficiency to detect some particles, such as protons, was very low. This limited the number of exclusive measurements that these experiments could study.

2.4.3 The modern era

The modern era aims to solve some of the outstanding questions in neutrino physics. DUNE [2] and Hyper-Kamiokande [105] will lead the effort on the measurement of the CP-violating phase in the lepton sector and the SBN program will test the three neutrino flavor paradigm [106]. In order to accomplish such goals, there is an extensive cross-section program aiming to measure neutrino interactions with unprecedented statistics. For instance, SBND, which is the near detector in the SBN program at a 110 m from the neutrino beam, expects more than 2,000,000 ν_μ -⁴⁰Ar CC0π events [107]. Consequently, the greatest challenge of the modern era is to reduce systematic uncertainties to the few percent level.

T2K ND280 [108], MINERvA [109] and MicroBooNE [110] are some of the

neutrino experiments leading the cross-section measurement effort at the moment. These experiments study neutrino interactions on heavy nuclei, such as ^{12}C , ^{16}O or ^{40}Ar , at different neutrino energies. These experiments provide flux-integrated cross sections for inclusive [111–116] as well as exclusive topologies: CC0 π [117–119], CCNp0 π [119–122] and CC1 π [116, 123–125]. All measurements are released as a function of the final particles kinematics.

In this era, there is a substantial change in the analysis procedure. The main difference is the definition of the event topology, which is not corrected in any way in order to extract information of a specific interaction. In addition, model-dependent corrections or cuts are avoided. Instead, measurements are restricted to specific regions of the phase space to guarantee good detection efficiency. Systematic errors due to flux uncertainties, smearing or efficiency corrections are included in the bin-to-bin covariance matrices, which are always provided in the data release. Correlation between datasets from the same experiment is usually not reported, with the exception of T2K ND280 ν_μ CC0 π measurement, which released such information. In addition, some analyses do not apply smearing corrections to their data. Instead, they provide the smearing matrices that allow comparing their data to MC predictions. This technique, known as *forward folding*, avoids the introduction of uncertainties due to unfolding [122].

Another improvement with respect to the early modern era is the spatial resolution of the experiments. For instance, MicroBooNE produced high statistics measurements without compromising the "bubble chamber"-like spatial resolution, see Fig. 2.9. These experiments are also able to measure low energy particles. For instance, MicroBooNE is able to reconstruct protons above 300 MeV/c [122]. This allows neutrino experiments to measure neutrino cross sections as a function of proton kinematics. Single- and double-differential ν_μ CCNp0 π cross-section measurements are provided by T2K ND280, MINERvA and MicroBooNE [119, 122, 126]. Another example is the recent triple-differential ν_μ CCNp0 π measurement from MINERvA as a function of muon and lepton kinematics [121]. All these measurements are extremely sensitive to nuclear effects.

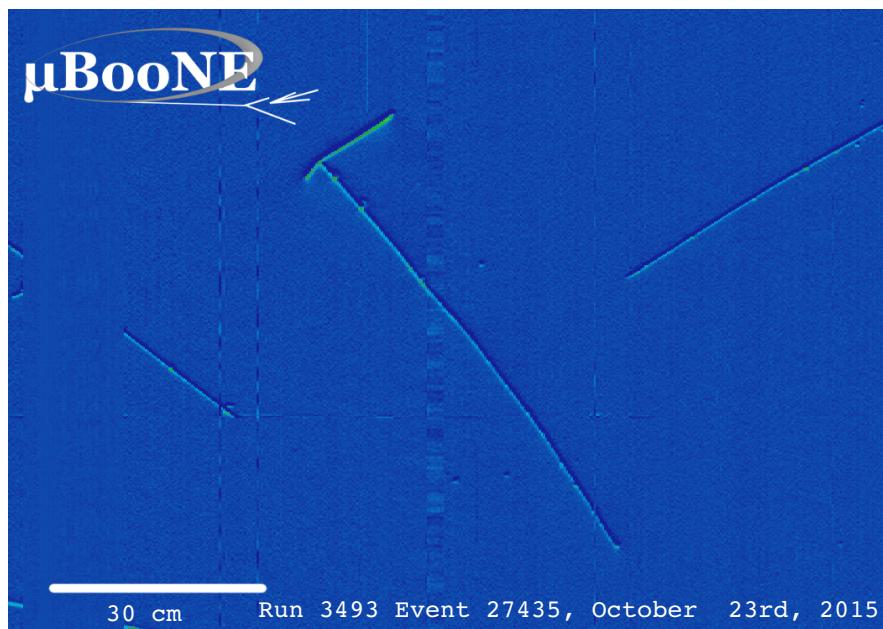


Figure 2.9: Example of a MicroBooNE ν_μ CC2p0 π event display [127]. The neutrino vertex can be identified in the junction of three tracks: the longest track corresponds to the muon candidate, whilst the shortest tracks correspond to two proton candidates.

Chapter 3

Neutrino-nucleon cross-section model tuning in GENIE v3

Abstract

In this Chapter¹, we summarize the results of a study performed within the GENIE global analysis framework, revisiting the GENIE bare-nucleon cross-section tuning and, in particular, the tuning of (a) the inclusive cross section, (b) the cross section of low-multiplicity inelastic channels (single-pion and double-pion production), and (c) the relative contributions of resonance and nonresonance processes to these final states. The same analysis was performed with several different comprehensive cross-section model sets available in GENIE Generator v3. In this work we perform a careful investigation of the observed tensions between exclusive and inclusive data, and install analysis improvements to handle systematics in historic data. All tuned model configurations discussed in this paper are available through public releases of the GENIE Generator. With this paper we aim to support the consumers of these physics tunes by providing comprehensive summaries of our alternate model constructions, of the relevant datasets and their systematics, and of our tuning procedure and results.

3.1 Introduction

GENIE is an international collaboration of scientists working on a global analysis of neutrino scattering data and on the incorporation of modern theoretical inputs and ex-

¹*This work has been published in Phys. Rev. D **104**, 072009 (2021). It follows the guidelines on presentation of publications within a PGR thesis requested by the University of Liverpool.*

perimental data into robust and predictive semiempirical comprehensive neutrino interaction simulations. GENIE develops and maintains a suite of well-known software products for the experimental neutrino community, which includes its popular Generator product [128]. With the recent release of the GENIE Generator v3, a substantial change in the way that the GENIE Collaboration approaches the process of developing, validating, characterizing, tuning and releasing comprehensive neutrino interaction simulations came into sharp focus. The focus of the GENIE Collaboration has always been the development of universal comprehensive models, handling all probes and targets and simulating all processes across the entire kinematic phase space relevant for neutrino experiments. Previously, the GENIE Collaboration released a single, preferred (*default*) comprehensive model that reflected our current understanding on the most predictive, robust, and self-consistent model that could be built out of GENIE neutrino interaction modeling elements. Whereas many other alternative modeling elements were made available to users, they had to be enabled by individual users through an error-prone procedure that could bring substantial physics and logical inconsistencies, invalidate procedures for addressing double counting issues, and damage the level of agreement with data, often in ways that were unsuspected by users that had a narrow focus on some particular modeling aspect and lacked the GENIE tools and procedures to fully characterize a comprehensive model. To address this, and in response to the community demand for *alternative* models, GENIE has released a number of comprehensive model configurations (CMCs) and is in the process of constructing several more. All such configurations, that are easily invoked and run out of the box, combine modeling elements in a way that is as consistent as possible, and are validated, characterized and tuned as a whole. This important development was underpinned by a substantial upgrade of GENIE capabilities for systematic model validation, model characterization through comparisons to large collections of complementary scattering data with neutrino, charged lepton, and hadron probes, and the development of an advanced global analysis of scattering data.

The GENIE global analysis was made possible through the continued development of curated data archives, and the successful large-scale refactoring and interfacing to the Professor tool [129] of a very extensive set of GENIE codes that implement comparisons to data, within a framework that allows the efficient manipulation of large ensembles of simulated events produced from a constellation of alternative models. The interface to the Professor tool enabled the efficient implementation of complex multi-parameter brute-force scans and removed substantial global analysis limitations by decoupling it from event reweighting procedures that, for all but the most trivial aspects of our physics domain, require substantial development time and are not exact, or

even possible at all. Professor ‘reduces the exponentially expensive process of brute-force tuning to a scaling closer to a power law in the number of parameters, while allowing for massive parallelization’ [130]. The Professor package has been extensively used for the tuning of Monte Carlo (MC) generators in the collider community.

The above developments allowed the GENIE Collaboration to fulfil its *dual purpose* described in its mission statement: GENIE develops a popular MC event generation platform and implements, within its platform, universal and comprehensive physics simulations for lepton scattering, as well as simulations for several Beyond the Standard Model processes. But, in addition, and separately from the previous mission, GENIE develops a global analysis of scattering data for the tuning and uncertainty characterization of comprehensive neutrino interaction models. The GENIE Generator is the main outlet for the GENIE global analysis results, and our goal is that, for each supported comprehensive model, several selected tuned versions shall be made available.

Typically, nuclear modifications to the cross-section are computed separately, and the decomposition of the total cross section into the possible exclusive final states proceeds via separate hadronization, intranuclear rescattering and particle decay codes. Therefore, bare-nucleon cross-sections are a crucial first modeling component to tune in the process of building a global fit of all relevant scattering data. Tunes for several aspects of GENIE modeling, including neutrino-induced hadronization and nuclear cross-sections for low-multiplicity channels, are near completion and will be released and published in the future. This paper summarizes the results of the first analysis performed within the GENIE global analysis framework, revisiting the GENIE bare-nucleon cross-section tune and, in particular, the tuning of the empirical nonresonance background contribution to one- and two-pion final states. A similar, albeit much simpler, analysis underpinned the tune of the well-known and widely-used comprehensive model that was included as the default model throughout the very long GENIE v2 series of releases. At that time, not sufficiently explored and understood tensions between inclusive and exclusive data, and an executive decision to anchor the GENIE v2 model on inclusive data, led to some expected and well known discrepancies with exclusive data that were increasingly brought into focus as new experiments started performing increasingly precise measurements of low-multiplicity exclusive final states [131]. Here, we perform a careful investigation of the observed tensions between exclusive and inclusive data, retune the bare-nucleon cross-section model for all GENIE comprehensive models available in GENIE v3, and provide best-fit values and correlations for several parameters influencing the GENIE bare-nucleon cross sections. The work presented here was based on the model implementations of GENIE v3.0.6 (released

on 23 July 2019), and the results of this work will be included in the GENIE v3.2.0 release. Preliminary versions of this work appeared in earlier releases of the GENIE v3 series (v3.0.0 - v3.0.6).

In Sec. 3.2, we summarize relevant aspects of the free nucleon cross-section modeling in GENIE, while in Sec. 3.3 we provide further details for the construction of comprehensive GENIE models considered in this work. In Sec. 3.4 we provide details of the datasets, parametrization of the model and data uncertainties for this particular tune. Sec. 3.5 describes the tuning procedure as well as the statistical methodology used. Finally, our tuning results are presented in Sec. 3.6.

3.2 Bare nucleon cross-section modeling in GENIE

In very simplified terms, neglecting diffractive production, as well as $|\Delta S| = 1$ and $|\Delta C| = 1$ processes, the total inelastic differential cross section for neutrino scattering off bare nucleons, $d^2\sigma^{inel}/dQ^2dW$, is computed in GENIE as

$$\frac{d^2\sigma^{inel}}{dQ^2dW} = \begin{cases} \frac{d^2\sigma^{\text{RES}}}{dQ^2dW} + \frac{d^2\sigma^{\text{SIS}}}{dQ^2dW} & \text{for } W < W_{\text{cut}} \\ \frac{d^2\sigma^{\text{DIS}}}{dQ^2dW} & \text{for } W \geq W_{\text{cut}} \end{cases} \quad (3.1)$$

The term $d^2\sigma^{\text{RES}}/dQ^2dW$ represents the contribution from all low-multiplicity inelastic channels proceeding via resonant (RES) production and, in present versions of GENIE, it is computed as an *incoherent* sum over several resonances. The resonances included in GENIE v3 are the ones specified by the Rein-Seghal paper [34]. The nine lightest N^* and the eight lightest Δ labeled by the PDG with three or four stars are considered. The following resonances are included in GENIE v3: $N(1440)$, $N(1520)$, $N(1535)$, $N(1650)$, $N(1675)$, $N(1680)$, $N(1700)$, $N(1720)$, $N(1710)$, $\Delta(1232)$, $\Delta(1600)$, $\Delta(1620)$, $\Delta(1700)$, $\Delta(1905)$, $\Delta(1910)$, $\Delta(1920)$ and $\Delta(1950)$. W_{cut} is a free parameter that determines the end of the SIS region. The nominal value is set to $W_{\text{cut}} = 1.7 \text{ GeV}/c^2$.

In the version of GENIE used in this work, there is the option to select one of several neutrino-induced resonance production calculations performed by Rein and Sehgal [34], Kuzmin, Lyubushkin and Naumov [132, 133], and Berger and Sehgal [38]. The last two models are extensions of the first one, that account for nonzero lepton masses. Both models are based on the same formalism and the only difference between them is that the latter includes the pion-pole contribution to the hadronic axial current.

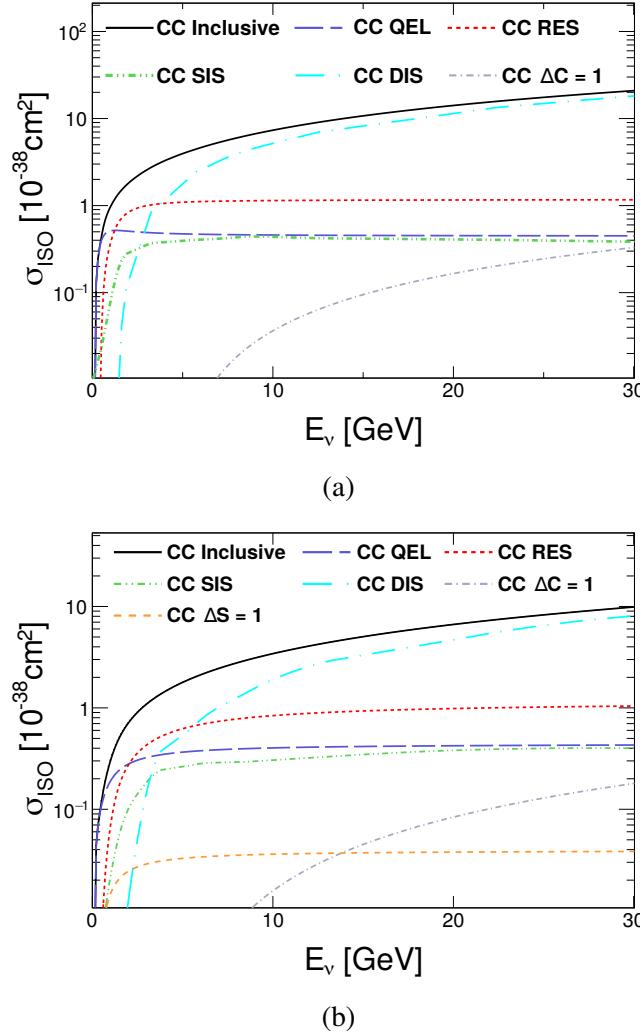


Figure 3.1: Summary of neutrino and anti-neutrino cross sections on isoscalar targets. (a) ν_μ CC on isoscalar targets. (b) $\bar{\nu}_\mu$ CC on isoscalar targets.

The term $d^2\sigma^{\text{DIS}}/dQ^2dW$ represents the GENIE calculation of the deep-inelastic scattering (DIS) cross-section that, in all relevant GENIE comprehensive model configurations, is carried out using an effective leading order model with the modifications suggested by Bodek and Yang [51] to describe scattering at low momentum-transfers. This model is the foundation of both the DIS model and the shallow-inelastic scattering (SIS) model in GENIE.

The term $d^2\sigma^{\text{SIS}}/dQ^2dW$ requires some elaboration. It represents the cross-section contribution from non-resonance SIS in the resonance region. In GENIE, this cross section is computed with an empirical model where the Bodeck and Yang inclusive deep-inelastic cross section is extrapolated into the resonance region and it is decomposed, via the GENIE AGKY [41] hadronization model, into the cross sections for different hadronic multiplicity channels. The extrapolation of the DIS model down to the

inelastic threshold, $W < W_{\text{cut}}$, includes, on average, the effect of the resonances [134]. Notice that, even though the Bodeck and Yang model is capable of describing the inclusive cross section at the inelastic threshold, we prefer to utilize an explicit resonance model. The contribution for hadronic multiplicities two and three, that are responsible for producing many final states similar to those produced via resonance excitation, are tuned to remove double counting. This tuning is the main topic of this work.

The nonresonance SIS cross section can be written as

$$\frac{d^2\sigma^{\text{SIS}}}{dQ^2 dW} = \frac{d^2\tilde{\sigma}^{\text{DIS}}}{dQ^2 dW} \cdot \Theta(W_{\text{cut}} - W) \cdot \sum_m f_m(Q^2, W) \quad (3.2)$$

where $\tilde{\sigma}^{\text{DIS}}$ represents the extrapolated deep-inelastic cross section into the resonance region, and m refers to the multiplicity of the hadronic system. The factor f_m relates the total calculated DIS cross section to the DIS contribution to this particular multiplicity channel. These factors are computed as

$$f_m(Q^2, W) = R_m P_m^{\text{had}}(Q^2, W) \quad (3.3)$$

where R_m is an adjustable parameter and P_m^{had} is the probability, taken from the GENIE hadronization model, that the DIS final state hadronic system multiplicity would be equal to m .

At the low W values relevant for SIS, P_m^{had} is constructed using the low-mass empirical AGKY model [41]. By construction, its distribution has an average $\langle m \rangle$ that increases logarithmically with W :

$$\langle m \rangle(Q^2, W) = \alpha + \beta \ln \left(\frac{W^2}{\text{GeV}^2/c^4} \right) + \beta' \ln \left(\frac{Q^2}{\text{GeV}^2/c^2} \right) \quad (3.4)$$

and the whole distribution is constructed from the average according to

$$P_m^{\text{had}}(Q^2, W) = \frac{1}{\langle m \rangle} \psi \left(\frac{m}{\langle m \rangle} \right) \quad (3.5)$$

where the function ψ (the Levy function) has the following form

$$\psi(z) = \frac{2e^{-c} c^{cz+1}}{\Gamma(cz+1)}, \quad z \equiv \frac{m}{\langle m \rangle} \quad (3.6)$$

In the above expressions, α , β , β' and c are adjustable parameters. In principle, α , β , β' , c and R_m , are different for each initial state ($\nu + p$, $\nu + n$, $\bar{\nu} + p$, $\bar{\nu} + n$) and are different for charged current and neutral current interactions. A new tune of the neutrino-induced hadronization models in GENIE is currently in progress and, in fu-

Parameter	Initial state			
	$\nu_\mu p$	$\nu_\mu n$	$\bar{\nu}_\mu p$	$\bar{\nu}_\mu n$
α	0.40	-0.20	0.02	0.80
β	1.42	1.42	1.28	0.95
c	7.93	5.22	5.22	7.93

Table 3.1: Relevant default GENIE v3 AGKY parameters for ν_μ and $\bar{\nu}_\mu$ CC interactions on proton and neutron. The parameter values are extracted from [136] neutrinos and from [137] for anti-neutrinos.

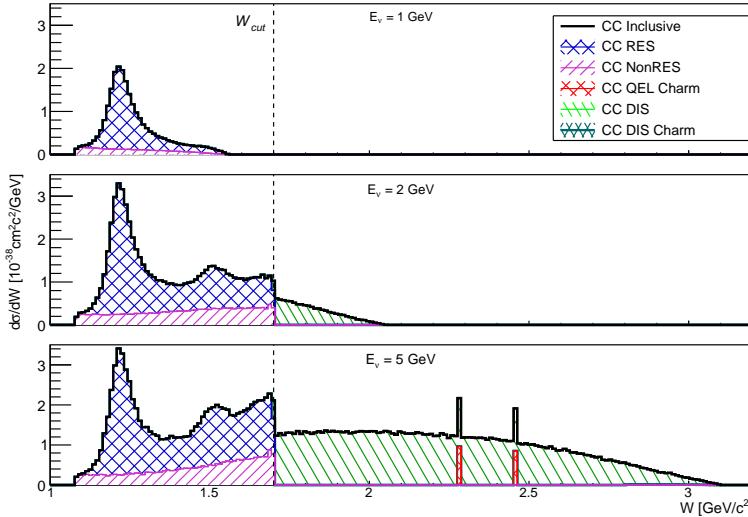


Figure 3.2: RES and SIS/DIS differential cross section as a function of the invariant mass for three different neutrino energies using a $1/E$ -like flux. W_{cut} divides the SIS and the DIS regions.

ture, it may be possible to perform a joint tuning of the GENIE cross section and hadronization modeling components for bare-nucleon targets. However, at this present work, the parameters α , β , β' and c were kept at the default values of the AGKY model in GENIE v3. For easy reference, the relevant values for the channels studied in this work are included in Tab. 4.1. No dependence on Q^2 has been observed in ν and $\bar{\nu}$ scattering data [135], hence $\beta' = 0$ for all channels.

For most inelastic processes simulated in neutrino-nucleus scattering by all current GENIE comprehensive model configurations, the total inelastic differential cross section for scattering off bare nucleons takes centre stage. In Fig. 3.1, the contribution to the ν_μ charged-current (CC) and $\bar{\nu}_\mu$ CC inclusive cross sections on isoscalar targets in GENIE is shown for the different interaction processes. The CC RES and SIS/DIS CC cross-section contribution for different neutrino energies is shown in Fig. 3.2.

3.3 Comprehensive model configurations in GENIE v3

GENIE has a large degree of configuration; for each process (RES, DIS, etc.) the system offers a number of alternative models to be used for event generation. In previous GENIE releases, only one model-process mapping was suggested by the out of the box configuration, despite the availability of alternative models. Yet, there was no guidance on how to correctly use different configurations according to the author and developers. In fact, the processes are not universal and their definitions are generator dependent. Hence, it was easy to come up with inconsistencies between the model configuration for different processes that were not supposed to be used together to get a correct comprehensive physics simulation.

This issue was addressed in GENIE v3 by introducing the concept of comprehensive model configuration (CMC) that is a consistent process-model association. Considering that GENIE already has about 20 different processes only for neutrinos, CMC definitions are quite complex objects and they need to be effectively named so that the community can use them unambiguously. For this purpose, the Collaboration developed a specific naming convention discussed in appendix A.1. Sec. 3.3.1 describes the models used in CMCs relevant for neutrino interactions.

3.3.1 CMCs available in GENIE v3

Several CMCs are available in GENIE v3, but they can be grouped together as their scopes are common. The first group of CMCs is historically motivated: it is based on the *default* configuration and simply provides updates for processes that were introduced later. The second family is an improvement of the first group in terms of the resonance model. The third one was constructed aiming to deliver the most up to date theoretical nuclear matter simulations. Out of these main ideas, a number of CMCs can be constructed simply changing minor aspects like final state interaction (FSI) modeling or form factors. Here, we briefly summarize the modeling components used in each comprehensive model configuration available in GENIE v3 that are relevant for this paper. More are available in Appendix. A.2.

G18_01a, G18_01b, G18_01c and G18_01d

These comprehensive models share an identical cross-section model construction, which is an adiabatic update of the historical default cross-section model of GENIE v2, now named as G00_00a CMC. For interactions on nucleons and nuclei, it relies on implementations of the following models: the Ahrens model [32] for neutral-current (NC)

elastic, the Llewellyn Smith model [31] for CC quasielastic (QEL), the Rein-Sehgal model [34] for NC and CC RES, the Rein-Sehgal model [138] for NC and CC coherent pion production, the Bodek-Yang model [51] for NC and CC DIS and nonresonance shallow inelastic scattering, the Kovalenko model [139] for QEL charm production, and the Aivazis-Olness-Tung slow rescaling model [140] for deep-inelastic charm production. Nuclear cross sections are calculated within the framework of a relativistic Fermi gas model, following the approach of Bodek-Ritchie [61]. Multi-nucleon processes in neutrino scattering off nuclear targets can be optionally enabled and simulated via an empirical GENIE model [128]. In addition, in GENIE v3, the adiabatic upgrade of the historical comprehensive model includes the simulation of processes that, previously, were either optional or missing. This includes both diffractive pion production based on an implementation of the Rein model [141], and quasielastic $|\Delta S| = 1$ hyperon ($\Lambda^0, \Sigma^-, \Sigma^0$) production based on the Pais model [33]. Single Kaon production, although optionally available for neutrinos in GENIE v3 [142], is not yet available for antineutrinos and inclusion in any published GENIE comprehensive configurations was postponed till an antineutrino implementation is available and the Kaon content of hadronic showers produced by GENIE has been retuned following the addition of the single-Kaon generator. Both G18_01a and G18_01b comprehensive models employ a revised resonance decay algorithm and an implementation of the AGKY [41] hadronization model that is unchanged with respect to that used at the latest releases of GENIE v2 series. Four comprehensive model variations are constructed by attaching different intranuclear hadron transport models to the same underlying cross section and hadronization models [143]. G18_01a uses an updated *INTRANUKE hA* effective intranuclear rescattering model which is unique to GENIE, G18_01b uses the new *INTRANUKE hN* model implementing a full intranuclear cascade including medium corrections, G18_01c uses an interface to the GEANT4 [144] Bertini intranuclear cascade [145] (version 4.10.2) and G18_01d uses an interface to the INCL++ (version 5.2.9.5) implementation of the Liège intranuclear model [146].

G18_02a, G18_02b, G18_02c and G18_02d

This is family of empirical models which is an evolved version of the G18_01[a-d] ones. The general construction of the cross-section model is similar to the one discussed above, with the exception that the implementations of the Rein-Sehgal models for CC and NC resonance neutrino-production, as well as for CC and NC coherent production of mesons, were replaced with updated models by Berger-Sehgal [38]. Similarly to G18_01[a-d], four comprehensive model variations are constructed by using alternative intranuclear hadron transport models on top of the same underlying

cross section and hadronization models (a: *INTRANUKE/hA*, b: *INTRANUKE/hN*, c: GEANT4/Bertini, and d: INCL++).

G18_10a, G18_10b, G18_10c and G18_10d

This is a family of models derived from the improved empirical ones (G18_02[a-d]) described above, by substituting both the Llewellyn Smith CC quasielastic model [31] and GENIE’s empirical multinucleon model with implementations of the corresponding Valencia models by Nieves et al. [70]. This family of comprehensive models provides a firmer theoretical basis for the simulation of neutrino-nucleus scattering around the quasielastic peak. Within this family of models, the nuclear environment is modeled using a Local Fermi Gas, matching the inputs used for the published Valencia calculations. Again, four comprehensive model variations (a-d) are constructed by using alternative intranuclear hadron transport models, following the same naming convention introduced above. The implementation of the Valencia model in GENIE does not predict the kinematics of the outgoing hadrons and its description needs to be accompanied by one of the FSI models available in GENIE (a-d) [71].

3.3.2 Free nucleons and CMCs

Although a large number (16) of CMCs were summarised above, with respect to the cross sections for (anti)neutrino scattering off bare nucleons, there are only two different model constructions: The one used in a) G18_01[a-d], and the one used in b) G18_02[a-d], G18_10[a-d] and G18_10[i-1]. The main difference between these two model constructions resides mainly in the treatment of the lepton mass. Although some differences can be expected between G18_10[a-d] and G18_10[i-1], because of different choices in the modeling of the axial form factor for quasielastic scattering, they do not manifest themselves in the context of this particular analysis.

Several variations of the tuning procedure were run and evaluated for testing purposes before converging to the procedure presented in this paper. Preliminary versions of this work were released in the GENIE v3 series (v3.0.0-v3.0.6) in a series of tunes carrying the 02_11a label. The final results presented in the paper will be made available in GENIE v3.2 in a series of 16 tunes, one for each of the 16 comprehensive model configurations summarised above, labeled as 02_11b. For example, the tune G18_10a_02_11b corresponds to the G18_10a comprehensive model with the parameters determined through the tuning procedure discussed in this paper (02_11b). The GENIE tune naming convention is discussed in appendix A.1. A full list of GENIE tunes is maintained in <http://tunes.genie-mc.org>. The preliminary versions

Model	CMC		
	G18_01*	G18_02*	G18_10*
Nuclear Model	Relativistic Fermi Gas [61]		Local Fermi Gas [70]
CC QEL	Llewellyn Smith model [31]		Valencia model [70]
CC RES	Rein-Sehgal [138]	Berger-Sehgal [38]	
CC DIS		Bodek-Yang [51]	

Table 3.2: List of relevant models for (anti)neutrino CC interactions on free nucleon.

(02_11a) of the tunes will be kept in GENIE v3.2, but they will be phased out in subsequent minor releases. It is useful to mention that CMCS whose parameters are not tuned using the procedures similar to those described in this paper are identified with the tune label 00_000, e.g. the configuration of G18_10a_02_11b before the tune is identified as G18_10a_00_000.

3.4 Data and model uncertainties review

The data used in this analysis are old and a careful review of the past analysis procedure is required in order to combine all the data together in a global analysis. This section summarises the data details and how the models used in the fit behave in the same energy region.

3.4.1 Datasets included in the fit and their systematics

In the current work, we consider Hydrogen and Deuterium data from the ANL 12 ft, BNL 7 ft, FNAL 15 ft and BEBC bubble chamber experiments. The data represent integrated cross sections for different incoming neutrino energy bins for

- ν_μ and $\bar{\nu}_\mu$ CC inclusive scattering [91, 147–173]
- ν_μ and $\bar{\nu}_\mu$ CC quasielastic scattering [91, 156, 170, 174–182]
- ν_μ and $\bar{\nu}_\mu$ CC single-pion production [93, 175, 183–190]
 - $\nu_\mu + n \rightarrow \mu^- + n + \pi^+$
 - $\nu_\mu + p \rightarrow \mu^- + p + \pi^+$
 - $\nu_\mu + n \rightarrow \mu^- + p + \pi^0$
 - $\bar{\nu}_\mu + p \rightarrow \mu^+ + p + \pi^-$
 - $\bar{\nu}_\mu + n \rightarrow \mu^+ + n + \pi^-$
- ν_μ CC two-pion production [191]

- $\nu_\mu + p \rightarrow \mu^- + n + 2\pi^+$
- $\nu_\mu + p \rightarrow \mu^- + p + \pi^+ + \pi^0$
- $\nu_\mu + p \rightarrow \mu^- + n + \pi^+ + \pi^-$

Not all of the available historical data has been used for the fit, as some datasets were superseded or reanalyzed, as in the case of ANL 12 ft and BNL 7 ft, datasets. The latest analysis are used. A detailed summary of the datasets used in the fit is shown in Tab. 4.5 and in Fig. 3.3. Some of the datasets included in the tune consider hydrogen-neon mixtures. The nuclear effects of the neon in the target mixture are shown to be negligible [192].

Low energy bins have a higher contribution to the χ^2 due to energy smearing and lack of unfolding in measurements. Hence, data points with $E_\nu < 0.5$ GeV are removed from the fit. In total, the tune is performed with 169 data points from bubble chamber experiments. Different analysis methods were implemented in each experiment, such as cuts applied on the W invariant mass, the outgoing muon momentum or the total longitudinal momentum of the final state. The associated GENIE prediction has been corrected by applying the same cuts to the generated events. Moreover, datasets from the same experiments are not independent as they share the same neutrino flux, detector, analysis methodology, etc. Although it is clear that some correlated uncertainties exist, the data releases do not contain any information about the correlation between them. In the GENIE database, we added a systematic error to the datasets of 15%. The methodology used to include them in the fit is detailed in Sec. 3.5.3. Other free nucleon data on heavier targets are available but used only for comparison with the GENIE prediction. No correction for nuclear effects is considered for deuterium targets.

3.4.2 Model uncertainties

The SIS cross section is tuned within the CMCs using either the Rein-Sehgal or Berger-Sehgal resonance models, see Sec. 3.3.1. The tuning main goal is the best value estimation for nine of the parameters that drive the GENIE predictions in the SIS region. These parameters are the W_{cut} as defined in Eq. 3.1, the four R_m coefficients for CC interactions on neutron/proton with $m = 2, 3$ from of the SIS region (Eq. 3.3), the axial masses used in the dipole form factors for RES and QEL interactions, and two global scaling factors, S_{RES} for the RES cross section and S_{DIS} for the DIS cross section. For clarity, we will refer to R_m parameters with the number of pions in the final state, namely $R_{vp}^{\text{CC}1\pi}$, $R_{vp}^{\text{CC}2\pi}$, $R_{vn}^{\text{CC}1\pi}$ and $R_{vn}^{\text{CC}2\pi}$.

Most of the parameters described in the previous paragraph are empirical and only

Experiment	N_p	Energy [GeV]	Target	Cuts	Ref.
$\nu_\mu + N \rightarrow \mu^- X$					
BNL 7 ft	13	0.6-10	^2H		[161]
BEBC	3	10-50	$^1\text{H}, \text{Ne}$		[159]
FNAL	6	10-110	^2H		[152]
	5	100-110	$^1\text{H}, \text{Ne}$		[164]
$\bar{\nu}_\mu + N \rightarrow \mu^+ X$					
BEBC	3	11-110	$^1\text{H}, \text{Ne}$		[147]
	1	10-50	$^1\text{H}, \text{Ne}$		[159]
	6	30-110	$^1\text{H}, \text{Ne}$		[148]
	1	10-110	$^1\text{H}, \text{Ne}$		[160]
BNL 7 ft	1	1-4	^1H		[170]
FNAL	5	10-110	$^2\text{H}, \text{Ne}$		[166]
	7	10-80	$^2\text{H}, \text{Ne}$		[171]
$\nu_\mu n \rightarrow \mu^- n \pi^+$					
ANL 12 ft	5	0.3-2	$^1\text{H}, ^2\text{H}$		[184]
ANL 12 ft,ReAna	7	0.3-3	^2H		[93]
BNL 7 ft,ReAna	11	0.1-4	^2H		[93]
$\nu_\mu p \rightarrow \mu^- p \pi^+$					
ANL 12 ft,ReAna	8	0-1.6	^2H		[93]
BNL 7 ft,ReAna	7	0-7	^2H		[93]
BEBC	7	1-30	^1H	$W < 1.4 \text{ GeV}$	[186]
	6	5-100	^2H	$W < 2 \text{ GeV}$	[175]
	5	10-80	^1H	$W < 2 \text{ GeV}$	[188]
FNAL	3	10-30	^1H	$W < 1.4 \text{ GeV}$	[193]
$\nu_\mu n \rightarrow \mu^- p \pi^0$					
ANL 12 ft	5	0.2-2	^2H		[184]
ANL 12 ft,ReAna	7	0.2-2	^2H		[93]
BNL 7 ft,ReAna	10	0.4-3	^2H		[93]
$\nu_\mu p \rightarrow \mu^- n \pi^+ \pi^+$					
ANL 12 ft	5	1-6	^2H		[191]
$\nu_\mu p \rightarrow \mu^- p \pi^+ \pi^0$					
ANL 12 ft	5	1-6	^2H		[191]
$\nu_\mu n \rightarrow \mu^- p \pi^+ \pi^-$					
ANL 12 ft	5	8-6	^2H		[191]
BNL 7 ft	10	0-20	^2H		[152]
$\bar{\nu}_\mu p \rightarrow \mu^+ p \pi^-$					
FNAL	1	5-70	^1H	$W < 1.9 \text{ GeV}$	[189]
$\nu_\mu + n \rightarrow \mu^- + p$					
ANL 12 ft	7	0-2	^2H		[174]
	8	0-2	$^1\text{H}, ^2\text{H}$		[91]
BNL 7 ft	4	0.2-2	^2H		[179]
BEBC	5	20-40	^2H		[175]
FNAL	2	0-50	^2H		[176]

Table 3.3: A summary of cross-section data used in this work. The number of data points released by each analysis (N_p), the neutrino energy range covered E_ν , the type of target and the cuts applied in the analysis procedure are specified in the table.

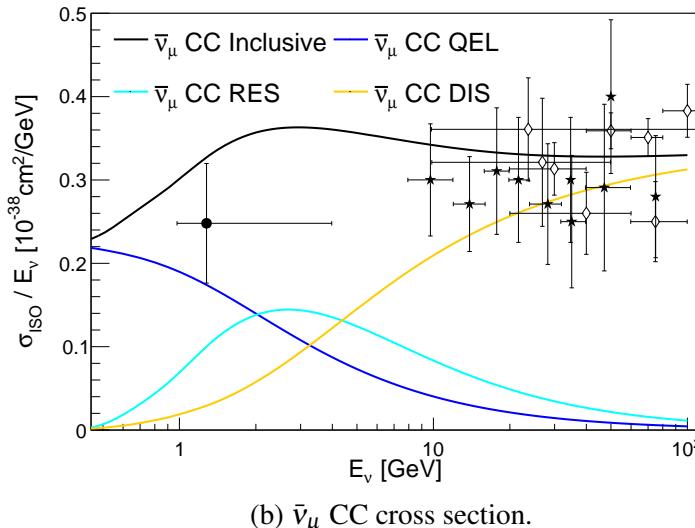
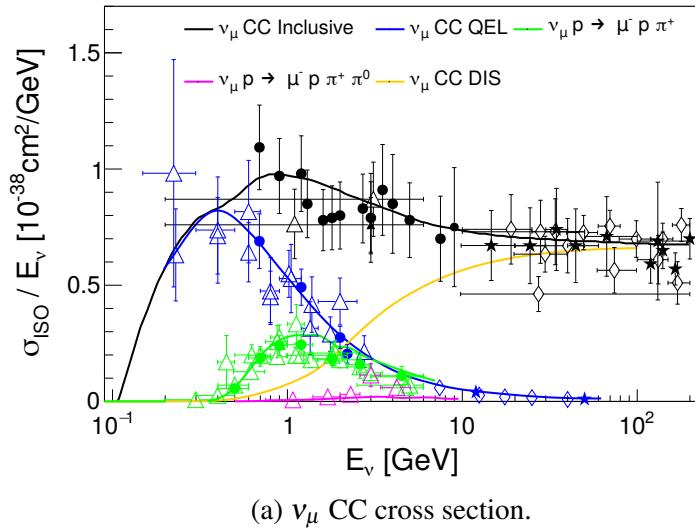


Figure 3.3: Charged current cross section on isoscalar targets as a function of the incoming neutrino energy. Breakdown of quasielastic, one and two pion production and deep-inelastic processes is shown. The predictions are computed using the G18_02a_00_000 configuration. The data on hydrogen and deuterium targets from Tab. 4.5 is shown if available from ANL 12 ft (triangle), BNL 7 ft (filled circle), BEBC (diamond) and FNAL (star).

3.4. DATA AND MODEL UNCERTAINTIES REVIEW

Parameter	GENIE parameter name	<i>Default</i> value	Min value	Max value	Prior
W_{cut} (GeV/c ²)	Wcut	1.7	1.5	2.3	
M_A^{QEL} (GeV/c ²)	QEL-Ma	0.999	0.75	1.10	1.014 ± 0.014 [26]
M_A^{RES} (GeV/c ²)	RES-Ma	1.12	0.8	1.3	1.12 ± 0.03 [35]
$R_{vp}^{\text{CC1}\pi}$	DIS-HMultWgt-vp-CC-m2	0.10	0.0	0.4	
$R_{vp}^{\text{CC2}\pi}$	DIS-HMultWgt-vp-CC-m3	1.00	0.0	2.0	
$R_{vn}^{\text{CC1}\pi}$	DIS-HMultWgt-vn-CC-m2	0.30	0.0	0.35	
$R_{vn}^{\text{CC2}\pi}$	DIS-HMultWgt-vn-CC-m3	1.00	0.8	3.0	
S_{RES}	RES-CC-XSecScale	1.0	0.6	1.2	
S_{DIS}	DIS-CC-XSecScale	1.032	0.9	1.15	1 ± 0.05

Table 3.4: Parameters of interest of the tunes and their statistical properties as used in the fitting procedures. The *Default* values correspond to the nominal values from GENIE v2 [128].

valid for GENIE. Therefore, there is no expectations about their values. The notable exceptions are the axial masses. For the G18_01a(/b/c/d) and G18_02a(/b/c/d) CMCs, the axial form factors are described using the dipole parameterisation which is a function of the invariant transferred momentum (Q^2):

$$F_A(Q^2) = F_A(0) \left(1 + \frac{Q^2}{M_A^2} \right)^{-2} \quad (3.7)$$

with $F_A(0) = g_A = -1.2695 \pm 0.002$ [29]. The axial mass, M_A , is extracted from data. There are different masses for every interaction type: M_A^{QEL} and M_A^{RES} . Both of these are evaluated from neutrino data on deuterium targets. The latest world average values for the axial masses are:

$$\begin{aligned} M_A^{\text{QEL}} &= 1.014 \pm 0.014 \text{ GeV}/c^2 & [26] \\ M_A^{\text{RES}} &= 1.12 \pm 0.03 \text{ GeV}/c^2 & [35] \end{aligned}$$

The extraction of these parameters requires neutrino differential cross sections as a function of Q^2 that are not used in this analysis.

Tab. 4.7 summarises the parameter pre-fit values and the allowed ranges. Previous fits to data are taken into account for the determination of the ranges [35, 194]. Each of the parameters have a different sensitivity to each dataset, as different scattering mechanisms are involved. The response of each parameter in the inclusive and exclusive cross sections is studied by varying each of them independently within the studied range. In Fig. 3.4, each parameter response is shown for inclusive and exclusive cross sections. When more than one parameter in the plot is impacting the same cross section, i.e. CC inclusive, the variations are added in quadrature.

At the MC level, where no correlation between the parameters is considered, the

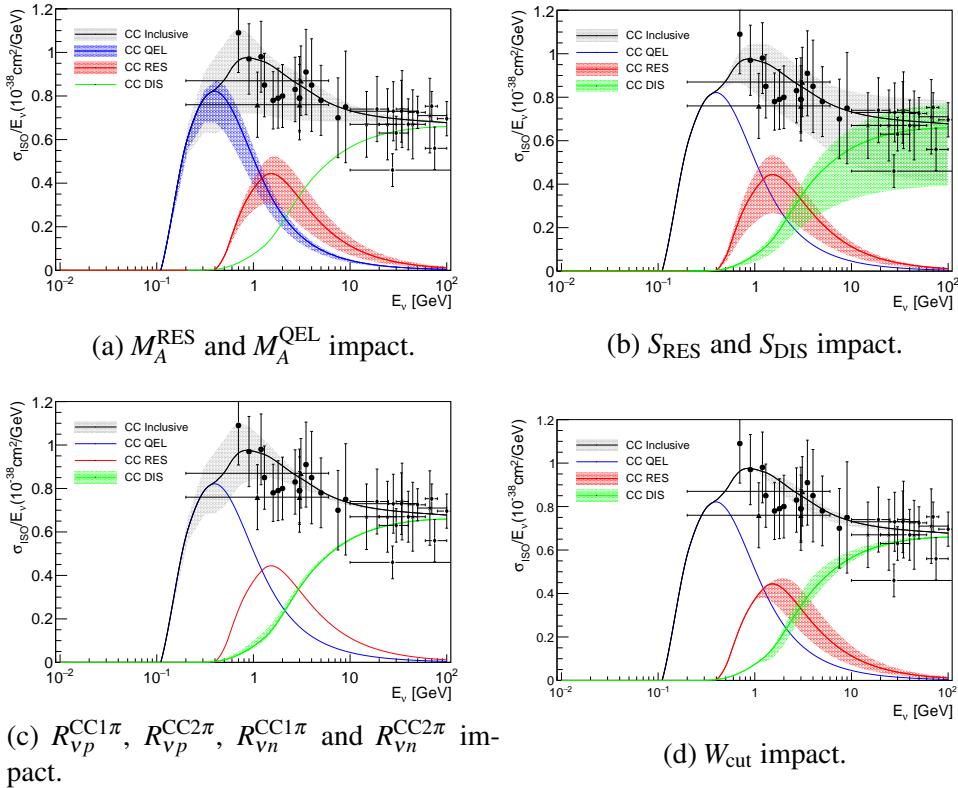


Figure 3.4: ν_μ CC inclusive G18_02a prediction against hydrogen and deuterium data. Each parameter response is characterized within the tuned region, specified in Tab. 4.7. Data corresponding to hydrogen and deuterium targets from ANL 12 ft (triangle), BNL 7 ft (filled circle), BEBC (diamond) and FNAL 15 ft (star).

impact of each of the parameters in the cross section can be classified as influencing a variation on

1. The CC QEL cross section
2. The CC RES cross section
3. The CC DIS cross section

For instance, M_A^{QEL} will only affect the QEL cross-section prediction, as summarized in Fig. 3.4a. Notice though that, at the tune level, this will no longer hold as the introduction of flux-nuisance parameters correlates exclusive channels. Hence, this will introduce a correlation between M_A^{QEL} and the SIS parameters.

The description of the CC RES cross section will be affected by the RES axial mass M_A^{RES} (Fig. 3.4a), the resonant scaling parameter S_{RES} (Fig. 3.4b), and W_{cut} (Fig. 3.4d). The G18_01a_00_000 and G18_02a_00_000 configurations overestimate one-pion production processes, and would favor a reduction in the CC RES cross section. Variations of M_A^{RES} have a huge impact on both exclusive and inclusive CC cross sections in the few-GeV region. However, as it is explained in Sec. 3.5.3, this parameter should agree with the world average extracted from fits to the axial form factor [35] and a deviation from this result is disfavoured by previous fits to data. Consequently, a reduction of S_{RES} is expected to improve the agreement with one-pion production data. On the other hand, W_{cut} will play an important role as it determines the number of resonances included in the CC RES calculation. The current default, $W_{\text{cut}} = 1.7 \text{ GeV}/c^2$, discards the resonances contributing at $W > W_{\text{cut}}$. Therefore, an increase on W_{cut} will incorporate new resonances in the calculation that were not taken into account in previous tunes. This increase is favoured by two-pion production data, as heavier resonances producing more than one pion are incorporated.

The SIS region is treated by combining two cross-section models, one for DIS and one for RES interactions. Thus, in that region, many parameters have a visible effect on the predictions as can be seen in Fig. 3.4: regardless of the parameter considered in the plot, there is always a visible error band in the few-GeV region. This is a clear hint for the presence of degeneracy that must be faced by our global tunes. An example of this is given by the R_m and the S_{DIS} parameters, which act as scaling factors for the DIS contribution at $W < W_{\text{cut}}$. As mentioned above, a desired result of the tune is to reduce the one-pion prediction and increase the two pion production. This can be accomplished via alterations of either the R_n and/or the S_{DIS} parameters.

3.5 Bare-nucleon cross-section tuning procedure

This section describes the core ideas behind the paper. Most of these are not specific for this work; they are general concepts developed within the GENIE tuning system and can therefore apply to future tune releases.

3.5.1 Likelihood construction

The GENIE integrated cross-section prediction is denoted with $\sigma_{\text{th}}^i(E_k|\boldsymbol{\theta})$, where E_k is the neutrino energy, $\boldsymbol{\theta}$ is a vector² of the adjustable physics parameters introduced in Sec. 3.4.2, and i is any of the ten reaction processes considered in the work presented in Tab. 4.5. Using $\sigma_{\text{th}}^i(E|\boldsymbol{\theta})$, we produce the corresponding prediction for the k -th energy bin of the j th dataset for the i th reaction type,

$$\sigma_{\text{th}}^{ij}(E_k|\boldsymbol{\theta}) = \varepsilon^{ij}(E_k, \boldsymbol{\theta}) \sigma_{\text{th}}^i(E_k|\boldsymbol{\theta}) \quad (3.8)$$

where $\varepsilon^{ij}(E_k, \boldsymbol{\theta})$ are dataset-dependent efficiencies expressing the fraction of events from the i th process that survive the kinematical cuts imposed by the experiment, see Tab. 4.5. The statistical error due to the MC sample size is also evaluated and this is denoted $\delta\sigma^{ij}(E_k|\boldsymbol{\theta})$.

Performing a multiparameter brute-force scan and tune using $\sigma_{\text{th}}^{ij}(E_k|\boldsymbol{\theta})$ is computationally inefficient. As was highlighted in the introduction, the GENIE global analysis framework relies on Professor [129] to reduce the computational complexity of brute-force scans while allowing for massive parallelisation. Using the values of $\sigma_{\text{th}}^{ij}(E_k|\boldsymbol{\theta})$ computed for a number (N_R) of randomised P-dimensional vectors $\boldsymbol{\theta}$, produced within the P-dimensional hyper-cube defined by the parameter ranges given in Tab. 4.7, we use Professor to generate a parameterisation of $\sigma_{\text{th}}^{ij}(E_k|\boldsymbol{\theta})$ and $\delta\sigma^{ij}(E_k|\boldsymbol{\theta})$ that will be denoted with $\tilde{\sigma}_{\text{th}}^{ij}(E_k|\boldsymbol{\theta})$ and $\delta\tilde{\sigma}^{ij}(E_k|\boldsymbol{\theta})$ respectively. As discussed in [129], the parameterisation is a generic polynomial of order M in the P-dimensional space, whose analytical form is

$$\begin{aligned} \tilde{\sigma}_{\text{th}}^{ij}(E_k|\boldsymbol{\theta}) &= \alpha_0^{ijk} + \sum_{n=1}^P \beta_n^{ijk} \theta_n + \sum_{n \leq m} \gamma_{nm}^{ijk} \theta_n \theta_m \\ &+ \dots + \sum_{n_1 \leq \dots \leq n_M} \xi_{n_1 \dots n_M}^{ijk} \prod_{\ell=1}^M \theta_{n_\ell} \end{aligned} \quad (3.9)$$

where θ_n is the coordinate of the n th parameter. The polynomial order M is set by the

² $\boldsymbol{\theta} = (W_{\text{cut}}, M_A^{\text{QEL}}, M_A^{\text{RES}}, R_{vp}^{\text{CC}1\pi}, R_{vp}^{\text{CC}2\pi}, R_{vn}^{\text{CC}1\pi}, R_{vn}^{\text{CC}2\pi}, S_{\text{RES}}, S_{\text{DIS}})$

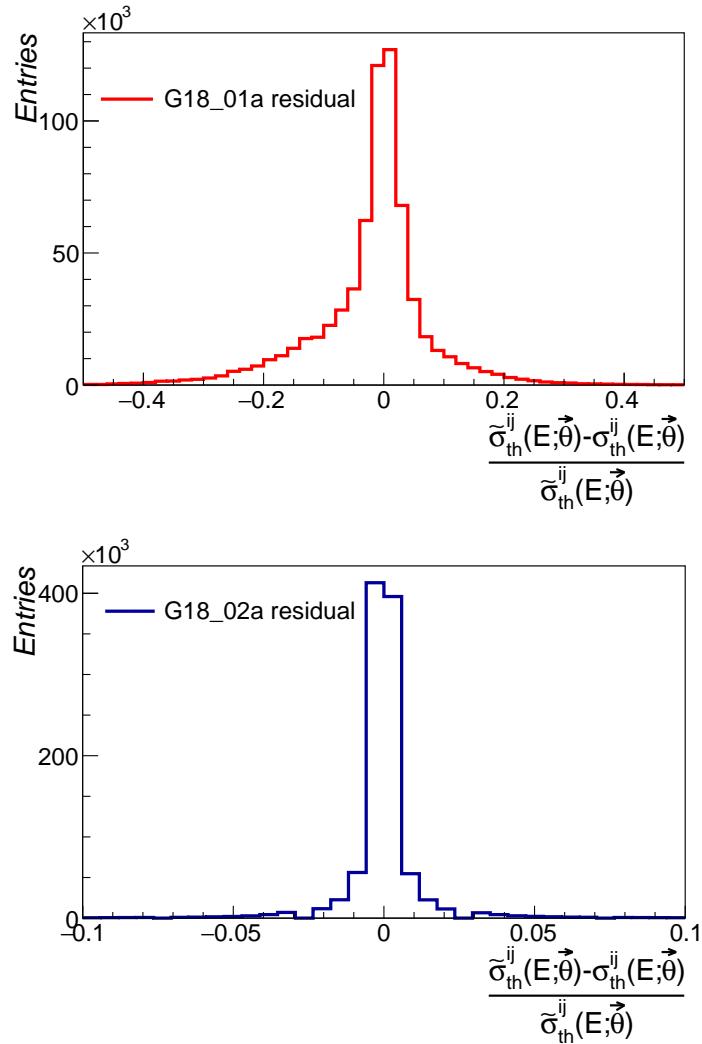


Figure 3.5: Fractional difference between true MC predictions calculated with a given $\vec{\theta}$ set vs the Professor parameterisation for both tunes. The G18_02a(b) parameterisation is improved as the number of scan points used and polynomial order M are higher.

user. The coefficients $\alpha_0^{ijk}, \beta_n^{ijk}, \gamma_{(nm)}^{ijk}, \dots, \xi_{(n_1\dots n_M)}^{ijk}$ are determined by Professor fitting the parameterisation against the computed $\sigma_{\text{th}}^{ij}(E_k|\boldsymbol{\theta})$. In the analysis presented here, a fourth-order polynomial was used for the G18_01a comprehensive model configuration while a fifth-order polynomial was used for G18_02a. Particularly, $N_R = 1500$ for G18_01a and $N_R = 2183$ for G18_02a. The accuracy of the parameterisation is demonstrated in the residual distributions shown in Fig. 3.5. fourth-order polynomials were the initial targets for both G18_01a and G18_02a fits. Some boundaries were relaxed during the analysis and more scan points were added allowing the usage of fifth-order polynomials for G18_02a. The parameterisation $\tilde{\sigma}_{\text{th}}^{ij}(E_k|\boldsymbol{\theta})$ is used instead of the exact predictions in order to estimate the best-fit parameters by minimizing the χ^2 .

3.5.2 Treatment of systematic uncertainties

A number of nuisance parameters, each with a corresponding prior, can also be used to tackle the problem of correlated datasets. As seen in Sec. 3.4.1, there are different datasets coming from the same experiments (ANL 12FT, BNL 7FT, BEBC, and FNAL 15FT). Each of these experiments share the same flux (from either a neutrino or an antineutrino beam), analysis procedure, etc. Therefore, there is a correlation between the datasets, even though it has not been quantified in the data releases. A possible approach is to add nuisance parameters that can connect datasets from experiments that used the same neutrino beam³. As the main systematic uncertainty comes from the fluxes, the nuisance parameters will act as scaling factors for our predictions ($\tilde{\sigma}_{\text{th}}^{ij}(E_k | \boldsymbol{\theta})$) and are same for datasets sharing the same flux.

Some of the ANL 12 ft and BNL 7 ft data were already corrected for the flux normalization [93]. Due to this correction, the associated systematic error is smaller and, accordingly, a more restricted nuisance parameter is applied to the reanalyzed datasets. These restricted parameters take into account other common systematics like reconstruction procedures, so they multiply all the predictions related to the same experiment. At the end of this procedure, each prediction can be scaled by up to two nuisance parameters, one for the flux and one for the remaining systematics. Thus, a total of nine independent nuisance parameters are used to account for the correlation between datasets. They are all the available combinations of experiment and neutrino flux exposure (v_μ and \bar{v}_μ) plus the restricted parameters for reanalyzed data:

$$\mathbf{f} = \left(f^{\text{ANL}}(v_\mu), f_{\text{Re}}^{\text{ANL}}(v_\mu), f^{\text{BNL}}(v_\mu), \right. \\ \left. f_{\text{Re}}^{\text{BNL}}(v_\mu), f^{\text{BNL}}(\bar{v}_\mu), f^{\text{BEBC}}(v_\mu), \right. \\ \left. f^{\text{BEBC}}(\bar{v}_\mu), f^{\text{FNAL}}(v_\mu), f^{\text{FNAL}}(\bar{v}_\mu) \right)$$

Quasielastic data for hydrogen and deuterium targets is included in the tune in order to constrain the nuisance parameters. Even though quasielastic data is not directly constraining the SIS parameters, it plays an important role to further constrain the fluxes of each experiment, as it is known at the 15 % level.

The main advantage of this method is the unbiased choice of the nuisance parameters, as their values will be determined by the minimization of the likelihood function. For the calculation of best-fit points and the calculation of intervals, these nuisance parameters are profiled (on every instance of our fit they are eliminated by substituting them with the value that minimizes χ^2).

³For instance, BEBC data on v_μ will have an associated nuisance parameter, which will be different from the one applied to BEBC \bar{v}_μ data.

Parameter	Prior
$f^{\text{ANL}}(\nu_\mu)$	1 ± 0.14
$f_{\text{Re}}^{\text{ANL}}(\nu_\mu)$	1 ± 0.05
$f^{\text{BNL}}(\nu_\mu)$	1 ± 0.14
$f_{\text{Re}}^{\text{BNL}}(\nu_\mu)$	1 ± 0.05
$f^{\text{BNL}}(\bar{\nu}_\mu)$	1 ± 0.15
$f^{\text{BEBC}}(\nu_\mu)$	1 ± 0.15
$f^{\text{BEBC}}(\bar{\nu}_\mu)$	1 ± 0.15
$f^{\text{FNAL}}(\nu_\mu)$	1 ± 0.15
$f^{\text{FNAL}}(\bar{\nu}_\mu)$	1 ± 0.15

Table 3.5: Nuisance parameters, f_j^j , per experiment (ANL 12 ft, BNL 7 ft, BEBC or FNAL 15 ft) and neutrino beam (ν_μ or $\bar{\nu}_\mu$). Priors consider the systematic uncertainty applied to each dataset as δf_j , where j is one of the datasets under study. The allowed range is $[0, 2]$ for nuisance parameters considered in the tune. See the text for the details on the prior values.

3.5.3 Discussion of priors

The likelihood is corrected using priors on parameters of interest ($\boldsymbol{\theta}$) and nuisance parameters (\mathbf{f}). Priors allow us to incorporate in this analysis the appropriate prefit uncertainties and correlations for the parameters of interest. Only Gaussian priors are considered at present.

The priors applied to each nuisance parameter f_j have a peak at 1 and different standard deviations δf_j . In general the total scaling factor applied to nonreanalyzed datasets are constrained by a conservative 15 % δf Gaussian prior, except for those nuisance parameters that act on the same experiment. Thus, the BEBC and FNAL 15 ft experiments have only one associated scaling factor $\delta f^{\text{BEBC}} = \delta f^{\text{FNAL}} = 0.15$ for both neutrino and anti-neutrino fluxes; the same is true for $f^{\text{BNL}}(\bar{\nu}_\mu)$. Up to two nuisance parameters can be applied to ANL 12 ft and BNL 7 ft data (i.e. $f^{\text{ANL}}(\nu_\mu)$ and $f_{\text{Re}}^{\text{ANL}}(\nu_\mu)$). The ANL 12 ft and BNL 7 ft restricted nuisance parameters, $f_{\text{Re}}^{\text{ANL}}(\nu)$ and $f_{\text{Re}}^{\text{BNL}}(\nu)$, have $\delta f = 5\%$. δf^{ANL} and δf^{BNL} are such that ANL 12 ft and BNL 7 ft non-reanalyzed datasets data are constrained by an overall 15% Gaussian. The full summary of the nuisance parameters is in Tab. 3.5.

Priors are applied to the parameters of interest to penalize disagreement with well-established parameter values. For instance, the description of neutrino CC quasielastic cross sections and single-pion production through baryon resonances is strongly determined by the shape of the weak axial and vector-transition form factors. As mentioned in Sec. 3.4.2, the form factors are extracted from datasets not used in this analysis. Our goal is not the extraction of the axial masses but the better estimation of the cross

section at the SIS region. For this reason, these values are used as priors in our global fits (see Tab. 4.7).

Another parameter of interest which is strongly constrained by data is the S_{DIS} parameter. This parameter dominates the cross-section behavior at high neutrino energies. Most of the data in that energy range comes from neutrino interactions with heavy nuclear targets and are therefore not included in the fit. A Gaussian prior is considered to ensure that agreement with these data are preserved⁴ by our tuning procedure. This would not be the case otherwise as the SIS region data would prefer much higher cross-section values for the DIS contribution. The prior on S_{DIS} provides a good solution for this problem because the degeneracy between DIS and nonresonant background parameters gives us multiple ways to accommodate good agreement between data and GENIE predictions in the SIS region. In other words, the introduction of the S_{DIS} prior breaks the degeneracy without adding more datasets to the fit.

3.5.4 Final form of the χ^2

Including all of the contributions from the previous sections and defining σ_d^{ijk} ($\delta\sigma_{\text{stat}}^{ijk}$) as the data central value (statistical error) corresponding to the $\tilde{\sigma}_{\text{th}}^{ij}(E_k|\boldsymbol{\theta})$ prediction, the complete form of our χ^2 distribution becomes:

$$\begin{aligned} \chi^2(\boldsymbol{\theta}, \mathbf{f}) = & \sum_{i,j,k} w^{ijk} \frac{(\phi_j(\mathbf{f}) \tilde{\sigma}_{\text{th}}^{ij}(E_k|\boldsymbol{\theta}) - \sigma_d^{ijk})^2}{(\delta\sigma_{\text{stat}}^{ijk})^2} \\ & + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \Sigma_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \\ & + \sum_j \frac{(f_j - 1)^2}{(\delta f_j)^2} \end{aligned} \quad (3.10)$$

where $\phi_j(\mathbf{f})$ is the product of the nuisance scaling factors that are relevant for j th dataset as described in Sec. 3.5.2. $\boldsymbol{\theta}_0$ and $\Sigma_{\boldsymbol{\theta}}$ are the central values and the covariance matrix of the priors for the parameters of interest, respectively. Equation 3.10 represents the full capability of our tuning machinery. However, the priors we applied for the present work were uncorrelated and so only the diagonal entries of $\Sigma_{\boldsymbol{\theta}}$ were used. The details on the priors applied in this analysis are described in Sec. 3.5.3.

The contribution of each point to the likelihood can be (de)emphasized using weights w^{ijk} to set the relative importance of different datasets (or of individual data points within a dataset). Such weighting schemes have been used extensively in general-purpose event generator tunes for the LHC (for an example, see [195]). In this particular analysis, the weights are used to include or exclude datasets only ($w^{ijk} \in \{0, 1\}$).

⁴The best agreement with all high energy data requires $S_{\text{DIS}} \sim 1$.

3.6 Tuning results

In order to properly understand the global tune, the tensions between datasets must be discussed. The identification of these tensions were instrumental to motivate a global tuning procedure and the usage of priors as described in previous sections to respect the agreement with all the datasets used in the fit at the same time. These tensions are studied by performing fits using a specific dataset to evaluate the impact of the partially-fitted predictions on the rest of the datasets included in the global tune.

In the following sections, different fit results are compared against a number of datasets, including those not used in the fit. Each dataset is identified by a label and Tab. 3.11-3.12 provides the proper references for each one of them. Some of those data are not extracted from (anti)neutrino interactions on hydrogen nor deuterium, yet they are included for historical reasons. This explains why some datasets are not compatible among themselves and reinforces the need for free nucleon tune only. Tab. 3.11-3.12 reports the target composition used for each dataset.

3.6.1 Partial fits

Two main subsets were identified in the global dataset in order to study tensions; inclusive and exclusive datasets. The fits consider the G18_02a CMC as the base configuration and include nuisance parameters to take into account the correlation between datasets from the same experiment, see Sec. 3.5 for more details. No priors on M_A^{RES} and M_A^{QEL} are applied as we are interested to see the impact of each subset on the prediction. The fit to inclusive data only is not sensitive to the scaling multiplicity parameters for the nonresonant background, therefore those parameters are fixed to their default values during the fit.

Partial fit results for inclusive and exclusive data are presented in Tab. 3.6. The tune against inclusive data only achieves much better agreement with inclusive data than GENIE G18_02a_00_000, see Fig. 3.6. This difference between the old and new inclusive tune is due to (1) the inclusion of only hydrogen and deuterium datasets, and (2) the effect of the nuisance parameters⁵. Particularly, without exclusive data, a small reduction of the resonant cross section is already observed in the CC RES region. The result for M_A^{RES} is consistent with previous results without the addition of priors [35]. W_{cut} is pulled to the lower edge of the parameter range: the parameter uncertainty could not be estimated as the χ^2 minimum was found on the contour.

As expected, the fit to exclusive data only is able to correctly describe exclusive

⁵Note that in the *default* GENIE tune, there was no method to include correlation between datasets coming from the same experiment.

Parameter	Inclusive	Exclusive
$W_{\text{cut}}(\text{GeV})$	1.52	2.00
$M_A^{\text{QEL}}(\text{GeV}/c^2)$	0.98 ± 0.01	1.003 ± 0.008
$M_A^{\text{RES}}(\text{GeV}/c^2)$	1.15 ± 0.02	0.88 ± 0.02
$R_{vp}^{\text{CC}1\pi}$	(0.10)	0.30 ± 0.02
$R_{vp}^{\text{CC}2\pi}$	(1.00)	1.28 ± 0.06
$R_{vn}^{\text{CC}1\pi}$	(0.30)	0.294 ± 0.002
$R_{vn}^{\text{CC}2\pi}$	(1.00)	3.19 ± 0.09
S_{RES}	0.87 ± 0.03	0.88 ± 0.02
S_{DIS}	1.027 ± 0.005	1.026 ± 0.007

Table 3.6: Parameter best-fit results for partial fits to inclusive and exclusive data using the G18_02a CMC as base configuration. Values within parentheses are kept fixed during the fit: they are the historical default values.

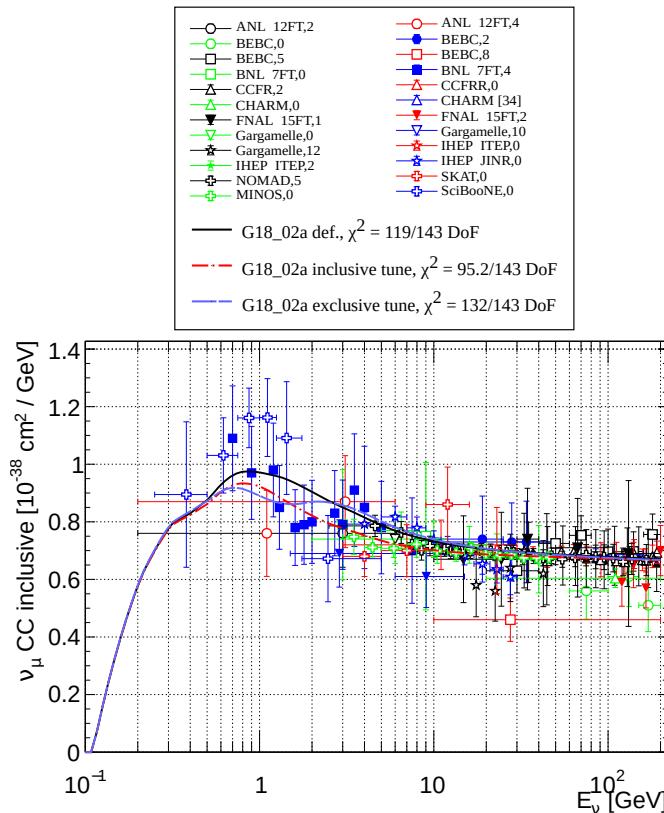


Figure 3.6: Comparison of ν_μ CC Inclusive cross section against bubble chamber data. The *default* GENIE configuration corresponds to the G18_02a_00_000 CMC. The inclusive tune is performed using the filled datapoints only. The predictions are computed with GENIE version 3.2 using the parameters specified in Tab. 3.6. The χ^2 values are calculated against all inclusive data available from bubble chamber experiments.

3.6. TUNING RESULTS

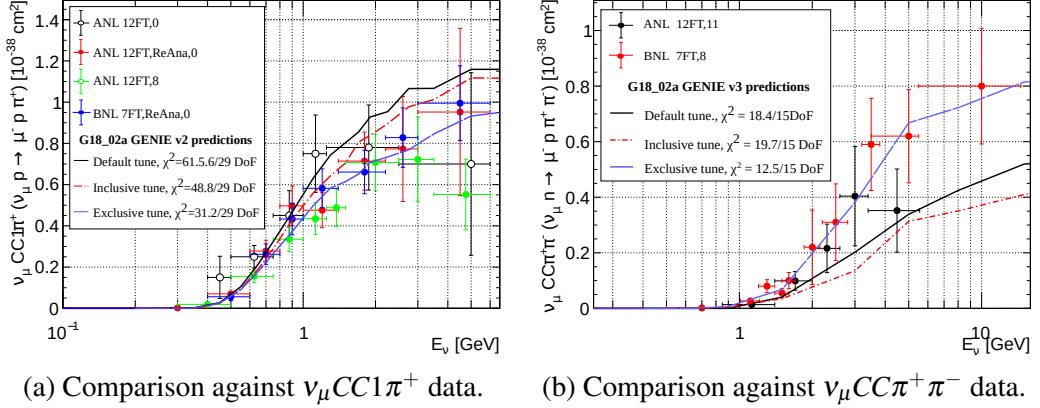


Figure 3.7: Comparison of v_μ CC exclusive channels against bubble chamber data. The *default* GENIE configuration corresponds to the G18_02a_00_000 CMC. The exclusive tune is sensitive to the exclusive datasets only, see Tab. 4.5. The predictions are computed with GENIE version 3.2 using the parameters specified in Tab. 3.6. The χ^2 values are calculated against all exclusive data shown in each plot.

datasets for one and two pion production. The low cross-section data for one pion production forces all the relevant parameters to decrease with respect to the default values see Fig. 3.7a. At the same time, two pion production data forces $R_{\nu p}^{CC2\pi}$, $R_{\nu n}^{CC2\pi}$ and W_{cut} to increase in order to match two pion production data, see Fig. 3.7b. The agreement with v_μ CC inclusive data is worse, see Fig. 3.6, but the compatibility is still acceptable given the large uncertainties on the data in that region. On the other hand, the partial fit does not obtain a good prediction for M_A^{RES} . W_{cut} is fixed to its maximum value of 2 GeV to avoid nonphysical regions.

The exclusive fit clearly shows a preference for a larger total cross section in the neutrino energy region between 1 and 10 GeV due to the high value of $R_{\nu p}^{CC2\pi}$ and $R_{\nu n}^{CC2\pi}$. This is a tension between exclusive and inclusive datasets as the inclusive prefer a lower value in that E_ν region. Since inclusive data constitute about 40% of all the data points, the inclusion of priors for the axial masses and S_{DIS} becomes crucial to overcome the tension [196].

3.6.2 Global fit

The analysis procedure outlined in previous sections was applied to the comprehensive model configurations listed in Sec. 3.3. The best-fit parameter values obtained from the GENIE analysis for each alternative CMC are shown in Tab. 3.7 and Tab. 3.8. The GENIE v3 cross-section curves that correspond to the two sets of tuned parameters are shown in Figs. 3.8- 3.13. For reference, we also show the cross-section predictions made by the G18_02a_00_000 CMC that are available in the last public release of the

Parameter	G18_01a(/b)	G18_02a(/b)
W_{cut}	1.94	1.81
M_A^{QEL}	1.00 ± 0.01	1.00 ± 0.013
M_A^{RES}	1.09 ± 0.02	1.09 ± 0.014
$R_{vp}^{\text{CC1}\pi}$	0.06 ± 0.03	0.008
$R_{vp}^{\text{CC2}\pi}$	1.1 ± 0.2	0.94 ± 0.075
$R_{vn}^{\text{CC1}\pi}$	0.14 ± 0.03	0.03 ± 0.010
$R_{vn}^{\text{CC2}\pi}$	2.8 ± 0.4	2.3 ± 0.12
S_{RES}	0.89 ± 0.04	0.84 ± 0.028
S_{DIS}	1.03 ± 0.02	1.06 ± 0.01
$\chi^2/157 \text{ DoF}$	1.84	1.64

Table 3.7: Best-fit parameter values and parameter ranges obtained by requiring that $\Delta\chi^2_{\text{profiled}} < \Delta\chi^2_{\text{critical}} = 1$. Results are shown for all alternative CMCs considered in this analysis. The best-fit values obtained for the G18_02a(/b) CMC can be used for the G18_10a(/b) as the same bare-nucleon underlying models are used.

GENIE v3 series (3.2).

For all CMCs the tune has the most impact on the SIS region. In the inclusive cross-section prediction, this translates into a decrease of both ν_μ and $\bar{\nu}_\mu$ CC inclusive cross section in the 0.5-10 GeV region, see Fig. 3.8 and Fig. 3.9. At the same time, the cross section at higher neutrino energies has barely changed, respecting the constraints of high-energy data. The agreement with quasielastic data, included in the tune in order to constrain the fluxes of each experiment, remained the same, see Fig. 3.10.

As discussed in Sec. 3.6, this decrease of the inclusive cross section at the SIS region is driven mainly by one-pion production data. The impact on one-pion exclusive channels is shown for (anti)neutrino on proton, Fig. 3.11, and neutrino on neutron, Fig. 3.12. The reduction of the one pion production cross section for neutrino on proton and neutron shows an improvement on $\nu_\mu \text{CC1}\pi^+$, ν_μ and $\nu_\mu \text{CC1}\pi^-$ and $\nu_\mu \text{CC1}\pi^0$ channels when comparing it with the available data.

Two-pion production exclusive cross sections are summarized in Fig. 3.13. This is the first time that two-pion production data are used to tune the SIS region, allowing the $R_{vp}^{\text{CC2}\pi}$ and $R_{vn}^{\text{CC2}\pi}$ parameters to be constrained. In this case, the two-pion exclusive cross section was underestimated by the *default* tune. For this particular exclusive process, comparisons are made against $\nu_\mu \text{CC}\pi^+\pi^+$, $\nu_\mu \text{CC}\pi^+\pi^0$ and $\nu_\mu \text{CC}\pi^+\pi^-$ data. The shape of the GENIE prediction for the $\nu_\mu \text{CC}\pi^+\pi^+$ and $\nu_\mu \text{CC}\pi^+\pi^0$ channels differs strongly from the data, and the models are not able to accommodate this behavior. However, the agreement with $\nu_\mu \text{CC}\pi^+\pi^-$ data has improved by increasing the cross

Parameter	G18_01a(/b)	G18_02a(/b)
$f^{\text{ANL}}(\nu_\mu)$	0.98 ± 0.05	0.89 ± 0.05
$f_{\text{Re}}^{\text{ANL}}(\nu_\mu)$	1.12 ± 0.05	1.2 ± 0.05
$f^{\text{BNL}}(\nu_\mu)$	1.01 ± 0.04	1.06 ± 0.04
$f_{\text{Re}}^{\text{BNL}}(\nu_\mu)$	1.08 ± 0.05	1.03 ± 0.04
$f^{\text{BNL}}(\bar{\nu}_\mu)$	1.00 ± 0.10	0.99 ± 0.10
$f^{\text{BEBC}}(\nu_\mu)$	0.91 ± 0.04	0.86 ± 0.03
$f^{\text{BEBC}}(\bar{\nu}_\mu)$	1.04 ± 0.04	0.99 ± 0.03
$f^{\text{FNAL}}(\nu_\mu)$	0.97 ± 0.04	0.94 ± 0.04
$f^{\text{FNAL}}(\bar{\nu}_\mu)$	0.95 ± 0.05	0.92 ± 0.05

Table 3.8: Best-fit nuisance parameters, f^j , per experiment (ANL 12 ft, BNL 7 ft, BEBC or FNAL 15 ft) and neutrino beam (ν_μ or $\bar{\nu}_\mu$). The nuisance parameters included in the fit are independent of GENIE.

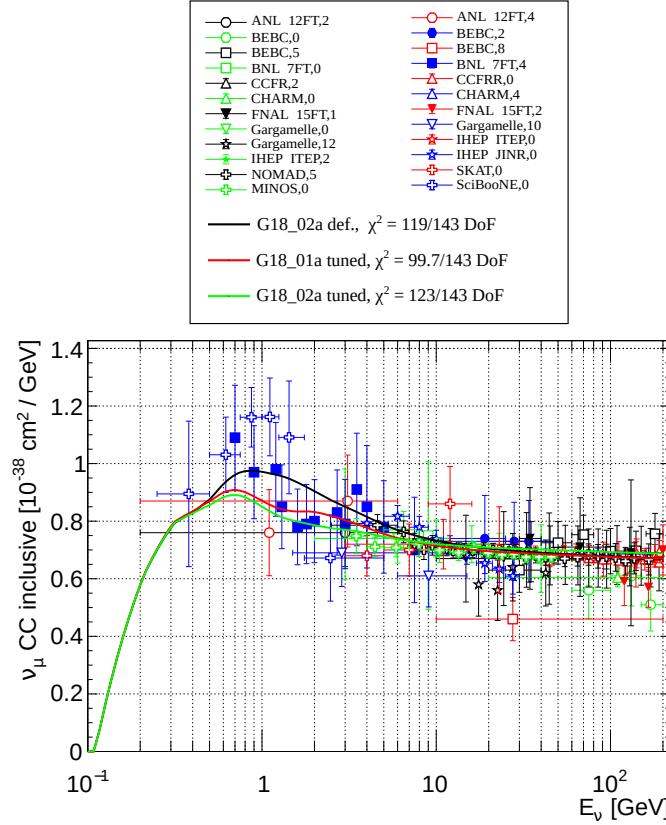


Figure 3.8: Best-fit prediction impact on muon neutrino CC inclusive cross sections as a function of the neutrino energy (E_ν). The associated predictions for the *default* G18_02a and tuned G18_01a and G18_02a are computed with GENIE v3.0.6. Predictions are compared against all the available data (anti-)neutrino interactions on H, ${}^2\text{H}$ and heavier targets. Both CMC have been tuned against some H, ${}^2\text{H}$ data (filled markers). Each χ^2 is computed using all data available. In Tab. 3.10, the χ^2 values per dataset are specified.

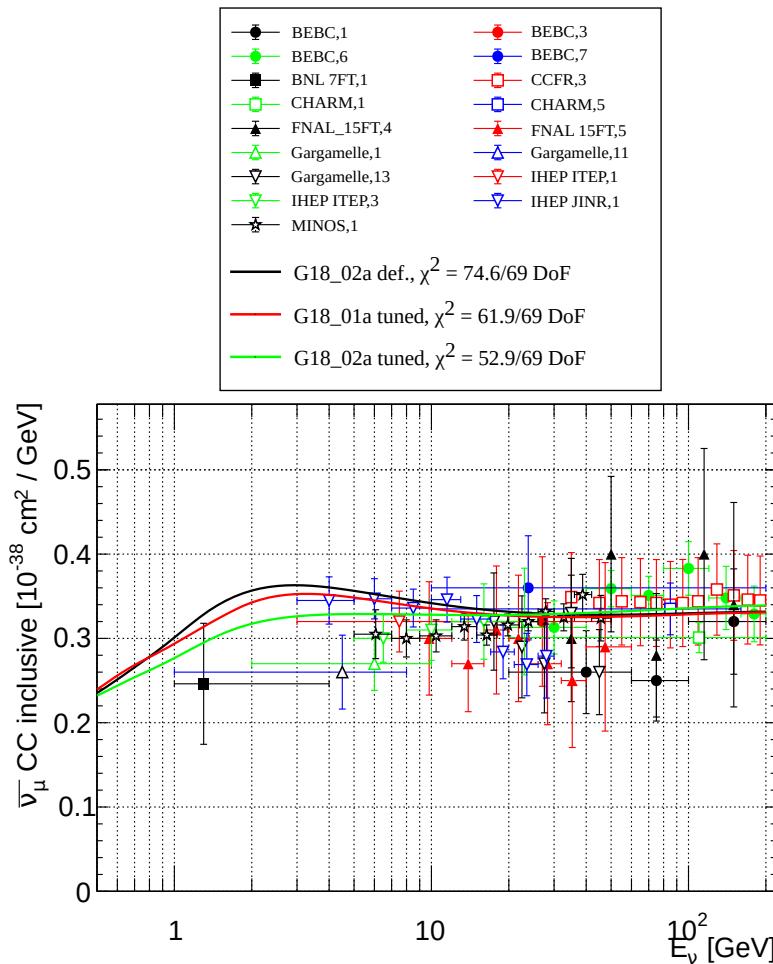
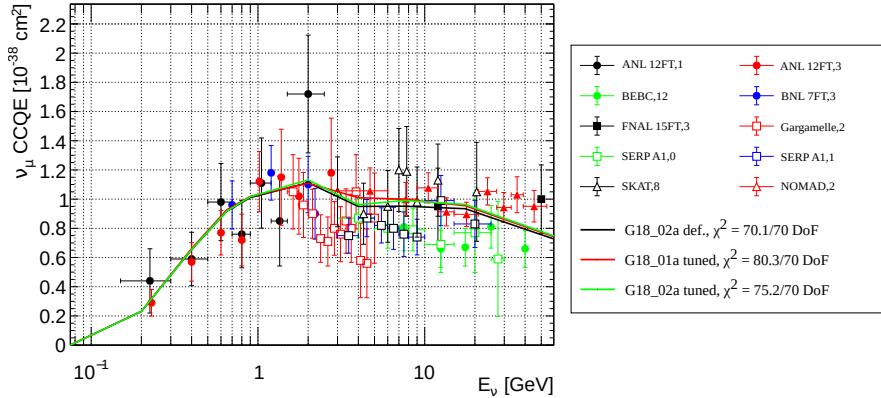
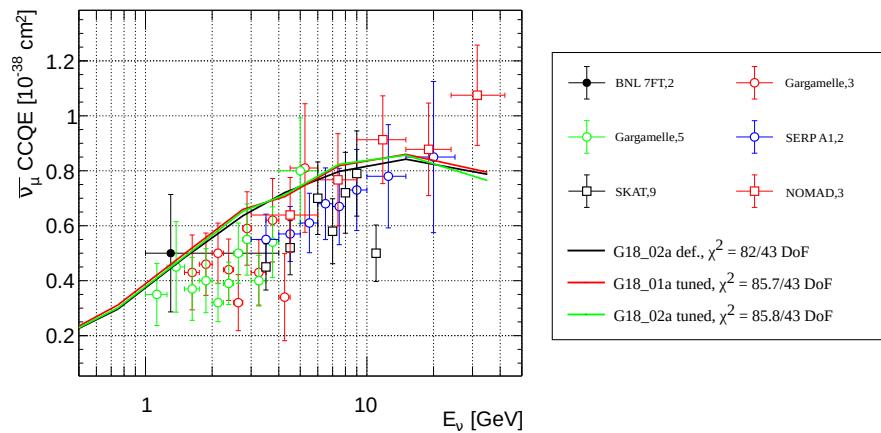


Figure 3.9: Best-fit prediction impact on muon anti-neutrino CC inclusive cross sections as a function of the neutrino energy (E_ν). The associated predictions for the *default* G18_02a and tuned G18_01a and G18_02a are computed with GENIE v3.0.6. Predictions are compared against all the available data (anti)neutrino interactions on H, ${}^2\text{H}$ and heavier targets. Both CMC have been tuned against some H, ${}^2\text{H}$ data (filled markers). Each χ^2 is computed using all data available. In Tab. 3.10, the χ^2 values per dataset are specified.



(a) Comparison of ν_μ CC quasielastic cross-section data against the *default* and tuned CMCs.



(b) Comparison of $\bar{\nu}_\mu$ CC quasielastic cross-section data against the *default* and tuned CMCs.

Figure 3.10: Best-fit prediction impact on muon (anti)neutrino CC quasielastic cross sections as a function of the neutrino energy (E_ν). The associated predictions for the *default* G18_02a, and tuned G18_01a and G18_02a are computed with GENIE v3.0.6. Predictions are compared against all the available data (anti)neutrino interactions on H, ${}^2\text{H}$ and heavier targets. Both CMC have been tuned against some H, ${}^2\text{H}$ data (filled markers). Each χ^2 is computed using all data available. In Tab. 3.10, the χ^2 values per dataset are specified.

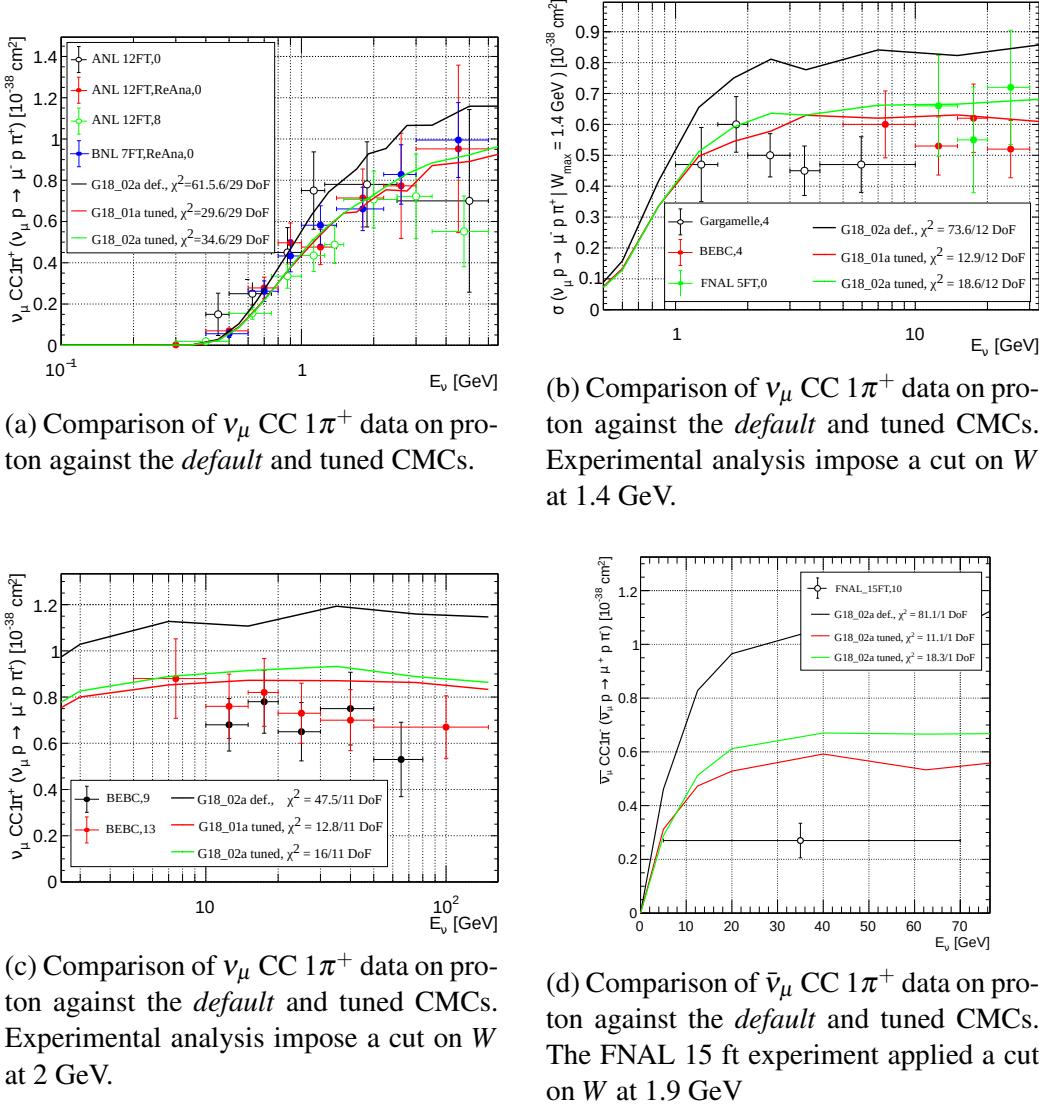
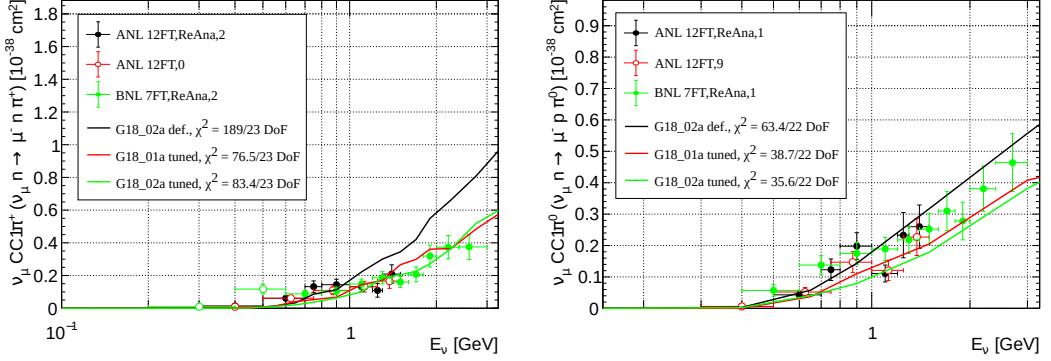


Figure 3.11: Best-fit prediction impact on muon neutrino on proton CC one pion production cross sections as a function of the neutrino energy (E_ν). The associated predictions for the *default* G18_02a and tuned G18_01a and G18_02a are computed with GENIE v3.0.6. Experimental cuts are also applied to the predictions when needed. Predictions are compared against all the available data (anti)neutrino interactions on H, ${}^2\text{H}$ and heavier targets. Both CMC have been tuned against some H, ${}^2\text{H}$ data (filled markers). Each χ^2 is computed using all data available. In Tab. 3.10, the χ^2 values per dataset are specified.

3.6. TUNING RESULTS



(a) Comparison of ν_μ CC $1\pi^+$ data on neutron against the *default* and tuned CMCs.

(b) Comparison of ν_μ CC $1\pi^0$ data on neutron against the *default* and tuned CMCs.

Figure 3.12: Best-fit prediction impact on muon neutrino on neutron CC one pion production cross sections as a function of the neutrino energy (E_ν). The associated predictions for the *default* G18_02a and tuned G18_01a and G18_02a are computed with GENIE v3.0.6. Predictions are compared against the original and reanalyzed ANL 12 ft and BNL 7 ft data [93, 184]. Only reanalyzed data with $E_\nu > 0.5$ GeV is used in the tune (filled markers). Each χ^2 is computed using all data available.

section with respect to the *default* CMC.

Despite the tensions between inclusive and exclusive data discussed in Sec. 3.4.2, the overall agreement for both cross-section model constructions has improved, see Tab. 3.10. Particularly, $\bar{\nu}_\mu$ CC inclusive predictions show better agreement after the tune, and the same is observed for ν_μ CC inclusive predictions for the G18_01a free nucleon tune. Although the impact on the cross-section prediction of the tune is similar for the existing configurations, the response of each model at the parameter level is not expected to be the same. Therefore, each tune is strongly affected by how the model is able to accommodate the data by modifying the tuned parameters. This reflects on the R_m parameters and W_{cut} which best-fit values are incompatible in some cases, such as for $R_{\nu n}^{\text{CC}1\pi}$, see Tab. 3.7. Particularly, the behaviour of $R_{\nu p}^{\text{CC}1\pi}$ on the G18_02a(/b) tune was showing preference for nonphysical regions of the tune, forcing us to fix this value to $R_{\nu p}^{\text{CC}1\pi} = 0.008$. On the other hand, the remaining parameters, such as M_A^{RES} and M_A^{QEL} , show agreement between the tunes and respect the applied priors.

3.6.3 Parameter error estimation

An estimate of the parameter uncertainties is shown in Tab. 3.7. For each parameter of interest allowed to float in the fit, the table shows the range of values that satisfies the condition $\Delta\chi^2_{\text{profile}}(\theta_i) < \Delta\chi^2_{\text{critical}} = 1$. In the previous expression, the function $\Delta\chi^2_{\text{profile}}(\theta_i)$ is constructed by fixing θ_i to a desired value and minimising the quantity $\Delta\chi^2(\boldsymbol{\theta}, \mathbf{f}) = \chi^2(\boldsymbol{\theta}, \mathbf{f}) - \chi^2_{\min}$ with respect to all other parameters that were allowed to

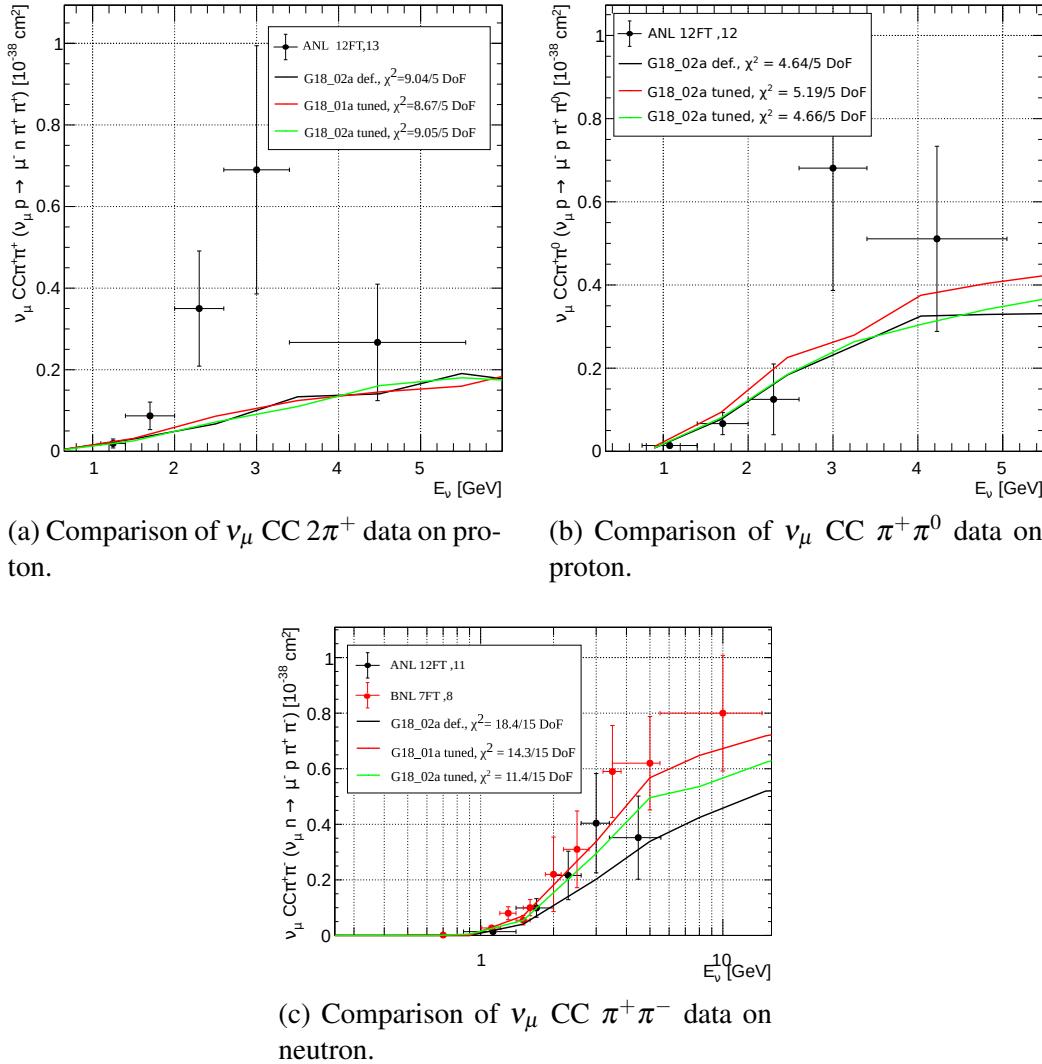


Figure 3.13: Best-fit prediction impact on muon neutrino CC two-pion production cross sections as a function of the neutrino energy (E_ν). The comparisons to two-pion production data are shown against the *default* and tuned CMCs. The associated predictions for the *default* G18_02a and tuned G18_01a and G18_02a are computed with GENIE v3.0.6. Predictions are compared against ANL 12 ft and BNL 7 ft data.

3.6. TUNING RESULTS

	M_A^{RES}	M_A^{QEL}	$R_{vp}^{\text{CC}1\pi}$	$R_{vp}^{\text{CC}2\pi}$	$R_{vn}^{\text{CC}1\pi}$	$R_{vn}^{\text{CC}2\pi}$	S_{RES}	S_{DIS}
M_A^{RES}	5.3E-4	-7E-5	5E-5	-8E-4	2.2E-4	-2.4E-3	-4.3E-4	-9E-5
M_A^{QEL}	-7E-5	1.2E-4	-6E-5	-1.2E-4	-5E-5	-7.6E-4	1.2E-4	1E-5
$R_{vp}^{\text{CC}1\pi}$	5E-5	-6E-5	9.3E-4	-1.6E-3	2.6E-4	5.4E-4	-2.8E-4	-6E-5
$R_{vp}^{\text{CC}2\pi}$	-8E-4	-1.2E-4	-1.6E-3	2.7E-2	2.0E-5	-2.5E-4	2E-3	-6.2E-4
$R_{vn}^{\text{CC}1\pi}$	2.2E-4	-5E-5	2.6E-4	2E-5	7.1E-4	2.3E-3	-5.3E-4	-8E-5
$R_{vn}^{\text{CC}2\pi}$	-2.4E-3	-7.6E-4	5.4E-4	-2.5E-4	2.3E-3	9.6E-2	-2.5E-3	-1.4E-3
S_{RES}	-4.3E-4	1.2E-4	-2.8E-4	2E-3	-5.3E-4	-2.5E-3	1.3E-3	1.8E-4
S_{DIS}	-9E-5	1E-5	-6E-5	-6.2E-4	-8E-5	-1.4E-3	1.8E-4	5.1E-4

(a) G18_01a(/b) covariance matrix.

	M_A^{RES}	M_A^{QEL}	$R_{vp}^{\text{CC}2\pi}$	$R_{vn}^{\text{CC}1\pi}$	$R_{vn}^{\text{CC}2\pi}$	S_{RES}	S_{DIS}
M_A^{RES}	1.7E-4	2.0E-5	-1.9E-4	-6.0E-5	4.4E-4	6.0E-5	-4.0E-5
M_A^{QEL}	2.0E-5	1.8E-4	-7.0E-5	3.0E-5	-2.1E-4	1.5E-4	1.0E-5
$R_{vp}^{\text{CC}2\pi}$	-1.9E-4	-7.0E-5	5.5E-3	1.5E-4	-2.4E-3	-6.9E-4	3.0E-5
$R_{vn}^{\text{CC}1\pi}$	-6.0E-5	3.0E-5	1.5E-4	1.1E-4	-1.0E-4	-6.0E-5	6.0E-5
$R_{vn}^{\text{CC}2\pi}$	4.4E-4	-2.1E-4	-2.4E-3	-1.0E-4	1.3E-2	2.3E-4	-8.0E-5
S_{RES}	6.0E-5	1.5E-4	-6.9E-4	-6.0E-5	2.3E-4	6.0E-4	-4.0E-5
S_{DIS}	-4.0E-5	1.0E-5	3.0E-5	6.0E-5	-8.0E-5	-4.0E-5	8.0E-5

(b) G18_02a(/b) covariance matrix.

Table 3.9: Parameter covariance matrices extracted the GENIE fit for the tuned CMCs.

float in the fit. See Sec. 3.5.1 for the definition of $\chi^2(\boldsymbol{\theta}, \mathbf{f})$. The constant χ^2_{\min} corresponds to the minimum value of $\chi^2(\boldsymbol{\theta}, \mathbf{f})$ obtained from the global fit. The $\Delta\chi^2_{\text{profile}}(\theta_i)$ functions we derive from our analysis are shown in Fig. 3.14, for all parameters θ_i that were allowed to float in the fit, up to $\Delta\chi^2_{\text{profile}}$ values of 2. Particularly, W_{cut} is fixed to the best-fit value during this approach, as it is an ad-hoc parameter introduced by the generator: by fixing it, its uncertainty will be reflected on the other parameters. It is important to emphasize that the uncertainties quoted relate only to $\Delta\chi^2_{\text{critical}} = 1$. However, this region is strongly determined by the underlying model used in the tune.

A covariance matrix is also obtained through the inversion of the Hessian of the log-likelihood function at the best-fit parameter point. The corresponding correlation matrices are presented in Tab. 3.9a and Tab. 3.9b for the tunes of all 4 different cross-section model constructions used in this work (see the correlation matrices in Fig. 3.15 and Fig. 3.16 for a graphical interpretation). An example of the propagation of model uncertainties from the Professor output to the GENIE Comparisons framework is shown in Fig. 4.11.

Joint $\Delta\chi^2_{\text{profile}}(\theta_i, \theta_j)$ functions, constructed by fixing two parameters at a grid of

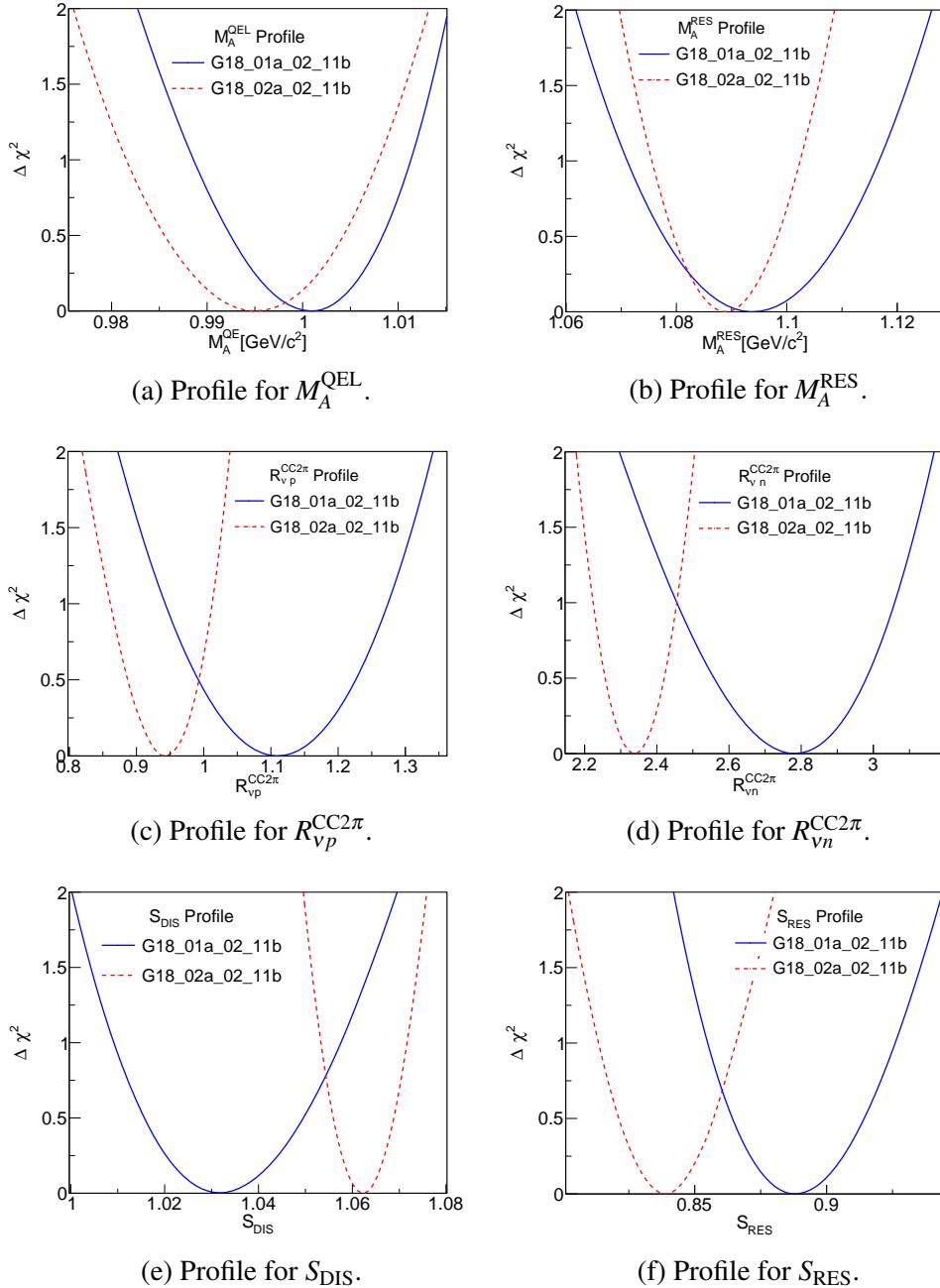


Figure 3.14: $\Delta\chi^2_{\text{profile}}(\theta_i)$ functions obtained fixing the parameter under study and minimizing $\Delta\chi^2(\vec{\theta}, \vec{f}) = \chi^2(\vec{\theta}, \vec{f}) - \chi^2_{\min}$ respect the other parameters in the tune. The profiles for both tunes are shown for each case. The G18_02a profiles show that this configuration is less able to accommodate in the model variations on each parameter best-fit values.

3.6. TUNING RESULTS

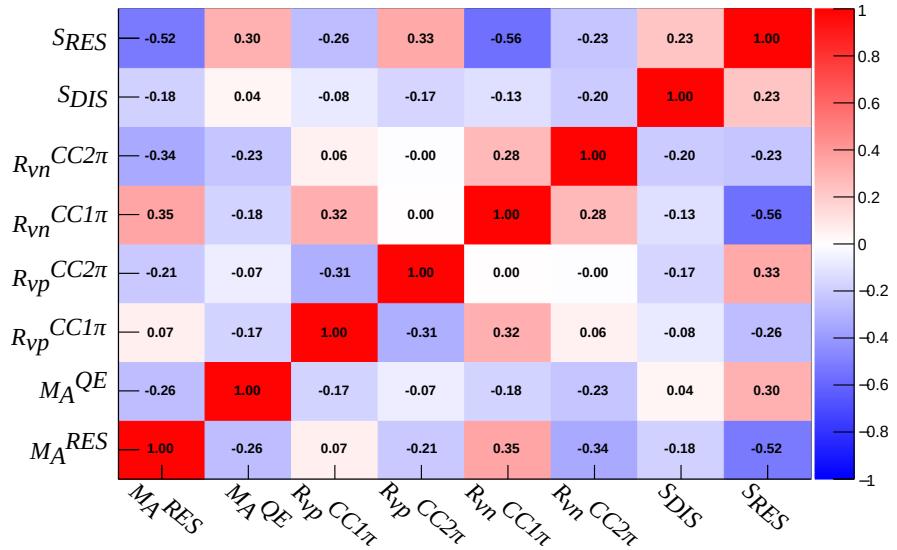


Figure 3.15: Parameter correlation matrix from the GENIE fit using the G18_01a(/b) CMC correlation matrix.

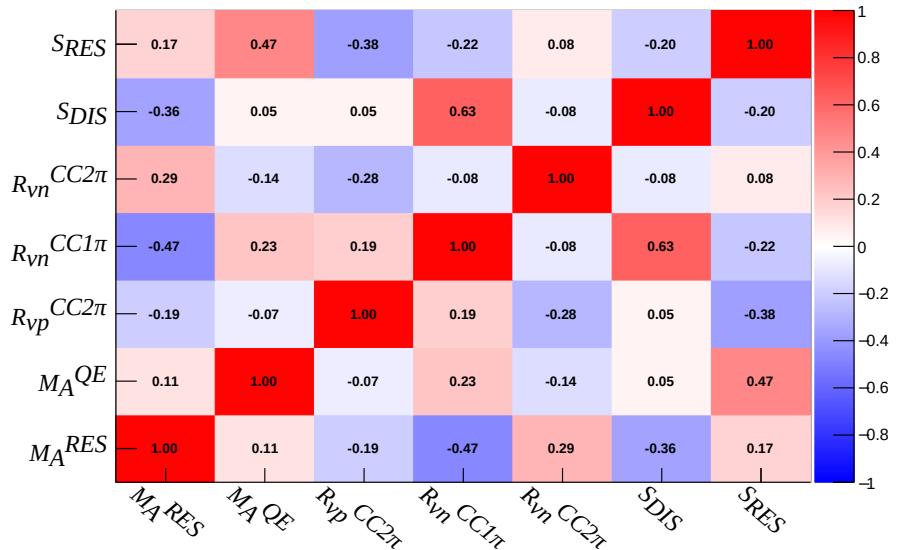


Figure 3.16: Parameter correlation matrix from the GENIE fit using the G18_02a(/b) CMC correlation matrix.

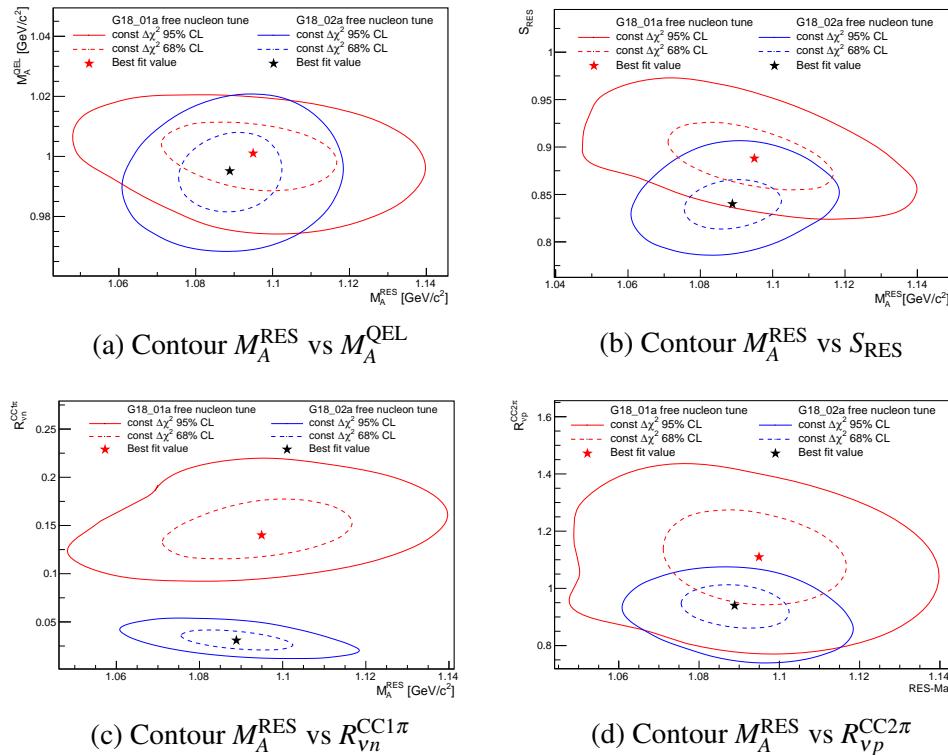


Figure 3.17: Joint $\Delta\chi^2_{\text{profile}}(\theta_i, \theta_j)$ functions obtained fixing the two parameters under study and minimizing $\Delta\chi^2(\vec{\theta}, \vec{f})$ respect the other parameters in the tune. The contours for both tunes are shown for each case as well as the best-fit values of each tune.

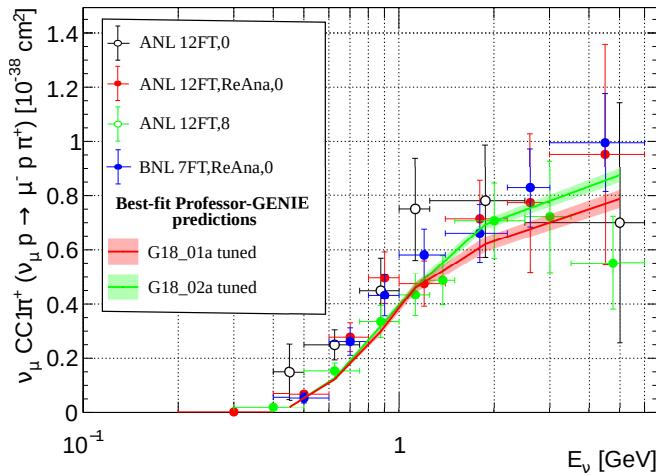
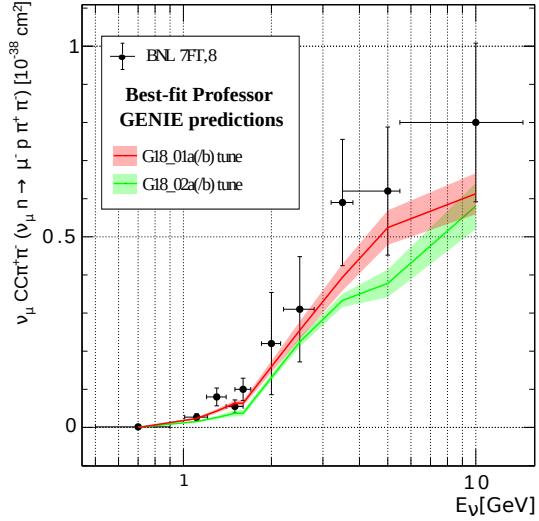

 (a) $\nu_\mu \text{CC}1\pi^+$ comparison.

 (b) $\nu_\mu \text{CC}\pi^+\pi^-$ comparison.

Figure 3.18: Comparisons of neutrino data against predictions obtained from the Professor-GENIE parametrization at the best-fit value. The uncertainties of the tune are propagated to the prediction considering the full covariance matrix.

values and minimizing and $\Delta\chi^2(\boldsymbol{\theta}, \mathbf{f})$ with respect to all other new parameters, are shown in Fig. 3.17 for selected sets of parameters. In Figs.3.17, we can see that the coverage of the parameter space for the 68% and 95% confidence level lines is wider for the G18_01a(/b) tunes. This characteristic is again not related with how well we can constrain the parameters from the data, but with the capability of the models to accommodate this data in each model implementation.

3.7 Conclusions

GENIE has released a number of comprehensive model configurations (CMCs) which consist of different modeling aspects combined altogether. In previous GENIE versions, there was a preferred *default* comprehensive model which failed to describe both inclusive and exclusive channels due to unresolved tensions between the data. These tensions, which are crucial to understand for the new generation of neutrino experiments, motivated a careful investigation and retune of the bare-nucleon cross-section model for all GENIE comprehensive models available in GENIE v3. Best-fit values and correlations for several parameters influencing the GENIE bare-nucleon cross sections are released in this paper.

In GENIE v3, we focus on improving understanding of the SIS region by tuning the GENIE CMC predictions on hydrogen and deuterium data from the ANL 12 ft, BNL 7 ft, BEBC and FNAL 15 ft bubble chamber experiments. The tuning of the nonresonant background takes a central stage in this work in order to remove double counting issues. The SIS region has been tuned against ν_μ and $\bar{\nu}_\mu$ CC inclusive, quasielastic, one-pion and two-pion integrated cross sections as a function of E_V . Quasielastic data has been introduced to the fit to better constrain the flux of each experiment.

The global fit describes both inclusive and exclusive cross sections simultaneously. Tensions were identified between datasets and interpreted as a need for a better tuning procedure, since difficulties in trying to extract information from these datasets were highlighted in previous attempts [131, Fig. 8]. The tensions motivated the systematic treatment of correlations between datasets and the inclusion of priors as discussed in the paper.

After the global tune, the inclusive cross section at the 1–10 GeV energy region decreased with respect to the historical *default* predictions. The tuned GENIE predictions of one pion production cross sections on free nucleons (ν_μ CC $p\pi^+$, $n\pi^+$, $p\pi^0$ and $p\pi^+\pi^-$) show a decrease in the nonresonant background contribution, improving the agreement with the data. The prediction for two-pion production mechanisms is also in better agreement with data for the $\nu_\mu n \rightarrow \mu^- p\pi^+\pi^-$ channel by increasing the

3.7. CONCLUSIONS

two pion production nonresonant background contribution.

This is only the first GENIE paper about a systematic tuning procedure. In the future the collaboration plans to release tunes against hadronisation and nuclear data. Both will address issues that near-future neutrino oscillation experiments will have to face.

Dataset	NDOF	χ^2 values for G18_01a		χ^2 values for G18_02a	
		Default	Best Fit	Default	Best Fit
ν_μ CC Inclusive					
BNL 7 ft [161]	13	11.1	9.95	14.7	7.75
BEBC [159]	3	0.215	0.101	0.067	0.045
FNAL 15 ft [164, 190]	10	3.85	3.92	4.04	4.98
$\bar{\nu}_\mu$ CC Inclusive					
BEBC [148, 159, 160]	11	11.17	11.5	9.79	9.8
BNL 7 ft [170]	1	1.83	1.51	1.96	0.827
FNAL 15 ft [166, 171]	13	3.86	4.12	4.32	4.13
$\nu_\mu n \rightarrow \mu^- n\pi^+$					
ANL 12 ft [184]	5	11.6	9.88	27.3	14.3
ANL 12 ft, ReAna [93]	7	31.3	21.0	48.8	25.2
BNL 7 ft, ReAna [93]	11	103	45.7	112	43.9
$\nu_\mu p \rightarrow \mu^- p\pi^+$					
ANL 12 ft, ReAna [93]	8	11	8.71	17.8	9.64
BNL 7 ft, ReAna [93]	7	6.16	3.11	9.71	3.9
BEBC [175, 186, 188]	15	33.98	15.9	82.6	21.0
FNAL [193]	3	1.11	0.74	2.87	0.66
$\nu_\mu n \rightarrow \mu^- p\pi^0$					
ANL 12 ft [184]	5	4.89	4.98	7.57	4.63
ANL 12 ft, ReAna [93]	7	12.6	12.0	17.4	11.5
BNL 7 ft, ReAna [93]	10	31.8	21.7	38.4	19.4
$\nu_\mu p \rightarrow \mu^- n\pi^+\pi^-$					
ANL 12 ft [191]	5	9.23	8.67	9.04	9.05
$\nu_\mu p \rightarrow \mu^- p\pi^+\pi^0$					
ANL 12 ft [191]	5	4.28	5.19	4.64	4.66
$\nu_\mu n \rightarrow \mu^- p\pi^+\pi^-$					
ANL 12 ft [191]	5	8.24	8.36	8.09	4.95
BNL 7 ft [152]	10	11.6	5.96	10.3	6.46
ν_μ CC QE					
ANL 12 ft [91, 174]	15	11.7	12.2	11.75	11.58
BNL 7 ft [179]	4	6.88	6.91	6.98	7.58
BEBC [175]	5	8.18	9.45	8.21	9.54
FNAL [176]	2	0.886	0.951	0.992	0.893
$\bar{\nu}_\mu$ CC QE					
BNL 7 ft [170]	1	0.161	0.135	0.078	0.106
Total	182	400.6	229.5	459.4	236.5

Table 3.10: Contributions to the *default* and tuned χ^2 for the datasets included in the fit. The data points with $E_\nu < 0.5$ GeV, a total of 10 points, are considered in the χ^2 calculations of this table, but were not used in the fit. For the calculation of the χ^2 , the covariance matrix between the datasets is used instead of Eq. 3.10, which incorporates nuisance parameters which are not implemented in GENIE. This explains the difference when comparing with the χ^2 out of Professor from Tab. 3.7.

3.7. CONCLUSIONS

Experiment	Target	Tag	Ref.
ν_μ CC Inclusive			
ANL 12 ft	$^1\text{H}, ^2\text{H}$	ANL 12FT,2	[91]
BEBBC	$^2\text{H}, \text{Ne}$	BEBBC,0	[147]
BEBBC	$^2\text{H}, \text{Ne}$	BEBBC,5	[148]
BNL 7 ft	$^2\text{H}, \text{Ne}$	BNL 7FT,0	[149]
CCFR	Fe	CCFR,2	[150]
CHARM	C	CHARM,0	[151]
FNAL 15 ft	^2H	FNAL 15FT,1	[152]
Gargamelle	$\text{C}_3\text{H}_8, \text{CF}_3\text{Br}$	Gargamelle,0	[153]
Gargamelle	$\text{C}_3\text{H}_8, \text{CF}_3\text{Br}$	Gargamelle,12	[154]
IHEP_IETEP	Fe	IHEP_IETEP,2	[155]
NOMAD	C	NOMAD,5	[156]
MINOS	C	MINOS,0	[157]
ANL 12 ft	$^1\text{H}, ^2\text{H}$	ANL 12FT,4	[158]
BEBBC	$^1\text{H}, \text{Ne}$	BEBBC,2	[159]
BEBBC	$^1\text{H}, \text{Ne}$	BEBBC,8	[160]
BNL 7 ft	^2H	BNL 7FT,4	[161]
CCFRR	Fe	CCFRR,0	[162]
CHARM	C	CHARM,4	[163]
FNAL 15 ft	$^1\text{H}, \text{Ne}$	FNAL 15FT,2	[164]
Gargamelle	$\text{C}_3\text{H}_8, \text{CF}_3\text{Br}$	Gargamelle,10	[165]
IHEP_IETEP	Fe	IHEP_IETEP,0	[155]
IHEP_JINR	Al	IHEP_JINR,0	[167]
SKAT	CF_3Br	SKAT,0	[168]
SciBooNE	C	SciBooNE,0	[169]
$\bar{\nu}_\mu$ CC Inclusive			
BEBBC	$^1\text{H}, \text{Ne}$	BEBBC,1	[147]
BEBBC	$^1\text{H}, \text{Ne}$	BEBBC,6	[148]
BNL 7 ft	^1H	BNL 7FT,1	[170]
CHARM	C	CHARM,1	[151]
FNAL 15 ft	$^2\text{H}, \text{Ne}$	FNAL 15FT,4	[171]
Gargamelle	$\text{C}_3\text{H}_8, \text{CF}_3\text{Br}$	Gargamelle,1	[153]
Gargamelle	$\text{C}_3\text{H}_8, \text{CF}_3\text{Br}$	Gargamelle,13	[154]
IHEP_IETEP	Fe	IHEP_IETEP,3	[155]
MINOS	C	MINOS,1	[157]
BEBBC	$^1\text{H}, \text{Ne}$	BEBBC,3	[159]
BEBBC	$^1\text{H}, \text{Ne}$	BEBBC,7	[160]
CCFR	Fe	CCFR,3	[150]
CHARM	C	CHARM,5	[163]
FNAL 15 ft	$^2\text{H}, \text{Ne}$	FNAL 15FT,5	[166]
Gargamelle	$\text{C}_3\text{H}_8, \text{CF}_3\text{Br}$	Gargamelle,11	[172]
IHEP_IETEP	Fe	IHEP_IETEP,1	[173]
IHEP_JINR	Fe	IHEP_JINR,1	[167]

Table 3.11: Summary of data used for comparisons in Figs. 3.6, 3.7a, 3.7b, 3.8, 3.9, 3.10, 3.11, 3.13, 3.12 and 4.11. This table links the experiment and the tag used for the legend in each plot to the corresponding reference. It also provides the composition of the active volume of the experiment.

Experiment	Target	Tag	Ref.
ν_μ CC Quasielastic			
ANL 12 ft	² H	ANL 12FT,1	[174]
BEBBC	² H	BEBBC,12	[175]
FNAL 15 ft	² H	FNAL 15FT,3	[176]
SERP A1	Al	SERP A1,0	[177]
SKAT	CF ₃ Br	SKAT,8	[178]
ANL 12 ft	¹ H, ² H	ANL 12FT,3	[91]
BNL 7 ft	² H	BNL 7FT,3	[179]
Gargamelle	C ₃ H ₈ , CF ₃ Br	Gargamelle,2	[180]
SERP A1	Al	SERP A1,1	[181]
NOMAD	C	NOMAD,2	[156]
$\bar{\nu}_\mu$ CC Quasielastic			
BNL 7 ft	¹ H	BNL 7FT,2	[170]
Gargamelle	C ₃ H ₈ , CF ₃ Br	Gargamelle,5	[182]
SKAT	C ₃ H ₈	SKAT,9	[178]
Gargamelle	CF ₃ Br	Gargamelle,3	[180]
SERP A1	Al	SERP A1,2	[181]
NOMAD	C	NOMAD,3	[156]
ν_μ CC1 π^+ ($\nu_\mu p \rightarrow \mu^- p \pi^+$)			
ANL 12 ft	² H	ANL 12FT,0	[184]
ANL 12 ft	² H	ANL 12FT,ReAna,0	[93]
ANL 12 ft	² H	ANL 12FT,8	[93]
BNL 7 ft	² H	BNL 7FT,ReAna,0	[93]
Gargamelle	C ₃ H ₈ , CF ₃ Br	Gargamelle,4	[185]
BEBBC	¹ H	BEBBC,4	[186]
FNAL 15 ft	¹ H	FNAL 15FT,0	[187]
BEBBC	¹ H	BEBBC,9	[188]
BEBBC	² H	BEBBC,13	[175]
$\bar{\nu}_\mu$ CC1 π^- ($\bar{\nu} p \rightarrow \mu^+ p \pi^-$)			
FNAL 15 ft	¹ H	FNAL 15FT,10	[189]
ν_μ CC1 π^+ ($\nu n \rightarrow \mu^- n \pi^+$)			
ANL 12 ft	¹ H, ² H	ANL 12FT,ReAna,2	[93]
BNL 7 ft	² H	BNL 7FT,ReAna,2	[93]
ANL 12 ft	¹ H, ² H	ANL 12FT,10	[184]
ν_μ CC1 π^0 ($\nu n \rightarrow \mu^- p \pi^0$)			
ANL 12 ft	¹ H, ² H	ANL 12FT,ReAna,1	[93]
BNL 7 ft	² H	BNL 7FT,ReAna,1	[93]
ANL 12 ft	¹ H, ² H	ANL 12FT,9	[184]
ν_μ CC1 $\pi^+ \pi^+$ ($\nu p \rightarrow \mu^- n \pi^+ \pi^+$)			
ANL 12 ft	² H	ANL 12FT,13	[191]
ν_μ CC1 $\pi^+ \pi^0$ ($\nu p \rightarrow \mu^- p \pi^+ \pi^0$)			
ANL 12 ft	² H	ANL 12FT,12	[191]
ν_μ CC1 $\pi^+ \pi^-$ ($\nu n \rightarrow \mu^- p \pi^+ \pi^-$)			
ANL 12 ft	² H	ANL 12FT,12	[191]
BNL 7 ft	² H	BNL 7FT,8	[190]

Table 3.12: Summary of data used for comparisons in Figs. 3.6, 3.7a, 3.7b, 3.8, 3.9, 3.10, 3.11, 3.13, 3.12 and 4.11. This table links the experiment and the tag used for the legend in each plot to the corresponding reference. It also provides the composition of the active volume of the experiment.

Chapter 4

Hadronization model tuning in GENIE v3

Abstract

The GENIE neutrino Monte Carlo describes neutrino-induced hadronization with an effective model, known as Andreopoulos-Gallagher-Kehayias-Yang (AGKY), which is interfaced with PYTHIA at high invariant mass. Only the low-mass AGKY model parameters were extracted from hadronic shower data from the FNAL 15 ft and BEBC experiments. In this chapter¹, the first hadronization tune on averaged charged multiplicity data from deuterium and hydrogen bubble chamber experiments is presented, with a complete estimation of parameter uncertainties. A partial tune on deuterium data highlights the tensions between hydrogen and deuterium datasets.

4.1 Introduction

The next generation of neutrino oscillation experiments will rely on the precise understanding of neutrino interactions at the percent level. Experiments such as T2K [197], NOvA [198], MINERvA [199] and MicroBooNE [200] study neutrino interactions over a broad energy range. In the few GeV region, 0π and 1π contributions dominate the event rate. Hence, most of the effort has been focused on the theoretical understanding of these interactions [201–204] as well as the precise measurement of quasielastic [99, 122, 205–209] and pion production cross sections [99, 210–216]. Pions, before

¹This work has been published in Phys. Rev. D 105, 012009 (2022). It follows the guidelines on presentation of publications within a PGR thesis requested by the University of Liverpool.

final-state interactions (FSI), can be produced by either neutrino resonance interactions [217] or hadronization processes. Hadronization models provide information about the multiplicities and kinematics of the hadrons before FSI given the neutrino-nucleon interaction and the event kinematics. The knowledge of the exact mixture of hadrons in showers affects the efficiency to distinguish between neutral-current (NC) and charged-current (CC) events, the event topological characterization [205, 213], impacts the estimation of backgrounds [218] and the calorimetric energy reconstruction. FSI interaction modeling and detector efficiency corrections are also crucial to avoid confusion in measurements of neutrino-induced hadron production. Unfortunately, due to the lack of unified models for exclusive hadronic multiparticle production over the energy range of interest for neutrino experiments, one must resort to stitching together different modeling ingredients. The GENIE neutrino Monte Carlo (MC) event generator [128] uses the Andreopoulos-Gallagher-Kehayias-Yang (AGKY) hadronization model [219] whose validity extends down to the inelastic threshold. At low hadronic invariant mass W the model is based on the Koba-Nielsen-Olesen scaling law (KNO), while at high- W it is based on the PYTHIA MC [220].

Current and future experiments operate at high energies, where potential biases originating from hadronization mismodeling become important. For instance, DUNE [2], IceCube-Gen2 [221, 222] and ORCA [223] will focus on the 2 to 20 GeV energy range where deep-inelastic scattering (DIS) events are dominant. The neutrino energy dependence on the main inelastic components of the expected event rate for CC ν_μ - ^{40}Ar scattering is shown in Fig. 4.1. Some relevant neutrino fluxes of interest are shown in Fig. 4.1 (top). It is seen that the contribution to the event rate from GeV neutrinos is mainly driven by CC RES events as well as shallow-inelastic scattering (SIS) and DIS events from the low- W AGKY model, whereas PYTHIA events dominate at high neutrino energies.

The description of the AGKY hadronization model implementation in GENIE is described in Sec. 4.2. There is a separate hadronization model to simulate DIS charm production, the Aivazis, Olness, and Tung model [226]. Hadronic remnants produced in the interaction are hadronized with PYTHIA.

The AGKY model parameters controlling hadronization at low invariant masses were extracted from some of the FNAL 15 ft bubble chamber and the Big European Bubble Chamber (BEBC) analyses [136, 137]. PYTHIA has never been tuned to low-energy neutrino-induced hadronization data. In 2010, GENIE revisited the AGKY parameter values and modified a number of PYTHIA parameters using information from the NUX PYTHIA tune [227], as discussed in Sec. 4.2. We refer to this parameter set as the *2010* GENIE AGKY tune or *2010* GENIE tune. Despite the modifications,

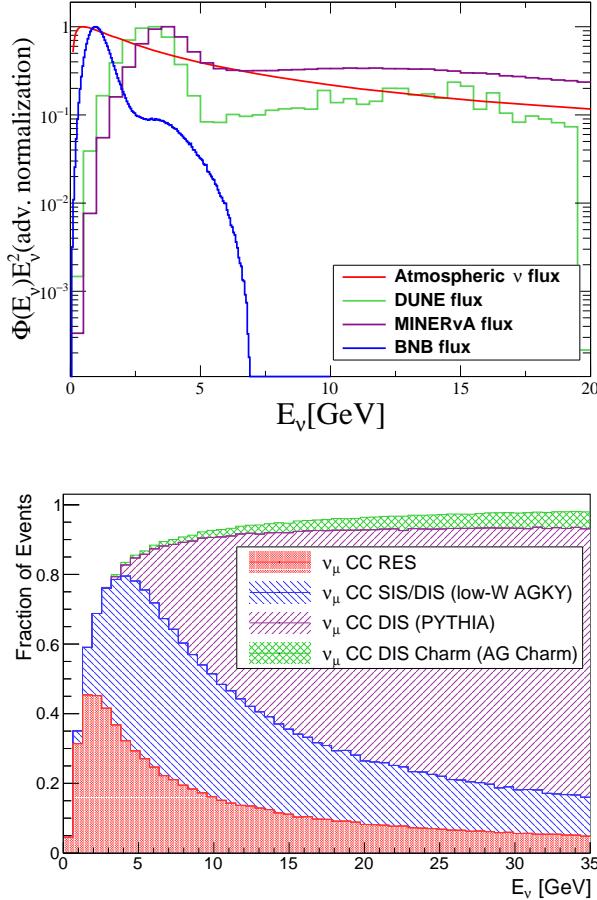


Figure 4.1: Normalized neutrino fluxes are shown for the atmospheric neutrino flux at Kamioka [224], DUNE [2], MINERvA [225] and BNB [100] flux predictions (top plot). Breakdown of CC events as a function of the neutrino energy from ν_μ scattering on ${}^{40}\text{Ar}$ (bottom plot). The plot was obtained with GENIE v3.00.06 using tune G18_02a_02_11a. The main components are: resonance (RES), shallow and deep inelastic scattering (SIS/DIS) and deep inelastic charm production (DIS Charm). DIS contributions are split according to the hadronization model used: low- W AGKY and PYTHIA.

several discrepancies between the model and neutrino-induced hadron shower data remained [228, 229].

This paper summarizes the results of the first tune of the AGKY hadronization model against averaged charged multiplicity data on hydrogen and deuterium targets from bubble chamber experiments. The analysis is performed within the GENIE v3.00.06 global analysis framework [217]. The base configuration used for all the plots presented here is the G18_02a_02_11a. This paper is organized as follows: the AGKY model specifics relevant for this work are described in Sec. 4.2, followed by an explanation of the analysis procedure applied to the hydrogen and deuterium datasets in Sec. 4.3. Section 4.5 discusses the free parameters in the model, and Sec. 4.6 presents the construction of the likelihood function used for fitting. The AGKY best-fit results are summarized in Sec. 4.7.

4.2 The AGKY model

The AGKY [219] model is the main hadronization model used in GENIE. As a function of hadronic invariant mass W , three different regimes are defined: an empirical model anchored to bubble chamber data at low W ($W < W_{\min}^{\text{tr}}$), a pure PYTHIA region for high- W ($W > W_{\max}^{\text{tr}}$) and a transition region that connects them. In the transition region, the probability to produce a PYTHIA event increases linearly with W , from zero at W_{\min}^{tr} to 1 at W_{\max}^{tr} . The values of the transition region limits are $W_{\min}^{\text{tr}} = 2.3 \text{ GeV}/c^2$ and $W_{\max}^{\text{tr}} = 3.0 \text{ GeV}/c^2$. The empirical low- W model and PYTHIA are valid in different mass ranges and they are combined accordingly.

The low- W AGKY and PYTHIA algorithms are described in the Sec. 4.2.1 and Sec. 4.2.2 respectively. The contribution of the main inelastic components as a function of W for events generated with the DUNE flux [2] is shown in Fig. 4.2. Most of the DIS/SIS events use the low- W AGKY model while the PYTHIA events are coming from the high-energy tail of the beam.

4.2.1 Effective low- W AGKY hadronization model

At low W , the showers are made of one baryon and any number of π or K consistent with momentum, charge, baryon and strange number, isospin and parity conservation

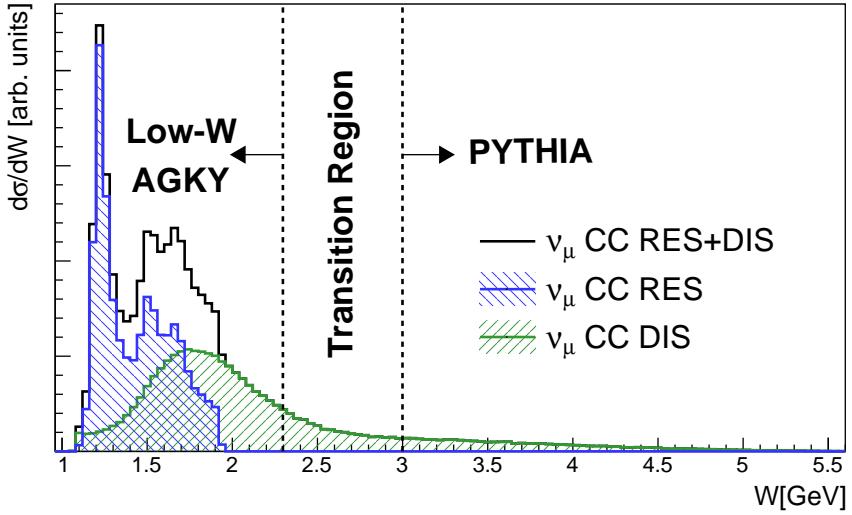


Figure 4.2: Flux integrated CC inelastic differential cross section as a function of the hadronic invariant mass for a DUNE ν_μ beam on ^{40}Ar , obtained with the G18_02a_02_11a tune. The distribution is decomposed in RES and DIS contributions. The DIS contribution to the total number of events is 38% and 36% for RES events. The ν_μ flux maximum is between 1 and 5 GeV.

laws:

$$\begin{aligned} \nu_\mu + p &\rightarrow \mu^- + X^{++}, \\ \nu_\mu + n &\rightarrow \mu^- + X^+, \\ \bar{\nu}_\mu + p &\rightarrow \mu^+ + X^0, \\ \bar{\nu}_\mu + n &\rightarrow \mu^+ + X^-. \end{aligned}$$

For instance, when approaching the pion production threshold, the $\nu_\mu p$ interaction would produce a shower made of a proton and a π^+ . In general, the hadron multiplicity at the lowest possible W is 2 as the hadronic final state can only be made of a pion and a nucleon.

As W increases, more possibilities are available. The model draws random integer numbers from the simulated hadronic multiplicity distribution to generate the number of particles in the shower, then the particles are labeled so that baryon number, charge, and strangeness are conserved. The particle content of a shower is selected so that the total mass is not exceeding W . The 4-momenta of the hadronic shower particles are generated by a weighted phase space decay of a particle of mass W to the selected hadronic-multiparticle state. There are many ingredients in the simulation of the hadronic probability distribution: average hadronic multiplicity data, the KNO

scaling law, particle content rules, phase space weighting and others, as discussed in detail in Ref. [219]. In this paper we focus on the description of the hadronic multiplicity. The hadronic multiplicity probability distribution depends on two ingredients: the measured average as a function of W , and an empirical parametrization of multiplicity dispersion. Both parametrizations must be extracted from data.

Empirical observation suggests that the average charged multiplicity is linear with $\ln W^2$:

$$\langle n_{\text{ch}} \rangle(W) = \alpha_{\text{ch}} + \beta_{\text{ch}} \ln \left(\frac{W^2}{\text{GeV}^2/c^4} \right). \quad (4.1)$$

The coefficients α_{ch} and β_{ch} depend on the initial state and their values can be extracted from neutrino-induced hadronization data, see Sec. 4.3. This behaviour has also been proved to be true for heavier nuclear targets [230, 231]. From fits to π^0 production data, it is known that $\langle n_{\text{ch}} \rangle \sim 0.5 \langle n_{\pi^0} \rangle$ [232]. Therefore, the total hadronic multiplicity is obtained from the charged one as

$$\langle n \rangle(W) \equiv 1.5 \langle n_{\text{ch}} \rangle(W). \quad (4.2)$$

Given the average $\langle n \rangle$, the hadronic multiplicity distribution, n , can be obtained from the KNO scaling law, which relates the dispersion of hadron multiplicities with a universal scaling function [233],

$$\langle n \rangle P(n) = f \left(\frac{n}{\langle n \rangle} \right). \quad (4.3)$$

The scaling function $f(n/\langle n \rangle)$ is parametrized with the Levy function $L(n/\langle n \rangle; c)$

$$L(n/\langle n \rangle; c) = \frac{2e^{-c} c^{c \frac{n}{\langle n \rangle} + 1}}{\Gamma(c \frac{n}{\langle n \rangle} + 1)}, \quad (4.4)$$

where Γ is the gamma function and c is the free parameter that has to be extracted from data and depends on the interaction isospin. By construction, the dispersion of the hadronic multiplicity distribution is independent from the average; see Fig. 4.3. The 2010 GENIE AGKY values of α_{ch} , β_{ch} and c are specified in Tab. 4.1.

4.2.2 PYTHIA in GENIE

The PYTHIA algorithm is well known for its wide use in high-energy collider experiments to simulate the evolution from a few-body hard process to a multihadronic final state [220, 234]. The PYTHIA hadronization model is based on the Lund string fragmentation framework, which describes the hadronization process as breakups in a

Parameter	$\nu_\mu p$	$\nu_\mu n$	$\bar{\nu}_\mu p$	$\bar{\nu}_\mu n$
α_{ch}	0.40	-0.20	0.02	0.80
β_{ch}	1.42	1.42	1.28	0.95
c	7.93	5.22	5.22	7.93

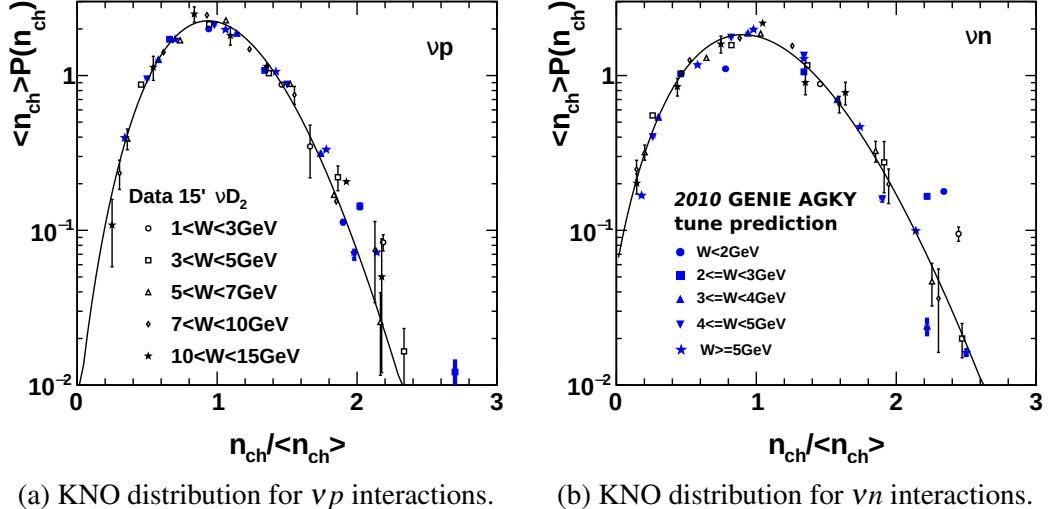
 Table 4.1: 2010 GENIE tune low W AGKY parameters.


Figure 4.3: KNO scaling distributions for neutrino interactions on deuterium [219]. The solid line is the best fit result of the Levy function to FNAL 15 ft bubble chamber data [136]. Blue dots show the 2010 GENIE AGKY prediction for a given W range specified in the legend.

Parameter	Name in PYTHIA	PYTHIA default	NUX tune	HERMES tune	2010 GENIE tune
$P_{s\bar{s}}$	PARJ(2)	0.30	0.21	0.25	0.30
$\langle p_\perp^2 \rangle [\text{GeV}^2/c^2]$	PARJ(21)	0.36	0.44	0.42	0.44
$E_{\text{CutOff}} [\text{GeV}]$	PARJ(33)	0.80	0.20	0.47	0.20
Lund a	PARJ(41)	0.30	0.30	0.68	0.30
Lund $b [c^4/\text{GeV}^2]$	PARJ(42)	0.58	0.58	0.35	0.58

Table 4.2: Summary of different PYTHIA parametrizations. The parameter configuration for PYTHIA, NUX, HERMES and 2010 GENIE tunes are specified. The details on the HERMES tune are given in Sec. 4.4.2.

string throughout production of new $q\bar{q}$ pairs [235]. Each string represents a color flux which is subject to a linear confined potential. In the Lund model, the $q\bar{q}$ pairs break by tunneling, which, together with causality, defines the Lund symmetric fragmentation function,

$$f(z) \propto \frac{(1-z)^a}{z} \exp\left(\frac{-bm_\perp^2}{z}\right) \quad (4.5)$$

with the transverse mass of the hadron defined as $m_\perp^2 \equiv m^2 + p_\perp^2/c$ and z being quantities that characterize the hadronic shower [236]. The transverse momentum is defined as $p_\perp^2 = p_x^2 + p_y^2$. z describes the fraction of available light-cone momentum $E + p_z$ transferred to the hadrons produced with energy E , and it is defined as $z = E/v$. The parameters a and b , known as Lund a and Lund b , are free parameters of the model that are responsible for distributing the longitudinal energy of the hadronic system after the interaction, and they should be tuned to reproduce experimental data [236]. In terms on the effect on $\langle n_{\text{ch}} \rangle$, as Lund a increases, the multiplicity increases as well, while the opposite is happening for Lund b .

In GENIE, PYTHIA is used to simulate the hadronization at high-energy invariant masses. Specifically, GENIE v3.00.06 uses PYTHIA 6. Future GENIE releases will slowly transition to PYTHIA 8. In particular, in v3.00.06, PYTHIA 8 is partially integrated in GENIE and it is fully integrated in the AGKY model. After the partial integration of PYTHIA 8, simulation outputs remained unchanged. Hence, the tune presented in this paper is also valid for PYTHIA 8. Moreover, different GENIE comprehensive model configuration (CMC) [217] have no impact on the hadronization predictions.

The default PYTHIA parameters shown in Tab. 4.2 come from fits to high-energy $e^+ - e^-$ experiments [237–248] ($\sqrt{s} \sim 35$ GeV). PYTHIA’s description to data at low energy, such as modern neutrino oscillation experiments (1 – 10 GeV) or even lower-energy $e^\pm - p$ experiments such as the HERMES experiment (at 27 GeV) [249], is not accurate, see Sec. 4.4. The first attempt to improve this disagreement was in 2010, where some of the PYTHIA parameters were tweaked according to a NUX PYTHIA

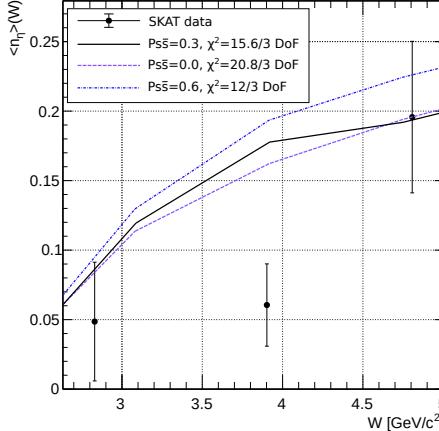


Figure 4.4: Parameter impact on the averaged η production data from SKAT [250].

tune [227]. The parameters modified by the NUX PYTHIA tune are as follows:

- $P_{s\bar{s}}$ controls the $s\bar{s}$ production suppression,
- $\langle p_\perp^2 \rangle$ determines the average hadron transverse momentum squared,
- E_{CutOff} is the energy cutoff for the fragmentation process.

These parameters are related to important hadron shower characteristics. The assumption of tunneling breakups implies the suppression of heavy-quark production, limiting its production in soft fragmentation processes. The suppression factor for heavy quarks is $u\bar{u}:d\bar{d}:s\bar{s}:c\bar{c} \sim 1:1:0.3:10^{-11}$ [236]. This is supported by η production data, Fig. 4.4. Previous tunes are in agreement with this fact, see Tab. 4.2. Each quark anti-quark pair receive opposite p_\perp kicks at each string breaking point according to a Gaussian distribution. The $\langle p_\perp^2 \rangle$ parameter controls the variance of the Gaussian distribution used at the breaking point. There are different datasets available to constrain this parameter [219], see for instance Fig. 4.5. Finally, E_{CutOff} determines the minimum energy at which the fragmentation of the parton system can occur, set to 0.8 GeV in PYTHIA. 2010 GENIE uses the best-fit value from the NUX PYTHIA tune, where $E_{\text{CutOff}} = 0.20$ GeV.

In GENIE v3 and previous releases, there is only one parameter set configuration for the low- W AGKY model (Tab. 4.1) and PYTHIA (Tab. 4.2) that is common for all CMC's.

4.3 Neutrino-induced hadronization data review

The characterization of the AGKY parameters relies on neutrino-induced hadronization data from BEBC and FNAL 15 ft experiments. These experiments published a variety of observables related to hadronization. This work is based mainly on charged

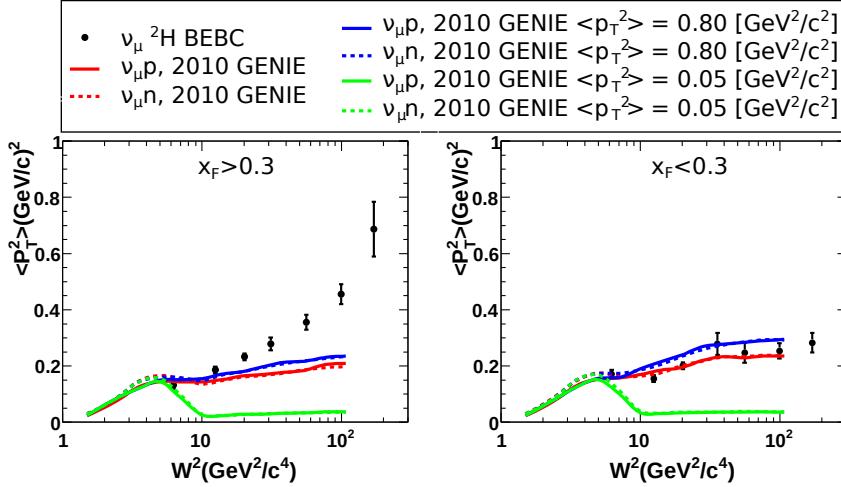


Figure 4.5: Effect of the $\langle p_T^2 \rangle$ parameter on the $\langle p_T^2 \rangle$ distributions as a function of W^2 for ν_μ data on ${}^2\text{H}$ from the BEBC experiment under different Feynman- x ($x_F = p_L/p_{L,\max}$) conditions [219]: $x_F > 0.3$ (left) and $x_F < 0.3$ (right). The 2010 GENIE parameter value is $\langle p_T^2 \rangle = 0.44 \text{ GeV}^2/\text{c}^2$. The validation range used for this plot is specified in the legend.

multiplicity data as a function of the hadronic invariant mass; hence, it is what this review is focusing on. The analyses procedure for both experiments are similar and depend on the target type that can be hydrogen or deuterium. The different analysis requirements need to be implemented in the GENIE hadronization analyses for a meaningful data/MC comparison; see Sec. 4.6. In this section, the analyses of interest for this work are discussed in detail.

4.3.1 Hydrogen data

The bubble chamber at Fermilab (FNAL 15 ft) and BEBC at CERN follow similar analyses procedures. The data considered in this work are those listed in Tab. 4.3.

Both experiments look for ν_μ and $\bar{\nu}_\mu$ CC interactions on hydrogen to study the averaged charged multiplicity of the final state as a function of the event invariant mass. The main requirement to select CC events is to detect a muon track. Muons are detected with a external-muon identifier (EMI), and a minimum muon momentum, p_μ , is usually required to guarantee good muon identification (ID). This is a consequence of the muon ID efficiency dependence on the muon momentum energy. For instance, in BEBC experiment, the muon ID efficiency varies from 40% to 100% in the muon momentum range of $3 \text{ GeV}/c \leq p_\mu \leq 10 \text{ GeV}/c$, with an average efficiency of 95%. The FNAL 15 ft experiment also uses a kinematic technique to identify negative muons in neutrino interactions [255]. Under this μ^- -ID method, only events in which the μ^-

Experiment	N_p	W^2 [GeV $^2/c^4$]	Cuts	Syst.	In Fit	Ref.
$\nu_\mu + p \rightarrow \mu^- + X^{++}$						
FNAL 15 ft (1976)	25	[1.5, 150]	$E_V^{\text{reco}} \geq 15$ GeV $p_L^{\text{visible}} \geq 10$ GeV/c $p^\mu \geq 5$ GeV/c $p_T^\mu \geq 1$ GeV/c	<i>Included</i>	$W^2 < 20$ GeV $^2/c^4$	[251]
$\bar{\nu}_\mu + p \rightarrow \mu^+ + X^0$						
BEBC (1983)	11	[9, 121]	$p^\mu \geq 3$ GeV/c $E_{\text{visible}}^{\text{visible}} \geq 5$ GeV $W^2 \geq 9$ GeV $^2/c^4$	3–5%	\times	[135]
BEBC (1990)	6	[6, 150]	$Q^2 \geq 1$ (GeV/c) 2 $p^\mu \geq 3$ GeV/c $W^2 \geq 4$ GeV $^2/c^4$	<i>Statistical</i>	$W^2 < 9$ GeV $^2/c^4$	[252]
BEBC (1992)	5	[12, 144]	$p^\mu \geq 3$ GeV/c	<i>Included</i>	✓	[253]
$\bar{\nu}_\mu + p \rightarrow \mu^- + X^{++}$						
FNAL 15 ft (1981)	10	[16, 100]	$p_{\text{ch}} \geq 5$ GeV/c $p_{\text{FW}}^{\text{tot}} \geq 2$ GeV/c $y_B \geq 0.1$ $y_B \leq 0.8$ $E_V^{\text{reco}} \geq 5$ GeV	<i>Statistical</i>	$W^2 < 30$ GeV $^2/c^4$	[254]
BEBC (1983)	10	[9, 121]	$p^\mu \geq 3$ GeV/c $E_{\text{visible}}^{\text{visible}} \geq 5$ GeV $W^2 \geq 9$ GeV $^2/c^4$	3–5%	\times	[135]
BEBC (1990)	6	[6, 144]	$Q^2 \geq 0.1$ (GeV/c) 2 $p^\mu \geq 3$ GeV/c $W^2 \geq 4$ GeV $^2/c^4$	<i>Statistical</i>	$W^2 < 10$ GeV $^2/c^4$	[252]
BEBC (1992)	5	[12, 144]	$p^\mu \geq 3$ GeV/c	<i>Included</i>	$W^2 < 60$ GeV $^2/c^4$	[253]

Table 4.3: Compilation of historical data from the BEBC and FNAL 15 ft bubble chamber experiments on averaged charged hadron multiplicity in muon (anti)neutrino on hydrogen interactions. Information about the number of points in each dataset, N_p , the W^2 range covered; and the cuts applied in each analysis is provided. Unless specified, the systematic errors were not included in the data release error bands and have been added in quadrature by the amount specified in this table, see details in Sec. 4.3.3. The sixth column specifies whether a dataset is included, discarded or partially included in the fit, see Sec. 4.4.3. The complete list of data points removed in this analysis is specified in the Sec. 4.7.

candidate has transverse momentum, p_\perp^μ , of at least $1 \text{ GeV}/c$ are accepted.

Selected events, which satisfy the conditions specified above, are analyzed to reconstruct the event topology and kinematics. In particular, BEBC uses the HYDRA program [135, 252, 253] and FNAL 15 ft a modified version of the TVGP program [256]. Only a small fraction of the charged final state hadrons is identified by using energy loss, range in hydrogen, break point probability and kinematic fits [135]. If left unidentified, the remaining charged hadrons are assumed to be pions: this assumption can cause migration of particles from the backward to the forward going hemisphere. For instance, the BEBC experiment is able to identify about 30% of the protons using the HYDRA algorithm, while the rest are classified as pions [252].

For ν_μ CC interactions, because of charge conservation, the experiments scan for events with three or more charged particles in the final state.

The topology of neutrino and antineutrino events is expected to be different. In antineutrino events, interactions with only one charged track can occur ($n_{\text{ch}} = 0$). Such events are not negligible at low E_ν and low W . However, these are removed due to low scanning efficiency and poor antineutrino energy reconstruction. Both BEBC and FNAL 15 ft correct for the effect of removing one-prong contributions in antineutrino samples using MC calculations [135, 252–254]. One-prong MC events are weighted so that the fraction of one-prong events agree with the experimental estimate. The scanning efficiencies for three-prong events are higher than 90%, improving as the number of charged secondaries increases ($\geq 95\%$).

In hydrogen and deuterium bubble chambers, the identification of neutral particles, such as π^0 , is difficult due to the low Z of the medium. As a consequence, the transverse momentum balance method is used to estimate the neutrino energy by assuming undetected neutral particles in the event [257],

$$E_\nu^{\text{reco}} = p_L^\mu c + p_L^{\text{ch}} c \left(1 + \frac{|\mathbf{p}_\perp^\mu + \mathbf{p}_T^{\text{ch}}|}{\sum_{i=1}^{n_{\text{ch}}} |\mathbf{p}_{\perp i}|} \right). \quad (4.6)$$

The subscript L and \perp refer to longitudinal and transverse components of the momenta relative to the neutrino direction, whereas the ch and μ labels denote the charged-hadron system and the muon, respectively. The index i runs over the charged hadrons in the hadronic system. By using this method, there is a non-negligible bias for the neutrino energy reconstruction. For instance, the BEBC experiment estimated the reconstructed neutrino energy to differ from the true energy by $\sim 10 - 15\%$ [135, 136]. Both bubble chambers corrected for this effect; see Sec. 4.3.3. In some analyses, cuts on the reconstructed neutrino energy, E_ν^{reco} , are applied [135, 136, 251].

Backgrounds from NC events, quasielastic (QEL) CC events or neutral-particle-

induced events are removed from the final sample using kinematic cuts that depend on each analysis. NC events can mimic CC events as a consequence of muon-hadron misidentification. On the one hand, for the FNAL 15 ft experiment, the muon-hadron misidentification increases at high Bjorken inelasticity values (y_B), and a cut on y_B is required to guarantee a good efficiency in selecting CC events [254]. In $\bar{\nu}_\mu$ events, backgrounds from low-energy neutrons as well as events caused by incoming hadron tracks that rescatter within the chamber are controlled by requiring the total momentum in the forward hemisphere, p_{FW}^{tot} , to be greater than 2 GeV/c [254, 256]. Moreover, FNAL 15 ft removes backgrounds from K_L^0 mesons by requiring the minimum total momentum from charged particles, p_{ch} , to be higher than 5 GeV/c [254]. On the other hand, the BEBC experiment applies kinematic cuts on either W or/and Q^2 to remove QEL events [135, 252, 253]. All cuts applied to the different analyses are shown in Tab. 4.3.

4.3.2 Deuterium data

The analyses algorithm followed by the FNAL 15 ft and BEBC bubble chamber experiments operating with deuterium aims to discriminate between interactions on proton and neutron. The data on deuterium considered in this work are those listed in Tab. 4.4.

Before classifying the event as a neutrino interaction on either proton or neutron, the analyses procedure is equivalent to the one described in Sec. 4.3. Each event has to contain a muon, identified with the EMI, that satisfies the cuts summarized in Tab. 4.4. The information about the event topology and kinematics is obtained using the TVGP-SQUAW or HYDRA algorithms for FNAL 15 ft [136] and BEBC respectively [137, 258, 259]. Particles are classified as pions if the algorithm fails to identify them as any other particle. The neutrino energy is reconstructed using the transverse momentum balance method. Similar kinematic cuts to those specified for the hydrogen analyses are applied.

The main difference between both analyses is the particle identification of struck nucleons in the event. A neutrino event is classified as a neutrino interaction on proton if the event topology has an odd number of prongs. Alternatively, the event is classified as an interaction on neutron if the event has an even number of prongs with no visible spectator or an odd number of prongs that include a visible proton. See a graphical interpretation in Fig. 4.6. The antineutrino case is similar except that the minimum prong multiplicity on proton is 1, instead of 3. Because of the selection criteria explained in Sec. 4.3, interactions with $n_{ch} = 0$ are not considered, effectively making the selection criteria for antineutrinos the same as for neutrinos.

In the analyses, a prong is classified as a proton if it corresponds to a particle

Experiment	N_p	W^2 [GeV $^2/c^4$]	Cuts	Syst.	In Fit	Ref.
$\bar{\nu}_\mu + p \rightarrow \mu^- X^{++}$						
FNAL 15 ft (1983)	14	[1, 225]	$p_\mu \geq 5$ GeV/ c $p_\mu^\perp \geq 1$ GeV/ c $p_{ch}^L \geq 5$ GeV/ c $p_p \leq 340$ MeV/ c $p_p \geq 200$ MeV/ c $W \geq 1.5$ GeV/ c^2 $E_v^{\text{reco}} \geq 10$ GeV	10%	$W^2 > 4$ GeV $^2/c^4$ †	[136]
$\nu_\mu + n \rightarrow \mu^- X^+$						
BEBC (1989)	6	[4, 196]	$p_\mu^\perp \geq 4$ GeV/ c $p_p \leq 300$ MeV/ c	ϵ_{cut} Not Included	✗	[258]
FNAL 15 ft (1983)	14	[1, 225]	$p_\mu^\perp \geq 1$ GeV/ c $p_{ch}^L \geq 5$ GeV/ c $E_v^{\text{reco}} \geq 10$ GeV $p_p \leq 340$ MeV/ c $p_p \geq 200$ MeV/ c	10%	✓	[136]
BEBC (1984)	8	[6, 112]	$p_\mu \geq 4$ GeV/ c $Q^2 \geq 1$ (GeV/ c) 2 $W^2 \geq 5$ GeV $^2/c^4$ $p_p \leq 300$ MeV/ c	ϵ_{cut} Statistical	✓	[259]
BEBC (1989)	6	[4, 196]	$p_\mu \geq 4$ GeV/ c $p_p \leq 300$ MeV/ c $W \geq 5$ GeV/ c^2	ϵ_{cut} Not Included	✗	[258]
$\bar{\nu}_\mu + p \rightarrow \mu^+ X^0$						
BEBC (1982)	8	[5, 75]	$p_\mu \geq 4$ GeV/ c $p_p \leq 300$ MeV/ c	Statistical	✓	[137]
BEBC (1989)	6	[4, 196]	$p_\mu^\perp \geq 4$ GeV/ c $p_p \leq 300$ MeV/ c	ϵ_{cut} Not Included	✗	[258]
$\bar{\nu}_\mu + n \rightarrow \mu^+ X^-$						
BEBC (1982)	8	[1.5, 56]	$p_\mu \geq 4$ $p_p \leq 300$ MeV/ c	Statistical	✓	[137]
BEBC (1989)	6	[4, 196]	$p_\mu \geq 4$ GeV/ c $p_p \leq 300$ MeV/ c	ϵ_{cut} Not Included	✗	[258]

Table 4.4: Compilation of historical data from the BEBC and FNAL 15 ft bubble chamber experiments on averaged charged hadron multiplicity in muon (anti)neutrino on deuterium interactions. Information about the number of points in each dataset, N_p , the W^2 range covered and the cuts applied in each analysis is provided. Unless specified, the systematic errors were not included in the data release error bands and have been added in quadrature by the amount specified in this table, see details in Sec. 4.3.3. The sixth column specifies whether a dataset is included, discarded or partially included in the fit, see Sec. 4.4.3.

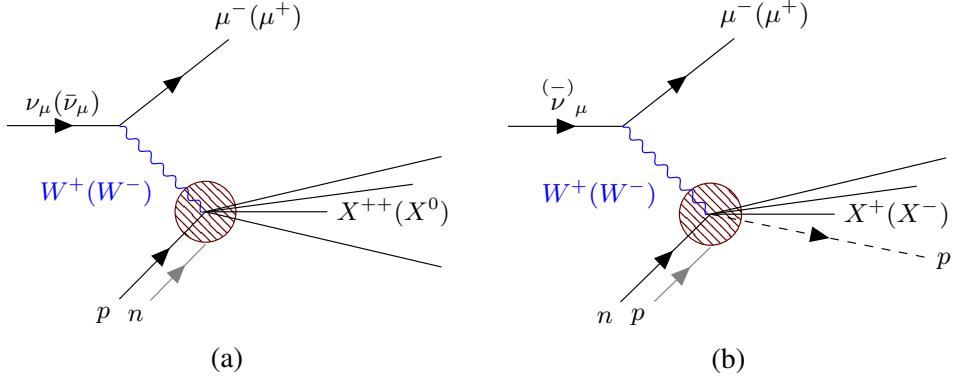


Figure 4.6: Bubble chamber analyses of ν_μ and $\bar{\nu}_\mu$ on ^2H data schematic procedure. The topology definition is based mainly on the number of prongs in each event. Possible visible proton spectators that satisfy the momentum requirements specified in Tab. 4.4 are represented with dashed lines. $\bar{\nu}_\mu$ ^2H one-prong events are not considered. (a) $\nu_\mu p$ topology. (b) $\nu_\mu n$ topology.

moving backward relatively to the beam direction ($\cos \theta_p < 0$) or a forward-going particle with low momentum. The maximum momentum cut is dataset dependent; see Tab. 4.4. If these conditions are not satisfied, the proton is not reconstructed, and for the purpose of the analyses, it is considered invisible. In the FNAL 15 ft analyses, for a proton to be detected as a prong, its momentum has to be $p_p > 200 \text{ MeV}/c$.

The deuterium target can induce rescattering of the hit nucleon with the spectator; this can increase the number of hadrons in the final state [255]. An odd number of prongs can occur in any possible neutrino interaction because of rescattering, independently of the hit nucleons, so the $\nu_\mu p$ sample will contain $\nu_\mu n$ events. In contrast, the $\nu_\mu n$ sample can only contain $\nu_\mu p$ events because of detector inefficiencies. Rescattering events have an impact on the event kinematics, which can be quantified defining an energy balance as

$$\varepsilon \equiv \sum_i (E_i - p_{L_i}c) - Mc^2, \quad (4.7)$$

where E_i and p_{L_i} are the i th charged particle energy and longitudinal momentum component relative to the neutrino direction respectively while M is the mass of the target nucleon assumed in the selection sample. Eq. 4.7 assumes that the nucleon is at rest and that the neutrino direction is known. In an ideal detector where all final state particles are identified, $\varepsilon = 0$ [260]. In a bubble chamber experiment, where only charged particles are detected, $\varepsilon < 0$. Possible particle misidentification reduces the ε value further, as particles are assigned to be pions as a default, unless identified otherwise. Rescattering events have a $\varepsilon > 0$ with a maximum value of $M_{^2\text{H}}c^2 - M_n c^2$, see Fig. 4.7.

The BEBC experiment eliminates rescattering events from the sample by imposing a cut on the energy balance [137, 258, 259]. An event is rejected due to rescattering if

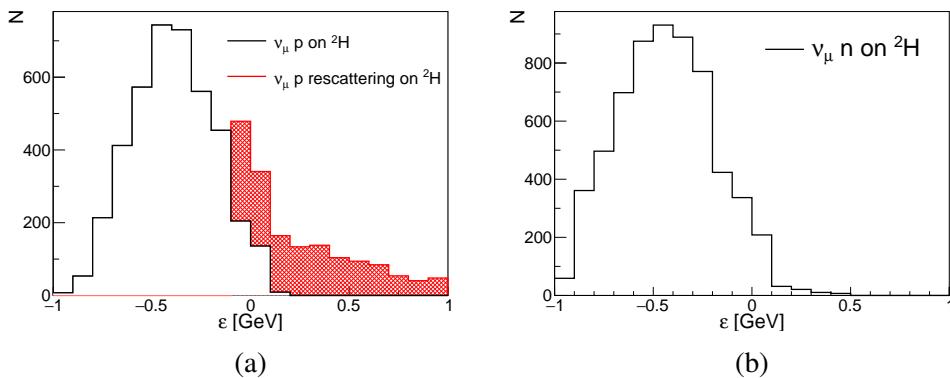


Figure 4.7: Energy balance distribution for ν_μ events on proton and neutron candidates digitized from the BEBC analyses paper [259]. Events that do not satisfy the ϵ_{reco} correspond to rescattering events and are highlighted in red. No rescattering contribution were observed in the $\nu_\mu n$ sample. (a) $\nu_\mu p$ events under the odd prong topology assumption. (b) $\nu_\mu n$ events under the even prong topology assumption. Neutron events with a spectator proton are not included.

the following conditions are satisfied:

- $\epsilon > 0.1 \text{ GeV}$,
- $\epsilon > -0.1 \text{ GeV}$ and the transverse missing momenta squared differs from zero, $(p_{\perp}^{\text{miss}})^2 > 0.075 (\text{GeV}/c)^2$.

The FNAL 15 ft experiment did not correct for rescattering events. In some of the analyses, additional cuts are considered for the deuterium analyses to remove backgrounds. For instance, the FNAL 15 ft bubble chamber reduces the background from neutral hadron-induced events by applying a cut on the total charged-particle longitudinal momentum in the final-state system (p_{ch}^L) [136].

4.3.3 Sources of systematic uncertainties in the FNAL 15 ft and BEBC experiments

MC studies were preformed by the FNAL 15 ft and BEBC bubble chamber experiments to correct for possible sources of errors. In particular, the different analyses correct for the following effects:

- EMI geometrical inefficiency [137, 252, 253, 259].
- Efficiency losses due to possible hadron miss-ID and migration of particles from the forward to backward hemispheres [135, 252, 253, 259].
- W^2 smearing due to the uncertainty in the neutrino energy reconstruction [135–137, 251–254, 258, 259].

- Neutrino energy uncertainty associated the transverse balance method [135–137, 251–254, 258, 259].
- Neutral particle decays (γ , K^0 and Δ) into charged particles that can lead to a higher charged multiplicity if the decay vertex is close to the primary one [135–137, 251–254].
- One prong event corrections [135, 252–254, 259]. This kind of events occur for low- W $\bar{\nu}p$ interactions in which only the μ^+ is observed.
- Efficiency to detect CC events [136, 254].
- Corrections due to the Fermi motion in Deuterium [259].
- Possible measurement errors [135, 136, 252, 253, 259].

The information about the BEBC systematic errors was obtained by using two MC programs: the LUND MC and a longitudinal phase-space model (LPS) [261]. Both MC were tuned to describe the BEBC experiment. From the MC generations, two samples are created: the initial, $d_{\text{initial}}^{\text{MC}}$, and modified, $d_{\text{modified}}^{\text{MC}}$, samples. The initial sample contains the truth information of the event. The modified sample includes modifications to mimic the analyses procedure. The ratio between the samples provides with a correction factor that it is applied to the data.

BEBC systematic errors are obtained from the difference between both MC calculations. The FNAL 15 ft corrected for some of the effects, but no clear information about the methodology followed to estimate the systematics is provided. Some of these experiments provide error bars which already include an estimation of systematics; however, this is not the case for most of the data. In particular, there are three different ways in which the BEBC and FNAL 15 ft experiments quote the systematic errors: (1) The systematic errors are already included in the total error, (2) The systematic uncertainty was quoted as a percentage with respect to the central value, (3) The systematic error is considered to be approximately of the same size of the statistical error or (4) No information is provided in the data release. For the cases (2) and (3), the systematic errors are added in quadrature to the statistical ones in this analysis. Particularly, for the datasets from Ref. [258], information on systematic errors is not provided in the data release. In Tab. 4.3 and Tab. 4.4, the information on the systematic error is provided. We label the different categories as (1) included, (2) with the percentage, (3) statistical, and (4) not included, respectively. No correlation matrices are provided by any of these experiments.

4.4 Review of previous tunes to hadronization data

While summarising the experimental fits to averaged charged multiplicity data, this section also explains the origin of the *2010* GENIE tune parameters. This is necessary to define proper selection criteria for a dataset to be included in a global fit.

4.4.1 Fits to bubble chamber data

Both BEBC and FNAL 15 ft experiments provided estimations of the α_{ch} and β_{ch} parameters for every released dataset. The individual fits were performed by fitting Eq. 4.1 in each channel. Fit results are summarized in Tab. 4.5.

There are six channels in total: ν_μ or $\bar{\nu}_\mu$ on proton or neutron, while the information on interactions on proton can be from data with hydrogen or deuterium targets. Information about neutrino interaction on neutron can only be extracted from deuterium samples. The BEBC and/or FNAL 15 ft experiments performed individual fits to each of the available channels.

From the best-fit values extracted for each dataset, we observe clear discrepancies for the α_{ch} and β_{ch} values between data releases and between the BEBC and FNAL 15 ft data (e.g. for $\nu_\mu p$ interactions on hydrogen). Discrepancies between hydrogen and deuterium samples are also present. This target-related discrepancy can also be observed in fits to OPERA and CHORUS data [230, 262]. These discrepancies could have different origins: the W^2 range, the beam energy or the kinematic cuts applied in the analyses.

The *2010* GENIE AGKY parameter values presented in Tab. 4.1 correspond to the analyses on deuterium targets highlighted in Tab. 4.5. Notice that the parameters used in the *2010* GENIE prediction come from fits to Eq. 4.1 over the whole W^2 range. This procedure is not adequate as the α_{ch} and β_{ch} should be extracted from a fit to data over the low- W validity range given that the AGKY model differs from the simplified linear behaviour.

The description of the shower particle content is linked to several observables whose correlation is still unknown. For instance, the averaged charged multiplicity and dispersion observables can be correlated. The full list of available hadronization data is shown in Ref. [219]. Ideally, the AGKY tune should improve the agreement with all hadronization related observables. The extraction of the averaged charged multiplicity parameters, such as α_{ch} and β_{ch} , strongly relays on the precise understanding of the datasets described in Sec. 4.3 and Sec. 4.3.2. However, the analyses of historical averaged charged multiplicity datasets already show clear disagreements between each of the different data releases, as summarized in Tab. 4.5. For these reasons, on this

Experiment	W^2 [GeV $^2/c^4$]	Target	α_{ch}	β_{ch}	Ref.
$\nu_\mu + p \rightarrow \mu^- X^{++}$					
FNAL 15 ft (1976)	[1.5, 150]	H	1.09 ± 0.38	1.09 ± 0.03	[251]
BEBC (1983)	[12, 112]	H	-0.05 ± 0.11	1.43 ± 0.04	[135]
FNAL 15 ft (1983)	[1.5, 160]	2H	0.05 ± 0.07	1.42 ± 0.03	[136]
BEBC (1990)	[6, 150]	H	0.911 ± 0.224	1.131 ± 0.086	[252]
BEBC (1992)	[12, 144]	H	0.40 ± 0.13	1.25 ± 0.04	[253]
$\nu_\mu + n \rightarrow \mu^- X^+$					
BEBC (1984)	[6, 112]	2H	1.75 ± 0.12	1.31 ± 0.04	[259]
FNAL 15 ft (1983)	[1.5, 160]	2H	-0.20 ± 0.07	1.42 ± 0.03	[136]
$\bar{\nu}_\mu + p \rightarrow \mu^+ X^0$					
FNAL 15 ft (1982)	[1.7, 74]	H	-0.44 ± 0.13	1.48 ± 0.06	[254]
BEBC (1982)	[5, 75]	2H	0.02 ± 0.20	1.28 ± 0.08	[137]
BEBC (1983)	[12, 96]	H	-0.56 ± 0.25	1.42 ± 0.08	[135]
BEBC (1990)	[6, 144]	H	0.222 ± 0.362	1.117 ± 0.100	[252]
BEBC (1992)	[12, 144]	H	-0.44 ± 0.20	1.30 ± 0.06	[253]
$\bar{\nu}_\mu + n \rightarrow \mu^+ X^-$					
BEBC (1982)	[1.5, 56]	2H	0.80 ± 0.09	0.95 ± 0.04	[137]
$\nu_\mu + A$					
OPERA (2018)	[1.6, 54.6]	Pb	-0.19 ± 0.18	0.76 ± 0.07	[230]
CHORUS (2007)	[1, 148]	Fuji ET-B7	1.07 ± 0.05	1.32 ± 0.11	[262]

Table 4.5: Compilation of best fit values for the intercept α_{ch} and slope β_{ch} obtained from individual fits to Eq.(4.1) against mean charged hadron multiplicity data as a function of W^2 . The parameters for charged-current ν_μ and $\bar{\nu}_\mu$ scattering data on hydrogen, deuterium, ^{207}Pb and the Fuji ET-B7 emulsion are shown in the table. 2010 GENIE parameters are extracted from the analyses highlighted in bold.

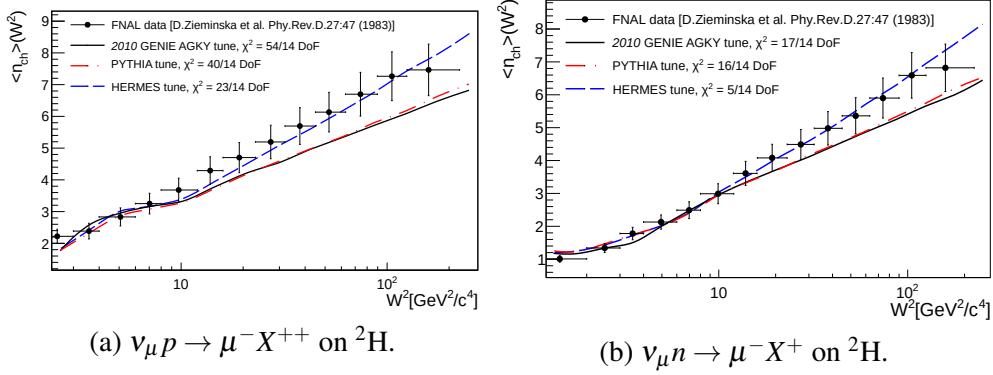


Figure 4.8: Comparison of FNAL average charged multiplicity deuterium data against GENIE predictions obtained with the parametrizations specified in Tab. 4.2.

work, we focus on the description and tune of averaged charged multiplicity data on hydrogen and deuterium samples.

4.4.2 The HERMES tune

The PYTHIA parameters are extracted from high-energy e^-e^+ experiments; see Sec. 4.2.2. From Fig. 4.8, we see that PYTHIA underestimates the averaged charged multiplicity. However, PYTHIA has not been tuned using data from neutrino experiments.

The 2010 GENIE tune, summarized in Sec. 4.2.2, aimed to improve the agreement with different hadronization observables by incorporating the results from a NUX PYTHIA tune [227]. However, this was not sufficient to improve the agreement of PYTHIA with average charged multiplicity data from bubble chambers experiments. Moreover, the tune lacked of information about the uncertainties of the fit parameters.

Information on PYTHIA parameters at lower energy was provided by the HERMES experiment, which tuned PYTHIA using $e^\pm p$ data at 27 GeV [249]. Ref. [228] suggests that the HERMES tune improves the agreement with neutrino data, as summarized in Fig. 4.8. The main differences between the HERMES tune and the GENIE 2010 retune are the modification of the Lund a and Lund b parameters, suggesting higher (lower) values of Lund a (Lund b).

The PYTHIA parameters with most impact on the average charged multiplicity for the 2010 GENIE AGKY and HERMES tunes are summarized in Tab. 4.2.

4.4.3 Requirements for including a dataset in the AGKY multiplicity tune

Only the averaged charged multiplicity data on hydrogen and deuterium is taken into account in this AGKY fit. Tables 4.3–4.4 summarize the information about which datasets are included in the tune. If possible, only the latest analysis of each experiment is included. Previous analyses are only considered if:

1. Its reanalyses did not cover all the original W^2 range,
2. The prediction interpolation by Professor fails to describe the GENIE prediction (see Sec. 4.6.1),
3. The data release lack of sufficient information about systematic errors.

In requirement (1), previous analyses are used to complement the covered W^2 range, as those points were not documented in the revisited ones. If the datasets are only included partially, the approximate W^2 range used is provided. An example is the BEBC $\bar{\nu}_\mu$ on H data, in particular datasets (BEBC,1) and (BEBC,2). In this case, the data point at $W^2 < 10 \text{ GeV}^2/c^4$ from the earlier release (BEBC,1) is included in the fit, while the others are not because the later (BEBC,2) is covering the same W region. This approach has already been implemented in other studies [229]. The exact W^2 range after requirement (2) is given in Sec. 4.7.

Global fits can be used to expose datasets that pull the results in different directions. This is the case of the most recent $\bar{\nu}_\mu$ measurement by BEBC experiment [258], which did not provide information on systematic errors and, consequently, the total error on this data tends to be much smaller than the rest; see Sec. 4.7. Such small errors give a strong preference to this dataset and, as a consequence, this measurement is in tension with other data, including older $\bar{\nu}_\mu$ BEBC measurements [137] for which information on the systematic uncertainty was provided, see Tab. 4.4. Given that the BEBC [258] analyses did not provide enough information on the systematic errors and they are in clear disagreement with the other ones, these are not considered in the tune and are shown for comparison only.

4.5 Parametrization of model uncertainties

This section discusses the impact on AGKY parameters on the predictions. The predictions are generated with the G18_02a_02_11a tune of GENIE version 3.0.6. This tune was previously obtained to improve the agreement with pion production data on free nucleon [217]. The complete model list for this tune is summarized in Tab. 4.6. As introduced in Sec. 4.2, hadronization is modeled with the AGKY model [219].

Simulation domain	Model
Nuclear model	Fermi Gas [61]
QEL	Llewellyn Smith [31]
QEL Charm	Kovalenko [139]
QEL $\Delta S = 1$	Pais [33]
RES	Rein-Sehgal [34]
SIS/DIS	Bodek-Yang [38]
DIS $\Delta S = 1$	Aivazis-Olness-Tung [140]
Coherent π production	Rein-Sehgal [34]
Hadronization	AGKY [219]
FSI	INTRANUKE hA [37]

Table 4.6: Complete list of models used for the G18_02a_02_11a/b CMC in GENIE v3.

Interactions with nuclei are calculated within the Relativistic Fermi Gas framework, using the Bodek-Ritchie model [61], and hadronic reinteractions are simulated using INTRANUKE hA. The main contributions to the averaged charged multiplicity predictions come from CC DIS and non-resonance SIS [38]. As the DIS and models are common for all GENIE v3 tunes, the choice of the base configuration does not affect the hadronization predictions. An updated version of the G18_02a_02_11a tune, named G18_02a_02_11b, has been recently released in Ref. [217]. In terms of the hadronization predictions, these CMC's are interchangeable and the results of this work are valid within the updated version.

The subset of parameters controlling the averaged charged hadron multiplicity is the target of our tune. The list contains the parameters α_{ch} and β_{ch} defined in Eq. 4.1 and the five PYTHIA parameters discussed before in Sec. 4.2.2. The ranges for α_{ch} and β_{ch} parameters are chosen in a such a way that they cover the values reported by experimental fits, see Tab. 4.5. The same approach is followed to define the PYTHIA parameters range from the HERMES tune, see Tab. 4.2.

The impact of each parameter range on the predictions of averaged charge multiplicity for ν_μ CC interactions on proton is shown in Fig. 4.9. As expected both α_{ch} and β_{ch} act on low W and their effect vanishes gradually over the transition region. In the PYTHIA region, the largest contribution comes from the Lund a and Lund b parameters. In the transition region, the prediction will be determined by both sets of parameters: as a consequence we anticipate a correlation between PYTHIA and the low- W AGKY parameters after the fit. The parameter ranges that defines the parameter space are defined in Tab. 4.7.

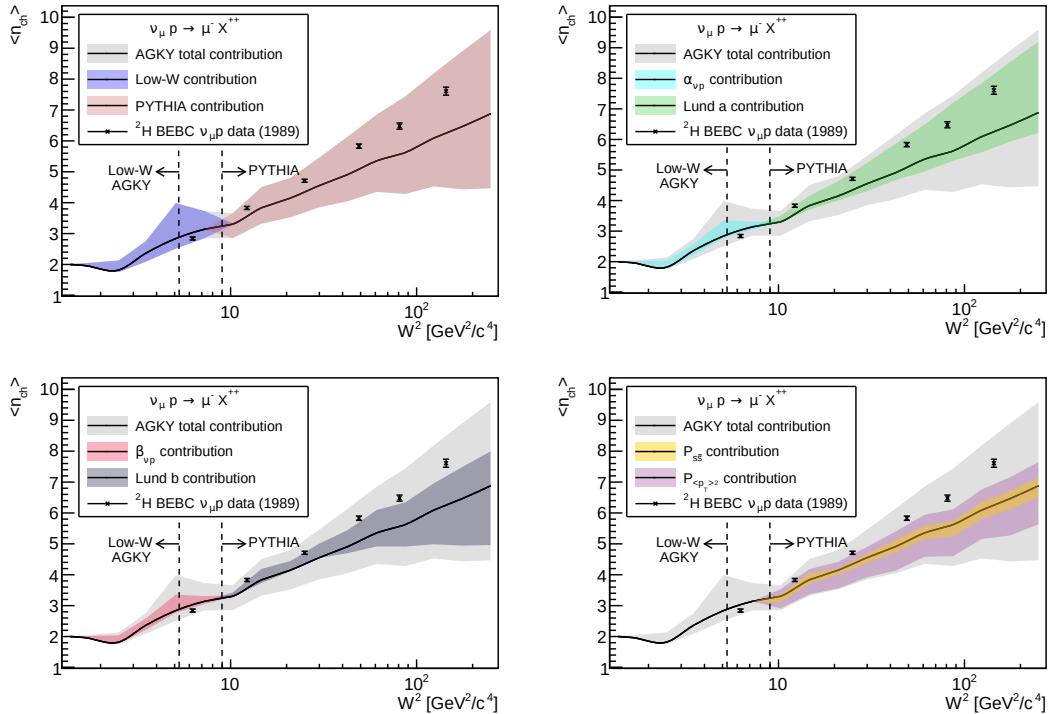


Figure 4.9: Impact of fit parameters on the prediction of the averaged charged multiplicity, as a function of W , for $\nu_\mu p \rightarrow \mu^- X^{++}$ interaction on a deuterium target. Each parameter has been varied within the range of study specified in Tab. 4.7. The top left plot shows the total contributions from low- W and PYTHIA parameters. All the other plots specify the contribution from specific parameters compared to the total, which is always rendered with the grey area. Dashed lines correspond to W_{\min}^{tr} and W_{\max}^{tr} , defining the transition region. BEBC data [258] are shown for reference.

4.6 Construction of the GENIE predictions and evaluation of the likelihood

To build the hadronization prediction for the data described in Sec. 4.3, v_μ and \bar{v}_μ CC events on H and ${}^2\text{H}$ are simulated. Events are generated using a "1/ E "-like flux, with a 0.1 – 200 GeV energy range. This is sufficient as the observables are given in terms of W ; hence, the neutrino flux is factorized out.

In order to compute the prediction associated to the i th dataset from Tab. 4.3 and Tab. 4.4, we select events simulated with the neutrino flux and target of the corresponding experiment and processed using the same experimental cuts. For each selected event we reconstruct E_V and W following the recipes described in Sec. 4.3. The events are classified in bins according to the reconstructed W , and for each bin, we evaluate the average charged multiplicity $\langle n_{\text{ch}} \rangle_i(W)$. This operation is repeated for a number of points in the parameter space $\boldsymbol{\theta}$ defined in Tab. 4.7. Each experiments has a different binning system, and therefore we identify the W bins using two indices: one for the dataset (i) and the other one for the bin index inside the dataset (j). Labeling with $\boldsymbol{\theta}$ the vector of coordinates of a point belonging to the parameters space, we can define our predictions associated to the i th dataset and a given j th W bin as $\langle n_{\text{ch}} \rangle_i(W_{ij}|\boldsymbol{\theta})$. The statistical error due to the MC sample size is also evaluated and this is referred to as $\sigma_{ij}(\boldsymbol{\theta})$.

We use Professor [129] to generate a parametrization denoted as $\tilde{n}_{ij}(\boldsymbol{\theta})$ and $\tilde{\sigma}_{ij}(\boldsymbol{\theta})$ interpolating the values of $\langle n_{\text{ch}} \rangle_i(W_{ij}|\boldsymbol{\theta})$ and $\sigma_{ij}(\boldsymbol{\theta})$ as a function of $\boldsymbol{\theta}$. The parametrization is a generic polynomial of order M in the P -dimensional space [129], whose analytical form is:

$$\begin{aligned}\tilde{n}_{ij}(\boldsymbol{\theta}) &= \alpha_0^{ijk} + \sum_{n=1}^P \beta_n^{ijk} \theta_n + \sum_{n \leq m} \gamma_{nm}^{ijk} \theta_n \theta_m \\ &+ \dots + \sum_{n_1 \leq \dots \leq n_M} \xi_{n_1 \dots n_M}^{ijk} \prod_{\ell=1}^M \theta_{n_\ell},\end{aligned}\quad (4.8)$$

where θ_n is the coordinate of the n -th parameter and M is the polynomial order, set to fourth order in this work. The coefficients α_0^{ijk} , β_n^{ijk} , $\gamma_{(nm)}^{ijk}$, ..., $\xi_{(n_1 \dots n_M)}^{ijk}$ are determined by Professor fitting the parametrization against the computed $\langle n_{\text{ch}} \rangle_i(W_{ij}|\boldsymbol{\theta})$ obtained by generating $O(10^4)$ points uniformly spread within parameter space defined in Tab. 4.7. Nonphysical regions in the sampled parameter space are avoided applying a veto function. In particular, every combination of $\boldsymbol{\theta}$ has to verify that $\langle n_{\text{ch}} \rangle \geq 0$ at the pion production threshold. The parametrization $\tilde{n}_{ij}(\boldsymbol{\theta})$ is used instead of the exact

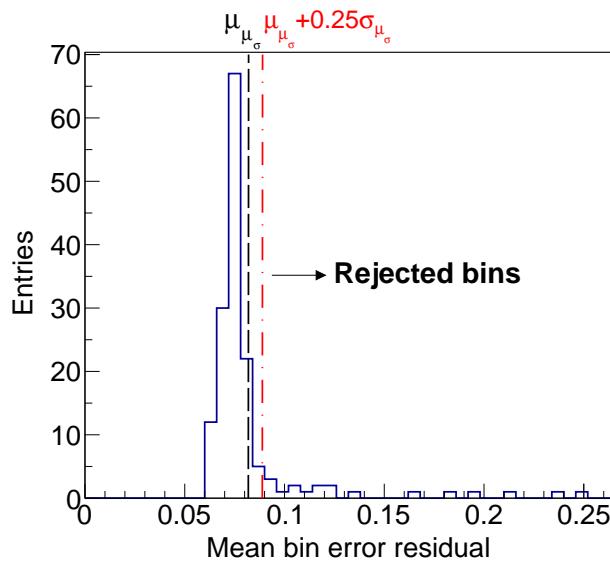


Figure 4.10: Distribution of mean bin error residual for all data points. The distribution mean value (μ_{μ_σ}) is shown with a dashed black line. Data points with a mean value higher than 0.25 of the mean bin error variance (σ_{μ_σ}) are rejected. This cut-off value is shown with the dashed red line.

predictions in order to estimate the best-fit parameters by minimising the χ^2 . The main advantage of this method is the reduction of the brute-force scans computational complexity while allowing for massive parallelization.

4.6.1 Professor interpolation cutoff condition

As mentioned in Sec. 4.4.3, in view of the fact that the Professor interpolation is just an approximation, it can fail to describe the actual prediction. When this happens, we remove from the analysis data points whose Professor interpolation, of the predicted mean value or predicted error, disagree too much with the GENIE prediction corresponding to that data point. The relative difference between the interpolation and the GENIE prediction is known as residual. For each data point, we calculate the bin central value and bin error residuals for all the points in our parameter space. The distributions of the residual for central values and errors are monitored, and whenever the means or the variances of a bin are too far from the average values among all bins, the corresponding data point is removed. In this analysis, the cutoff condition requires that any data points with a mean central value or error that exceeds the average values among all bins by $0.25\sigma_{\mu_\mu}$ or $0.25\sigma_{\mu_\sigma}$ respectively is removed from the analysis. An example of the cutoff condition on the error residual distribution for all data points is shown in Fig. 4.10.

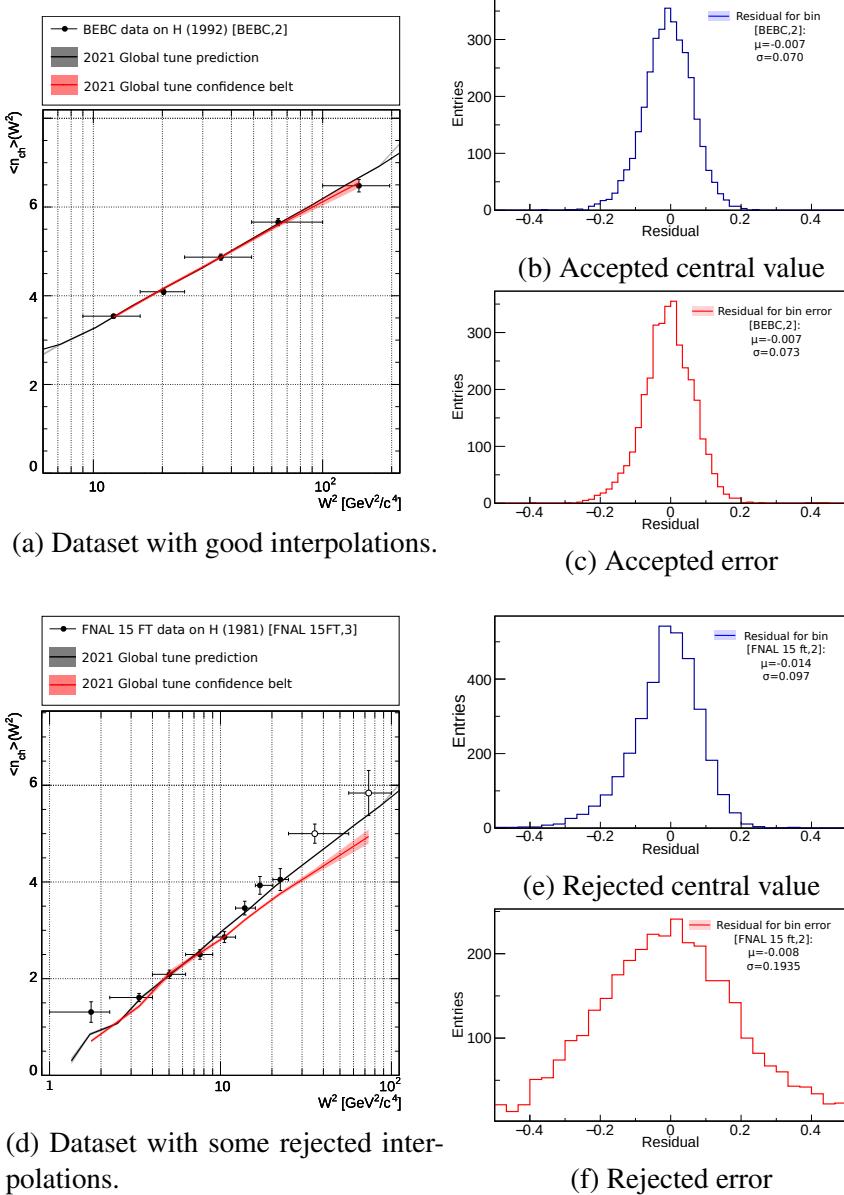


Figure 4.11: Comparisons of v_μ data on H against predictions obtained from the Professor parametrization (red confidence belt) at the best-fit value for the AGKY global tune and the actual GENIE prediction (black line). Datapoints that do not satisfy our selection criteria are shown in empty markers. For the Professor parametrization, the uncertainties of the tuned parameters are propagated to the prediction considering the full covariance matrix. For three selected bins (the bin with highest W for (BEBC,2) (b-c) and the two higher W bins for the (FNAL 15FT,3) plot (e-f) on the left column), the central values and error residual distributions are shown, blue and red respectively: accepted parametrizations at the top, rejected parametrization at the bottom. It can be seen that the residual distribution of the rejected bins is wider than its accepted counterpart. In this particular case, the two data points with higher W are neglected as the parametrization of the bin value and error do not satisfy the required criteria.

The corresponding distribution associated to the bin central value and bin error residual for the last bin of the BEBC [253] and the two lasts bins of the FNAL 15 ft [136] datasets are shown in Fig. 4.11 (b-c) and (e-f) respectively. Two examples are given in Fig. 4.11: a dataset in which the interpolation is accurate for all the W^2 range and a dataset in which the interpolation fails for some of the dataset points, highlighted with empty markers. This criteria allows us to ensure that the Professor parametrization does not fail for the data considered in the tune. A total of $\sim 18\%$ of the data points have been removed due to this requirement. In this work, it has been observed that the residual variance increases with W^2 , with few exceptions. The complete list of removed datapoints is specified in Sec. 4.7.

The variance of the residual distribution for a given data point can be improved by increasing the order of the polynomial used for the Professor interpolation. In this case, a polynomial of order 4 is used. However, specifically in this particular tune where 13 parameters are tuned, an increase of the order requires the generation of a much higher number of MC samples, which can be very computationally demanding.

4.6.2 Parameter priors

Our parameters of interest affect other hadronization observables and not only the averaged multiplicity. This is taken into account by using Gaussian priors.

The $s\bar{s}$ suppression factor not only impacts the averaged multiplicity data but also the η multiplicity production, see Fig. 4.4. A prior of 0.30 ± 0.05 is considered in order to preserve a good agreement with the SKAT data [250].

Variations of E_{CutOff} affect the shape of $F(x_F)$ invariant distribution, defined as

$$F(x_F) = \frac{1}{N_{\text{ev}}} \cdot \frac{1}{\pi} \cdot \frac{E}{p^{L \max} c} \cdot \frac{dN}{dx_F}, \quad (4.9)$$

where x_F is the Feynman variable, N_{ev} the total number of events, and E and $p^{L \max}$ the energy and maximum longitudinal momentum of the final-state hadron in the hadronic center of mass respectively. The $F(x_F)$ invariant distribution describes the fragmentation process for the forward and backward hemispheres, and it allows to study the symmetry between this two fragmentation regions. In Fig. 4.12, the $F(x_F)$ invariant distribution for $\bar{\nu}_\mu$ data on ${}^2\text{H}$ [259] is compared against GENIE predictions obtained by varying the E_{CutOff} within a $[0, 2]$ GeV range. The main conclusion is that small values of this parameter preserve the agreement with data. In order to avoid an increase of $F(x_F)$ at $|x_F| \sim 1$, we apply a prior on E_{CutOff} of 0.25 ± 0.05 GeV.

Another parameter that has a strong impact on other observables is $\langle p_\perp^2 \rangle$. As demonstrated in Fig. 4.5, low values of $\langle p_\perp^2 \rangle$ are not in agreement with data for $\langle p_T^2 \rangle$

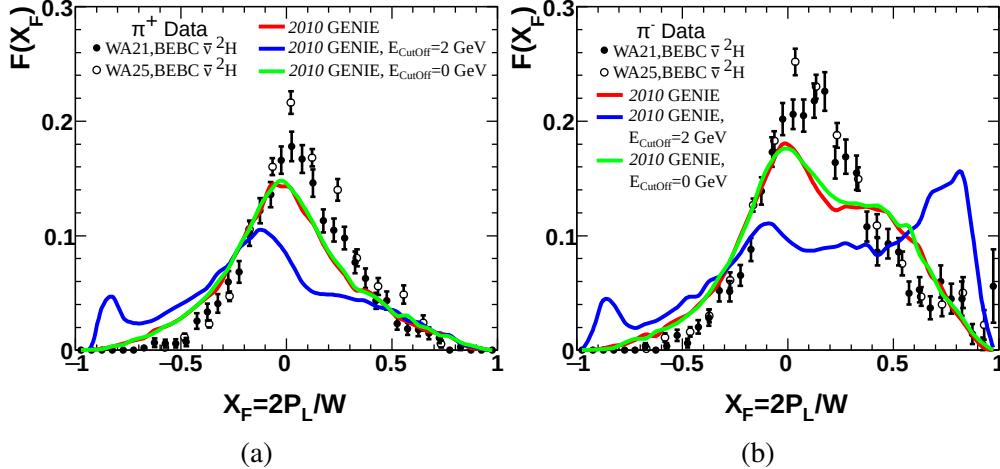


Figure 4.12: Effect of E_{CutOff} on the x_F invariant distributions for π^+ (a) and π^- (b) in $\bar{\nu}_\mu$ data on ${}^2\text{H}$ from the BEBC experiment [259]. The 2010 GENIE tune value for the energy cut-off is $E_{\text{CutOff}} = 0.2 \text{ GeV}$.

distributions. Thus, we also apply a prior on the parameter to guarantee the agreement with this data of $0.44 \pm 0.05 \text{ (GeV}/c)^2$. No priors are applied to the remaining parameters.

4.6.3 Final form of the χ^2

Using the parametrization and the corresponding set of points belonging to the i th dataset, $\mathcal{D}_{ij} \pm \delta \mathcal{D}_{ij}$, we seek to estimate the best-fit parameters $\hat{\boldsymbol{\theta}}$ by minimising the quantity:

$$\chi^2(\boldsymbol{\theta}) = \sum_{i,j} w_{ij} \frac{(\tilde{n}_{ij}(\boldsymbol{\theta}) - \mathcal{D}_{ij})^2}{\tilde{\sigma}_{ij}^2(\boldsymbol{\theta}) + \delta \mathcal{D}_{ij}^2} + \sum_l \frac{(\theta_l - \mu_l)^2}{\sigma_l^2}. \quad (4.10)$$

The first term allows the minimization between data and prediction while applying weights, w_{ij} , that allow to consider only specific data points in the fit. The second term adds uncorrelated Gaussian priors for a given parameter; the vectors of central values and variances are denoted μ_l and σ_l respectively.

4.7 AGKY tune results

Starting from ν_μ and $\bar{\nu}_\mu$ hadronization data, two tunes were considered: a global tune (2021 GENIE global) and a deuterium only tune (2021 ${}^2\text{H}$). The reason for a deuterium-only fit is because other studies showed tensions between data on H and ${}^2\text{H}$ targets on bubble chamber experiments [229]. The goal of the global tune is to improve the agreement with hydrogen and deuterium targets, regardless of these tensions,

4.7. AGKY TUNE RESULTS

Parameter	GENIE parameter name	2010 GENIE	Allowed range	2021 Global Fit	2021 ^2H Fit
Low- W empirical model					
α_{vp}	KNO-Alpha-vp	0.40	$[-1.0, 2.0]$	1.1 ± 0.3	1.2 ± 0.4
α_{vn}	KNO-Alpha-vn	-0.20	$[-1.0, 2.0]$	$1.75^{+0.14}_{-0.11}$	-0.58 ± 0.07
$\alpha_{\bar{v}p}$	KNO-Alpha-vbp	0.02	$[-1.0, 2.0]$	$1.32^{+0.16}_{-0.14}$	1.9 ± 0.08
$\alpha_{\bar{v}n}$	KNO-Alpha-vbn	0.80	$[-1.0, 2.0]$	1.11 ± 0.09	1.07 ± 0.3
β_{vp}	KNO-Beta-vp	1.42	$[0.0, 2.5]$	0.79 ± 0.15	0.9 ± 0.3
β_{vn}	KNO-Beta-vn	1.42	$[0.0, 2.5]$	0.5 ± 0.1	1.9 ± 0.3
$\beta_{\bar{v}p}$	KNO-Beta-vbp	1.28	$[0.0, 2.5]$	0.8 ± 0.1	0.3 ± 0.1
$\beta_{\bar{v}n}$	KNO-Beta-vbn	0.95	$[0.0, 2.5]$	$0.88^{+0.09}_{-0.08}$	0.9 ± 0.2
PYTHIA					
$P_{s\bar{s}}$	PYTHIA-SSBarSuppression	0.30	$[0.0, 1.0]$	0.27 ± 0.04	0.29 ± 0.05
$\langle p_\perp^2 \rangle [\text{GeV}^2/c^2]$	PYTHIA-GaussianPt2	0.44	$[0.1, 0.7]$	0.46 ± 0.05	0.43 ± 0.04
$E_{\text{CutOff}} [\text{GeV}]$	PYTHIA-RemainingEnergyCutoff	0.20	$[0.0, 1.0]$	0.30 ± 0.04	0.24 ± 0.05
Lund a	PYTHIA-Lunda	0.30	$[0.0, 2.0]$	1.53 ± 0.13	1.85 ± 0.15
Lund b [c^4/GeV^2]	PYTHIA-Lundb	0.58	$[0.0, 1.5]$	1.16 ± 0.09	1.0 ± 0.2
		$\chi^2 =$	87.9/62 DoF	29.5/32 DoF	

Table 4.7: Best-fit result parameters for the AGKY 2021 global tune and ^2H -only tune. The range of study and priors used in the tune are specified in the table. The 2010 GENIE AGKY parameter values are also specified for reference, as well as the parameter name used in the GENIE software. See Sec. 4.7.1 and Sec. 4.7.2 for the details on the error estimation of each tune. Posterior distributions are not always symmetric; in that case the interval is reported accordingly. The total χ^2 obtained from each fit is obtained from the minimization of Eq. 4.10.

while the deuterium-only tune was performed to quantify the tensions within the same framework. An hydrogen-only fit was not considered because it could not constrain the neutron related parameters of the low- W empirical model.

The analyses procedure outlined in the previous sections was applied to both tunes. The likelihood function, Eq. 4.10, was minimized against the averaged charged multiplicity data that satisfies our selection criteria; see Sec. 4.3. The best-fit parameter set for both tunes and the χ^2 values obtained using the Professor parametrizations and Eq. 4.10 are summarized in Tab. 4.7.

GENIE predictions for all the averaged charged multiplicity data available are shown in Figs. 4.13–4.16 before and after the AGKY tunes. The results show the prediction for the 2021 GENIE global tune in red and 2021 GENIE ^2H tune in green. To distinguish data points used in the analyses from those that were not, the used points have completely black markers, the others are represented as empty circles. Each dataset is associated to a tag, defined in Tab. 4.7. Vertical error bars include statistical and systematic uncertainties following our data review. Horizontal bars correspond to the bin width used in the data release, and are only shown if those are available in the original paper.

Experiment	Target	Tag	Ref.
$\nu_\mu + p \rightarrow \mu^- X^{++}$			
FNAL 15 ft (1976)	H	FNAL 15 ft,0	[251]
BEBC (1983)	H	BEBC,0	[135]
BEBC (1990)	H	BEBC,1	[252]
BEBC (1992)	H	BEBC,2	[253]
FNAL 15 ft (1983)	² H	FNAL 15 ft,1	[136]
BEBC (1989)	² H	BEBC,3	[258]
$\nu_\mu + n \rightarrow \mu^- X^+$			
FNAL 15 ft (1983)	² H	FNAL 15 ft,2	[136]
BEBC (1984)	² H	BEBC,4	[259]
BEBC (1989)	² H	BEBC,5	[258]
$\bar{\nu}_\mu + p \rightarrow \mu^+ X^0$			
FNAL 15 ft (1981)	H	FNAL 15 ft,3	[254]
BEBC (1983)	H	BEBC,6	[135]
BEBC (1990)	H	BEBC,7	[252]
BEBC (1992)	H	BEBC,8	[253]
BEBC (1982)	² H	BEBC,9	[137]
BEBC (1989)	² H	BEBC,10	[258]
$\bar{\nu}_\mu + n \rightarrow \mu^+ X^-$			
BEBC (1982)	² H	BEBC,11	[137]
BEBC (1989)	² H	BEBC,12	[258]

Table 4.8: Summary of data used for comparisons in Figs. 4.7, pPredictionsHadro , 4.14, and 4.16. This table links the experiment and the tag used for the legend in each plot to the corresponding reference.

4.7. AGKY TUNE RESULTS

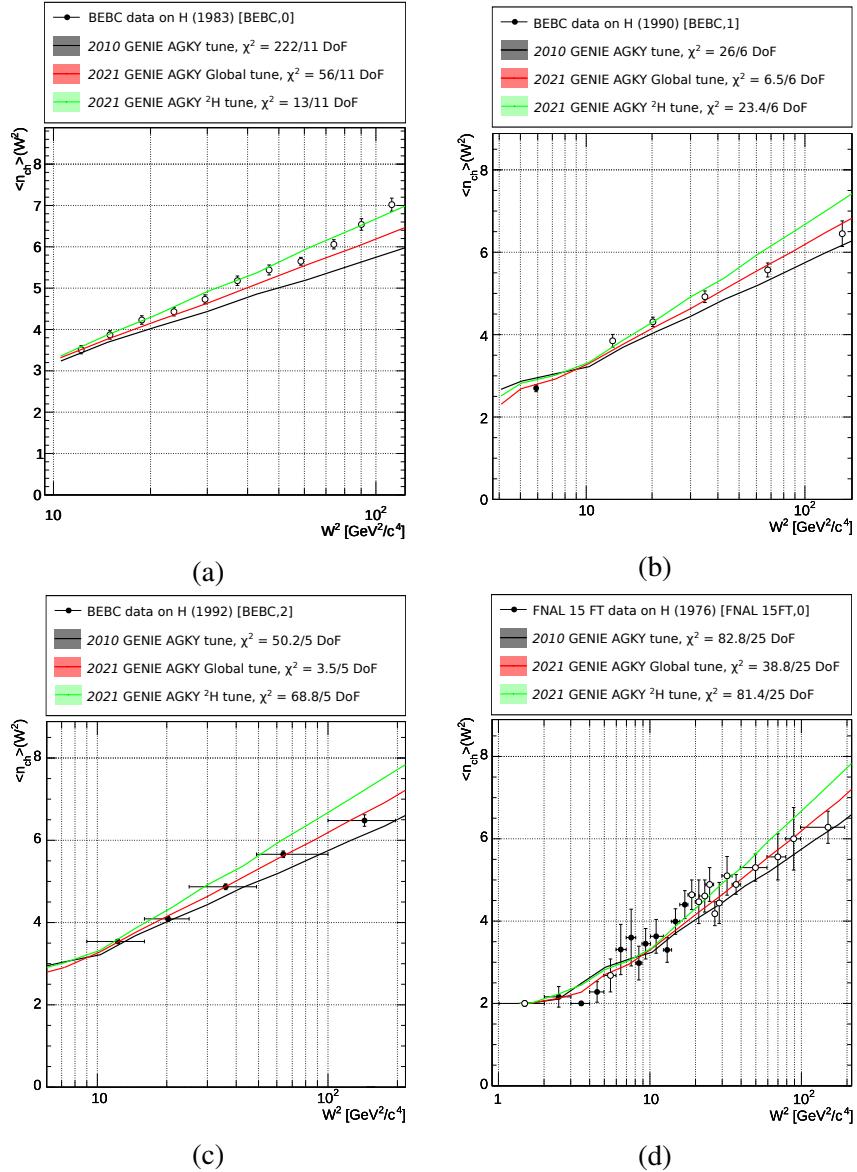


Figure 4.13: Comparison of $\langle n_{ch} \rangle$ against neutrino-induced hadronization data on $\nu_\mu + p$ interactions on H from BEBC [135, 252, 253] and FNAL [251] bubble chamber experiments filled with H. Datapoints used in the AGKY 2021 global tune analysis are shown as filled black markers. Discarded datapoints are represented using empty markers. The ${}^2\text{H}$ tune prediction is shown for comparison only. The predictions are computed using the parameters specified in Tab. 4.7. The χ^2 values are calculated against all the data from each experiment. See definition of *tags* in Tab. 4.7.

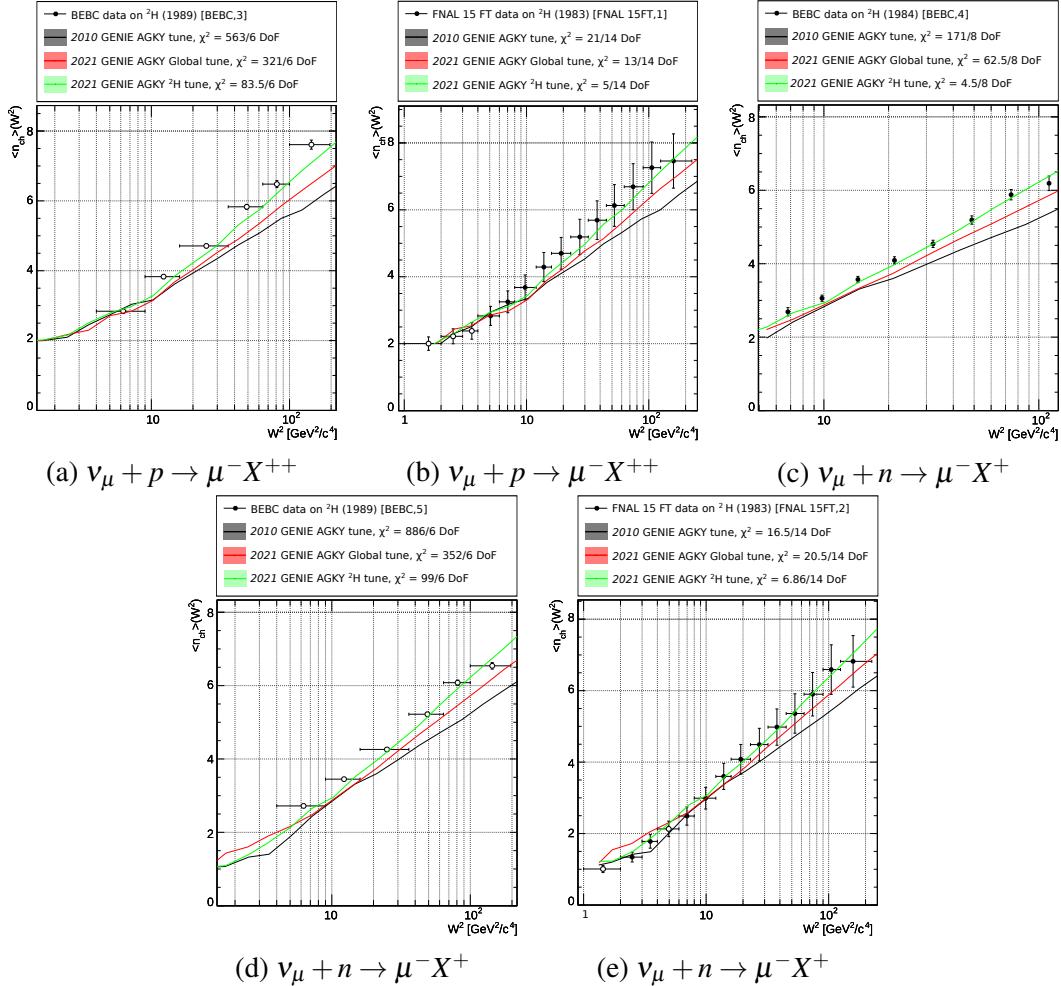


Figure 4.14: Comparison of $\langle n_{\text{ch}} \rangle$ against neutrino-induced hadronization data on ν_μ interactions on p and n from the BEBC bubble chamber experiment filled with ${}^2\text{H}$ [137, 258]. Datapoints used in the AGKY 2021 global tune analysis are shown as filled black markers. Discarded datapoints are represented using empty markers. The predictions are computed using the parameters specified in Tab. 4.7. The χ^2 values are calculated against all the data from each experiment. See definition of *tags* in Tab. 4.7.

4.7. AGKY TUNE RESULTS

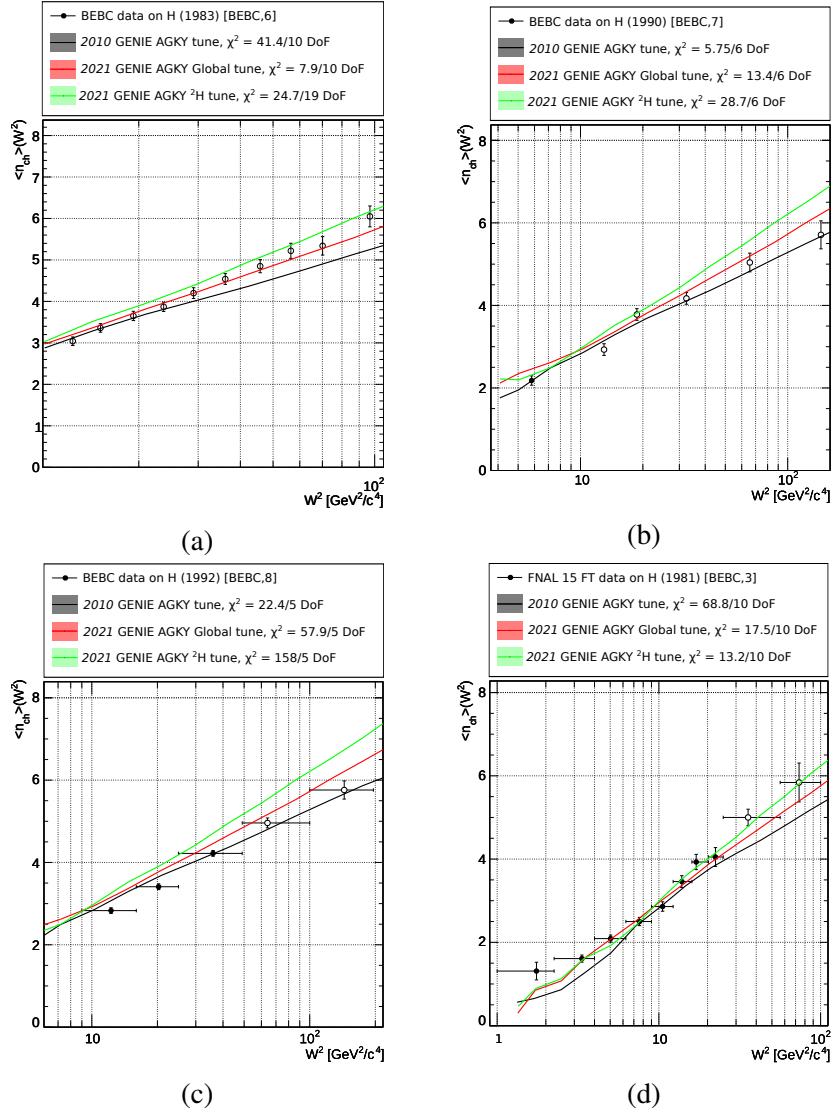


Figure 4.15: Comparison of $\langle n_{\text{ch}} \rangle$ against neutrino-induced hadronization data on $\bar{\nu}_\mu + p$ interactions on H from the BEBC [135, 252, 253] and FNAL [254] bubble chamber experiment filled with H. Datapoints used in the AGKY 2021 global tune analysis are shown as filled black markers. Discarded datapoints are represented using empty markers. The predictions are computed using the parameters specified in Tab. 4.7. The χ^2 values are calculated against all the data from each experiment. See definition of tags in Tab. 4.7.

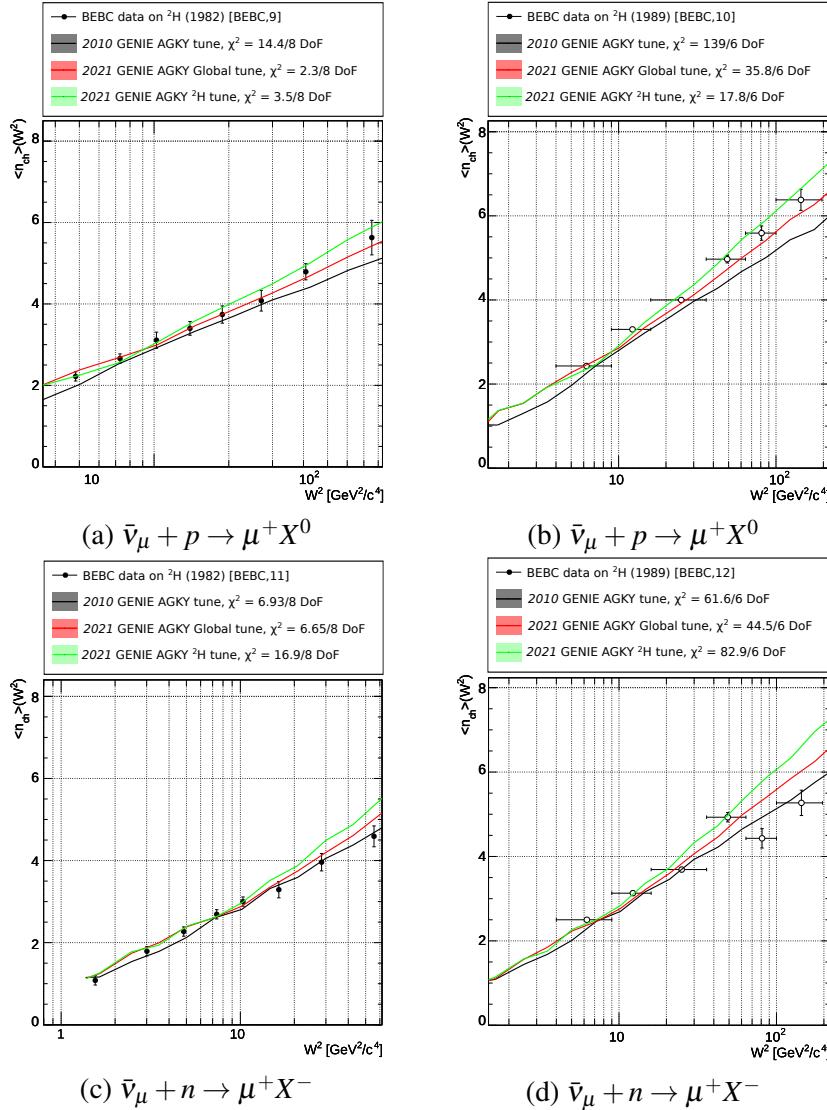


Figure 4.16: Comparison of $\langle n_{ch} \rangle$ against neutrino-induced hadronization data on $\bar{\nu}_\mu$ interactions on p and n from the BEBC bubble chamber experiment filled with ^2H [137, 258]. Datapoints used in the AGKY 2021 global tune analysis are shown as filled black markers. Discarded datapoints are represented using empty markers. The predictions are computed using the parameters specified in Tab. 4.7. The χ^2 values are calculated against all the data from each experiment. See definition of *tags* in Tab. 4.7.

In terms of prediction differences, the *2021 GENIE* global tune tends to underpredict deuterium data whereas the *2021 GENIE* ^2H tune overpredicts the hydrogen data. This especially true for the PYTHIA region, at high W . This is translated in the parameters with an increase (decrease) of Lund a (Lund b) for the deuterium tune with respect to the global tune.

The summary of the χ^2 values per dataset as well as the total contributions are shown in Tab. 4.9. Three different χ^2 values are presented: χ^2_{2010} , $\chi^2_{2021(\text{Global})}$ and $\chi^2_{2021(^2\text{H})}$ using, respectively, the *2010 GENIE*, *2021 GENIE* global and *2021 GENIE* ^2H tune parameters. The χ^2 values per dataset are computed by comparing the GENIE predictions against all the data points in each dataset, regardless of the point being used in the fit or not. Differences between the χ^2 obtained with Eq. 4.10 and the one calculated using the GENIE predictions directly are expected. One of the reasons is that Eq. 4.10 only considers the datapoints included in the tune. Moreover, further differences arise from the fact that the Professor parametrization \tilde{n}_{ij} is not exact, as explained in Sec. 4.6.

It is important to stress that the total χ^2 from Tab. 4.9 are not providing any information related to goodness of fit, but it simply shows the general agreement with respect to available datasets. A sense of the goodness of fit can be obtained looking at the total χ^2 calculated with the datasets included in the fit only; see Tab. 4.10.

The parameters covariance matrices for both tunes are obtained by inverting the Hessian of the log-likelihood function at the best fit point; see Tab. 4.11 and Tab. 4.12. As expected, the low- W AGKY and PYTHIA parameters are now correlated in both tunes because of the interplay of the models in the transition region, with a number of parameters showing a correlation above 50%. See a graphical representation of the correlation matrix in Fig. 4.17.

The results form the *2021 GENIE* AGKY tunes will available in GENIE v3.2.0. Users can run the *2021 GENIE* tunes global and ^2H tunes out of the box using the G18_02a_03_330 and G18_02a_03_320 comprehensive configurations respectively.

4.7.1 The *2021 GENIE* AGKY global tune

After the AGKY global tune, the GENIE predictions show a better agreement to the data. In particular, for the datasets included in the *2021 GENIE* global tune, the χ^2 associated to the prediction is $\chi^2_{2010} = 486/109$ degrees of freedom (DoF). After the tune, the $\chi^2_{2021(\text{Global})}$ is $242/109$ DoF. This is clearly an improvement although the agreement is not completely satisfactory since the p value is 4×10^{-12} .

The improvement in the data description is general and both deuterium and hydrogen samples have a better agreement. Moreover, after the tune both hydrogen and

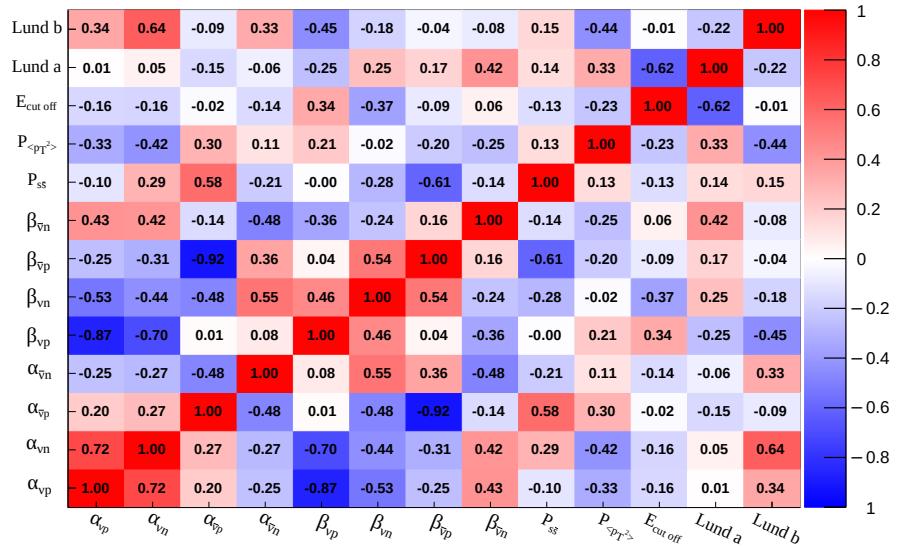
Experiment	χ^2_{2010}	$\chi^2_{2021(\text{Global})}$	$\chi^2_{2021(^2\text{H})}$	DoF	In tune
$\nu_\mu + p \rightarrow \mu^- X^{++}$					
Data on hydrogen					
FNAL 15 ft,0	83	39	81	25	Partially
BEBC,0	222	56	13	11	\times
BEBC,1	26	7	23	6	Partially
BEBC,2	50.2	3.5	68.8	5	\checkmark
Data on deuterium					
FNAL 15 ft,1	21	13	5	14	Partially
BEBC,3	563	321	84	6	\times
Total for $\nu_\mu p$	965	447	275	67	
$\nu_\mu + n \rightarrow \mu^- X^+$					
FNAL 15 ft,2	17	21	7	14	Partially
BEBC,4	171	6	5	8	\checkmark
BEBC,5	886	352	99	6	\times
Total for $\nu_\mu n$	1,074	435	111	28	
$\bar{\nu}_\mu + p \rightarrow \mu^+ X^0$					
Data on hydrogen					
FNAL 15 ft,3	69	18	13	10	Partially
BEBC,6	41	8	25	10	\times
BEBC,7	5.8	13.4	28.7	6	Partially
BEBC,8	22.4	57.9	158.0	5	\checkmark
Data on deuterium					
BEBC,9	14	2	4	8	\checkmark
BEBC,10	139	36	18	6	\times
Total for $\bar{\nu}_\mu p$	292	135	246	45	
$\bar{\nu}_\mu + n \rightarrow \mu^+ X^-$					
BEBC,11	6.9	6.7	16.9	8	\checkmark
BEBC,12	61.6	44.5	82.9	6	\times
Total for $\bar{\nu}_\mu n$	69	51	100	14	
χ^2 Summary					
All data	2,398	1,068	731	154	
All ${}^2\text{H}$ data	1,879	858	320	76	
All H data	519	202	411	78	

Table 4.9: Summary of χ^2 values for the datasets shown in Figs. 4.14, 4.16, 4.13, and 4.15. The table shows the χ^2 per dataset and interaction channel as well as the total and per channel χ^2 . The χ^2 values are calculated using the GENIE predictions for each tune: 2010 GENIE, χ^2_{2010} , 2021 GENIE, $\chi^2_{2021(\text{Global})}$, and 2021 GENIE, $\chi^2_{2021(^2\text{H})}$.

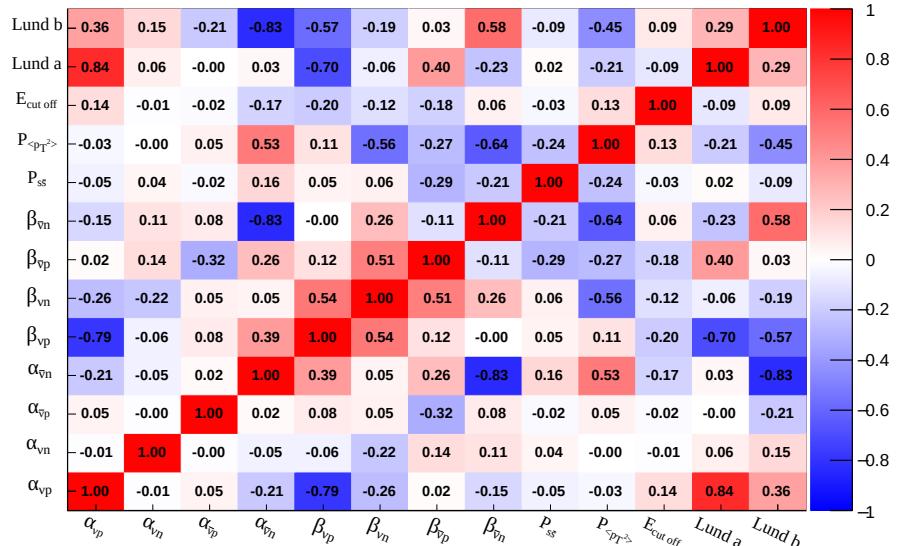
4.7. AGKY TUNE RESULTS

Datasets	χ^2_{2010}	$\chi^2_{2021(\text{Global})}$	$\chi^2_{2021(2\text{H})}$	DoF
All Data in tune	486	242	410	109
${}^2\text{H}$ Data in tune	230	105	37	52
H Data in tune	256	138	374	57

Table 4.10: Total χ^2 calculated with the datasets included in each fit: 2010 GENIE, χ^2_{2010} , 2021 GENIE, $\chi^2_{2021(\text{Global})}$, and 2021 GENIE, $\chi^2_{2021(2\text{H})}$.



(a) Global tune correlation matrix.



(b) Deuterium only tune correlation matrix.

Figure 4.17: Parameter correlation matrix for the 2021 GENIE AGKY tunes against averaged charged multiplicity data.

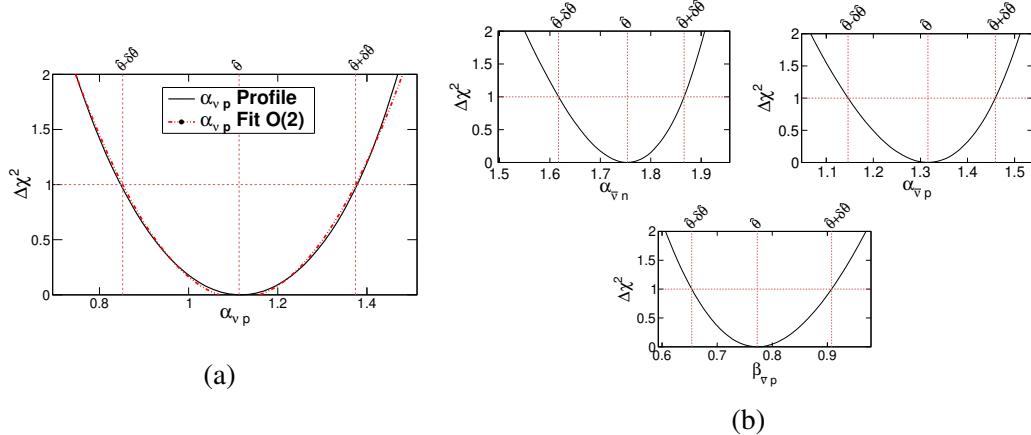


Figure 4.18: Joint function obtained fixing the two parameters under study and minimizing $\Delta\chi^2_{\text{profile}}(\boldsymbol{\theta})$ respect the other parameters in the 2021 GENIE global tune. The dashed lines represent the parameter range that satisfies the condition $\Delta\chi^2_{\text{profile}}(\theta_i) < 1$. This is also denoted as $\hat{\theta} \pm \delta\hat{\theta}$. (a) Symmetric parameter profile, (b) asymmetric parameter profiles.

deuterium samples have similar goodness of fit, and in general, the level of agreement is the same. This can be noted from the χ^2 contributions from Tab. 4.10.

The agreement with the datasets not included in the tune has also improved, as shown in Tab. 4.9. The total χ^2 computed using all available data is reduced significantly for both H and ${}^2\text{H}$ datasets. Particularly, the global tune shows a better agreement against all hydrogen data. As expected from Sec. 4.4.3, the datasets with highest contribution to the total χ^2 after the global tune are (BEBC,3) and (BEBC,5).

The main effect of the tune is observed in the PYTHIA region, at $W > 3 \text{ GeV}/c^2$, where the prediction of $\langle n_{\text{ch}} \rangle$ increased. This is a direct consequence of the increase on Lund a and Lund b . This behaviour is consistent with the HERMES tune, summarized in Sec. 4.2.2.

For each parameter, the corresponding uncertainty is obtained with the profiling method under the condition $\Delta\chi^2_{\text{profile}}(\theta_i) < 1$. The profiles are calculated by fixing the value of the parameter under study θ_i to a desired value and minimizing the quantity $\Delta\chi^2(\boldsymbol{\theta}) = \chi^2(\boldsymbol{\theta}) - \chi^2_{\min}$ with respect to all others parameters that were allowed to float in the fit. The constant χ^2_{\min} corresponds to the global minimum value of $\chi^2(\boldsymbol{\theta})$. Some parameters have a good Gaussian behaviour and a symmetric profile. For some others this is not true and this gives rise to asymmetric uncertainties for the parameters. Example of a symmetric parameter profile compared to the non-Gaussian ones is shown in Fig. 4.18. The contours for some pairs of the AGKY parameters are shown in Fig. 4.19.

The fit covariance matrix can be propagated back to the GENIE predictions giving

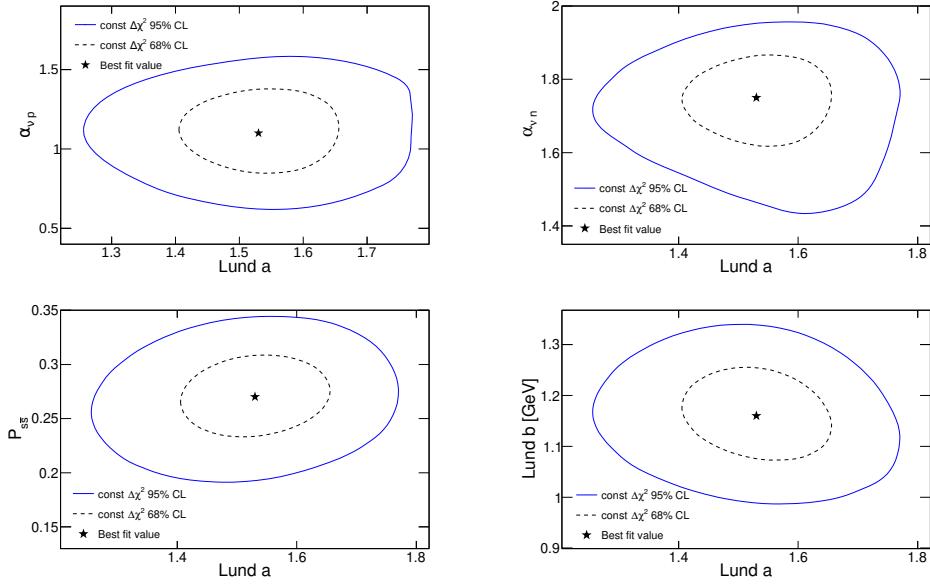


Figure 4.19: Joint $\Delta\chi^2_{\text{contour}}(\theta_i, \theta_j)$ function obtained fixing the two parameters under study and minimizing $\Delta\chi^2_{\text{contour}}(\boldsymbol{\theta})$ respect the other parameters in the 2021 GENIE global tune. The 95% and 68% contour lines are shown as well as the best fit values for the global tune.

a posterior confidence belt for the prediction associated to the tune. As an example, a comparison of the global tune prediction and the associated posterior confidence belt is shown in Fig. 4.11.

4.7.2 The 2021 GENIE AGKY ${}^2\text{H}$ tune

For the datasets included in the deuterium only tune, the χ^2 associated to the 2010 GENIE AGKY prediction is $\chi^2_{2010} = 230/52$ DoF. After the tune, the total $\chi^2_{2021({}^2\text{H})}$ is 37/52 DoF that corresponds to a p value of 0.94. Being the deuterium only goodness of fit so much better than the global tune confirms the high tension between H and ${}^2\text{H}$ datasets.

Surprisingly, the deuterium-only tune shows a better agreement than the global tune when all neutrino-induced hadronization data are considered, see Tab. 4.9. Yet, this does not imply that the deuterium-only fit is a better tune, it simply reinforces that the discarded dataset are not compatible with the data used in the fit.

The tension between hydrogen and deuterium data was already observed in other studies where a modified KNO-based model was tuned to averaged charged multiplicity data from bubble chamber experiments [229]. Those studies suggest that the origin of tensions between H and ${}^2\text{H}$ could be due to rescattering effects on deuterium. As explained in Sec. 4.3, the bubble chamber experiments claim that rescattering effects

have a smaller effect on neutron samples as a consequence of the classification into ν_μ on p or ν_μ on n events. This is a consequence of the neutron reinteraction with the proton from the deuterium, which is then kicked out and, therefore, miss-identified as a $\nu_\mu p$ event. If the disagreements were only due to rescattering, the global tune would have a better agreement than the deuterium-only tune on ν_μ and $\bar{\nu}_\mu$ on neutron data. However, a better agreement of the global tune on neutron samples is not observed.

4.7.3 AGKY global and deuterium only tunes impact on other neutrino-induced hadronization observables

The analyses procedure discussed in this paper focuses on the description of the charged averaged multiplicity. However, as discussed in Sec. 4.4.1, different observables are linked with the shower particle content description. In this section, the effect of the *2021* global tune on different hadronization observables is discussed.

A wider comparison against all available hadronization observables for the G00_00a_00_00a AGKY predictions is reported in [219]. Some information provided by these observables is included in the tune using priors; see Sec. 4.6. The agreement of the *2021* GENIE AGKY global tune with these observables is not compromised.

There are, however, other observables that show tensions with the averaged charged multiplicity data. The neutral pion averaged multiplicity is related with the charged hadron multiplicity via Eq. 4.2; an increase on the charged averaged multiplicity is equivalent to a higher neutral pion averaged multiplicity. This result is incompatible with the data, as demonstrated in Fig. 4.20. Another example is the dispersion observable, defined as $D = \sqrt{\langle n^2 \rangle - \langle n \rangle^2}$. The comparison of data on the ratio of $D/\langle n_{\text{ch}} \rangle$ vs W^2 for the different tunes is shown in Fig. 4.21. In this case, the disagreement also increases with the invariant mass.

The tension between charged averaged multiplicity with $\langle n_{\pi^0} \rangle$ and dispersion data was already observed when using the HERMES parametrization described in Sec. 4.4.2. The origin of these tensions is beyond the scope of this paper as we aim for a better description of the charged averaged multiplicity data only. The further understanding of the connection between the different observables would require to repeat the analyses procedure of this paper including other hadronization related observables. Yet, it is important to understand how the *2021* GENIE AGKY global tune impacts other hadronization related observables.

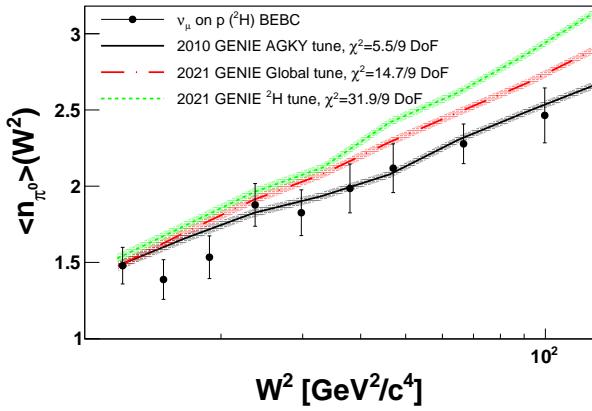


Figure 4.20: Comparison of the predicted $\langle n_{\pi^0} \rangle$ against neutrino-induced hadronization data on ν_μ interactions on p from the BEBC bubble chamber experiment filled with 2H [137, 258]. The predictions shown correspond to the 2010 GENIE AGKY (black), the 2021 GENIE AGKY global (red) and the 2021 GENIE AGKY 2H (green) tunes.

4.7.4 2021 GENIE AGKY global tune impact at the SIS region

Other nonhadronization observables can be affected by this tune. The main impact is on the description of the SIS region in GENIE, since it is linked with final state multiplicities [217]. In GENIE, the SIS is modeled applying scaling factors to the DIS cross section. These factors depend on the multiplicity of the process. Hence, variations on the final state multiplicity probabilities [Eq. 4.3] change the scaling applied to the DIS cross section, affecting the DIS contribution to the SIS. The P_n^{had} probability distributions for the 2010 GENIE AGKY tune and for the AGKY global tune are shown in Fig. 4.22.

The impact of the AGKY tune on CC inclusive cross sections is summarized in Fig. 4.23. When applying the AGKY global tune to the SIS region, an increase of CC inclusive cross section is observed, for both ν_μ and $\bar{\nu}_\mu$. The exclusive cross sections for different pion multiplicities show that the AGKY tune enhances the 2π production whilst the 1π production barely changes; see Fig. 4.24. As a consequence, the agreement with inclusive and ν_μ CC $\pi^+\pi^-$ data is lost.

Both, the bare nucleon tune [217] and the 2021 GENIE global tune, show a preference to increase the two-pion production, suggesting that a joint tune could preserve the agreement with inclusive and exclusive data at low W . This was neglected in previous analyses to minimise the tune's complexity, but this analysis clearly suggests otherwise. The high- W AGKY parameters do not need anymore refinements. On the contrary, the low- W parameters require a joint tune in order to have a satisfactory result that can be used to extract data-driven parameter uncertainties.

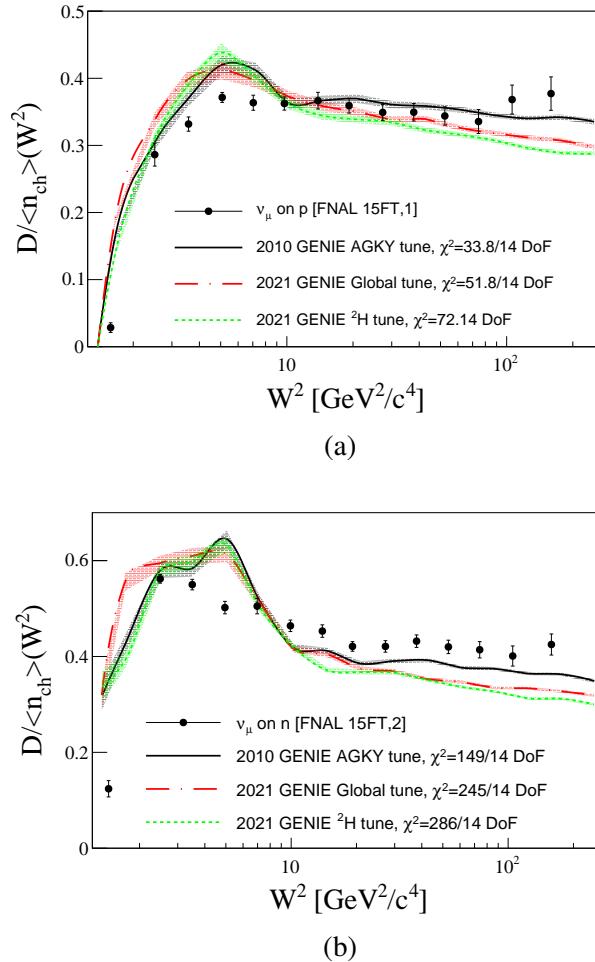


Figure 4.21: Comparison of the predicted $D/\langle n_{ch} \rangle$ against neutrino-induced hadronization data on ν_μ interactions on p (a) and n (b) from the FNAL 15 ft bubble chamber experiment filled with ${}^2\text{H}$ [136]. The predictions shown correspond to the 2010 GENIE AGKY (black), the 2021 GENIE AGKY global (red) and the 2021 GENIE AGKY ${}^2\text{H}$ (green) tunes. (a) Comparison against ν_μ on p data and (b) comparison against ν_μ on n data.

4.8. CONCLUSIONS

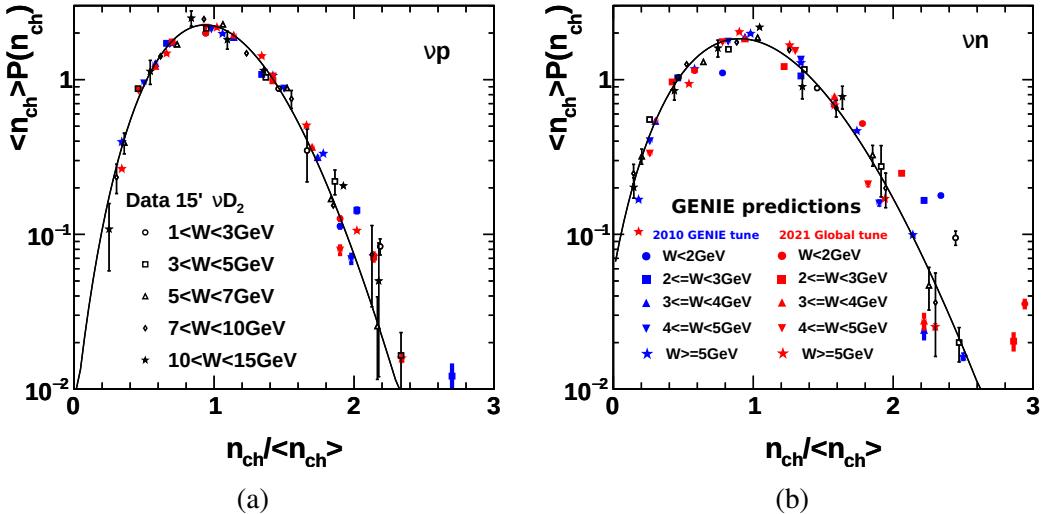


Figure 4.22: Comparison of the KNO scaling distributions for neutrino interactions on deuterium against the predictions for 2010 GENIE tune (blue) and the 2021 GENIE global tune (red). The solid line is the best fit result of the Levi function to FNAL 15 ft bubble chamber data [136]. The W range used for each data and predicted point is specified in the legend of Fig. 4.22(a) and 4.22 (b). (a) $\nu_\mu p$ KNO scaling distribution. (b) $\nu_\mu n$ KNO scaling distribution.

4.8 Conclusions

In this paper, we present the first GENIE tune of the AGKY model [37, 219] which was possible thanks to the Professor framework [129]. The analysis goal was to improve the GENIE agreement with neutrino charged averaged multiplicity data and to provide with the first data-driven constraints on hadronization parameters. Specifically, we constrained parameters of both low- W empirical model and PYTHIA using data from the BEBC and FNAL 15 ft bubble chamber experiments filled with hydrogen and deuterium.

Tensions between hydrogen and deuterium data were observed and two separate tunes were performed: a global and a deuterium-only tune. In particular, the 2021 AGKY global tune prediction underpredicts the deuterium data at the PYTHIA region whereas the deuterium-only tune over-predicts the hydrogen data.

Further investigations on hadronization samples are needed in order to clarify the origin of this discrepancy. A possible solution could come from more recent neutrino experiments that released data on neutrino-induced hadronization. This is the case of NOMAD [263, 264] for ν_μ on mainly carbon target, CHORUS for ν_μ and $\bar{\nu}_\mu$ on Fuji ET-7B emulsion [231, 265], OPERA for ν_μ on lead [230, 266] and MicroBooNE for ν_μ on argon [267]. But of course these samples include nuclear effects and therefore are not in the scope of this work.

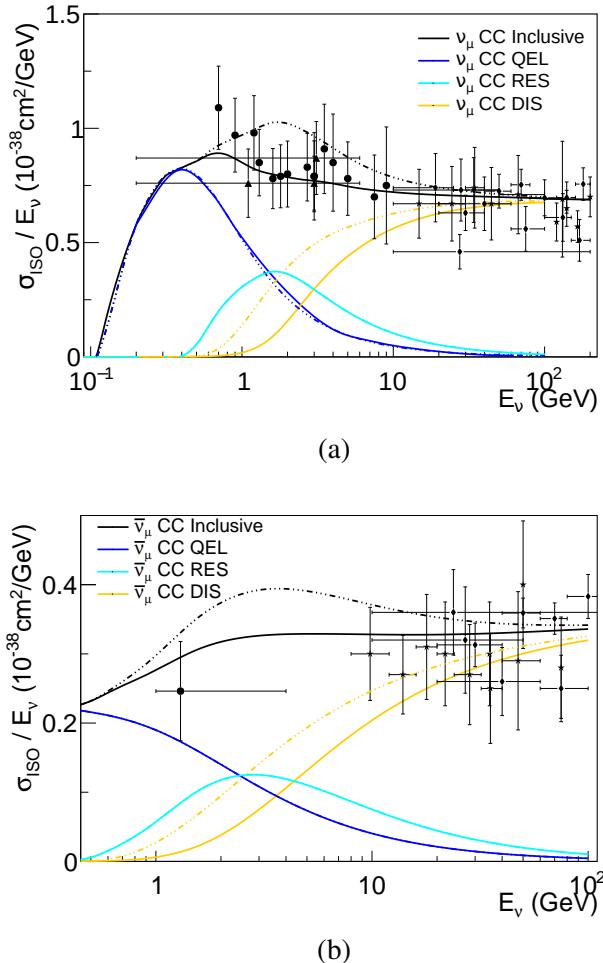


Figure 4.23: Comparison of the ν_μ and $\bar{\nu}_\mu$ CC inclusive cross section on free nucleon for the 2010 GENIE AGKY tune [217] (continuous lines) and the 2021 GENIE global tune (dashed lines) against hydrogen and deuterium data from ANL 12 ft (triangle), BNL 7 ft (filled circle), BEBC (diamond) and FNAL (filled star). The breakdown of the CC QEL, CC RES and CC DIS contributions is shown for before and after the 2021 GENIE AGKY global tune. (a) ν_μ CC inclusive cross section. (b) $\bar{\nu}_\mu$ CC inclusive cross section.

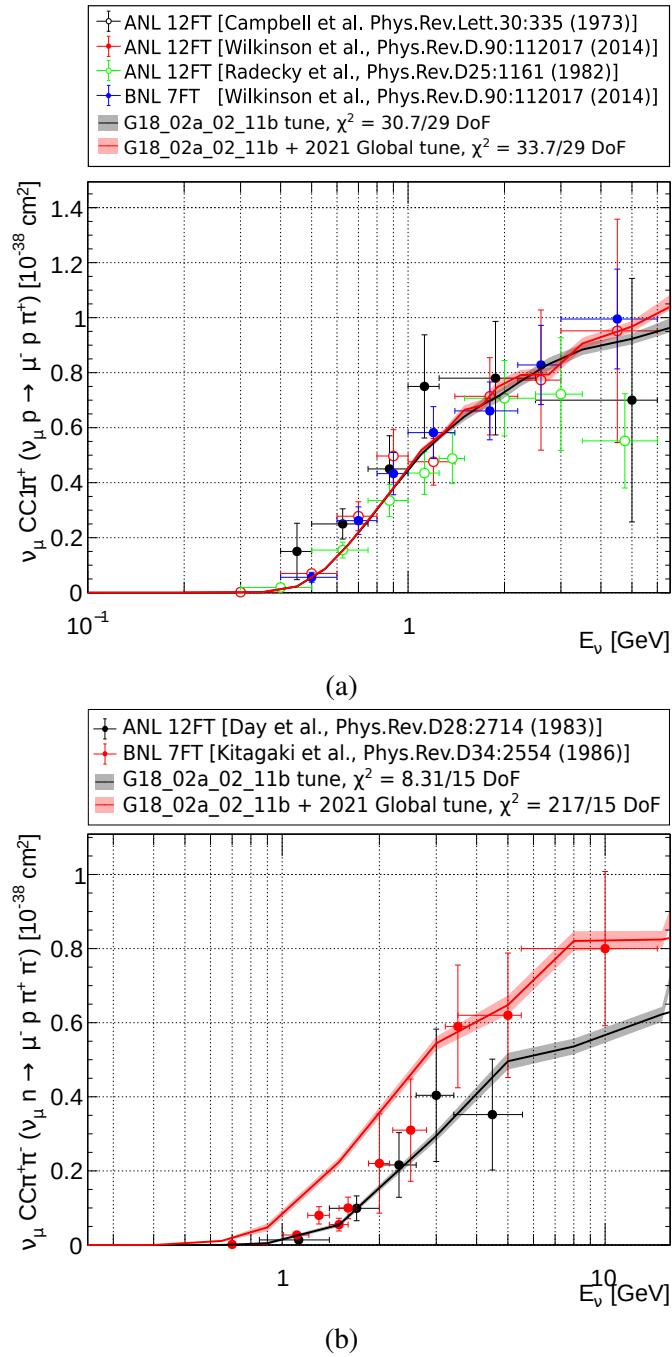


Figure 4.24: Comparison of ν_μ CC exclusive cross section data on free nucleon for the 2010 GENIE AGKY tune [217] (black) and the 2021 GENIE global tune (red) against ANL 12 ft and BNL 12 ft data. (a) One-pion production. (b) Two-pion production.

	α_{vp}	α_{vn}	$\alpha_{\bar{v}p}$	$\alpha_{\bar{v}n}$	β_{vp}	β_{vn}	$\beta_{\bar{v}p}$	$\beta_{\bar{v}n}$	P_{ss}	$\langle p_{\perp}^2 \rangle$	E_{CutOff}	Lund <i>a</i>	Lund <i>b</i>
α_{vp}	1.8E-1	-2.2E-4	1.5E-3	-2.8E-2	-1.1E-1	-1.9E-2	9.2E-4	-1.5E-2	-1.0E-3	-5.8E-4	2.7E-3	5.6E-2	3.4E-2
α_{vn}	-2.2E-4	5.4E-3	-3.0E-5	-1.3E-3	-1.4E-3	-2.9E-3	1.1E-3	1.8E-3	1.4E-4	0.0	-3.0E-5	7.2E-4	2.4E-3
$\alpha_{\bar{v}p}$	1.5E-3	-3.0E-5	6.2E-3	5.7E-4	2.2E-3	7.7E-4	-2.9E-3	1.4E-3	-8.0E-5	1.8E-4	-6.0E-5	-3.0E-5	-3.7E-3
$\alpha_{\bar{v}n}$	-2.8E-2	-1.3E-3	5.7E-4	1.0E-1	4.3E-2	2.9E-3	9.5E-3	-6.1E-2	2.4E-3	7.1E-3	-2.6E-3	1.6E-3	-5.9E-2
β_{vp}	-1.1E-1	-1.4E-3	2.2E-3	4.3E-2	1.2E-1	3.3E-2	4.7E-3	-6.0E-5	7.7E-4	1.6E-3	-3.3E-3	-3.8E-2	-4.4E-2
β_{vn}	-1.9E-2	-2.9E-3	7.7E-4	2.9E-3	3.3E-2	3.2E-2	1.1E-2	1.1E-2	4.9E-4	-4.2E-3	-1.1E-3	-1.8E-3	-7.5E-3
$\beta_{\bar{v}p}$	9.2E-4	1.1E-3	-2.9E-3	9.5E-3	4.7E-3	1.1E-2	1.3E-2	-3.0E-3	-1.6E-3	-1.3E-3	-9.9E-4	7.2E-3	7.7E-4
$\beta_{\bar{v}n}$	-1.5E-2	1.8E-3	1.4E-3	-6.1E-2	-6.0E-5	1.1E-2	-3.0E-3	5.2E-2	-2.4E-3	-6.1E-3	6.3E-4	-8.5E-3	2.9E-2
P_{ss}	-1.0E-3	1.4E-4	-8.0E-5	2.4E-3	7.7E-4	4.9E-4	-1.6E-3	-2.4E-3	2.3E-3	-4.8E-4	-7.0E-5	1.2E-4	-9.6E-4
$\langle p_{\perp}^2 \rangle$	-5.8E-4	0.0	1.8E-4	7.1E-3	1.6E-3	-4.2E-3	-1.3E-3	-6.1E-3	-4.8E-4	1.8E-3	2.7E-4	-1.4E-3	-4.2E-3
E_{CutOff}	2.7E-3	-3.0E-5	-6.0E-5	-2.6E-3	-3.3E-3	-1.1E-3	-9.9E-4	6.3E-4	-7.0E-5	2.7E-4	2.3E-3	-7.0E-4	9.2E-4
Lund <i>a</i>	5.6E-2	7.2E-4	-3.0E-5	1.6E-3	-3.8E-2	-1.8E-3	7.2E-3	-8.5E-3	1.2E-4	-1.4E-3	-7.0E-4	2.5E-2	1.0E-2
Lund <i>b</i>	3.4E-2	2.4E-3	-3.7E-3	-5.9E-2	-4.4E-2	-7.5E-3	7.7E-4	2.9E-2	-9.6E-4	-4.2E-3	9.2E-4	1.0E-2	5.0E-2

Table 4.11: Parameter covariance matrix extracted from the 2021 GENIE AGKY global tune.

	α_{vp}	α_{vn}	$\alpha_{\bar{v}p}$	$\alpha_{\bar{v}n}$	β_{vp}	β_{vn}	$\beta_{\bar{v}p}$	$\beta_{\bar{v}n}$	P_{ss}	$\langle p_{\perp}^2 \rangle$	E_{CutOff}	Lund <i>a</i>	Lund <i>b</i>
α_{vp}	7.7E-2	2.5E-2	8.6E-3	-6.3E-3	-3.6E-2	-1.6E-2	-8.7E-3	1.0E-2	-1.1E-3	-2.9E-3	-1.7E-3	3.8E-4	8.8E-3
α_{vn}	2.5E-2	1.5E-2	5.2E-3	-3.1E-3	-1.3E-2	-5.8E-3	-4.8E-3	4.4E-3	1.4E-3	-1.6E-3	-7.8E-4	7.9E-4	7.2E-3
$\alpha_{\bar{v}p}$	8.6E-3	5.2E-3	2.4E-2	-6.9E-3	3.0E-4	-8.1E-3	-1.8E-2	-1.9E-3	3.4E-3	1.5E-3	-1.2E-4	-2.9E-3	-1.2E-3
$\alpha_{\bar{v}n}$	-6.3E-3	-3.1E-3	-6.9E-3	8.5E-3	1.1E-3	5.5E-3	4.1E-3	-3.8E-3	-7.1E-4	3.3E-4	-5.0E-4	-7.1E-4	2.8E-3
β_{vp}	-3.6E-2	-1.3E-2	3.0E-4	1.1E-3	2.2E-2	7.4E-3	7.2E-4	-4.5E-3	-2.0E-5	1.0E-3	1.9E-3	-4.6E-3	-6.2E-3
β_{vn}	-1.6E-2	-5.8E-3	-8.1E-3	5.5E-3	7.4E-3	1.2E-2	7.2E-3	-2.1E-3	-1.1E-3	-6.0E-5	-1.5E-3	3.3E-3	-1.8E-3
$\beta_{\bar{v}p}$	-8.7E-3	-4.8E-3	-1.8E-2	4.1E-3	7.2E-4	7.2E-3	1.6E-2	1.7E-3	-2.8E-3	-8.0E-4	-4.2E-4	2.6E-3	-4.9E-4
$\beta_{\bar{v}n}$	1.0E-2	4.4E-3	-1.9E-3	-3.8E-3	-4.5E-3	-2.1E-3	1.7E-3	7.1E-3	-4.3E-4	-6.6E-4	1.8E-4	4.4E-3	-6.1E-4
P_{ss}	-1.1E-3	1.4E-3	3.4E-3	-7.1E-4	-2.0E-5	-1.1E-3	-2.8E-3	-4.3E-4	1.4E-3	1.6E-4	-1.9E-4	6.3E-4	5.3E-4
$\langle p_{\perp}^2 \rangle$	-2.9E-3	-1.6E-3	1.5E-3	3.3E-4	1.0E-3	-6.0E-5	-8.0E-4	-6.6E-4	1.6E-4	9.9E-4	-2.8E-4	1.3E-3	-1.3E-3
E_{CutOff}	-1.7E-3	-7.8E-4	-1.2E-4	-5.0E-4	1.9E-3	-1.5E-3	-4.2E-4	1.8E-4	-1.9E-4	-2.8E-4	1.5E-3	-2.9E-3	-2.0E-5
Lund <i>a</i>	3.8E-4	7.9E-4	-2.9E-3	-7.1E-4	-4.6E-3	3.3E-3	2.6E-3	4.4E-3	6.3E-4	1.3E-3	-2.9E-3	1.5E-2	-2.5E-3
Lund <i>b</i>	8.8E-3	7.2E-3	-1.2E-3	2.8E-3	-6.2E-3	-1.8E-3	-4.9E-4	-6.1E-4	5.3E-4	-1.3E-3	-2.0E-5	-2.5E-3	8.5E-3

Table 4.12: Parameter covariance matrix extracted from the 2021 GENIE AGKY ${}^2\text{H}$ tune.

Despite the tensions, the global tune shows a better agreement with the charged averaged multiplicity data and provides the first data driven analysis of this kind using neutrino interactions. This statistical analysis can be a useful input for proper systematic studies of modern neutrino experiments. The main effect of the tune is the increase of the averaged charged multiplicity for $W^2 > 10 \text{ GeV}^2/c^4$, modeled with PYTHIA. The low- W region is also affected but constraints due to energy, momentum, charge, baryon number and strangeness conservation laws reduce the available phase space and the effect of the tuning procedure.

The effect of the 2021 GENIE AGKY global tune at the SIS region is an increase on the two pion production cross section, which affects the current agreement with CC inclusive data [217]. Therefore, we conclude that this tune is more appropriate at higher energies where the contribution of the SIS region is not relevant. The information on the systematic uncertainties coming from the low- W AGKY parameters is still valuable for neutrino experiments interested in the $W < 2 \text{ GeV}/c^2$ region. A joint tune of the SIS region and hadronization datasets would address this disagreement.

Chapter 5

Neutrino-nuclei CC0 π cross section tuning in GENIE v3

Abstract

This article¹ summarizes the state of the art of ν_μ and $\bar{\nu}_\mu$ CC0 π cross-section measurements on carbon and argon and discusses the parametrization of relevant modeling uncertainties in GENIE v3. The CC0 π event topology is common in experiments at a few-GeV energy range. Although its main contribution comes from quasielastic interactions, this topology is still not well understood. The GENIE global analysis framework is exploited to analyze CC0 π data from MiniBooNE, T2K and MINERvA. A partial tune for each experiment is performed, providing a common base for the discussion of tensions between datasets. The results offer an improved description of nuclear CC0 π data as well as data-driven uncertainties for each experiment. This work is a step towards a GENIE global tune that improves our understanding of neutrino interactions on nuclei. It follows from earlier GENIE work on the analysis of neutrino scattering data on hydrogen and deuterium.

5.1 Introduction

A major experimental program aims to measure neutrino-nucleus interactions at the few-GeV region. MiniBooNE was the first neutrino experiment to provide a double-differential flux-integrated CC0 π cross-section measurement with high statistics on

¹This chapter is written as a paper. We aim to publish it in Phys. Rev. D after submission.

carbon [98]. Since then T2K [108], MicroBooNE [110] and MINERvA [113] have produced a large body of measurements on different nucleus, such as carbon or argon. However, a detailed quantitative understanding of neutrino-nucleus interactions is still missing.

In order to avoid biases in cross-section measurements due to theory assumptions, neutrino experiments focus on the study of specific typologies instead of interaction processes, such as quasielastic (QEL). The most dominant event topology below the 1 GeV region is CC 0π . This topology is usually defined as an event with a muon and no pions in the final state. As a consequence of the nuclear medium, different interaction processes contribute to the CC 0π measurement. Neutrino charged-current (CC) QEL interactions are the dominant contribution to this topology at the few-GeV energy range. Two-particles-two-holes (2p2h) have been shown to be crucial for the correct description of the data. The shallow-inelastic scattering (SIS) region is also relevant. For instance, pions produced after a CC resonant (RES) interaction can be absorbed due to final-state interactions (FSI). Similarly, deep-inelastic scattering (DIS) events can also contribute. In GENIE, SIS corresponds to the cross-section contribution from the nonresonant backgrounds in the RES region. Fig. 5.1 summarizes the interaction processes and typologies of interest at the few GeV region as a function of the neutrino energy. In addition, the flux predictions used for the cross-section measurements of MiniBooNE, MicroBooNE, T2K ND280 and MINERvA are also provided.

The GENIE Collaboration is building a global analysis of neutrino, charged-lepton and hadron-scattering data. The analysis is constructed in a staged manner: the early efforts focused on the analysis of neutrino data on hydrogen and deuterium for the purpose of tuning nonresonant backgrounds in the SIS region [217], as well as tuning of hadronic multiplicities relevant for neutrino-induced hadronization models [269]. However, no explicit tune to recent nuclear data has been performed by GENIE.

There are outstanding discrepancies between GENIE predictions and more recent datasets, which use heavy nuclei as a target. Several neutrino collaborations, such as MicroBooNE and MINERvA, tried to address these discrepancies by tuning GENIE against ν_μ CC 0π T2K data and inclusive ν_μ CC MINERvA data respectively [270–272]. All these tunes simulate 2p2h interactions with the Valencia model [70]. In both cases, the results suggest an enhancement of the 2p2h cross section. These tunes are not available for wider use within GENIE, and in some cases, these were performed with obsolete GENIE versions which differ substantially from the latest one.

In this paper, we describe the GENIE analysis of the available ν_μ and $\bar{\nu}_\mu$ CC 0π data from MicroBooNE, T2K, MINERvA and MicroBooNE. The main goal is to provide improved simulations tuned to nuclear data and quantify the major sources of

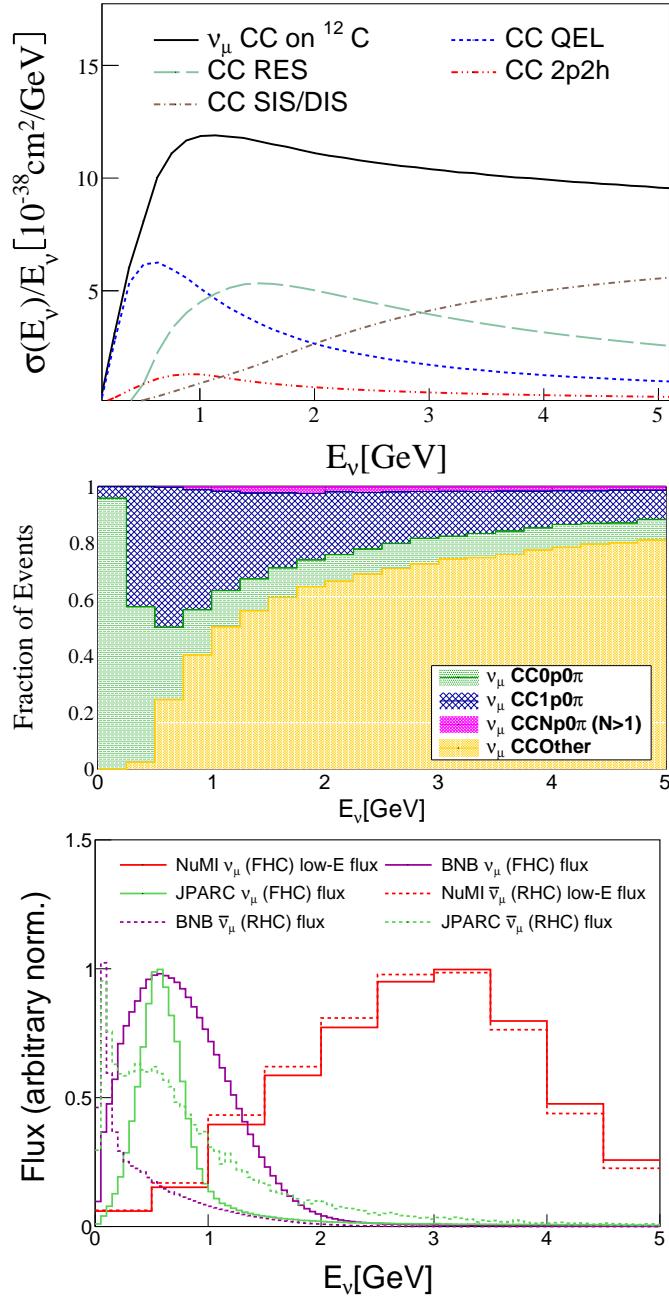


Figure 5.1: (Top) Summary of ν_μ ^{12}C CC cross section as a function of the neutrino energy. (Middle) Fraction of ν_μ ^{12}C typologies a function of the neutrino energy. This plot considers a momentum threshold for protons of 450 MeV/c. The GE-NIE predictions are obtained with the G18_10a_02_11b tune. (Bottom) Summary of ν_μ (continuous lines) and $\bar{\nu}_\mu$ (dashed lines) normalized flux distributions for T2K ND280 JPARC flux [78], MiniBooNE and MicroBooNE BNB flux [100] and MINERvA NuMI flux [268] detectors. The flux predictions for neutrino and antineutrino modes are referred to as "FHC" and "RHC" respectively.

uncertainties in CC 0π measurements. In order to do so, new degrees of freedom are developed within the GENIE Monte Carlo (MC) event generator for the variation of the nominal models. Most of the new degrees of freedom can be used to tune other available comprehensive model configuration (CMC) in GENIE. In this analysis we focus on the ‘retuning’ of the G18_10a_02_11b tune against ν_μ - ^{12}C CC 0π data from MicroBooNE, T2K and MINERvA. The G18_10a_02_11b was previously tuned against free nucleon data [217]. In this paper, we refer to the G18_10a_02_11b tune as *nominal*.

G18_10a_02_11b is one of the many CMC that can be tuned within GENIE. The main motivations behind this choice are: (1) we can use data-driven constraints from previous GENIE tunes on hydrogen and deuterium [217]. (2) QEL and 2p2h is modeled with the Valencia model. This is a theory-based model which is used in most neutrino analyses. (3) FSI interactions are modeled with the INTRANUKE hA model. This is an empirical model which is driven by hadron-nucleus data and it is easy to tune. Other CMC will be considered in future iterations of this work.

The GENIE global analysis software is used to perform a partial tune for each experiment using double-differential flux-integrated CC 0π cross-section measurements as a function of muon kinematics. Only carbon datasets are considered for this work. This provides a common ground for the discussion of tensions. The analysis software is build on top of the GENIE Comparisons and Tuning products, which is interfaced to Professor [129]. This work is a step closer to a global tune with neutrino-nucleus cross-section data, which can be performed using the same analysis strategy once all the tensions are well understood. Future iterations of this work will also incorporate measurements on different typologies, such as CC 1π .

This work is organized as follows: Sec. 5.2 provides an overview of the available CC 0π data to date. The newly developed GENIE parameters are discussed in Sec. 5.3. This is followed by a description of the tuning procedure in Sec. 5.4 and a discussion of the tune results, Sec. 5.5, and tensions between CC 0π and CCNp 0π datasets, in Sec. 5.6. The main conclusions of this paper are highlighted in Sec. 5.7.

5.2 Review of neutrino nucleus CC 0π measurements

Cross-section measurements of the CC 0π topology were carried out by Sci-BooNE [102], NOMAD [104], MiniBooNE [99, 273], T2K ND280 [119], MINERvA [117, 126, 274] and MicroBooNE [122, 209].

The definition of CC 0π topology varies between the different published measurements as a consequence of the different detection capabilities of each experiment. In some analyses, the topology definition is optimized to study more exclusive final states

with a specific proton multiplicity. The following nomenclature is adopted to avoid confusion for the reader: analyses that require at least a proton in the final state are referred to as CCNp 0π . If the analysis requires only one proton in the final state, N is substituted by the number of protons, i.e. CC0p 0π for events with no visible protons. In some cases, the topology definition requires at least two protons in the final state. This is denoted as CC2p 0π . When there is no requirement on the number of protons, the topology is defined as CC 0π .

This section offers with comparisons of GENIE against all CC 0π and CCNp 0π data available from MiniBooNE, T2K, MINERvA and MicroBooNE. The corresponding GENIE predictions are obtained by replicating the analysis within GENIE: neutrino interaction events are simulated for each experiment given the neutrino flux, target material, and analysis cuts. With this information, the GENIE prediction for the corresponding differential flux-integrated cross section is evaluated. Tab. 5.1 summarizes the information used in the evaluation of the GENIE predictions: the target type, neutrino flux mean energy and event topology definition. The neutrino flux spectrum associated with each experiment is provided in Fig. 5.1 (bottom) [78, 100, 268]. Both MiniBooNE and MicroBooNE use the same neutrino flux, hence the second is not specified in Fig. 5.1. The *kinematic quantity* column in Tab. 5.1 lists the observed kinematic quantities used to extract the cross-section measurements. The definition of each kinematic quantity is given in Sec. 5.2.1. Some of the available measurements are double-differential or triple-differential ones. This is indicated by a comma separated list for the kinematic quantities used in the data analysis. In addition, the year of the data release and the number of bins (N_{Bins}) for each dataset are specified. The details on the analysis requirements for each experiments are discussed in the corresponding subsections. For completeness, Tab. 5.1 includes measurements from SciBooNE and NOMAD which are not discussed further in this paper as their analysis strategy is limited with respect to the other measurements discussed in this work.

This work focuses on the tuning of double-differential flux-integrated CC 0π and CC0p 0π cross-section measurements. Double differential flux-integrated cross-section measurements are the most adequate for tuning as they provide the most complete information to characterize the final state. In addition, such measurements are the most common across the different experiments, except for MicroBooNE. Single-differential flux-integrated CCNp 0π measurements, available from T2K, MINERvA and MicroBooNE, are really valuable as these are really sensible to nuclear effects. However, these have a reduced N_{bins} , which limits the analysis. CCNp 0π measurements are mainly used for the investigation of tensions. For this purpose, MINERvA CCNp 0π data is used in the analysis. T2K ND280 provides a triple-differential CC1p 0π mea-

Experiment	$\langle E_\nu \rangle$	Target	Topology	Kinematic quantity	N_{Bins}	Year	Ref.	
ν_μ -A measurements								
SciBooNE	700 MeV	^{12}C	CC0 π	E_ν^{QEL}	5	2006	[102]	✗
NOMAD	23 GeV	^{12}C	CC0 π	E_ν^{QEL}	10	2009	[104]	✗
MiniBooNE	788 MeV	^{12}C	CC0 π	$T_\mu, \cos\theta_\mu$ Q_{QEL}^2 E_ν^{QEL}	137 17 14	2010	[99]	✓
T2K ND280	600 MeV	^{12}C	CC0p0 π	$p_\mu, \cos\theta_\mu$	60	2018	[119]	✓
	600 MeV	^{12}C	$\bar{\text{CC}}1\bar{p}0\pi$	$\cos\theta_\mu, \cos\theta_p, p_p$	40	2018	[119]	✗
	600 MeV	^{12}C	$\bar{\text{CC}}2\bar{p}0\pi$	—	1	2018	[119]	✗
	600 MeV	^{12}C	$\bar{\text{CCN}}\bar{p}0\pi$	δp_T $\delta\phi_T$ $\delta\alpha_T$ $\Delta p_p, \cos\theta_\mu, p_\mu$ $ \Delta p_p , \cos\theta_\mu, p_\mu$ $\Delta\theta_p, \cos\theta_\mu, p_\mu$	8 8 8 49 49 35	2018	[119]	✗
MINERvA	3.5 GeV	^{12}C	CC0 π	p_T^μ, p_L^μ Q_{QEL}^2 E_ν^{QEL}	144 16 12	2019	[117]	✓
	3.5 GeV	^{12}C	$\bar{\text{CCN}}\bar{p}0\pi$	p_p θ_p δp_T $\delta\alpha_T$ $\delta\phi_T$	25 26 24 12 18	2018	[126]	✓
	3.5 GeV	^{12}C	$\bar{\text{CCN}}\bar{p}0\pi$	δp_{T_x} δp_{T_y}	32 33	2020	[274]	✗
	6 GeV	^{12}C	CC0 π	p_T^μ, p_L^μ Q_{QEL}^2	184 19	2020	[118]	✗
	6 GeV	^{12}C	$\bar{\text{CCN}}\bar{0}\pi$	$\bar{p}_T^\mu, \bar{p}_L^\mu, \sum \bar{T}_p$ $E_\mu, q_0^{\text{QEL}}, \sum T_p$	660 540	2022	[121]	✗
MicroBooNE	800 MeV	^{40}Ar	CC1p0 π	p_μ $\cos\theta_\mu$ p_p Q_{QEL}^2 E_ν^{cal}	7 7 7 7 7	2020	[209]	✓
	800 MeV	^{40}Ar	$\bar{\text{CCN}}\bar{p}0\pi$	p_μ^{reco} $\cos\theta_\mu^{\text{reco}}$ p_p^{reco} $\cos\theta_p^{\text{reco}}$ $\theta_{\mu p}^{\text{reco}}$	10 12 10 9 6	2020	[275]	✗
$\bar{\nu}_\mu$ -A measurements								
NOMAD	23 GeV	^{12}C	CC0 π	E_ν^{QEL}	6	2009	[104]	✗
MiniBooNE	665 MeV	^{12}C	CC0 π	$T_\mu, \cos\theta_\mu$ Q_{QEL}^2 E_ν^{QEL}	78 16 14	2013	[273]	✓
MINERvA	3.5 GeV	^{12}C	CC0p0 π	p_T^μ, p_L^μ Q_{QEL}^2 E_ν^{QEL}	60 8 10	2013	[276]	✓

Table 5.1: Summary of CC0 π analyses of ν_μ and $\bar{\nu}_\mu$ interactions on nuclei. For each analysis, information on the neutrino flux mean energy, target type and event topology is provided. The *kinematic quantity* column specifies the list of kinematic quantities used in the cross-section measurement. Integrated cross-section measurements are denoted with a —. All kinematic quantities are defined in Sec. 5.2.1. The last column specifies whether the analysis is considered in the analysis.

Simulation domain	Model
Nuclear model	Local Fermi Gas [40]
QEL and 2p2h	Nieves Simo Vacas [70]
QEL Charm	Kovalenko [139]
QEL $\Delta S = 1$	Pais [33]
RES	Rein-Sehgal [34]
SIS/DIS	Bodek-Yang [38]
DIS $\Delta S = 1$	Aivazis-Olness-Tung [140]
Coherent π production	Rein-Sehgal [34]
Hadronization	AGKY [219]
FSI	INTRANUKE hA [37]

Table 5.2: Complete list of models used for the G18_10a_02_11a tune in GENIE v3 [217].

surement as a function of proton and muon kinematics. In addition, MINERvA recently² released a triple-differential measurement as well [121]. Additional single- and triple-differential CCN0 π datasets not considered in the first iteration of this work will be included in future iterations.

In this paper, all predictions are calculated using the G18_10a_02_11b tune. The G18_10a_02_11b uses the Nieves Simo Vacas or Valencia model to simulate QEL and 2p2h events in the nuclear medium. FSI are modeled using the *hA* model. The nuclear ground state is described with the Local Fermi gas (LFG) model [40]. The other interaction processes are common with the free-nucleon recipe described in Ref. [217]. Tab. 5.2 details the full list of interaction processes associated to this CMC. This configuration was previously tuned against hydrogen and deuterium data [217].

The format of all the comparisons with data reported in this section is common: the data and differential cross-section prediction are represented in black. In addition, the contribution from different interaction models is shown for CCRES, CC2p2h and CCDIS/SIS. The contribution to the G18_10a_02_11b predictions from CCDIS/SIS events is really small at the neutrino energies considered in this work. For this reason, the contribution is grouped into a single category (DIS). The 2p2h contribution is divided further into four categories that depend on the event invariant mass, W . The W regions are:

- $W < M_N = 938 \text{ MeV}/c^2$
- $M_N < W < W_{\text{Dip}} = \text{MeV}/c^2$
- $W_{\text{Dip}} < W < M_\Delta = 1232 \text{ MeV}/c^2$.
- $W > M_\Delta$

²This dataset was released while this paper was getting finalized.

The data error bars include statistical and systematic uncertainties. The errors on the x-axis represent the bin width used in the original analysis.

5.2.1 Kinematic quantities of interest for CC0 π measurements

Differential neutrino cross-section measurements are given as a function of different kinematic quantities. In this paper, these kinematic quantities are classified into *direct*, *inferred with an underlying process hypothesis* or *inferred without an underlying process hypothesis*.

Direct

Kinematic quantities that can be measured by the detector are classified as *direct*. For instance, an example of *direct* quantity would be the muon momentum, p_μ , or angle with respect to the beam-line axis, θ_μ . In some cases, the muon kinetic energy, T_μ , is used instead. All cross-section measurements specified in Tab. 5.1 released data as a function of the muon kinematics. Depending on the detector capabilities, *direct* quantities can also be related to proton kinematics. We refer to proton momentum and angle as p_p and θ_p respectively.

Muon *direct* observables depend strongly on the neutrino energy and these are less sensitive to nuclear effects. This motivated the recent efforts on the study of more exclusive typologies that allow measurement of the cross section as a function of the outgoing-proton kinematics [119, 122, 126, 209]. These depend weakly on the neutrino energy and are significantly altered by nuclear effects [277].

Inferred kinematic with an underlying process hypothesis

This category includes measurements that rely on the reconstruction of neutrino properties assuming a specific interaction type. For instance, the kinematics of a CC0 π event can be reconstructed under the hypothesis that the initial nucleon was at rest and that there is no inelastic production of mesons in the final state (QEL hypothesis). Under this hypothesis, the reconstructed neutrino energy (E_{ν}^{QEL}) and squared four-momentum transferred (Q_{QEL}^2) are:

$$E_{\nu}^{QEL} = \frac{M_f^2 - (M_i - E_b)^2 - M_\mu^2 + 2(M_i - E_b)E_\mu}{2(M_i - E_b - E_\mu + p_\mu \cos \theta_\mu)} \quad (5.1)$$

$$Q_{QEL}^2 = 2E_{\nu,QEL}(E_\mu - p_\mu \cos \theta_\mu) - M_\mu^2 \quad (5.2)$$

where M_i (M_f) is the initial (final) nucleon mass³, M_μ is the muon mass, and E_μ is the muon energy. The binding energy, E_B , depends on the target type. Its specific value is provided in each analysis.

The main disadvantage of using these quantities is that the underlying hypothesis is uncertain. The presence of the nuclear environment complicates the characterization of event typologies: no single event topology is produced only by a single underlying process. This is highlighted by the T2K ND280 results on inferred kinematics [119]. In their analysis, they reconstruct the energy and momentum of the outgoing proton assuming a QEL interaction. Instead of presenting the cross-section measurement as a function of the *inferred with kinematic with an underlying process* quantities, they used the difference between the *direct* and the *inferred* one. In particular, T2K ND280 explored this quantity for the proton kinematics:

$$\Delta p_p \equiv |\mathbf{p}_p^{direct}| - |\mathbf{p}_p^{QEL}| \quad (5.3)$$

$$\Delta\theta_p \equiv |\theta_p^{direct}| - |\theta_p^{QEL}| \quad (5.4)$$

$$|\Delta\mathbf{p}_p| \equiv |\mathbf{p}_p^{direct} - \mathbf{p}_p^{QEL}| \quad (5.5)$$

These are referred to as *proton inferred kinematics* quantities. Here, the superscript indicates whether the kinematic quantity is *direct* (i.e \mathbf{p}_p^{direct}) or *inferred with an underlying process hypothesis* (i.e. \mathbf{p}_p^{QEL}). The reconstructed proton energy and momentum under the QEL hypothesis are:

$$E_p^{QEL} = E_v^{QEL} - E_\mu + m_p \quad (5.6)$$

$$\mathbf{p}_p^{QEL} = (-p_\mu^x, -p_\mu^y, -p_\mu^z + E_v^{QEL}) \quad (5.7)$$

These observables can be used to highlight nuclear effects in CC0 π measurements, as the quantities defined in Eq. 5.5 deviate from zero when nuclear effects are present.

Inferred without an underlying process hypothesis observables

This category includes those kinematical quantities which are inferred from *direct* ones but do not assume a specific underlying interaction process. An example of interest for this work is the single-transverse kinematic invariance (STKI) variables [277]. STKI provide direct constraints on nuclear effects that, in some cases, have a weak dependence on the neutrino energy. STKI quantities are inferred from the muon and primary state hadron kinematics and only detectors capable of measuring low energy hadrons

³For instance, for the neutrino analysis, $M_i = M_n$ and $M_f = M_p$, whereas $M_i = M_p$ and $M_f = M_n$ for the antineutrino case.

can provide such information. So far, just the T2K ND280 and MINERvA released single-differential flux-integrated cross-section measurements as a function of these quantities [119, 126, 274].

The *transverse momentum imbalance*, $\delta\mathbf{p}_T$, is defined as the sum of the transverse muon and proton momentum:

$$\delta\mathbf{p}_T \equiv \mathbf{p}_T^\mu + \mathbf{p}_T^p . \quad (5.8)$$

As the neutrino travels in the longitudinal direction, the transverse muon momentum is related to the transverse momentum transferred as $\mathbf{p}_T^\mu = -\mathbf{q}_T$. The angle between $\delta\mathbf{p}_T$ and $-\mathbf{p}_T^\mu$ is known as *boosting angle*, $\delta\alpha_T$:

$$\delta\alpha_T \equiv \arccos \frac{-\mathbf{p}_T^\mu \cdot \delta\mathbf{p}_T}{p_T^\mu \delta p_T} . \quad (5.9)$$

The deflection of the nucleon with respect to \mathbf{q}_T is measured with the $\delta\phi_T$ angle:

$$\delta\phi_T \equiv \arccos \frac{-\mathbf{p}_T^\mu \cdot \mathbf{p}_T^p}{p_T^\mu p_T^p} . \quad (5.10)$$

A more recent study investigates the CC π cross-section dependency on the muon-proton momentum imbalances parallel (δp_{Ty}) and longitudinal (δp_{Tx}) to the momentum transfer in the transverse plane [278]. These quantities are mathematically defined as:

$$\delta p_{Tx} = (\hat{\mathbf{p}}_v \times \hat{\mathbf{p}}_T^\mu) \cdot \delta\mathbf{p}_T \quad (5.11)$$

$$\delta p_{Ty} = -\hat{\mathbf{p}}_\mu \cdot \delta\mathbf{p}_T \quad (5.12)$$

given the Cartesian coordinate system defined with respect to the neutrino and muon kinematics. The neutrino direction is given by $\hat{\mathbf{p}}_v$. All these quantities define what experiments refer to as STKI variables.

A graphical representation of the definition of the STKI variables for a neutrino interaction with and without nuclear effects is shown in Fig. 5.3. When the interaction occurs with a static free nucleon, i.e. no nuclear effects, $\mathbf{p}_T^\mu = -\mathbf{p}_T^p$, $\delta\mathbf{p}_T = 0$ and $\delta\phi_T = 0$, see Fig. 5.2a. If FSI effects and nucleon correlations are neglected, $\delta\mathbf{p}$ coincides with the initial nucleon momentum \mathbf{p}_{Ni} . Moreover, $\delta\alpha_T$ is uniform due to the isotropic nature of the Fermi motion. Nuclear effects cause an imbalance between the initial neutrino momentum and $\delta\mathbf{p}_T$. This imbalance is caused by Fermi motion, nucleon correlations, non-QEL interactions and FSI. For instance, FSI effects smear the $\delta\mathbf{p}_T$ distribution and modify the shape of the $\delta\alpha_T$ distribution. In GENIE, the

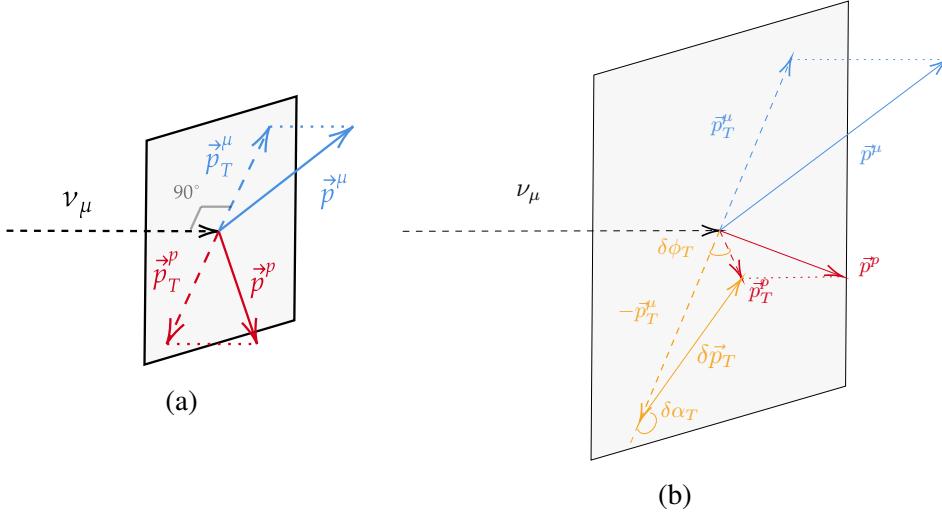


Figure 5.2: Graphical definition of the STKI variables in a ν_μ CCQEL neutrino interaction on a nuclear target. The incoming neutrino, represented as a dashed arrow, interacts with a free nucleon at rest (a) or with a bound nucleon subject to Fermi motion (b). The outgoing muon (proton) is represented in blue (red). The transverse plane is represented in grey. The incoming neutrino is perpendicular to the transverse plane. Nuclear effects distort the free nucleon picture (a) creating an imbalance between the muon and nucleon transverse momentum (b). The STKI variables that define this imbalance are highlighted in orange.

hA FSI model enhances the cross section at $\delta p_T > 0.2 \text{ GeV}/c$ and $\delta \alpha_T \sim 180^\circ$; see Fig. 5.3. This region is referred to as high-transverse kinematic imbalance region.

Ref. [277] demonstrated that the δp_T and $\delta \alpha_T$ dependence on the neutrino energy is smaller than possible uncertainties due to FSI modeling. The $\delta \phi_T$ variable has a stronger dependence on the neutrino energy as it scales with $\delta p_T/p^\mu$: at higher neutrino energies, the distribution at small angles becomes narrower. The dependency of the STKI variables in GENIE with the neutrino energy is shown in Fig. 5.4. Changes in the neutrino energy affect mostly the tail of the δp_T distribution and the $\delta \alpha_T$ distribution at backward angles.

5.2.2 MiniBooNE CC π cross-section measurement

The MiniBooNE experiment studies neutrinos produced with the Booster Neutrino Beam (BNB) [100]. MiniBooNE provided the first-high statistics ν_μ and $\bar{\nu}_\mu$ CC π double-differential flux-integrated cross-section measurement on carbon, at $\langle E_\nu \rangle \sim 800 \text{ MeV}$ and $\langle E_{\bar{\nu}} \rangle \sim 500 \text{ MeV}$ respectively [99, 273]. Flux-unfolded total cross section, $\sigma(E_\nu^{QEL})$, and single-differential flux-integrated cross section as a function of the squared four-momentum transferred, $d\sigma/dQ_{QEL}^2$, were also reported.

Both MiniBooNE analyses study CC π events with a muon in the final state and

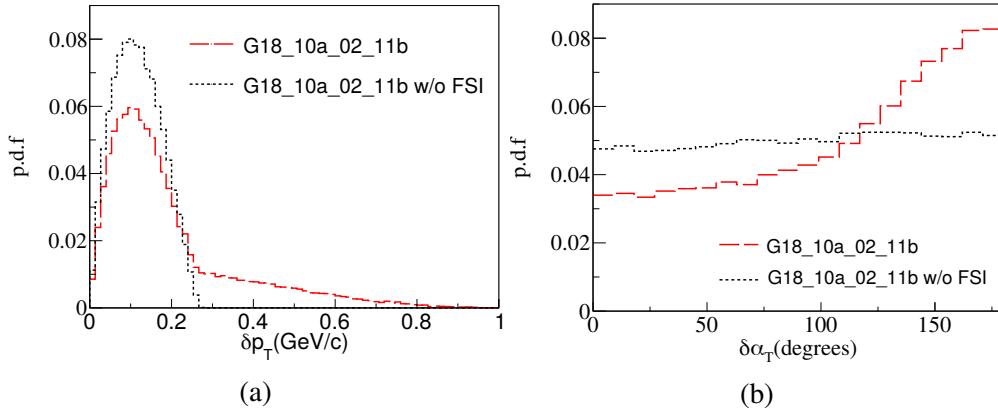


Figure 5.3: Probability density function of δp_T and $\delta \alpha_T$ for the G18_10a_02_11b tune with (a) and without (b) FSI. Both predictions are obtained simulating ν_μ CCQEL interactions only on ^{12}C at 1 GeV with the G18_10a_02_11b tune. The GENIE version used in this plots is the GENIE v3.00.06.

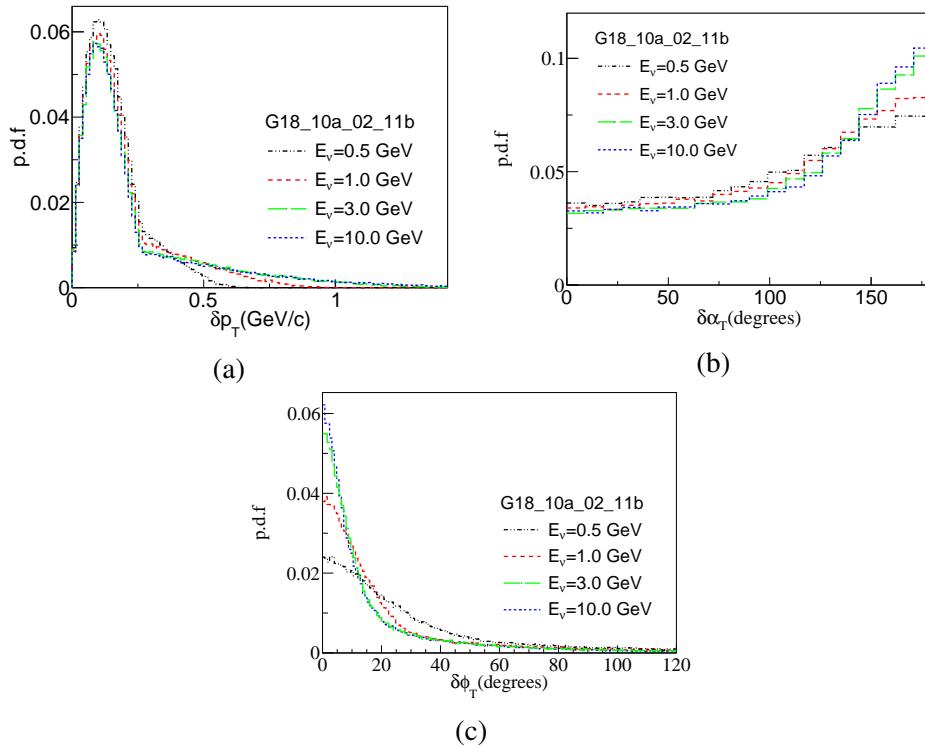


Figure 5.4: Probability density function of STKI variables. The predictions are obtained simulating ν_μ CCQEL interactions only on ^{12}C with the G18_10a_02_11b tune at different neutrino energies. The plots use the GENIE v3.00.06 version.

no pions. The signal topology of a muon in the detector is described in two sub-events: the first one associated with the primary Cherenkov light from the muon, and the second one, produced by the Cherenkov light from the Michel electron, which is produced in the muon decay. The MiniBooNE’s Michel electron detection efficiency is $\sim 83\%$ [205]. This requirement provides a sample of mostly CC events, as neutral-current events only have one sub-event.

Positively charged pions produced in the detector leave a distinct signature in the detector, as the π^+ decays immediately into a muon and a muon neutrino. The Cherenkov light from the π^+ contributes to the total light of the primary muon. This process can be distinguished from a CCQEL interaction as the muon produced from the pion decay will also decay into a Michel electron (three sub-events). Negatively charged pions are absorbed by the nuclear environment and contribute to the CC0 π topology. In the GENIE predictions, pion production events are removed by requiring no pions in the final state.

Recoil protons also emit scintillation light. However, such scintillation light signal produced is either indistinguishable from the muon signal or its momentum below the Cherenkov threshold⁴. For this reason, no requirements based on the recoil proton are considered in the MiniBooNE analyses.

The analysis considers further model-dependent cuts to correct for backgrounds and extract the CCQEL cross-section from the CC0 π sample. In the original publication, these are referred to as *irreducible backgrounds*. An example of irreducible background is CC1 π events that were not removed by the cut on the pion subevent topology or pion production events in which the pion is absorbed. This is corrected using a MC simulation tuned to v_μ CC1 π MiniBooNE data. Information on v_μ CC1 π^+ sample is used to characterize this background and correct for single-pion events which were not removed by the CC0 π selection criteria in the neutrino and antineutrino analyses. This procedure is one of the main limitations of this dataset as it incorporates strong biases in the reported measurement. The contribution to the cross-section measurement from irreducible backgrounds is also reported, allowing the comparison against CC0 π data.

The quality of the MiniBooNE CC0 π data release is poor in comparison with the rest. The MiniBooNE collaboration provided measurements in bins of T_μ and $\cos \theta_\mu$, but did not provide the bin-to-bin covariances for either of the two measurements. Instead, they quoted a normalization systematic uncertainty of $\sim 10.7\%$. As suggested by Ref. [99], this error is added as a systematic in our database, effectively including a correlation between the bins.

In Fig. 5.5 and Fig.5.6, the flux-integrated double differential v_μ and \bar{v}_μ CC0 π

⁴Charged particles with $\beta > 0.68$ produce Cherenkov radiation in mineral oil.

cross section data as a function of p_μ and T_μ are compared against GENIE. The main observation is that the GENIE tune under-predicts the data. In particular, the G18_10a_02_11b disagreement with the data is more significant at backward angles, where the cross-section is determined by CCQEL events only. The disagreement is also observed at forward angles, where there is a significant contribution from non-QEL events.

5.2.3 T2K ND280 CC 0π cross-section measurements

The Tokai-to-Kamioka (T2K) experiment is an accelerator-based long-baseline experiment that studies neutrino oscillations. Neutrinos are generated at the Japan Proton Accelerator Research Complex (J-PARC) facility [78], 280 m away from the T2K near detectors [108]: INGRID and ND280. The T2K ND280 detector is used to measure neutrino interactions on carbon at $\langle E_\nu \rangle \sim 600$ MeV. Details on the detector setup can be found in Ref. [119]. The measurements described here use the detector central tracker region, composed of three time projection chambers (TPC) and two fine-grained detectors (FGD1 and FGD2). The FGDs are the target mass and are also used to track charged particles. Carbon measurements use the FGD1 as the target mass. The central region is surrounded by an electromagnetic calorimeter (ECal), which is contained within a magnet. This setup allows measuring the particle charge and momentum. This information, together with energy deposition, is used to identify charged particles.

The first double-differential flux-integrated CC 0π cross-section measurement provided by T2K ND280 was released back in 2015 [279]. This measurement is surpassed by Ref. [119], which considers improved constraints on systematic uncertainties. Ref. [119] provides additional measurements including double- and triple-differential measurements for different proton multiplicities as well as two CCNp 0π single-differential cross-section measurements as a function of STKI and *proton inferred kinematics* quantities.

All measurements from Ref. [279] require one muon and no pions in the final state, regardless of the number of nucleons in the event. Any event must contain at least one track in the TPC, which must be either a muon or a proton. If it is a proton, they look for a muon-like track in the FGD1 or ECal. Other events with tracks that are not consistent with the muon-like or proton-like signature are rejected. Events with low-momentum charged or neutral pions are removed by requiring no Michel electrons or photons. At the MC level, this is implemented by removing events with pions or photons in the final state, respectively.

The selected sample is divided further depending on the number of protons above the detection threshold of 500 MeV/c: no protons (CC0p 0π), one proton (CC1p 0π) or

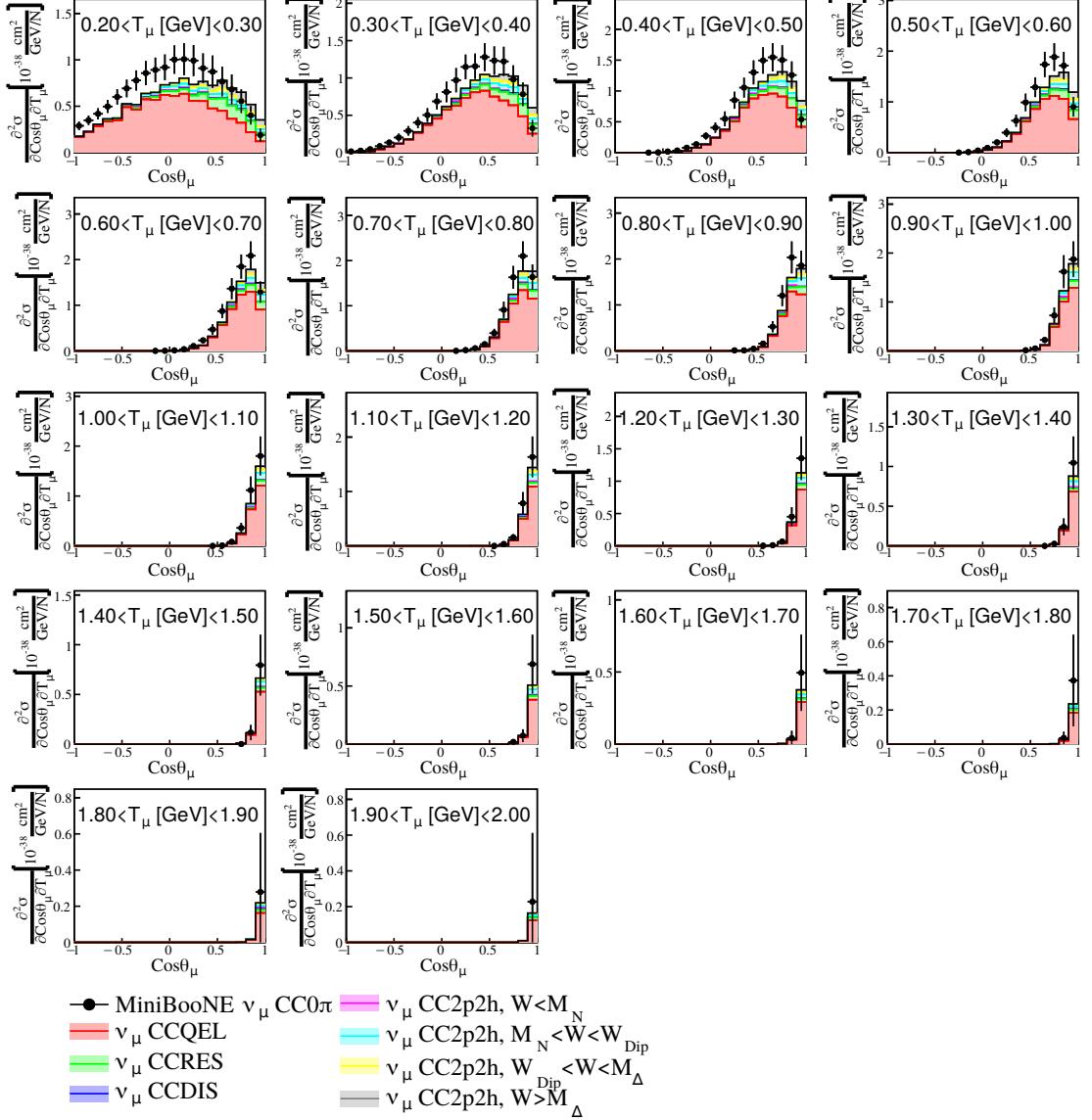


Figure 5.5: MiniBooNE ν_μ CC0 π double-differential flux-averaged cross section as a function of the muon angle (θ_μ) and kinetic energy (T_μ) [99]. The corresponding slices on T_μ are compared against the G18_10a_02_11b tune. The GENIE prediction is divided into different categories: CCQEL, CCRES, CC2p2h and CCDIS.

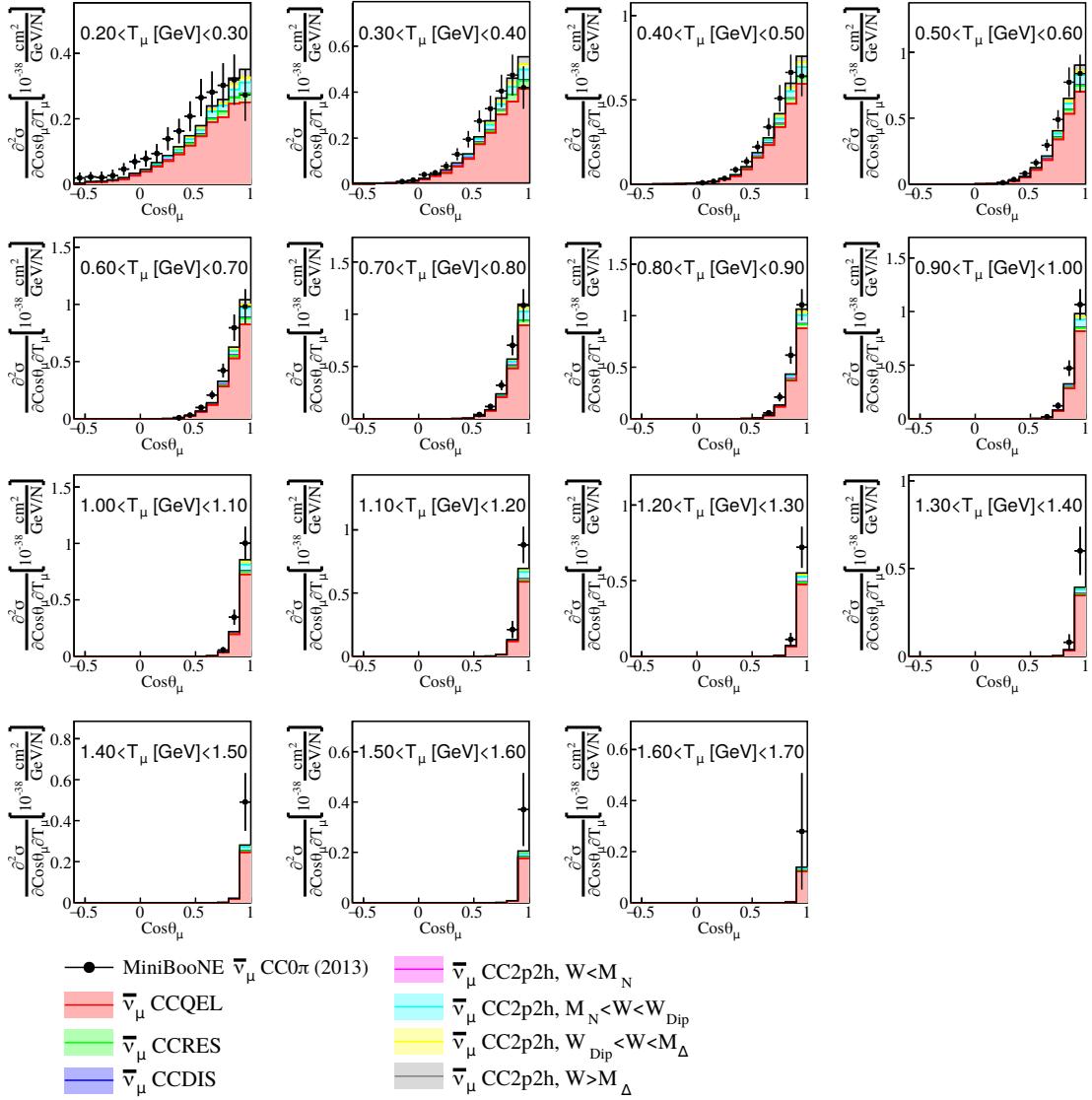


Figure 5.6: MiniBooNE $\bar{\nu}_\mu$ CC0 π double-differential flux-averaged cross section as a function of the muon angle, θ_μ , and kinetic energy, T_μ [273]. The corresponding slices on T_μ are compared against the G18_10a_02_11b tune. The GENIE prediction is divided into different categories: CCQEL, CCRES, CC2p2h and CCDIS.

5.2. REVIEW OF NEUTRINO NUCLEUS CC 0π MEASUREMENTS

T2K ND280 Analysis	p_p	$\cos \theta_p$	p_μ	$\cos \theta_\mu$
CC0p0 π	<500 MeV/c (or no proton)			
CC1p0 π	>500 MeV/c			
CCNp0 π , STKI	450 < p_p < 1000 MeV/c	>0.4	>250 MeV/c	>-0.6
CCNp0 π , proton inferred kinematics	>450 MeV/c	>0.4		

Table 5.3: Phase-space restrictions for the T2K ND280 analyses from Ref. [119]. The proton cuts are only applied to the highest energy proton.

more than one visible proton (CC2p0 π) in the final state. The CC0p0 π and CC1p0 π are double- and triple-differential cross-section measurements as a function of the muon and muon and proton kinematics respectively. The total CC2p0 π cross-section is also reported. The STKI and *proton inferred kinematics* are obtained with the CCNp0 π sample: they require the presence of at least one visible proton ($p_p > 500$ MeV/c).

Efficiency corrections for CCNp0 π events can be very model dependent. To avoid this, different kinematical restrictions are considered for each analysis, selecting regions in which the efficiency is flat or well understood. These are specified in Tab. 5.3. Events with more than one proton are reconstructed using the information from the highest energy one, which has to satisfy the kinematical limits of Tab. 5.3. The samples are not corrected for events with protons below the detection threshold or any of the kinematical cuts considered in the analysis. The same cuts are applied at the generator level when evaluating the GENIE predictions.

The GENIE comparison against the ν_μ CC0p0 π double-differential cross section are presented in Fig. 5.7. The main contribution to the CC0p0 π topology comes from CCQEL events. The second contribution is from CC2p2h events with $M_N < W < W_{\text{Dip}}$. The contribution from 2p2h events with $W < M_N$ or $W > W_{\text{Dip}}$ is negligible for the CC0p0 π measurement. GENIE is under-predicting the data at backward angles.

This disagreement in the overall normalization is also observed in Fig. 5.8, which compares GENIE against the cross section as a function of the proton multiplicity. This observation conflicts with ν_μ CC1p0 π data, which is not under-predicted. There are some outstanding differences between the GENIE predictions for the CC0p0 π and CC1p0 π topologies. Whilst the total contribution from 2p2h events is similar, the main 2p2h contribution comes from 2p2h events with $W > W_{\text{Dip}}$. In addition, the fraction from RES events is higher with respect to the CC0p0 π one.

Fig. 5.9 provides comparisons against ν_μ CCNp0 π data as a function of the STKI variables, concluding that non-QEL interactions are essential to describe this data within regions of high transverse kinematic imbalance.

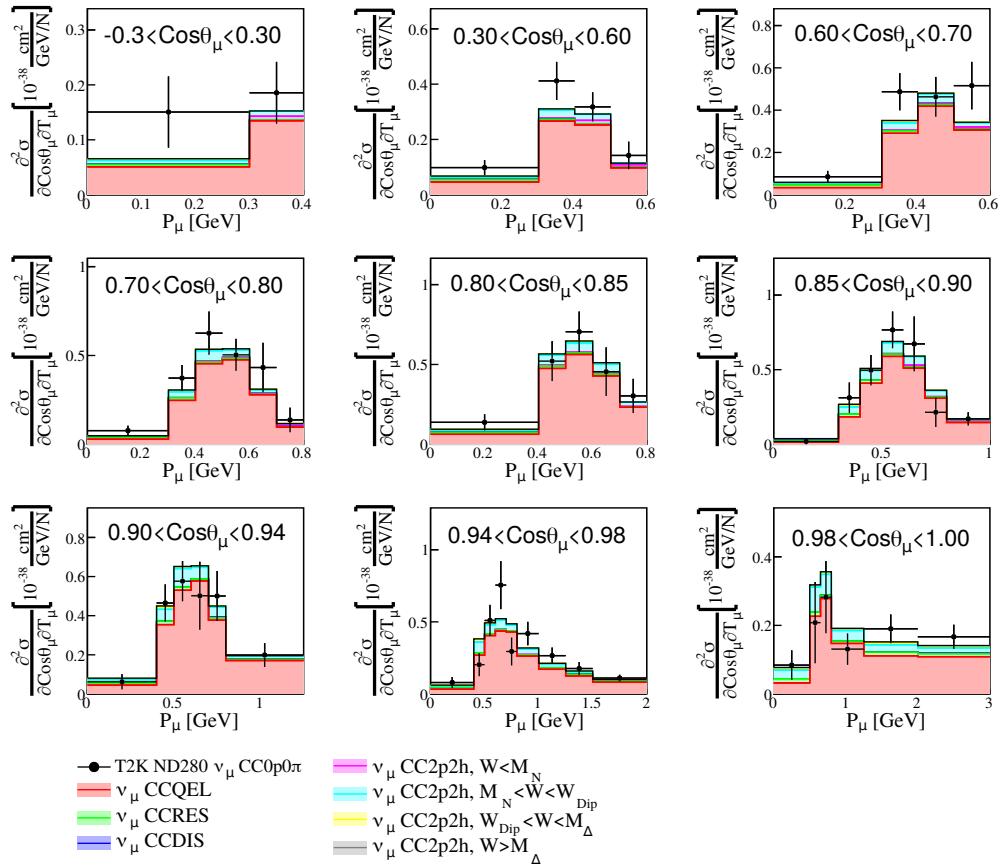


Figure 5.7: T2K ND280 double-differential flux-averaged ν_μ CC0p0 π cross section as a function of the proton multiplicity [279]. The data is compared against the G18_10a_02_11b tune. The GENIE prediction is divided into different categories: CCQEL, CCRES, CC2p2h and CCDIS. For readability, the CC0p0 π high energy bins are not included in these plots.

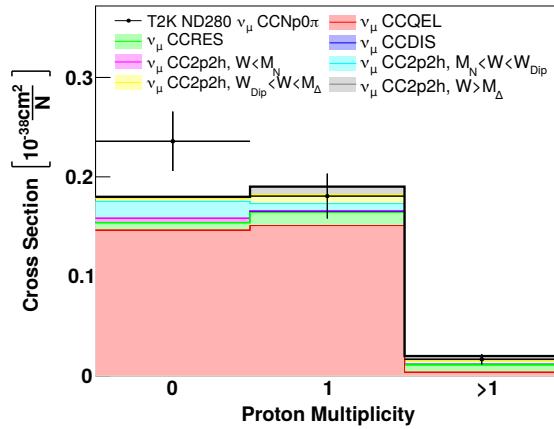


Figure 5.8: T2K ND280 flux-averaged ν_μ CCNp0 π cross section as a function of the proton multiplicity [279]. The data is compared against the G18_10a_02_11b tune. The GENIE prediction is divided into different categories: CCQEL, CCRES, CC2p2h and CCDIS.

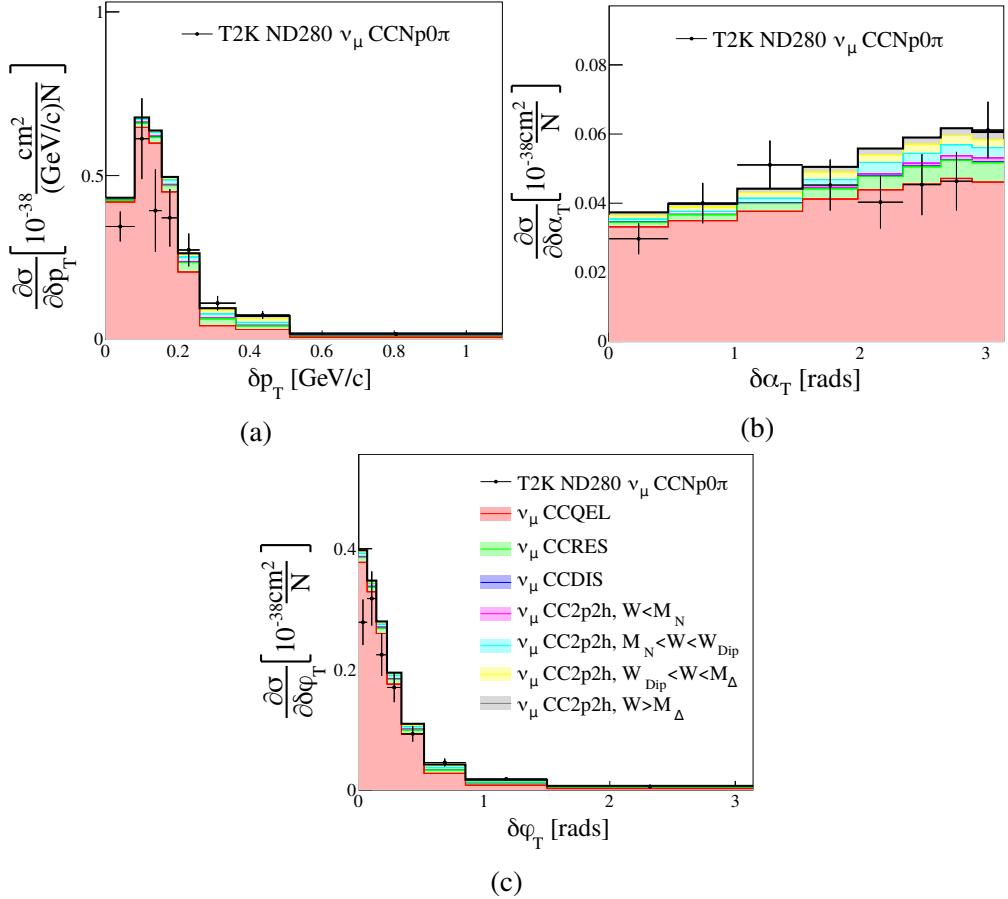


Figure 5.9: T2K ND280 single-differential flux-averaged ν_μ CCNp0 π cross section as a function of STKI variables [279]. The data is compared against the G18_10a_02_11b tune. The GENIE prediction is divided into different interaction categories: CCQEL, CCRES, CC2p2h and CCDIS.

5.2.4 MINERvA CC 0π cross-section measurements

MINERvA studies neutrino interactions on nuclear targets for neutrino and antineutrino interactions at \sim 1-10 GeV at Fermilab [199]. MINERvA’s detector is composed of a segmented scintillator detector surrounded by electromagnetic and hadronic calorimeters. The detector is situated 2.1 m upstream of the MINOS near detector [280], which is a magnetized iron spectrometer. MINOS is used to reconstruct the muon momentum and charge.

Neutrinos are generated at the Neutrino at the Main Injector (NuMI) beamline [268]. This beam has two configurations: *low energy* flux ($\langle E_\nu \rangle \sim 3.5$ GeV) and *medium energy* flux ($\langle E_\nu \rangle \sim 6$ GeV). The NuMI beam can operate in neutrino mode (FHC) and antineutrino mode (RHC). The FHC and RHC *low energy* flux predictions are shown in Fig. 5.1.

MINERvA extracted several CC 0π and CCNp 0π measurements using the NuMI *low-energy* flux [117, 126, 274, 276]. A CC 0π measurement using the NuMI *medium energy* flux is also available [118]. This review focuses on the CC 0π and CCNp 0π measurements obtained with the *low-energy* flux.

The exact target mixture is composed of carbon (88.51%), hydrogen (8.18%), oxygen (2.5%), titanium (0.47%), chlorine (0.2%), aluminium (0.07%), and silicon (0.07%). In the calculation of the GENIE predictions, only the three most abundant targets are considered. The relative mass abundances are renormalized to take this approximation into account. This simplifies the computing power and has a negligible effect on our predictions.

MINERvA ν_μ and $\bar{\nu}_\mu$ CC 0π cross-section measurement

MINERvA reported the CC 0π double-differential flux-integrated cross section as a function of muon momentum in the transverse (\perp) and longitudinal (\parallel) direction relative to the neutrino beam [117, 276]. The differential flux-integrated cross section as a function of E_ν^{QEL} and Q_{QEL}^2 are reported as well [207]. The neutrino energy and the momentum transferred are reconstructed under the QEL hypothesis, described in Sec. 5.2.1. The binding energy used to reconstruct E_ν^{QEL} according to Eq. 5.1 in their neutrino and antineutrino analysis is $E_b = 34$ MeV and $E_b = 30$ MeV respectively.

The ν_μ CC 0π topology is defined as an event with one muon, μ^- , any number of protons and neutrons, any photons below nuclear de-excitation energies, $E_\gamma \leq 10$ MeV, no mesons and no heavy or excited baryons in the final state. Muons are identified by looking for tracks that have a match with the MINOS detector, which is used to

determine the muon momentum and charge⁵. Because of geometric acceptance, both analyses require $\theta_\mu < 20^\circ$. Events containing low-energy photons are accepted as they can arise from nuclear de-excitation. Pions are removed by applying a cut on the recoil energy, $E_{\text{recoil}} \leq 500$ MeV, defined as the activity that is not coming from a muon or any tracked protons. E_{recoil} is corrected for the calorimetric detector response [276]. The recoil energy does not include energy deposited at less than 150 mm from the neutrino vertex as it could be due to proton absorption nearby the vertex. Moreover, events with Michel electrons are removed, as they assume they come from a π decay chain ($\pi \rightarrow \mu \rightarrow e$).

The $\bar{\nu}_\mu$ CC0 π topology [276] is similar to the ν_μ one, with some differences. Due to the nature of this interaction, the muon must be positively charged. Moreover, the analysis requires to have no visible protons in the final state, i.e. protons with kinetic energy above 120 MeV. Finally, mesons are removed using the information on the recoil energy deposited outside the vertex region only.

The GENIE prediction is evaluated with MC events that satisfy the criteria specified above with few exceptions: the removal of events with mesons in the final state is based on true information only. Baryons, are short living and decayed into mesons using the GENIE particle decayer. The requirements on the removal energy are not implemented in our MC analysis either as the data was already corrected for this effect.

GENIE comparisons against the double-differential ν_μ CC0 π measurement is shown in Fig. 5.10 and in Fig. 5.12 for $\bar{\nu}_\mu$ CC0p0 π data. For both ν_μ and $\bar{\nu}_\mu$ data, the G18_10a_02_11b underestimates the data. This is true especially in the phase-space regions in which 2p2h events dominate. In high p_T regions, where the contribution of 2p2h events is negligible, the agreement improves. This can be seen for the $0.85 < p_T < 2.5$ GeV/c slices in Fig. 5.10. Consequently, the reconstructed neutrino energy is also under-predicted, as observed in Fig. 5.11.

MINERvA ν_μ CCNp0 π production cross-section measurement

The MINERvA Collaboration released two CCNp0 π analysis: single-differential cross-section measurements as a function of δp_T , δp , $\delta\alpha_T$ and $\delta\phi_T$ [126], and a single-differential measurement as a function of δp_{Ty} or δp_{Tx} respectively [274].

In both analyses, the topology is defined as events with a muon, no mesons and at least one proton in the final state that satisfy the following conditions:

$$1.5 \text{ GeV}/c < p_\mu < 10 \text{ GeV}/c \text{ and } \theta_\mu < 20^\circ \quad (5.13)$$

⁵The MINERvA detector is not able to measure the muon charge as it does not have a magnetic field.

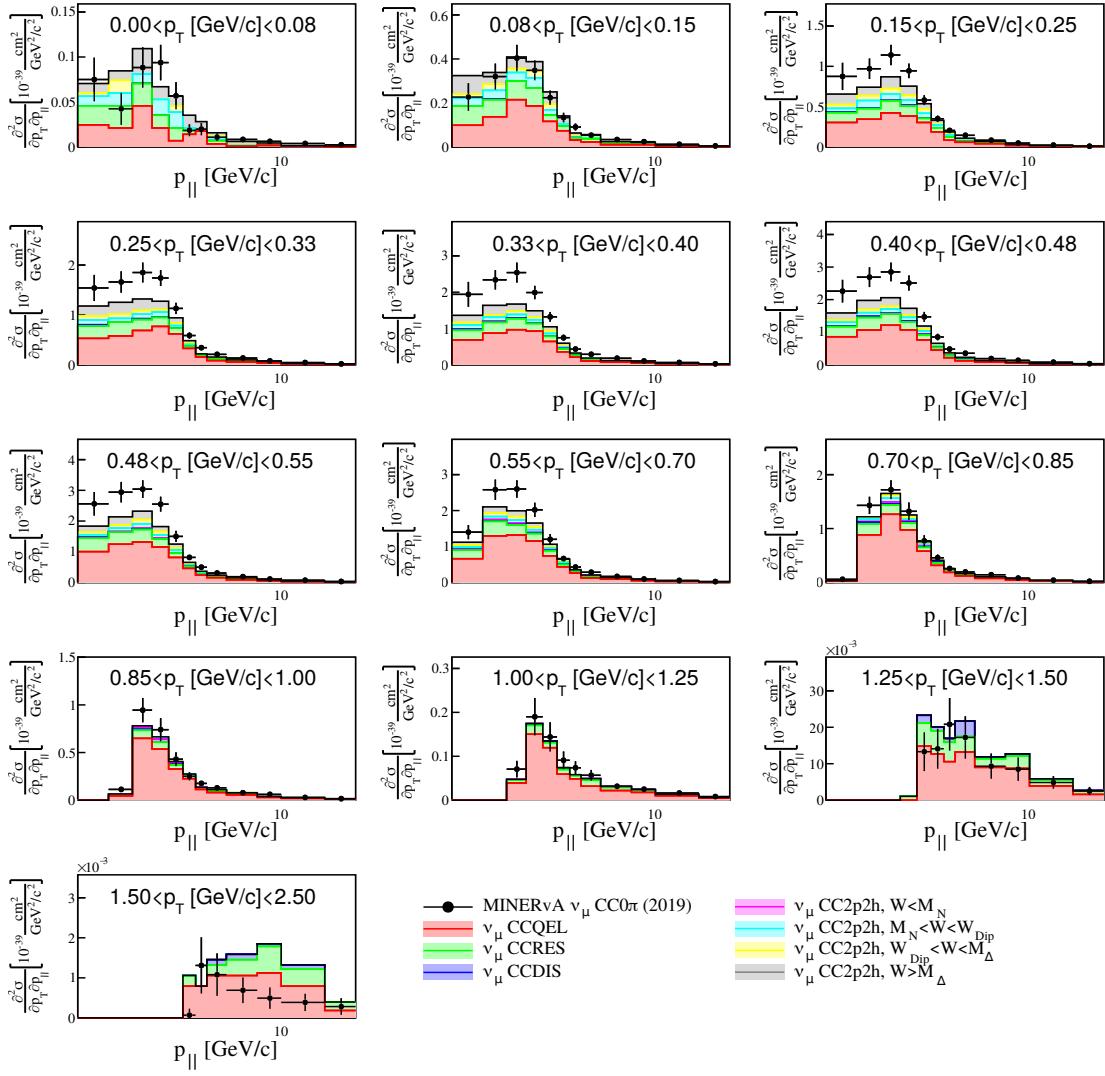


Figure 5.10: MINERvA v_μ CC0 π double-differential flux-averaged cross-section as a function of the muon longitudinal momentum, p_{\parallel} , and transverse momentum, p_T [117]. The corresponding slices on p_T are compared against the G18_10a_02_11b tune. The GENIE prediction is divided into different interaction categories.

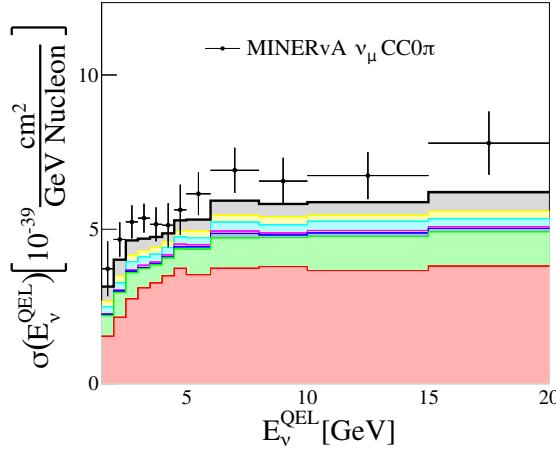


Figure 5.11: MINERvA v_μ CC0 π flux-averaged cross-section as a function of reconstructed neutrino energy, E_ν^{QEL} [117]. The data is compared against the G18_10a_02_11b tune.

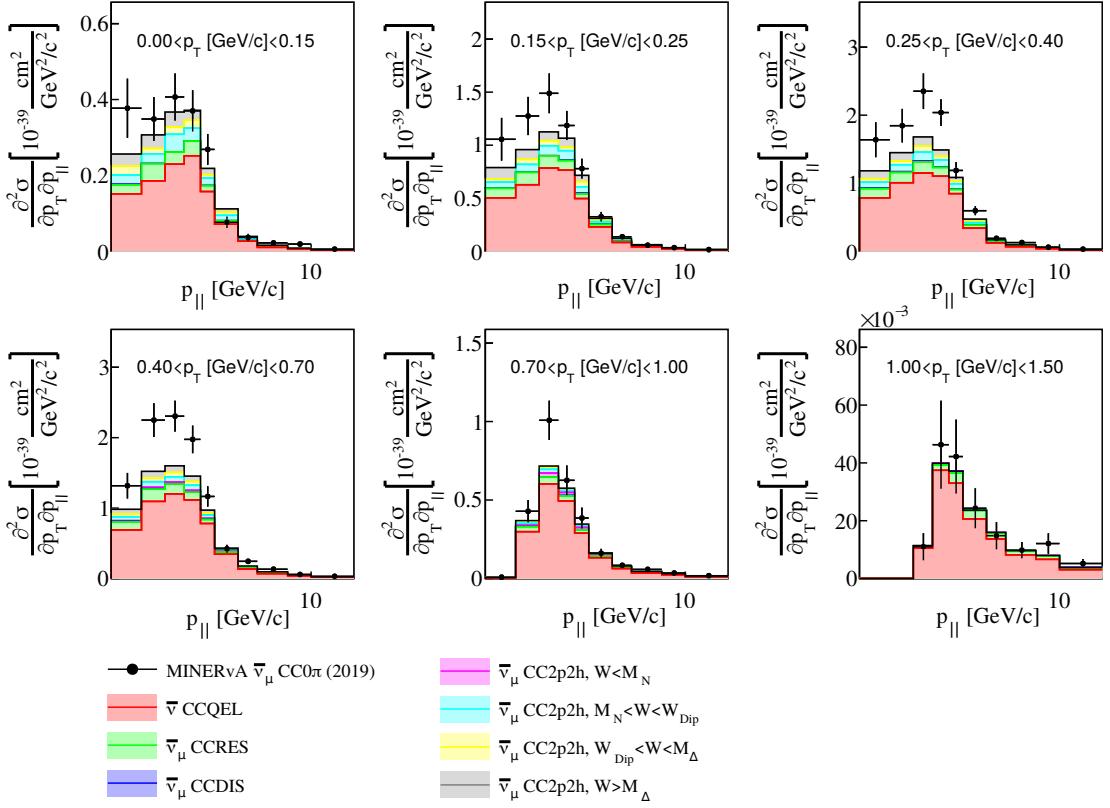


Figure 5.12: MINERvA \bar{v}_μ CC0 π double-differential flux-averaged cross-section as a function of the muon longitudinal momentum, $p_{||}$, and transverse momentum, p_T [117]. The corresponding slices on p_T are compared against the G18_10a_02_11b tune. The GENIE prediction is divided into different interaction categories.

$$0.45 \text{ GeV}/c < p_p < 1.2 \text{ GeV}/c \text{ and } \theta_p < 70^\circ \quad (5.14)$$

where p_μ (p_p) and θ_μ (θ_p) are the muon (lead proton) momentum and opening angle with respect of the neutrino direction. The lead proton is defined as the proton with the highest energy that satisfies the phase space cuts mentioned above.

In this case, the G18_10a_02_11b tune is not under-predicting the data, see Fig. 5.13. This fact, already observed in T2K data, reflects a possible tension between CC 0π and CCNp 0π measurements. The break down into different interaction modes highlights the 2p2h model dependence with W at different proton momenta: 2p2h events with low (high) W dominate at low (high) proton momentum. The contribution from RES events is most significant low proton momentum, small proton angles and $\delta\alpha_T \sim 180^\circ$. In fact, non-QEL events dominate in regions of high transverse kinematic imbalance, such $\delta p_T > 0.2 \text{ GeV}/c$.

5.2.5 MicroBooNE CCNp 0π cross-section measurement

The MicroBooNE experiment is a Liquid Argon Time Projection Chamber (LArTPC) detector situated 500 m away from the BNB beam at Fermilab [100, 110]. LArTPC detectors use complex software algorithms to reconstruct the neutrino event topology with excellent spatial resolution in the detector [281–283]. For instance, MicroBooNE can reconstruct proton tracks of 2 cm with a $\sim 26\%$ efficiency [275]. Different particle identification (PID) algorithms, based on the characteristic signal of each particle in the detector, allow the identification of proton and μ/π candidates, but these methods fail to distinguish between muons and pions.

MicroBooNE provides the first high-statistics cross-section measurements on argon: ν_μ CC inclusive [116], ν_μ CC1p 0π [284], ν_μ CCNp 0π [275], and ν_μ CC π^0 production [285]. The detector is situated 500 m away from the BNB beam at Fermilab [100]. In this section, we focus on the description of the CCNp 0π measurement [275], given that the CC1p 0π measurement [284] is a subsample of the CCNp 0π one.

The CCNp 0π analysis presents a total of five single differential flux-integrated cross-section measurements. The single differential cross sections are given in terms of the muon momentum (p_μ), muon angle (θ_μ), leading proton momentum (p_p), leading proton angle (θ_p), and the angle between the muon and the leading proton ($\theta_{\mu p}$). The main limitation of this approach is that the number of bins in each measurement is insufficient to perform a tune with MicroBooNE data only.

The CCNp 0π topology is defined as an event with one muon, at least one visible proton, any number of neutrons and no pions in the final state. In the analysis, the

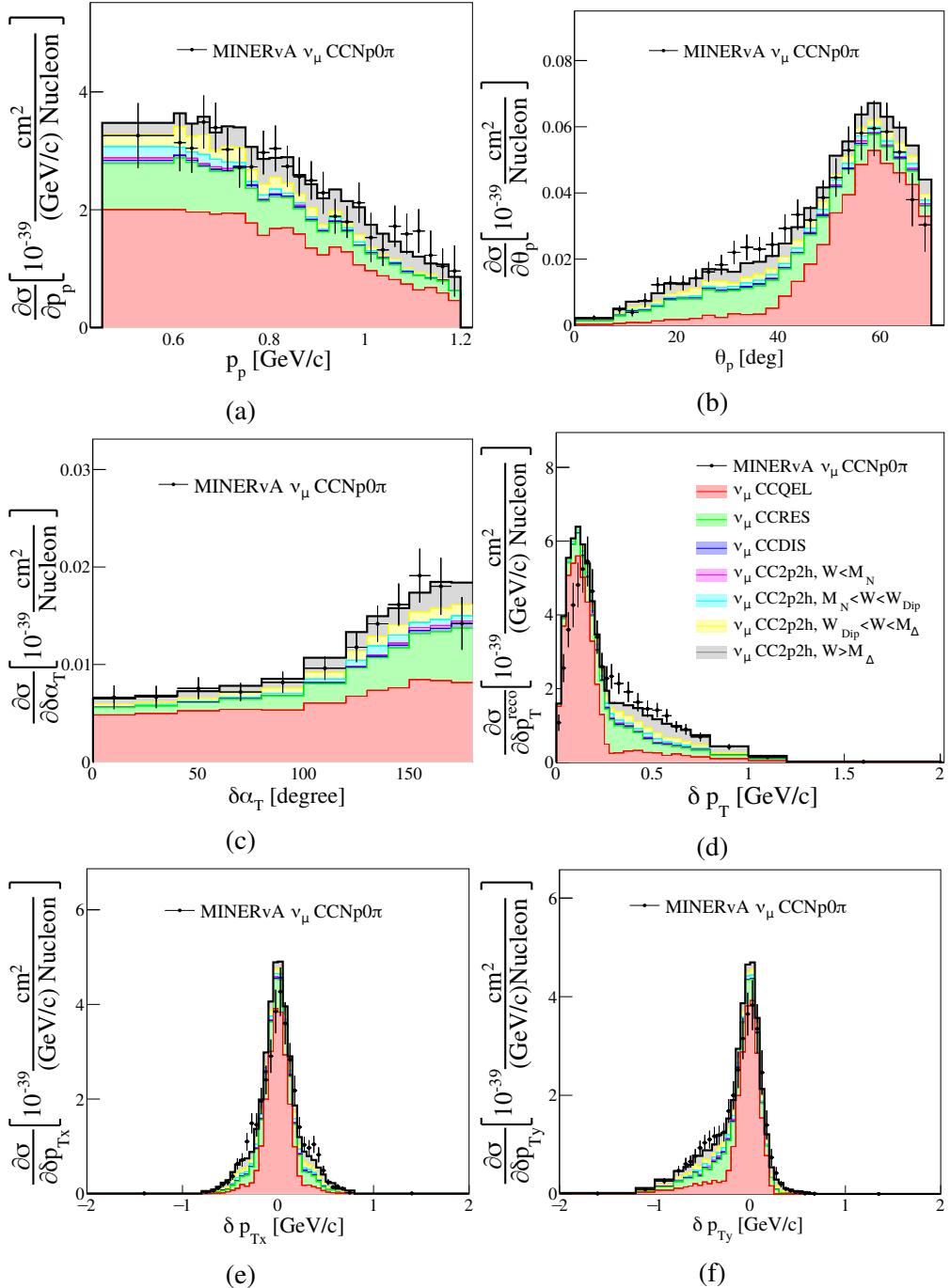


Figure 5.13: MINERvA ν_μ CCNp0 π single-differential flux-averaged cross-section as a function of STKI variables [126, 274]. The data is compared against the G18_10a_02_11b tune. The GENIE prediction is divided into interaction categories.

muon candidate is the longest track which is not identified as a proton. Other tracks in the event must be compatible with the proton PID hypothesis. In order to guarantee at least a 5% efficiency in the momentum reconstruction, they require the muon (proton) to have a momentum of at least 100 MeV/c (300 MeV/c). In addition, the leading proton candidate is must have a reconstructed momentum of less than 1.2 GeV/c. This cut avoids regions of the phase-space in which the proton candidate length is greater than the muon one. These analysis criteria remove visible pions in the event but pions below 30 MeV/c, which are not reconstructed. No corrections are applied to remove events with protons or pions below the detection threshold. The same requirements are applied to the corresponding MC predictions reported in this article.

The differential cross-section measurements were not unfolded to true muon momentum and muon angle. Instead, the results are presented in terms of the reconstructed quantities. The smearing matrices that convert from the reconstructed to the truth quantities are provided in the data release and are used for the evaluation of the GENIE predictions in the reconstructed space [275]. This method is known as *forward folding*. Kinematic quantities in the reconstructed space are denoted with a *reco* superscript. For instance, p_μ^{reco} stands for reconstructed muon momentum.

Fig. 5.14 presents the comparison between the MicroBooNE data and the GENIE predictions. The nominal agreement for the G18_10a_02_11b tune is reasonably good, except for the bin at highest $\cos \theta_\mu^{reco}$, which is largely over-predicted. The contribution of non-QEL interactions increases at forward muon and proton angles, see Fig. 5.14b and Fig. 5.14d. The G18_10a_02_11b dependency on 2p2h events at different W with the proton momenta is re-encountered.

5.2.6 Dataset overview and initial considerations

This subsection summarizes the main observations from this section.

Comparisons of G18_10a_02_11b against the available nuclear data show that CC π and CC0p0 π datasets are under-predicted, whilst CCNp0 π datasets are in good agreement with the *nominal* predictions. MiniBooNE CC π and T2K ND280 CC0p0 π data are under-predicted at muon backward angles, where the contribution to the prediction is mostly from CCQEL events. At forward angles, where the contribution from non-CCQEL events is significant, the data is also under-predicted. The disagreement with MINERvA CC π data is most significant in the region where 2p2h events dominate, $0.15 < p_T < 0.1$ GeV. Finally, non-QEL events dominate the region of high transverse kinematic imbalance. This contribution is essential to describe the data.

The G18_10a_02_11b predictions a function of the leading proton momentum show a dependency of 2p2h with W : at high proton momentum, 2p2h events with

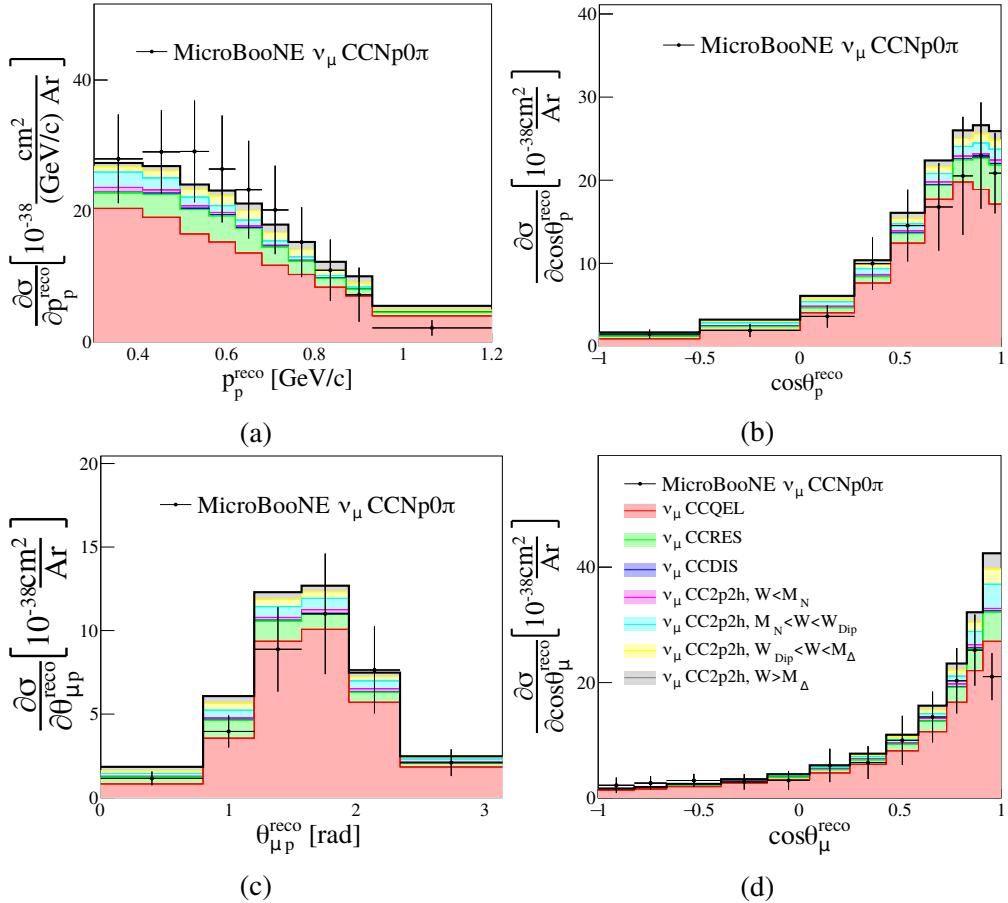


Figure 5.14: MicroBooNE v_μ CCNp0 π flux-averaged differential cross section on ^{40}Ar as a function of muon and proton kinematics. The GENIE prediction is obtained with the G18_10a_02_11b tune. The nominal prediction is divided into interaction categories.

$W > M_\Delta$ dominate, whilst the opposite is true at low momentum. 2p2h events contributing to the T2K ND280 CC0p0 π sample have $W < W_{\text{Dip}}$. Higher multiplicity samples have a significant contribution from 2p2h events with $W > W_{\text{Dip}}$. The contribution from 2p2h events with $W < M_N$ is negligible for all the analyses discussed in this paper.

5.3 Discussion of CC0 π model uncertainties in GENIE

This section describes the parameters available to tweak CC0 π predictions within G18_10a_02_11b. These can be grouped into five categories: CCQEL, CCSIS, CC2p2h, FSI or nuclear model parameters. No A-dependence is considered in any of the parametrizations described here. The study of possible A-dependencies is left to future iterations of this work.

The parameters selected for this analysis are optimized for the G18_10a_02_11a tune. The complete list of parameters is shown in Tab. 5.4. The parameter ranges of interest used for the Professor parametrization are also provided. Most of these parameters can be applied to other CMCs [217]. We strive to have as many common, model independent, parameters to allow for systematic comparison between CMCs, but this is not always possible. An extension of this work to other CMC will be a subject of a future paper.

5.3.1 Charged-current quasi-elastic uncertainty parametrizations

The QEL cross section at the free nucleon level is parametrized with the QEL axial mass, M_A^{QEL} , and a QEL scaling factor, S_{QEL} . Both parameters are common in the simulation of neutrino interactions on free nucleon and nuclei. M_A^{QEL} is used in the dipole parametrization of the QEL form factor. More elaborated CMC based on the z-expansion model [30] are available in GENIE. In this work, preference is given to tune M_A^{QEL} as data-driven uncertainties from hydrogen and deuterium can be used to constrain this parameter [217].

In the nuclear environment, the QEL cross section is affected by nucleon screening due to long-range nucleon-nucleon correlations. This is included into our calculations with the random-phase approximation (RPA) correction [286]. The main effect of the RPA correction is a suppression of the QEL cross section at low Q^2 . This correction is well supported by data and theory, but models differ in predicting its exact strength. This uncertainty is incorporated in GENIE with two parameters: one to scale the nominal QEL cross-section prediction with RPA corrections, ω_{RPA} , and the other one to scale the QEL cross section without RPA corrections, $\omega_{\text{No RPA}}$. The total QEL cross

Parameter	Nominal Value	Range	In Final Tune
M_A^{QEL} (GeV/c ²)	1.00 ± 0.01	[0.97, 1.18]	✓
S_{QEL}	1	—	✗
ω_{RPA}	1	[-0.5, 1.5]	✓
$\omega_{\text{No RPA}}$	0	[-0.5, 1.5]	✓
M_A^{RES} (GeV/c ²)	1.09 ± 0.014	—	✗
S_{RES}	0.84 ± 0.03	[0.5, 1.5]	✓
$R_{vp}^{CC1\pi}$	0.008	—	✗
$R_{vn}^{CC1\pi}$	0.94 ± 0.075	—	✗
$R_{vp}^{CC2\pi}$	0.03 ± 0.01	—	✗
$R_{vn}^{CC2\pi}$	2.3 ± 0.12	—	✗
S_N^{2p2h}	1	[0, 2]	✓
S_Δ^{2p2h}	1	[0, 2]	✓
S_{PL}^{2p2h}	1	[0, 2]	✓
$S_{Abs}^{\pi^\pm}$	1	(1)	✗
$S_{MFP}^{\pi^\pm}$	1	(1)	✗
f^{QEL}	0	(0)	✗
f^{2p2h}	0	(0)	✗

Table 5.4: Summary of parameters relevant for CC0 π analysis. The range of interest, nominal value in GENIE v3 is also shown. The range of interest corresponds to the parameter space used for the Professor parametrization [217]. (—) is used for parameters that are excluded in the analysis. The range for the parameters considered in the Professor parametrization but not used in the final tune is not reported. In such cases, the parameters are fixed to the corresponding nominal values (in parenthesis) in the final analysis, described in Sec. 5.4. The last column specifies whether the parameter is considered in the final analysis, which is described in Sec. 5.4.

section is calculated as a linear combination of the cross section with and without RPA corrections:

$$\sigma^{QEL} = \omega_{RPA} \cdot \sigma_{RPA}^{QEL} + \omega_{No\,RPA} \cdot \sigma_{No\,RPA}^{QEL}. \quad (5.15)$$

This parametrization can be used to scale the QEL cross section when $\omega_{RPA} + \omega_{No\,RPA} \neq 1$. If $\omega_{No\,RPA} = 0$, ω_{RPA} has the exact same effect as S_{QEL} . Therefore, S_{QEL} is not included in the tune. One benefit of this approach is that possible scaling factors on the RPA parametrization do not alter the agreement with free nucleon data. In addition, it reduces the analysis computing time. In Fig. 5.15, the CC QEL cross section as a function of the neutrino energy is shown for different combinations of ω_{RPA} and $\omega_{No\,RPA}$.

Choosing each parameter range of interest is crucial for the correct evaluation of the post-fit uncertainties. In some cases, such as $\omega_{No\,RPA}$, we sample negative values to allow the best-fit result to be at its physical limit of 0. In the case of the RPA parametrization, we impose the additional condition that $0.4 < \omega_{RPA} + \omega_{No\,RPA} < 1.6$ in the sampling on the phase space so that $\sigma^{QEL} > 0$. Fig. 5.16 shows the distribution of sampled parameter values for ω_{RPA} and $\omega_{No\,RPA}$. Notice that the two limit cases are at the centre of the phase space.

It is desirable to apply priors to ω_{RPA} and $\omega_{No\,RPA}$, as effectively, parameter combinations for which $\mathcal{S}_{RPA} \equiv \omega_{RPA} + \omega_{No\,RPA} \neq 1$ act as a scaling of the QEL cross section. Hydrogen and deuterium QEL cross-section measurements are compatible with $\mathcal{S}_{RPA} = 1$. However, nuclear effects might introduce an uncertainty in the scaling. A possible way to include this information is to consider uncorrelated priors on the sum, \mathcal{S}_{RPA} , and the difference, $\Delta_{RPA} \equiv \omega_{RPA} - \omega_{No\,RPA}$,

$$\mathcal{S}_{RPA} = 1 \pm \sigma_{\mathcal{S}} \quad (5.16)$$

$$\Delta_{RPA} = 1 \pm \sigma_{\Delta} \quad (5.17)$$

being $\sigma_{\mathcal{S}}$ and σ_{Δ} the variance associated to the prior on \mathcal{S}_{RPA} and Δ_{RPA} respectively. In terms of ω_{RPA} and $\omega_{No\,RPA}$, this approach includes a correlation between these parameters:

$$\Sigma_{RPA} = \frac{1}{4} \begin{pmatrix} \sigma_{\mathcal{S}}^2 + \sigma_{\Delta}^2 & \sigma_{\mathcal{S}}^2 - \sigma_{\Delta}^2 \\ \sigma_{\mathcal{S}}^2 - \sigma_{\Delta}^2 & \sigma_{\mathcal{S}}^2 + \sigma_{\Delta}^2 \end{pmatrix} \quad (5.18)$$

This correlation between ω_{RPA} and $\omega_{No\,RPA}$ is included in the tune. The corresponding central values are $\mu_{RPA} = 1$ and $\mu_{No\,RPA} = 0$ respectively. The $\sigma_{\mathcal{S}}$ and σ_{Δ} are determined from previous tune iterations, see Sec. 5.4.4. As concluded from Sec. 5.2, some flexibility in the QEL scaling may be required to describe the data, hence, in this analysis $\sigma_{\mathcal{S}} = 0.2$. This method requires to impose a prior on Δ_{RPA} as well. Such prior

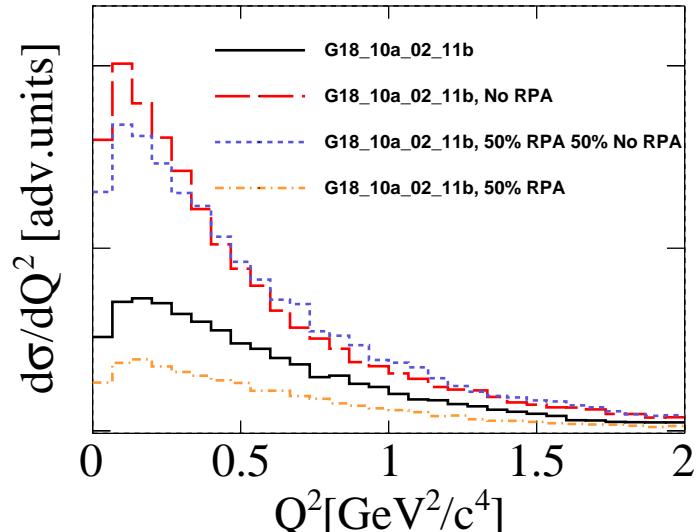
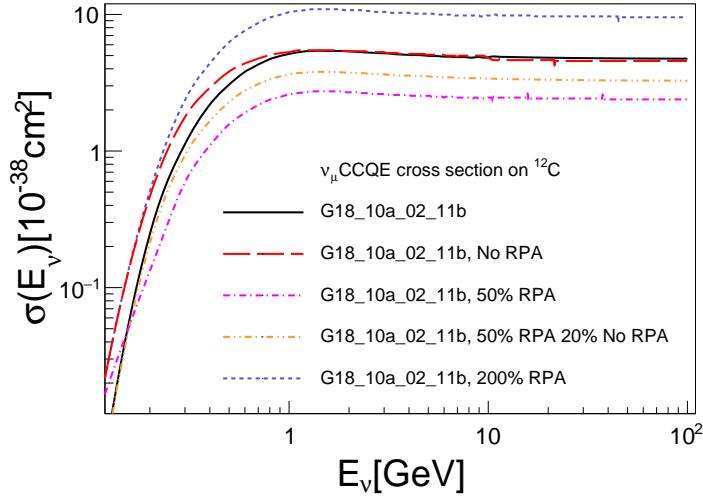


Figure 5.15: Impact of the RPA parametrization on the CCQEL cross section. The G18_10a_02_11b prediction is shown in black. The other predictions are obtained with the same tune while changing the RPA weight values. (a) Total CCQEL cross section for ^{12}C . The blue prediction is equivalent to $S_{QEL} = 2$. (b) Differential cross section as a function of Q^2 .

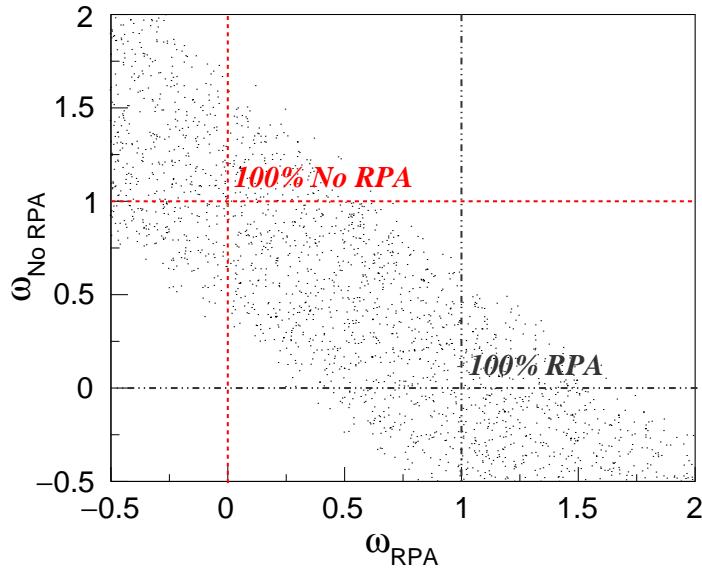


Figure 5.16: Distribution of scan points used for the GENIE tune in the ω_{RPA} vs ω_{NoRPA} phase space. The grey (red) line intersection highlights the limit case in which we consider 100% (0%) of the RPA strength. This plot shows a total of 2,050 scan points distributed uniformly. The parameter ranges considered are those from Tab. 5.4.

affects the strength of the RPA correction, which we aim to constrain from data. In order to avoid strong constraints on Δ_{RPA} , $\sigma_\Delta = 5$.

A similar RPA parametrization was used in the latest theory-driven MicroBooNE tune [270]. The MicroBooNE Collaboration employed the GENIE ReWeight package to parametrize the RPA effect as a linear combination from the QEL cross section with the RPA correction to the QEL cross section without RPA using a single parameter limited to [0,1]. Both approaches are equivalent when $\mathcal{S}_{RPA} = 1$.

5.3.2 Charged-current shallow-inelastic uncertainty parametrizations

SIS events also contribute to the CC π signal as pions can be absorbed by the nuclear medium. Therefore, SIS mismodeling impact the interpretation of the measurements and must be considered in the tune. The parameters available in GENIE to modify the RES and nonresonant background are:

1. RES axial mass, M_A^{RES} ,
2. RES scaling factor, S_{RES} ,
3. SIS scaling parameters that depend on the initial state, $R_{vp}^{CC1\pi}$, $R_{vn}^{CC1\pi}$, $R_{vp}^{CC2\pi}$,

$$R_{vn}^{CC2\pi}.$$

These parameters have been previously tuned against hydrogen and deuterium data [217].

The S_{RES} parameter is the only one included in the CC0 π tune. Nonresonant background parameters are not included: single pion nonresonant background parameters have a small impact on the CC0 π predictions, as observed in the comparisons from Sec. 5.2. In addition, higher multiplicity RES and SIS/DIS contributions are negligible.

5.3.3 Charged-current multi-nucleon uncertainty parameteterizations

The tuning of 2p2h models takes a central role in this work. As discussed in Sec. 5.2, untuned GENIE predictions with G18_10a_02_11b underestimate the data in regions where 2p2h events contribute.

Previous tuning attempts by other neutrino collaborations indicate a preference for a higher 2p2h cross section. The simplest approach to enhance 2p2h is to use a global scaling factor. We refer to this parameter as S_{2p2h} . MINERvA opted for an empirical approach where they add an extra Gaussian contribution to enhance 2p2h interactions in q_0 and q_3 . This is tuned to MINERvA CC inclusive data. This tune is known as MnvGENIE v1 tune [287, 288]. MicroBooNE parametrized the 2p2h cross-section uncertainty with a linear extrapolation between the GENIE 2p2h Empirical and Valencia model to account for possible shape differences. In addition, they also tweak S_{2p2h} . In this paper, we refer to this tune as μ BooNE tune [270].

Different GENIE 2p2h models predict a slightly different strength and shape for the 2p2h cross section [74]. These differences motivated the development of a new parametrization that is able to modify the strength as well as the shape of the cross section in the q_0 - q_3 space. This is accomplished by scaling the 2p2h cross section a function of W :

$$\sigma^{2p2h} \rightarrow S(W) \cdot \sigma^{2p2h} \quad (5.19)$$

$S(W)$ is the scaling function and σ^{2p2h} the *nominal* cross section calculation.

The scaling function, $S(W)$, depends linearly on W . In this work, the scaling function is optimized for the Valencia model which has two characteristic peaks in the q_0 - q_3 space, as it can be seen in Fig. 5.17. The peaks are situated at $W = M_N$ and $W = M_\Delta$. The dip between the two peaks is at W_{Dip} . This is implemented by imposing the following boundary conditions:

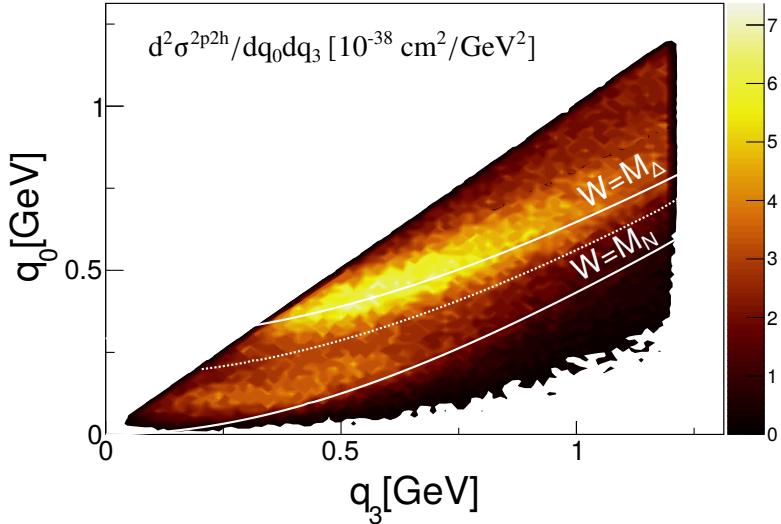


Figure 5.17: Double-differential ν_μ ^{12}C CC2p2h cross section from the Valencia model in GENIE. Lines of constant W at $W = M_N = 938 \text{ MeV}/c^2$, $W = M_{\text{Dip}} = 1120 \text{ MeV}/c^2$ (dotted line) and $W = M_\Delta = 1232 \text{ MeV}/c^2$ are also shown. The events are generated with the G18_10a_02_11b tune using 3 GeV neutrinos.

- $S_{PL,min}^{2p2h} \equiv S(W = W_{PL,min})$
- $S_N^{2p2h} \equiv S(W = M_N)$
- $S_{\text{Dip}}^{2p2h} \equiv S(W = W_{\text{Dip}})$
- $S_N^{2p2h} \equiv S(W = M_\Delta)$
- $S_{PL,max}^{2p2h} \equiv S(W = W_{PL,max})$

The S^{2p2h} parameters are referred to in this work as 2p2h scaling parameters. The limits of the 2p2h phase space are defined by $W_{PL,min}$ and $W_{PL,max}$. The upper limit is obtained by simply imposing $Q^2 = 0$. The lower limit is parametrized as a function of q_0 and q_3 . In all GENIE v3 CMCs, the 2p2h scaling parameters are set to 1.

Only three out of the five 2p2h scaling parameters are included in the tune: S_N^{2p2h} , S_Δ^{2p2h} and $S_{PL,max}^{2p2h}$. Events with $W < M_N$ are negligible for all CC π measurements of interest for this work, hence, $S_{PL,min}^{2p2h}$ is not included in the tune. In addition, S_{Dip}^{2p2h} is also not included as the region between N and Δ peaks is too narrow in W and the data cannot be sensitive to such parameter. In order to facilitate readability, the $S_{PL,max}^{2p2h}$ parameter is redefined as S_{PL}^{2p2h} .

The dependency of the scaling function with W for a particular set of parameters is shown in Fig. 5.18 (top). This particular example enhances (suppresses) the 2p2h cross-section peak at the $W = M_N$ ($W = M_\Delta$) region. The effect on the predictions depends on the neutrino energy, proton multiplicity and proton momenta, as discussed in Sec. 5.2.

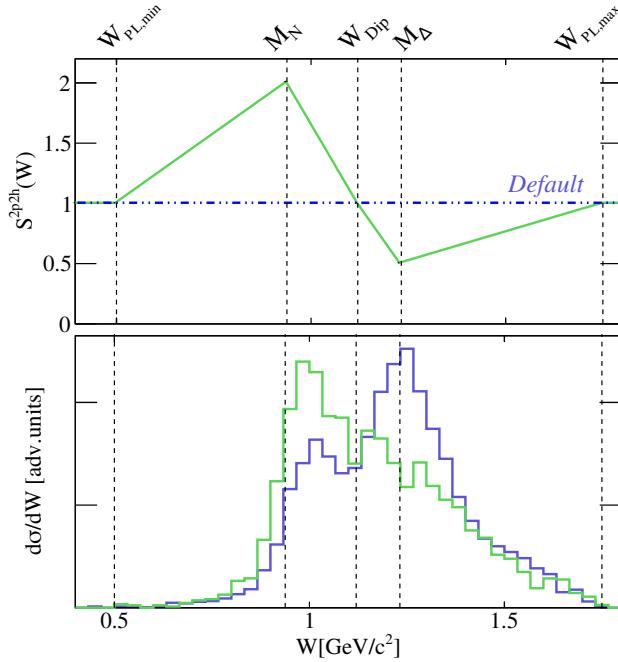


Figure 5.18: Graphic representation of the 2p2h scaling as a function of W . On the top, the *default* parametrization (dashed blue) and an example scaling function (green) are shown. The highlighted dashed vertical lines correspond to the tunable scaling parameters for $W = W_{PL,\min}$, M_N , W_{Dip} , M_Δ and $W_{PL,\max}$. The example scaling function considers $S^{2p2h}(M_N) = 2$, $S^{2p2h}(M_\Delta) = 0.5$ and $S^{2p2h}(W_{PL,\min}) = S^{2p2h}(W_{\text{Dip}}) = S^{2p2h}(W_{PL,\max}) = 1$. The bottom figure shows the Valencia 2p2h flux-integrated cross section as a function of W for the G18_10a_02_11b tune in blue, and the same prediction scaled with the example scaling function in green. This plot is obtained simulating ν_μ interactions on ^{12}C with the NuMI ν_μ low energy configuration [268].

5.3.4 Nuclear model uncertainty parametrizations

Uncertainties in the nuclear model affect the dynamics of the outgoing muon and nucleon after a QEL or a 2p2h interaction.

In the Valencia model implementation in GENIE, the differential cross section is evaluated at an effective energy transfer \tilde{q}_0 , which takes into account the nucleon removal energy. The implementation in the QEL and 2p2h processes is slightly different. The effective energy transfer \tilde{q}_0 used in the Valencia QEL model implementation is:

$$\tilde{q}_0 = q_0 + E_{N_i} - E_p = E_{N_f} - E_p \quad (5.20)$$

E_{N_i} is the energy of the off-shell initial nucleon, which is bound with a binding energy E_B . E_p is the energy of the initial nucleon on-shell with a momentum \mathbf{p} , $E_p = \sqrt{M_N^2 + \mathbf{p}^2}$. E_{N_f} is the energy of the nucleon produced after the QEL interaction, which is on-shell. In other words, the effective energy transfer is reduced relative to the ordinary one by the amount of energy needed to put the initial nucleon on the mass shell. The binding energy and initial nucleon momentum are determined by the corresponding nuclear model. In this work, for QEL interactions we refer to \tilde{q}_0 as \tilde{q}_0^{QEL} . Notice that \tilde{q}_0^{QEL} depends on the event kinematics.

In the Valencia 2p2h model implementation, the effective energy transfer is calculated as:

$$\tilde{q}_0 = q_0 - q_{shift}^{2p2h}, \text{ where } q_{shift}^{2p2h} \equiv M(A_{Z+1}) - M(A_Z) \quad (5.21)$$

In this case, q_{shift}^{2p2h} is independent of the event kinematics. For a carbon target, $q_{shift}^{2p2h}(^{12}C) = 16.8$ MeV, whilst $q_{shift}^{2p2h}(^{40}Ar) = 0.99$ MeV for argon.

Shifts on \tilde{q}_0 are effective modifications of the binding energy in the nuclear model. It is possible to apply relative shift to \tilde{q}_0 for both QEL and 2p2h calculations by modifying q_0^{QEL} and q_{shift}^{2p2h} . This modification translates as:

$$\tilde{q}_0^{QEL} \rightarrow \tilde{q}_0^{QEL}(1 + f^{QEL}) \quad (5.22)$$

$$q_{shift}^{2p2h} \rightarrow q_{shift}^{2p2h}(1 + f^{2p2h}) \quad (5.23)$$

f^{QEL} and f^{2p2h} are two dimensionless parameters. In the GENIE v3 version, both parameters default to 0. Both f^{QEL} and f^{2p2h} parameters are included in the initial iteration of this analysis.

Ref. [289] suggests that shifts on \tilde{q}_0^{QEL} (q_{shift}^{2p2h}) of 0-20 MeV (0-40 MeV) for QEL (2p2h) are in reasonable agreement with electro-scattering data. The effect of such variations on the 2p2h cross-section prediction is shown in Fig. 5.19. The biggest

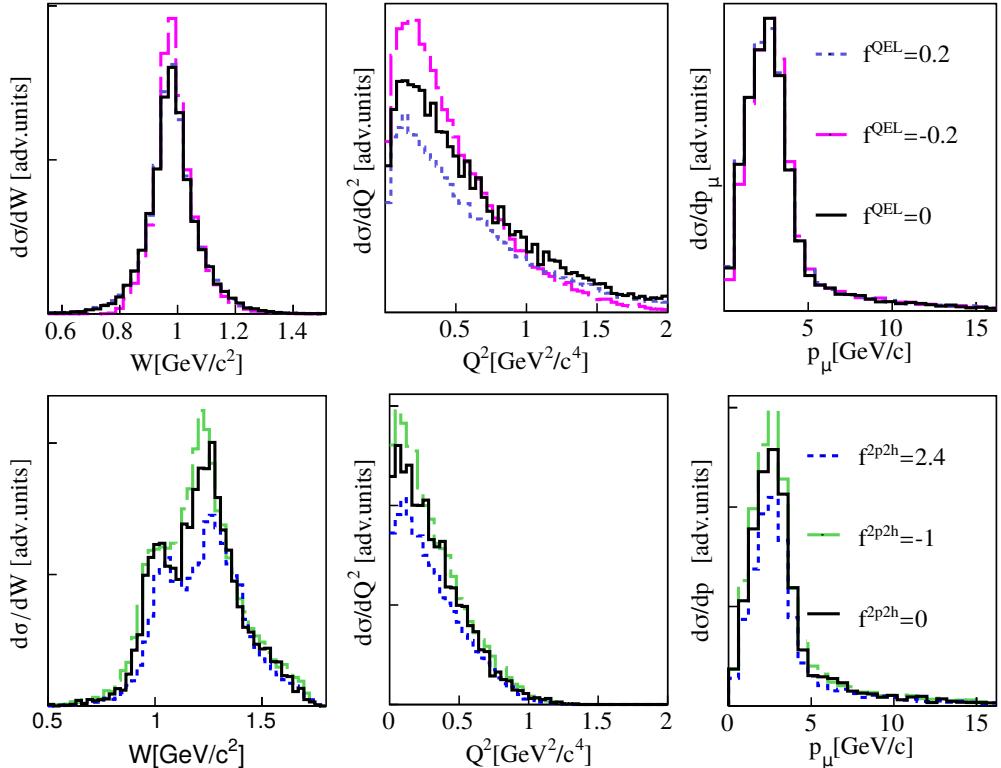


Figure 5.19: Flux-integrated differential $v_\mu^{12}\text{C}$ CC2p2h cross section dependence with W , Q^2 or p_μ . Events are generated with the G18_10a_02_11b tune and the NuMI v_μ low energy flux [268]. The top (bottom) three plots show the CCQEL (CC2p2h) differential cross section as a function of W , Q^2 or p_μ . The black prediction corresponds to the GENIE v3 case, where no shifts on \tilde{q}_0^{QEL} and q_{shift}^{2p2h} are considered. The variations considered for the f^{QEL} and f^{2p2h} parameters correspond to an absolute shift to \tilde{q}_0^{QEL} and q_{shift}^{2p2h} of 20 MeV for QEL interactions and of 40 MeV for 2p2h interactions.

variation is observed on $d\sigma/dQ^2$ for both QEL and 2p2h. For the 2p2h cross section, this systematic shifts peaks position in W .

5.3.5 Final state interaction uncertainty parametrizations

Final-state interactions (FSI) are crucial for modeling nuclear cross sections as they affect the event topology and kinematics of an event. There are different models available in GENIE to simulate FSI [40, 143]. In particular, G18_10a_02_11b models FSI with the INTRANUKE hA model [37].

INTRANUKE hA is an empirical model that considers a single interaction which is based on hadron-nucleus data [143]. In particular, pion-nucleus data is used to determine the inelastic (Inel), absorption (Abs), charge-exchange (CEx) and pion production (π Prod) fractions (f_i). The fractions depend on the pion kinetic energy and the nuclear atomic number. These fractions satisfy that $\sum_i f_i^{\pi^\pm} = 1$ (unitarity condition),

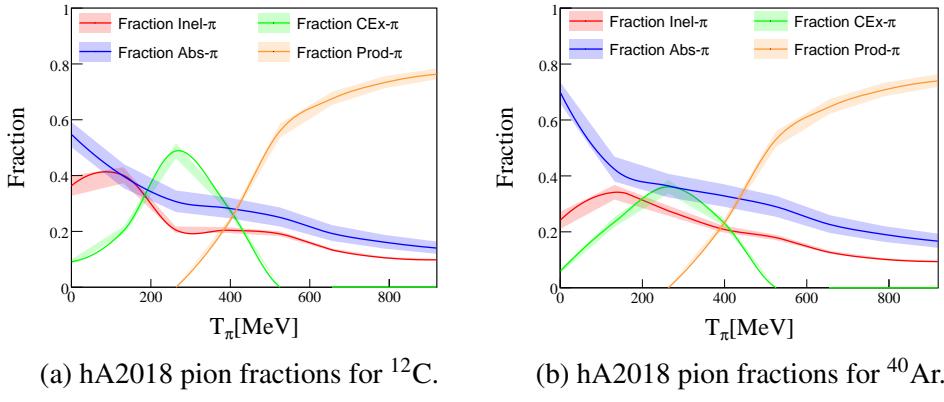


Figure 5.20: hA2018 FSI pion fractions as a function of the pion kinetic energy for a carbon target. The error bands represent the fraction variation when applying a $S_{\text{Abs}}^{\pi^\pm} = 1.2$ on the pion absorption fraction, which corresponds to a variation of $\sim 15\%$ for the pion absorption fraction on carbon at $T_\pi = 200$ MeV.

where i is an index that runs over the available processes aforementioned.

A parameter is introduced to be able to modify the $f_{\text{Abs}}^{\pi^\pm}$ while preserving unitarity:

$$f'_{\text{Abs}}^{\pi^\pm} = \frac{S_{\text{Abs}}^{\pi^\pm} \cdot f_{\text{Abs}}^{\pi^\pm}}{f_{\text{Inel}}^{\pi^\pm} + S_{\text{Abs}}^{\pi^\pm} \cdot f_{\text{Abs}}^{\pi^\pm} + f_{\text{CEx}}^{\pi^\pm} + f_{\text{Prod}}^{\pi^\pm}} \quad (5.24)$$

The other fractions are also modified as a consequence of this scaling. Notice that variations of $S_{\text{Abs}}^{\pi^\pm}$ do not scale $f'_{\text{Abs}}^{\pi^\pm}$ linearly.

Similarly, a scaling parameter is introduced to scale the charged pion mean-free path. This is referred to as $S_{\text{MFP}}^{\pi^\pm}$. The same approach can be applied to other processes and to nucleon fractions.

Fig. 5.20 shows the dependence of each hA fraction as a function of the pion kinetic energy (T_π) for carbon and argon targets. The FSI fractions and their uncertainty are extracted from fits to hadron-nucleus scattering data [143, 290]. The uncertainty associated to $f_{\text{Abs}}^{\pi^\pm}$ is 15%.

Variations of the FSI parameters considered in this work result in the migration of CC1 π events into the CC0 π sample. The effect on the prediction depends on the topology definition. For CC0 π samples, it mostly affects the overall normalization of the cross-section. The measurement most sensitive to this variation is the ν_μ CCNp0 π MINERvA differential cross section as a function of $\delta\alpha_T$, see Fig. 5.21. A decrease in $S_{\text{Abs}}^{\pi^\pm}$ reduces the cross section at $\delta\alpha_T \sim 180$. In addition, this model variation also affects the slope of the distribution.

In this tune, only parameters related to charged pion absorption are included: $S_{\text{Abs}}^{\pi^\pm}$ and $S_{\text{MFP}}^{\pi^\pm}$. Pion inelastic fractions are not relevant at the energies of interest for this

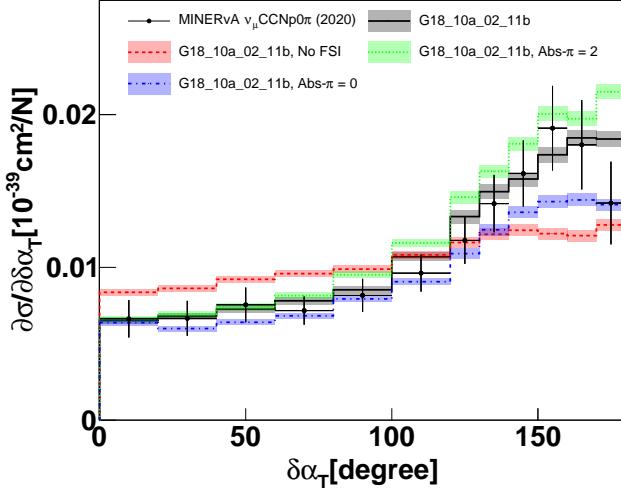


Figure 5.21: Impact of $S_{\text{Abs}}^{\pi^\pm}$ on MINERvA CCNp0 π flux-integrated differential cross section predictions as a function of $\delta\alpha_T$.

work. Nucleon FSI parameters are relevant for the study of exclusive cross-section measurements with protons in the final state. Ideally, to perform a global tune with CC0 π and CCNp0 π data, nucleon FSI parameters must be considered in the analysis. Including these parameters in the analysis substantially increases the computing time. In addition, it is desirable to first understand the tensions between CC0 π and CCNp0 π measurements. Therefore, it is therefore convenient to reduce the complexity of the analysis and focus on CC0 π datasets only. Nucleon FSI parameters will be included in future iterations of this work.

5.4 Tuning procedure

This section summarizes the tuning procedure for the analysis. The main goal is to tune GENIE against MiniBooNE, T2K ND280, MINERvA and MicroBooNE CC0 π data.

5.4.1 Construction of the GENIE prediction

In order to build the prediction associated to each dataset specified in Sec. 5.2, we generate v_μ and \bar{v}_μ CC events for the experiment target using the neutrino fluxes from Fig. 5.1. In this work, the events are generated with the G18_10a_02_11b tune [217].

To compute the prediction associated with the i th dataset, we generate N_i^{TOT} events. Events that do not satisfy the corresponding selection criteria specified in Sec. 5.2 are rejected. The number of accepted events in a given bin is $N_{ij}(\boldsymbol{\theta})$. $\boldsymbol{\theta}$ is the vector of tunable parameters specified in Tab. 5.4.

We build the corresponding n -differential flux-integrated cross section prediction for a given set of observables, O , as:

$$\left(\frac{\partial^n \sigma_{th}^i(\boldsymbol{\theta})}{\partial O^n} \right)_j = \frac{N_{ij}(\boldsymbol{\theta})}{\Phi_i N_i^{TOT} \Delta O_{ij}} \int dE_\nu \frac{d\phi_i}{dE_\nu} \sum_{T_i} R_{T_i} \sigma_{T_i}(E_\nu, \boldsymbol{\theta}) \quad (5.25)$$

Φ_i is the integrated flux for the i th dataset. ΔO_{ij} corresponds to the j th n-dimensional bin volume for the quantities used in the differential cross section calculation. $d\phi/dE_\nu$ is the expected flux at a given neutrino energy. For a target mix, the averaged cross section is evaluated by summing over the nucleus type in the target mix, T_i . The ratio of a specific nucleus type with respect to the total nuclei is R_{T_i} and $\sigma_{T_i}(E_\nu)$ is the total cross section for a given nucleus type.

5.4.2 Avoiding the Peele's pertinent puzzle

The bin to bin covariance matrix provided by each experiment is considered in the evaluation of the χ^2 . The T2K ND280, MicroBooNE and MINERvA datasets have highly correlated bin to bin covariance matrices. Previous attempts to fit neutrino-nucleus data using the full covariance matrices result in a significant reduction of the cross section [270, 291]. These results are not surprising in highly correlated bins ($\rho > 60\%$) in the Gaussian approximation [292]. This is known as Peele's pertinent puzzle (PPP) [292, 293].

To avoid PPP, we change our variables in order to reduce the correlation for the i th dataset using the following prescription:

$$Z_{ij} \equiv \begin{cases} \sum_k D_{ik} & j = 0 \\ \frac{D_{ij}}{\sum_k D_{ik}} & 0 < j < N_i \end{cases} \quad (5.26)$$

D_{ij} corresponds to the i th dataset mean value at the j th bin. The j th and k th indexes run over the number of bins associated with the i th dataset. This is known as norm-shape (NS) transformation. After the NS transformation, the integral is moved into the first bin of the i th dataset, whilst the rest describes the shape distribution. This transformation is applied to both data and predictions.

The bin to bin covariance associated with the i th dataset, $\Sigma_D(\mathbf{D})_{ijk}$, transforms as follows:

$$\Sigma_{NS}(\mathbf{Z})_{ijk} \equiv \left[\left(\frac{d\mathbf{Z}}{d\mathbf{D}} \right) \Sigma_D(\mathbf{D}) \left(\frac{d\mathbf{Z}}{d\mathbf{D}} \right)^T \right]_{ijk} \quad (5.27)$$

where,

$$\left(\frac{d\mathbf{Z}}{d\mathbf{D}} \right)_{ija} = \begin{cases} 1 & j = 0 \\ \frac{\delta_{ja} \left(\sum_k D_{ik} \right) - D_{ij}}{\left(\sum_k D_{ik} \right)^2} & 0 < j < N_i \end{cases}. \quad (5.28)$$

After the NS transformation the relative uncertainties are constant when the normalization changes.

The same transformation is applied to the prediction mean values and covariance. Before the NS transformation, the prediction covariance only has diagonal elements. This is not true after the NS transformation. However, the off-diagonal elements on the prediction covariance are small and are neglected in this work. The prediction central values and errors after the NS transformation are denoted as $Y_{ij}(\theta)$ and $\delta Y_{ij}(\theta)$ respectively.

5.4.3 Professor parametrization

Given that performing a multi-parameter brute-force scan is not feasible, we use Professor [129] to parametrize the behavior of our predicted cross section and error in each bin in the NS space. We refer to this quantities as $\tilde{Y}_j^i(\theta)$ and $\delta \tilde{Y}_j^i(\theta)$. In this particular tune, we opted for a fourth order parametrization. This work, where originally eleven parameters were included in the analysis, requires a total of 2k scan points sampled across the ranges specified in Tab. 5.4. The accuracy of the parametrization is shown in Fig. 5.22. This parametrization is used for the estimation of the best-fit values by minimizing the χ^2 .

5.4.4 Final choice of parameters for the CC0 π tune

A series of preliminary tunes were performed using different priors or parameter sets. The goal of this study is to determine which parameters to include in the final tune.

Nuclear effects in the QEL cross section are tweaked with the RPA parametrization. Free-nucleon cross-section data suggests that the QEL cross section should not be scaled. This condition can be incorporated in our analysis by imposing a more restrictive prior on S_{RPA} of $\sigma_S = 0.01$. Tunes performed using this prior result in worse goodness of fit, suggesting that a less restrictive prior on the sum is desired to improve the agreement with the data. This motivated our choice for a prior on the sum of $\sigma_S = 0.2$, as described in Sec. 5.3.1.

In the particular case of T2K ND280, variations of S_Δ^{2p2h} and S_{PL}^{2p2h} do not affect the

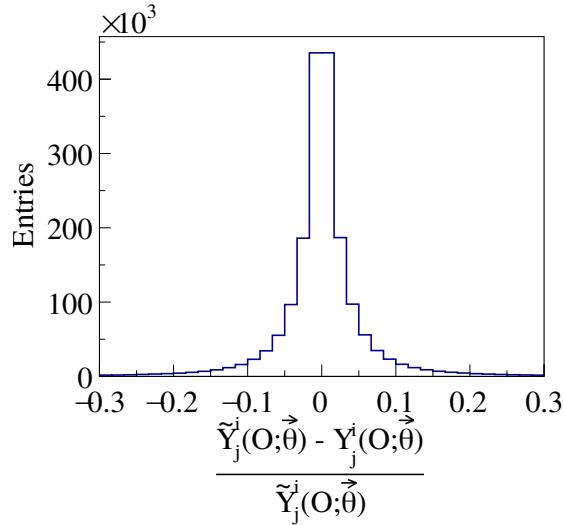


Figure 5.22: Fractional difference between true MC predictions in the NS space calculated with a given θ parameter set.

CC0p 0π predictions. This is highlighted in Fig 5.8, where only events with $W < W_{\text{Dip}}$ contribute to the 2p2h cross section prediction with no protons above the detection threshold. Therefore, these parameters are not included when tuning against T2K ND280 CC0p 0π data.

FSI interactions are important to describe CC 0π measurements. In particular, modification of $S_{\text{Abs}}^{\pi^\pm}$ and $S_{MFP}^{\pi^\pm}$ affect the RES contribution to the prediction. The *nominal* charged pion absorption fraction and mean-free path are extracted from pion-scattering data. Variations of these parameters must preserve the agreement with this data. Previous studies suggest that variations on these magnitudes should not exceed 20% [143]. This information is included in the partial tunes using priors. Two tunes are performed by including or excluding the FSI parameters in the tune⁶. The results suggest that variations of these parameter that respect pion-scattering data do not have a big impact on the tune results. Consequently, these parameters are not included in the final analysis.

A similar study is performed by including or excluding the f^{QEL} and f^{2p2h} parameters in the tune. Strong correlations are observed between the f^{QEL} , f^{2p2h} and ω_{RPA} and S^{2p2h} parameters. In some cases, these correlations lead to unphysical values for f^{QEL} and f^{2p2h} . For this reason, these parameters are excluded from the final analysis.

The final parameter set used in this work is summarized in Tab. 5.4.

⁶ f^{QEL} and f^{2p2h} parameters are set to the nominal values for this study.

Parameter	Prior	M_A^{QE}	S_{RES}
M_A^{QEL}	$1.00 \pm 0.01 \text{ GeV}/c^2$	M_A^{QE}	1.5×10^{-4}
S_{RES}	0.84 ± 0.028	S_{RES}	6.0×10^{-4}

Table 5.5: Priors (left) and covariance matrix (right) for M_A^{QE} and S_{RES} obtained to the free nucleon tune from Ref. [217].

5.4.5 Discussion of data driven priors

Some of the parameters of interest for this tune were already tuned to neutrino data. Particularly, the G18_10a_02_11b tune with hydrogen and deuterium data provided with data-driven constraints for M_A^{QE} and S_{RES} [217]. The information on the parameter priors central values as well as the correlation between the two parameters out of the free nucleon tune is included in the χ^2 minimization. The complete information on the priors is provided in Tab. 5.5. In this analysis we also include priors on ω_{RPA} and $\omega_{\text{No RPA}}$, as discussed in Sec. 5.3.1.

5.4.6 Evaluation of the χ^2

The complete form for our χ^2 is:

$$\chi^2(\boldsymbol{\theta}) = \sum_i \mathbf{w}_i \Delta \tilde{\mathbf{Y}}_i^T(\boldsymbol{\theta}) \Sigma_{NS,i}^{-1} \Delta \tilde{\mathbf{Y}}_i(\boldsymbol{\theta}) + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \Sigma_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \quad (5.29)$$

$\Delta \tilde{\mathbf{Y}}_{ij}$ is the difference between the NS parametrization prediction and the i th dataset at the j th bin, $\Delta \tilde{\mathbf{Y}}_{ij}(\boldsymbol{\theta}) = (\tilde{\mathbf{Y}}_{ij}(\boldsymbol{\theta}) - Z_{ij})$. The ω_{ij} is the weight applied to j th bin from the i th dataset. In this work, weights are used to include or exclude data from the analysis. In other words, they are either 1 or 0. The prediction errors, $\delta \tilde{\mathbf{Y}}_i(\boldsymbol{\theta})$, are added in quadrature to $\Sigma_{NS,i}$. The second term takes care of correlated priors in our fit. $\boldsymbol{\theta}_0$ and $\Sigma_{\boldsymbol{\theta}}$ are the central values vector and the covariance matrix of the priors for the parameters of interest. The details on the priors applied in this analysis are described in Sec. 5.4.5.

5.5 Tuning Results

For this paper, the following simple naming scheme is adopted for each of the partial GENIE tunes presented in this work⁷:

$$Gxx [a-d]$$

where

- G is a capital letter that stands for GENIE, highlighting the authorship of the tunes.
- xx is a number assigned to each experiment, i.e., MiniBooNE (10), T2K ND280 (20) or MINERvA (30). When using antineutrino datasets, xx is increased by one unit. For CCNp0 π datasets, xx is increased by five units.
- [a-d] refers to the alternative intranuclear hadron model used in the analysis: (a) INTRANUKE/hA, (b) INTRANUKE/hN, (c) GEANT4/Bertini and (d) INCL++.

A total of six partial tunes are performed: three tunes on neutrino CC0 π data, two tunes using antineutrino CC0 π data and one tune using ν_μ CCNp0 π data. The tunes on CC0 π data aim to explore avenues for improving the agreement between GENIE and data, consolidate the main elements of the GENIE CC0 π tuning methodology and provide a common ground for the discussion of tensions. The tune on CCNp0 π data aims to highlight tensions between CC0 π and CCNp0 π datasets. All of the tunes presented in this work consider carbon datasets only. Joint fits to all available data will be performed at a future iteration of this work, aiming to produce the tunes that will be publicly released through the GENIE platform.

In all CC0 π tunes, the analyses are carried out using double-differential flux-integrated CC0 π cross-section data as a function of muon kinematics. Preference is given to datasets that do not require a minimum number of protons above detection threshold in the final state. Whenever CC0 π datasets are not available for a particular experiment, the tune is performed using CC0p0 π datasets instead.

G18_10a_02_11b is the starting point for all these tunes and provides the *nominal* predictions. The corresponding names assigned to each tune prepared for the purposes of this paper are the following:

- G10a Tune: GENIE tune to MiniBooNE ν_μ CC0 π data [99].
- G11a Tune: GENIE tune to MiniBooNE $\bar{\nu}_\mu$ CC0 π data [273].

⁷Note that this is a different from the standard naming scheme used for the tunes released through the GENIE platform. The standard naming convention will be used if one or more of the tunes produced in this work or future iterations is prepared for release in GENIE.

5.5. TUNING RESULTS

Parameters	G10a Tune	G11a Tune	G20a Tune	G30a Tune	G31a Tune	G35a Tune
$M_A^{QEL}(\text{GeV}/c^2)$	1.02 ± 0.01	1.01 ± 0.01	1.00 ± 0.01	1.00 ± 0.02	1.00 ± 0.01	0.99 ± 0.01
ω_{RPA}	1.20 ± 0.03	1.14 ± 0.06	1.2 ± 0.2	0.9 ± 0.1	1.3 ± 0.2	0.75 ± 0.3
$\omega_{\text{No RPA}}$	0.05 ± 0.02	0.09 ± 0.05	-0.1 ± 0.1	0.2 ± 0.1	0.2 ± 0.2	0.09 ± 0.3
S_{RES}	0.85 ± 0.02	0.86 ± 0.05	0.84 ± 0.02	0.84 ± 0.03	0.84 ± 0.02	0.84 ± 0.02
S_N^{2p2h}	1.5 ± 0.4	2.3 ± 0.01	1.7 ± 0.3	1.2 ± 0.4	1.7 ± 0.5	0.33 ± 0.2
S_A^{2p2h}	0.7 ± 0.2	0.7 ± 0.3	(1.00)	2.1 ± 0.2	2.3 ± 0.2	0.5 ± 0.4
S_{PL}^{2p2h}	0.4 ± 0.1	0.4 ± 0.1	(1.00)	0.9 ± 0.2	0.4 ± 0.1	1.5 ± 0.4
χ^2	89/130	77/71	60/55	61/137	67/53	17/19

Table 5.6: Best-fit parameter values for the different partial tunes. Parameter values within parenthesis are kept fixed during the fit. The χ^2 values are calculated with the Professor parametrization, in accordance to Eq. 5.29.

- G20a Tune: GENIE tune to T2K ND280 ν_μ CC0p0 π data [119].
- G30a Tune: GENIE tune to MINERvA ν_μ CC0 π data [117].
- G31a Tune: GENIE tune to MINERvA $\bar{\nu}_\mu$ CC0p0 π data [276].
- G35a Tune: GENIE tune to MINERvA ν_μ CCNp0 π data [126].

Other measurements, including MicroBooNE datasets, are used for comparisons only. Each partial tune is performed following the recipe described in Sec. 5.4.

5.5.1 Discussion of partial CC0 π tune results

The best-fit parameter values and the χ^2 values calculated with the Professor parametrization at the best-fit point for each tune are summarized in Tab. 5.6. The *nominal* and best-fit predictions are presented in Figures 5.23 to 5.29. Tab. 5.7 provides the χ^2 values computed for each partial tune for all the available CC0 π datasets. In this case, the χ^2 values are calculated with the NS transformation with the GENIE predictions. The Professor parametrization is not exact; hence, the χ^2 values from Tab. 5.7 are different to the ones provided in Tab. 5.6.

It is observed that the description of the data after the tune improved substantially. For instance, the agreement with MINERvA ν_μ CC0 π before the tune is $\chi^2_{\text{Nominal}} = 626/144$ DoF. After the tune, $\chi^2_{\text{Nominal}} = 151/144$ DoF. This is a consequence of an improvement in the overall normalization for each partial tune.

All carbon tunes show similar trends; whilst the tunes are in good agreement with the priors on M_A^{QEL} and S_{RES} , the other parameters differ from the *nominal* parameter values. There is a clear inclination for QEL with RPA corrections. In addition, the tunes prefer a higher QEL, i.e. $\omega_{\text{RPA}} + \omega_{\text{RPA}} > 1$, and 2p2h cross section. Finally, the

Dataset	χ^2_{Nominal}	χ^2_{G10a}	χ^2_{G11a}	χ^2_{G20a}	χ^2_{G30a}	χ^2_{G31a}	χ^2_{G35a}	DoF
MiniBooNE ν_μ CC π	1817	121	160	314	379	1279	2727	137
MiniBooNE $\bar{\nu}_\mu$ CC π	444	208	214	246	403	491	879	60
T2K ND280 ν_μ CC0p0 π	139	447	600	123	237	916	239	60
MINERvA ν_μ CC π	626	252	202	270	151	360	953	144
MINERvA $\bar{\nu}_\mu$ CC π	2259	1837	1680	2232	1794	82	1810	78

Table 5.7: Summary of χ^2 values associated the CC π datasets specified in each row. The χ^2 values are calculated using the NS method for seven different GENIE predictions: G18_10a_02_11b, G10a, G11a, G20a, G30a, G31a and G35a. The values highlighted in bold correspond to the best-fit χ^2 for the partial tune predictions.

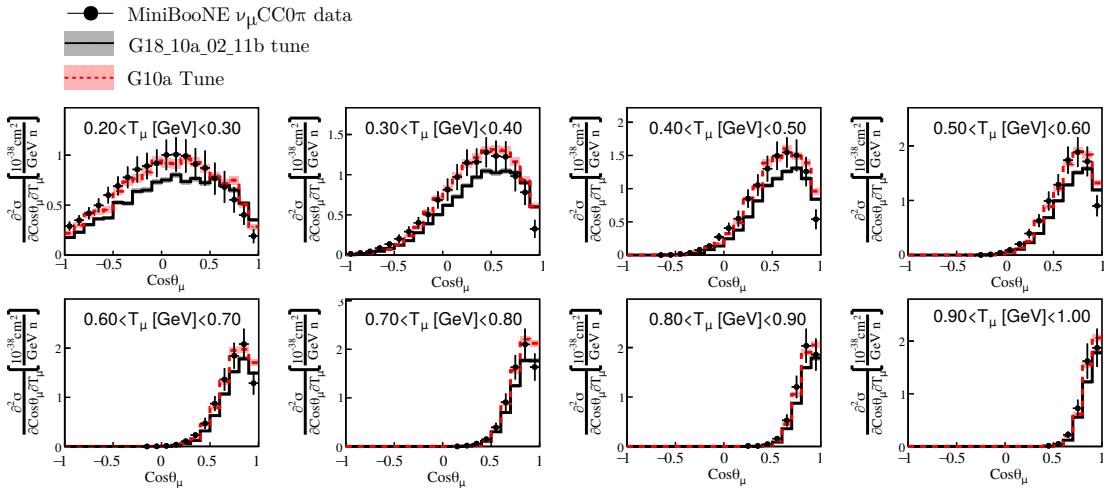


Figure 5.23: Comparison of the G18_10a_02_11b and G10a tunes against MiniBooNE ν_μ CC π double differential data [99]. The comparisons are restricted to the $0.2 < T_\mu < 1.0$ GeV phase space. The predictions are computed using the parameters specified in Tab. 5.6. The total χ^2 associated to this dataset before (after) the tune is reported in Tab. 5.7.

different tunes suggest an underlying energy dependence for $S^{2p2h}(W)$: the G10a, G11a and G20a tunes enhance (suppress) the Valencia 2p2h cross section at the nucleon (Δ) region. The G30a and G31a tunes enhance mostly the cross section at the Δ region.

The enhancement of the QEL cross section is crucial for the description of MiniBooNE CC π data at $\cos \theta_\mu < 0$. Particularly, the G10a and G11a tunes suggest an increase of the QEL cross section of about 20%. Similar QEL scalings have been observed by MicroBooNE [270] and recent Lattice Quantum-Chromodynamics (QCD) calculations [294]. The increase (decrease) of the S_N^{2p2h} (S_Δ^{2p2h} and S_{PL}^{2p2h}) is also crucial to correctly describe MiniBooNE ν_μ and $\bar{\nu}_\mu$ CC π data.

The G20a tune also offers a better description of T2K ND280 CC0p0 π data. This tune suggests a scaling of $S_N^{2p2h} = 1.7 \pm 0.3$. This value is compatible with the results presented by MicroBooNE [270]. In this particular case, the scaling of QEL is around

5.5. TUNING RESULTS

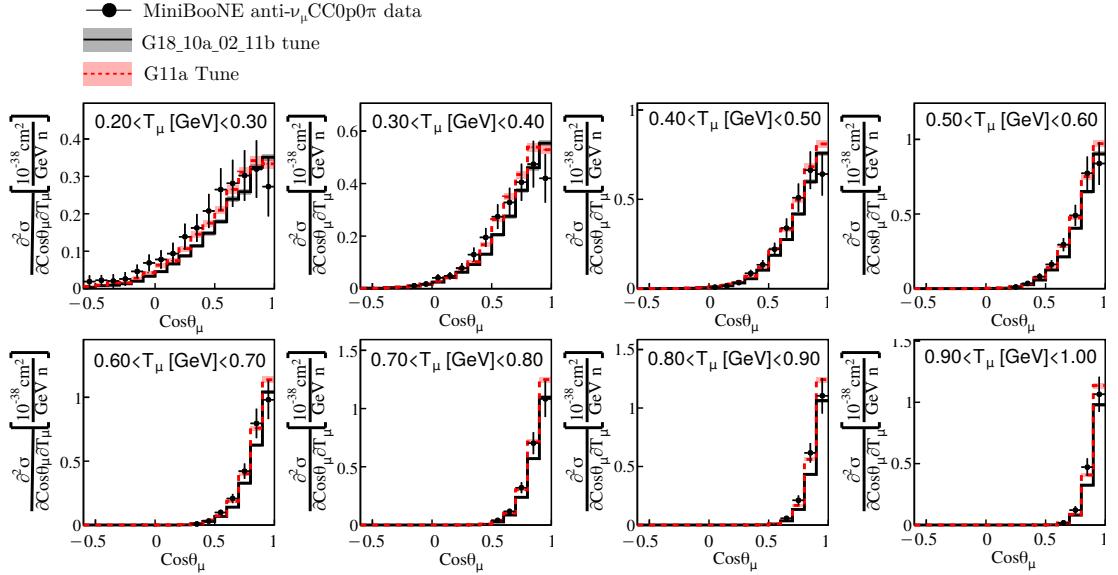


Figure 5.24: Comparison of the G18_10a_02_11b and G11a tunes against MiniBooNE $\bar{\nu}_\mu$ CC0 π double differential data [273]. The comparisons are restricted to the $0.2 < T_\mu < 1.0$ GeV phase space. The predictions are computed using the parameters specified in Tab. 5.6. The total χ^2 associated to this dataset before (after) the tune is reported in Tab. 5.7.

10%. The post-fit value of $\omega_{\text{No RPA}}$, although negative, is compatible with zero. This result is physical as $\omega_{\text{RPA}} + \omega_{\text{No RPA}} > 0$, hence the total cross-section is positive. This scenario can be avoided by reducing the $\omega_{\text{No RPA}}$ range to $[0, 1.5]$. However, the parameter range is not reduced further to allow a valid estimation of the error on $\omega_{\text{No RPA}}$.

Before the tune, the G18_10a_02_11b prediction under-predicted MINERvA CC0 π data in the phase space regions where 2p2h events dominate ($0.15 < p_T < 0.7$). The results suggest that an enhancement of QEL, as well as 2p2h, improves the agreement with data. In fact, the G30a and G31a tunes provide with a better description of ν_μ CC0 π and $\bar{\nu}_\mu$ CC0p0 π data respectively. The improvement in the normalization of the cross section is reflected in the post-fit χ^2 values from Tab. 5.7. The same is true for the cross section as a function of the reconstructed neutrino energy, Fig. 5.29, and single-differential cross section data, Fig. 5.28 and Fig. 5.30. Both tunes over-predict the data at very low Q_{QEL}^2 .

5.5.2 Tension between CC0 π partial tunes

Tensions between datasets can be explored by comparing the different tunes. Fig. 5.31 compares the G10a, G20a and G30a predictions against MiniBooNE ν_μ CC0 π data. Even though the normalization of the three tunes is similar, differences in the predicted cross-section shape exist. The G10a tune is the only one out of the three that

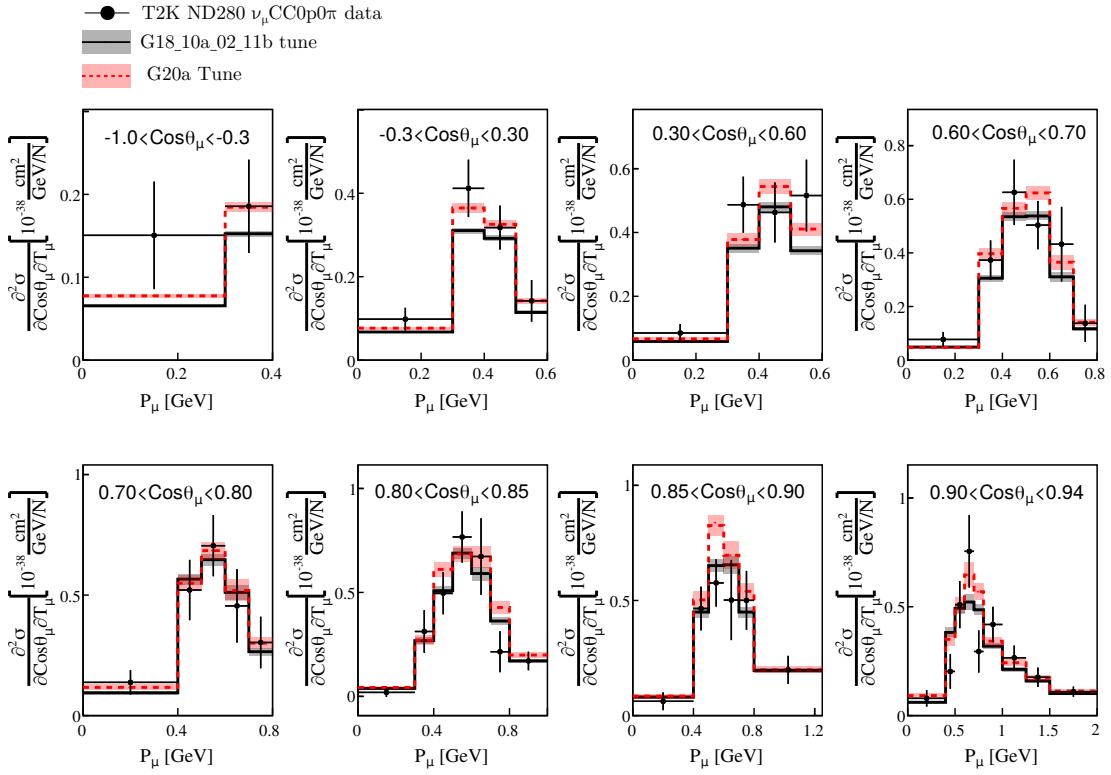


Figure 5.25: Comparison of the G18_10a_02_11b and G20a Tune against T2K ND280 ν_μ CC0p0 π double differential data [119]. The comparisons are restricted to the $-1.0 < \cos \theta_\mu < 0.94$ phase space. The predictions are computed using the parameters specified in Tab. 5.6. The total χ^2 associated to this dataset before (after) the tune is reported in Tab. 5.7.

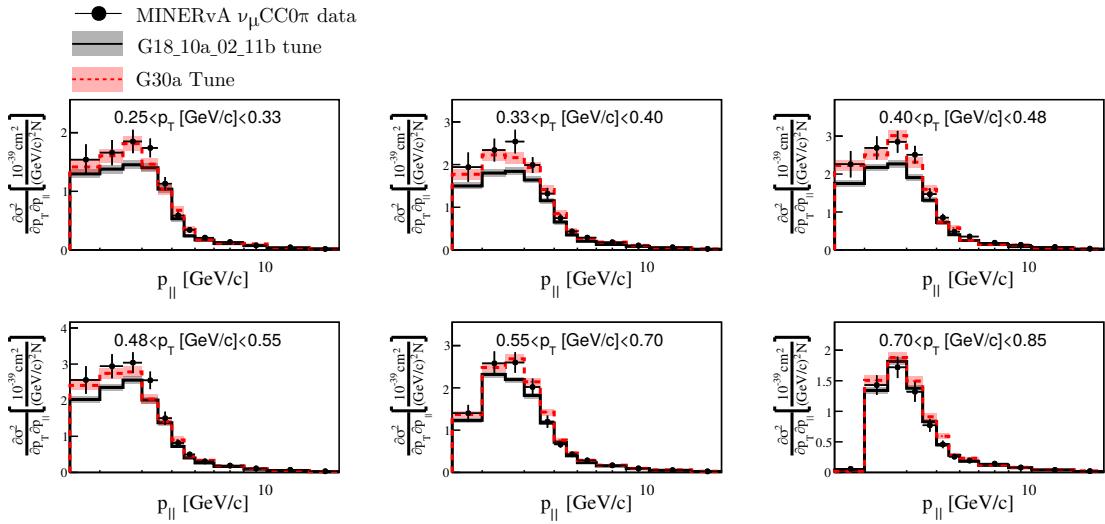


Figure 5.26: Comparison of the G18_10a_02_11b and G30a tunes against MINERvA ν_μ CC0 π double differential data [117]. The comparisons are restricted to the $0.25 < p_T < 0.85$ GeV/c phase space. The predictions are computed using the parameters specified in Tab. 5.6. The total χ^2 associated to this dataset before (after) the tune is reported in Tab. 5.7.

5.5. TUNING RESULTS

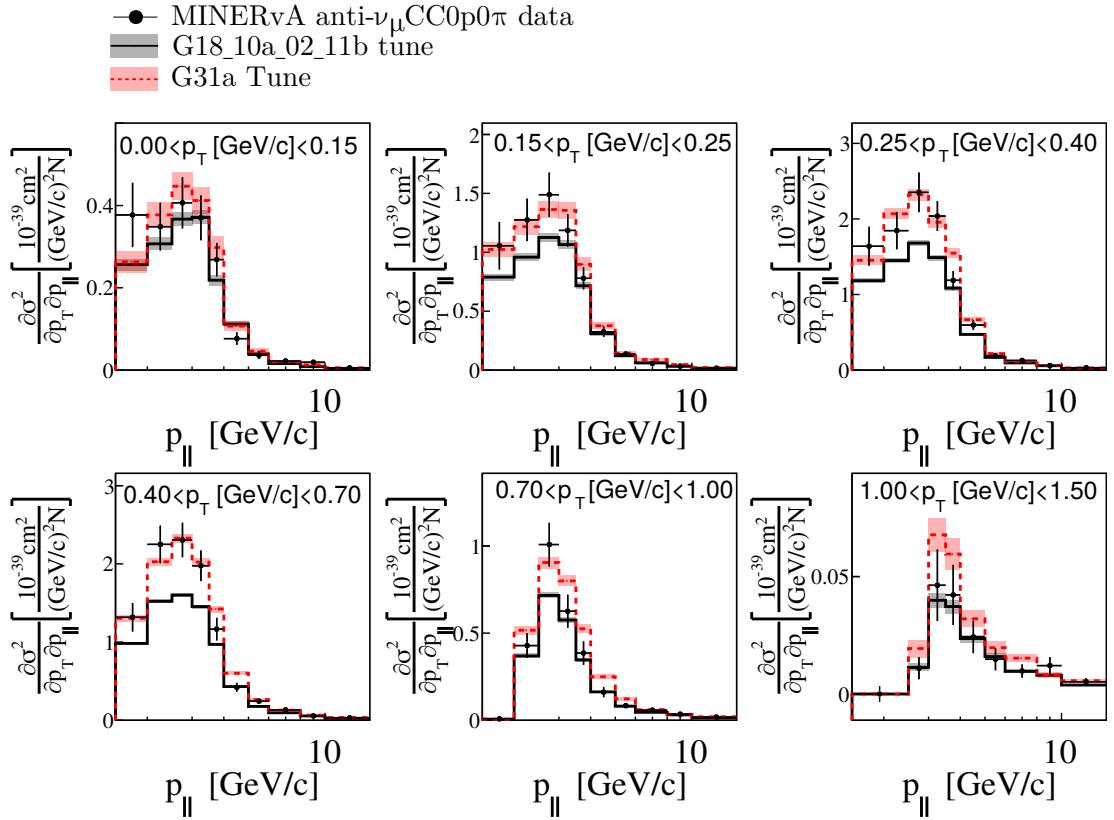


Figure 5.27: Comparison of the G18_10a_02_11b and G31a tunes against MINERvA $\bar{\nu}_\mu$ CC0p0 π double differential data [273]. The comparisons are restricted to the $0 < p_T < 1.5$ GeV/c phase space. The predictions are computed using the parameters specified in Tab. 5.6. The total χ^2 associated to this dataset before (after) the tune reported in Tab. 5.7.

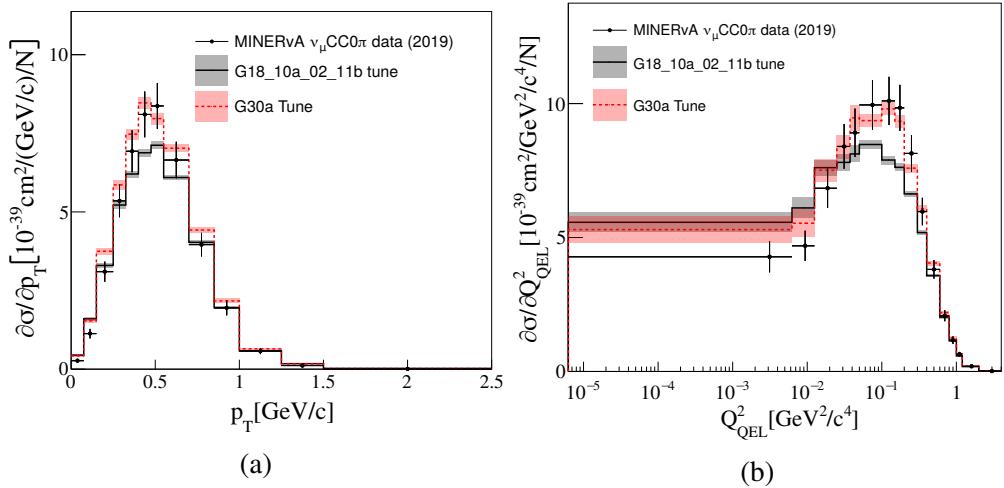


Figure 5.28: Comparison of the G18_10a_02_11b and G30a tunes against MINERvA ν_μ CC0p0 π single-differential data [117]. The predictions are computed using the parameters specified in Tab. 5.6.

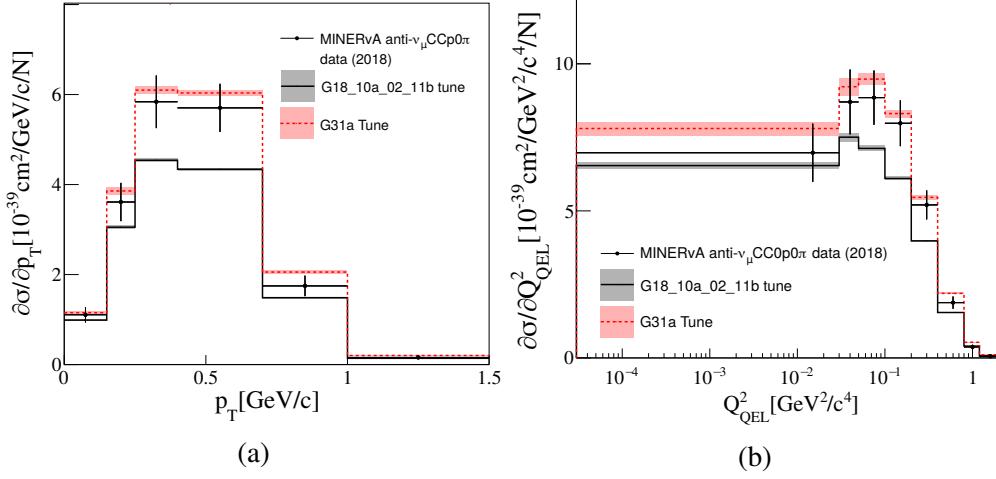


Figure 5.29: Comparison of the G18_10a_02_11b and G31a tunes against MINERvA $\bar{\nu}_\mu$ CC0p0 π single-differential data [273]. The predictions are computed using the parameters specified in Tab. 5.6.

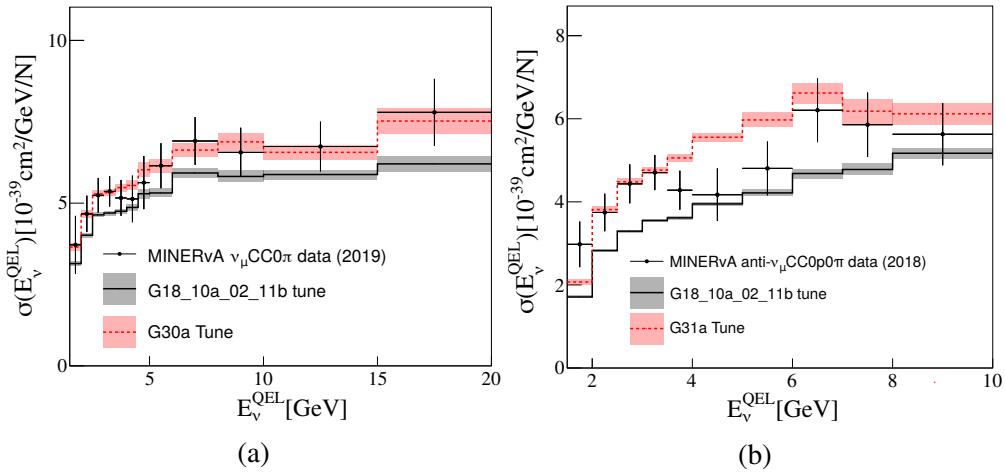


Figure 5.30: Comparison of G18_10a_02_11b, G31a (a) and G31a (b) against MINERvA ν_μ and $\bar{\nu}_\mu$ CC0p0 π integrated cross section data [117, 273]. The predictions are computed using the parameters specified in Tab. 5.6.

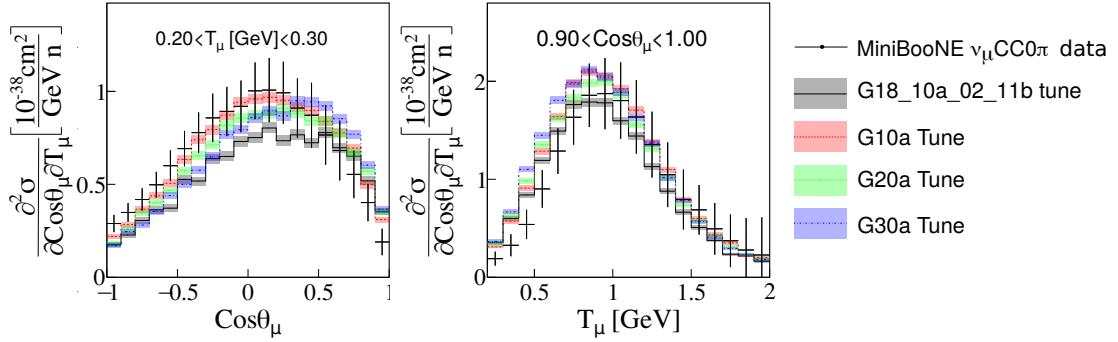


Figure 5.31: Comparison of the G18_10a_02_11b, G10a, G20a and G30a tunes against MiniBooNE ν_μ CC0p0 π integrated cross-section data [99]. The predictions are computed using the parameters specified in Tab. 5.6.

successfully describes the shape of the data, as it can be seen in Fig. 5.31 (left). The other tunes underestimate the cross-section at backward muon angles. In addition, G30a over-predicts the cross section at forward angles as a consequence of the enhancement of the 2p2h cross section at the Δ -region. All tunes overestimate the cross section at forward muon angles and low muon kinetic energies, as demonstrated in Fig. 5.31 (right).

The G31a tune is in clear tension with all the rest, including partial tunes performed with MINERvA neutrino data. The tension is a consequence of a higher QEL scaling, of around 50%, and a higher scaling for 2p2h parameters. These results lead to the over-prediction of all the other datasets. The comparison of G30a and G31a against MINERvA and MiniBooNE ν_μ CC0 π data are shown in Fig. 5.32. The effect of this tension on the χ^2 is reported in Tab. 5.7. This tension may be due to the different topology definitions of both samples, with requires no visible protons above $T_p = 120$ MeV for the antineutrino sample. This tension is not observed for the T2K ND280 dataset, which requires no visible protons above 450 MeV.

5.6 Tensions between ν_μ CC0 π and ν_μ CCNp0 π datasets

T2K ND280, MINERvA and MicroBooNE are the only experiments that released cross-section measurements for different proton multiplicities. The G18_10a_02_11b, tuned to hydrogen and deuterium data, does not have good agreement with CCNp0 π datasets.

After the partial tunes using ν_μ CC0 π data, the agreement with CCNp0 π data deteriorates. This is highlighted in Tab. 5.8, which summarizes the post-fit χ^2 values as

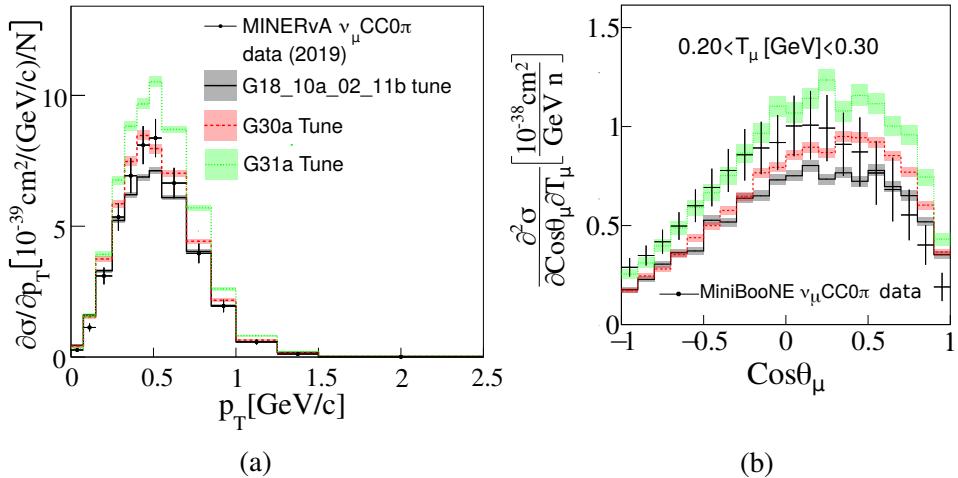


Figure 5.32: Comparison of the G18_10a_02_11b, G30a and G31a tunes against (a) MINERvA [117] and (b) MiniBooNE [99] v_μ CC0p0 π double-differential cross-section data. The predictions are computed using the parameters specified in Tab. 5.6.

sociated to CCNp0 π datasets. In all cases, the χ^2 computed with each partial tune prediction increases with respect to the χ^2 computed with the *nominal* G18_10a_02_11b tune, associated to the G18_10a_02_11b tune.

All G10a, G20a and G30a tunes overpredict v_μ CCNp0 π data. Fig. 5.33 shows a comparison of the partial tune predictions against different single-differential CCNp0 π cross-section measurements from MINERvA. Fig. 5.33 shows that none of the available tunes can describe the peak at low δp_T and that all partial tunes overestimate the cross section at low proton momentum and forward angles. The same observations are made when comparing the tunes against T2K ND280 and MicroBooNE CCNp0 π data, see Fig. 5.34 and Fig. 5.35 respectively.

To further explore this tension, an additional tune is performed using the MINERvA ν_μ CCNp0 π dataset as a function of the proton angle. Following the naming scheme described at the beginning of Sec. 5.5, this tune is referred to as G35a. The best-fit results are listed in Tab. 5.6.

The G35a tune suggests a significant reduction of the QEL cross section. In addition, the tune suppresses the Valencia cross-section peak prediction at the nucleon region and shift the Δ peak to a higher W . This result contradicts the rest of the partial tunes presented in this article, reinforcing the fact that there is a strong tension between CC 0π and CCNp 0π datasets. The summary of χ^2 is reported in Tab. 5.7 and Tab. 5.8.

An important observation is that the G35a tune also improves the agreement with MicroBooNE CCNp π data, suggesting that a possible A-dependency on the parameters does not play an important role.

The tension between CC0 π and CCNp0 π datasets needs to be resolved in order to

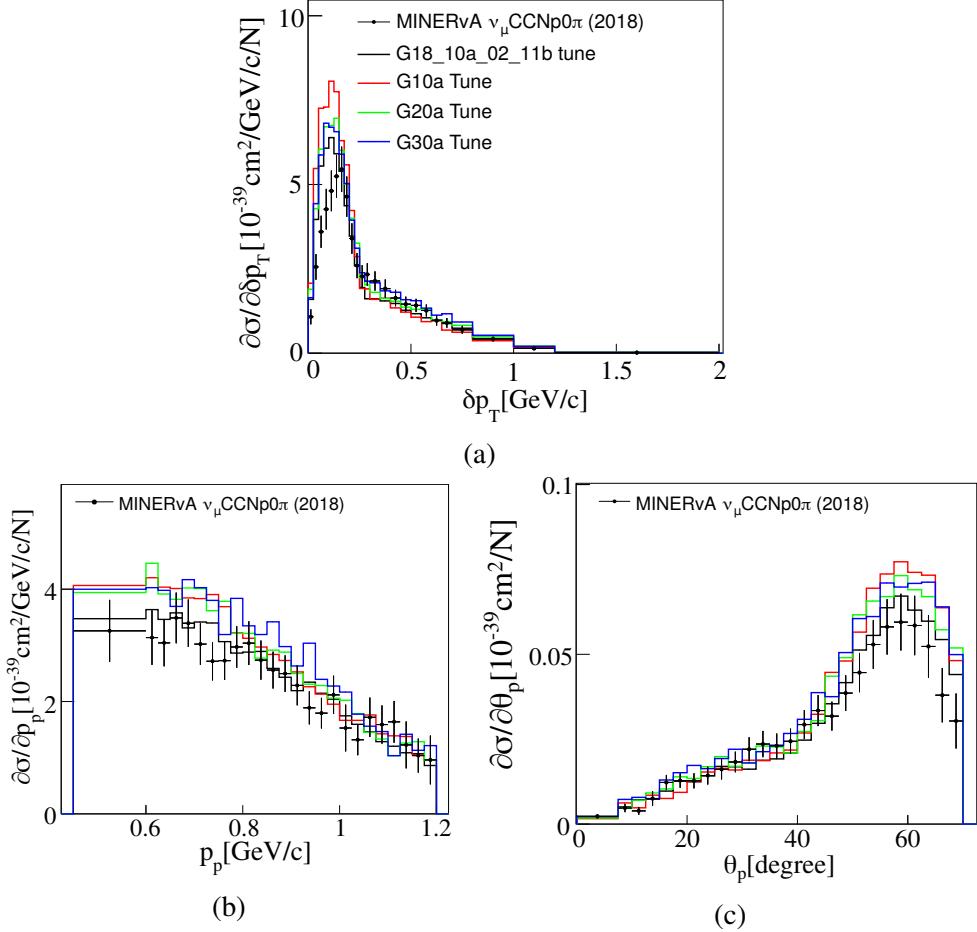


Figure 5.33: Comparison of the G18_10a_02_11b, G10a, G20a and G30a tunes against MINERvA ν_μ CCNP0 π single-differential cross-section data as a function of (a) δp_T , (b) p_p or (c) θ_p [126]. In order to ease the readability of these plots, no statistical errors are shown. The predictions are computed using the parameters specified in Tab. 5.6.

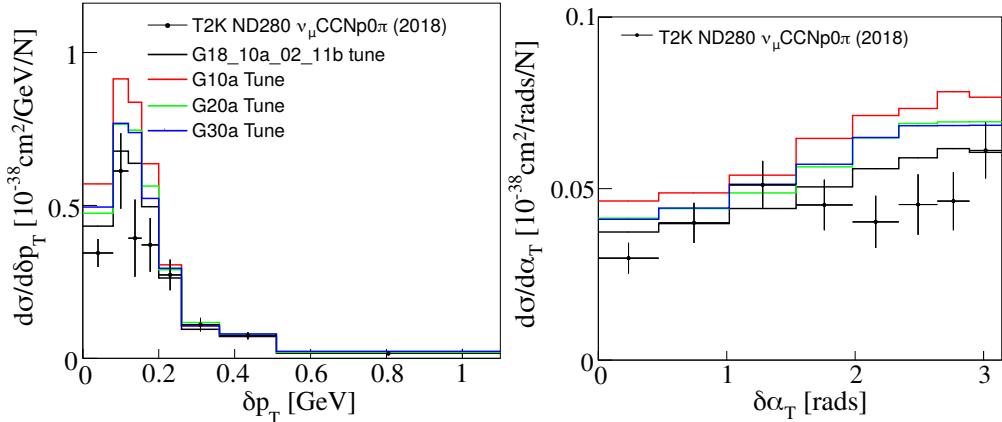


Figure 5.34: Comparison of the G18_10a_02_11b, G10a, G20a and G30a tunes against T2K ND280 ν_μ CCNP0 π single-differential cross-section data as a function of (a) δp_T or (b) $\delta \alpha_T$ [119]. In order to ease the readability of these plots, no statistical errors are shown. The predictions are computed using the parameters specified in Tab. 5.6.

Dataset	χ^2_{Nominal}	χ^2_{G10a}	χ^2_{G11a}	χ^2_{G20a}	χ^2_{G30a}	χ^2_{G31a}	χ^2_{G35a}	DoF
T2K ND280 CCNp π data								
$d\sigma/d\delta p_T$	228	1741	1499	883	759	95	25	8
$d\sigma/d\delta\phi_T$	292	2489	2117	1190	1049	1950	16	8
$d\sigma/d\delta\alpha_T$	27	58	53	42	41	95	21	8
MINERvA CCNp π data								
$d\sigma/dp_p$	21	22	25	32	36	58	27	25
$d\sigma/d\theta_p$	58	153	150	113	129	226	20	26
$d\sigma/d\delta p_T$	102	637	568	360	352	625	42	24
$d\sigma/d\delta\phi_T$	87	505	467	314	354	566	18	23
$d\sigma/d\delta\alpha_T$	15	21	29	24	30	57	17	12
$d\sigma/d\delta p_{Tx}$	159	727	710	467	555	768	62	32
$d\sigma/d\delta p_{Ty}$	127	832	776	553	599	792	51	33
MicroBooNE CCNp π data								
$d\sigma/dp_\mu^{reco}$	71	402	413	245	251	1186	40	10
$d\sigma/d\cos\theta_\mu^{reco}$	413	238	236	210	245	471	149	12
$d\sigma/dp_p^{reco}$	33	96	97	73	76	267	20	10
$d\sigma/d\cos\theta_p^{reco}$	100	176	179	135	139	393	33	9
$d\sigma/d\theta_{\mu p}^{reco}$	549	186	196	199	218	304	136	6

Table 5.8: Summary of χ^2 values associated the CCNp π datasets specified in each row. The χ^2 values are calculated using the NS method for seven different tunes: G18_10a_02_11b, G10a, G11a, G20a, G30a and G31a. The values highlighted in bold correspond to the best-fit χ^2_{G35a} for the partial tune using the specified dataset.

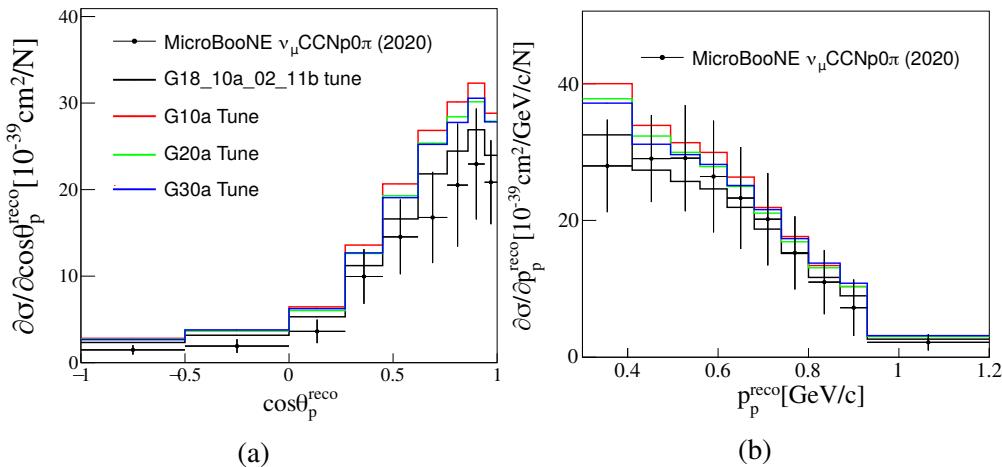


Figure 5.35: Comparison of the G18_10a_02_11b, G10a, G20a and G30a tunes against MicroBooNE v_μ CCNp π single-differential cross-section data as a function of (a) $\cos\theta_p^{reco}$ or (b) p_p^{reco} [275]. In order to ease the readability of these plots, no statistical errors are shown. The predictions are computed using the parameters specified in Tab. 5.6.

perform a global tune of CC0 π data that can describe all data available to date.

5.6.1 Investigation of tensions between CC0 π and CCNp0 π datasets

This section offers an insight into possible modeling uncertainties that may contribute to the tension between CC0 π and CCNp0 π datasets and explores avenues of accommodating both within future joint tunes. None of the uncertainties described in this section has a big impact on CC0 π datasets.

Nuclear model variations

The nuclear model determines the momentum and binding energy of the hit nucleon. In GENIE, three nuclear models are available: Relativistic Fermi gas (RFG), Local Fermi Gas (LFG) and Correlated Fermi gas (CFG) [40]. By default, G18_10a_02_11b uses the LFG.

The nuclear model choice affects the CCNp0 π predictions. Fig. 5.36 shows the impact of the underlying unclear model against CCNp0 π single-differential cross-section measurements as a function of δp_T . Differences between the models are significant for the cross-section peak prediction at low p_T . The RFG model is the only one out of the three that predicts the MINERvA data below the maximum. However, it still over-predicts the cross section at the peak. Alternatively, the CFG model successfully predicts the peak normalization. This is reflected in the χ^2_{CFG} , reported in Tab. 5.9.

The main characteristic of the RFG and the CFG implementations in GENIE is that nucleons can have a momentum above the Fermi momentum in its ground state. This tail in the momentum distribution is a consequence of nucleon correlations in the nuclear medium. As a consequence of including those effects in the nuclear model, the description of the tail of the δp_T distribution improves. This study suggests that using a more elaborate nuclear model is required to describe CCNp0 π measurements.

The differences between the three GENIE nuclear model predictions are not enough to explain the discrepancy between CC0 π and CCNp0 π data: all models predict a higher cross section for processes with protons in the final state concerning those with no protons in the final state. This is highlighted in Fig. 5.37.

Nucleon Final State Interaction model variations

Mismodeling of nucleon FSI can cause migration between CC0p0 π and CCNp0 π samples [295]. Ref. [296] suggests increasing nucleon FSI strength in cascade models

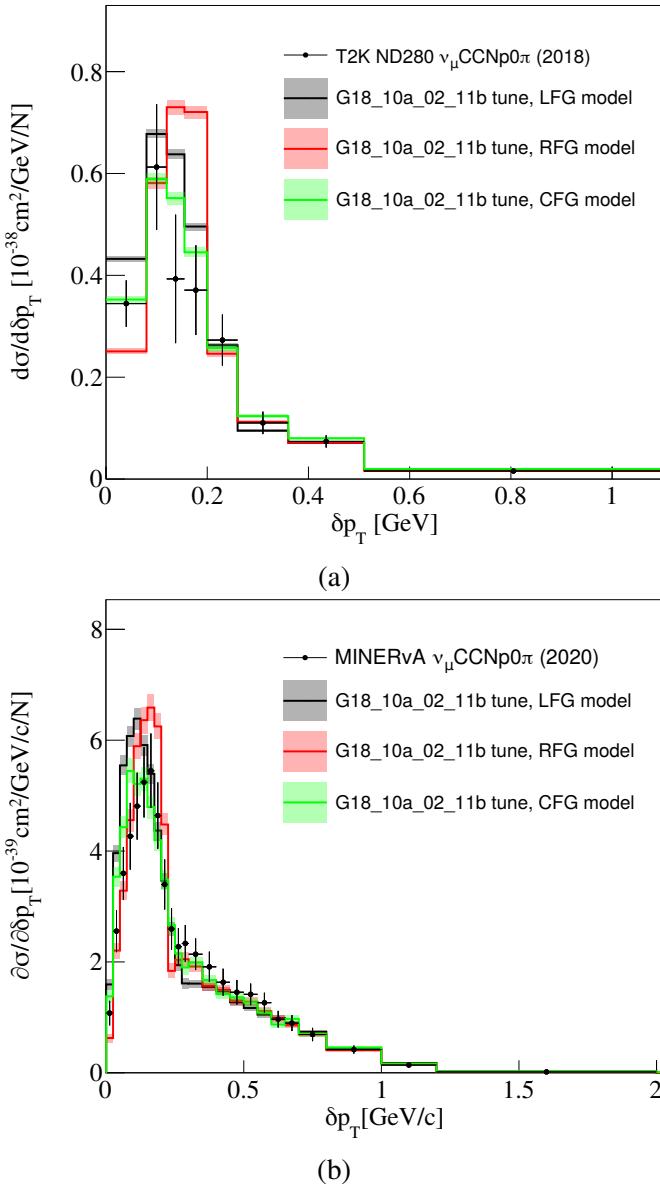


Figure 5.36: Comparison of the G18_10a_02_11b tune against (a) T2K ND280 [119] and (b) MINERvA [274] v_μ CCNp0 π single-differential cross-section data as a function of $\delta\alpha_T$. Three calculations are shown for different nuclear models: LFG (black), RFG (red) and CFG (green). The errors correspond to statistical uncertainties only.

Dataset	G18_10a_02_11b			
	χ^2_{LFG}	χ^2_{RFG}	χ^2_{CFG}	DoF
T2K ND280 CCNp0 π data				
$d\sigma/d\delta p_T$	228	149	27	8
$d\sigma/d\delta\phi_T$	292	29	20	8
$d\sigma/d\delta\alpha_T$	27	25	26	8
MINERvA CCNp0 π data				
$d\sigma/dp_p$	21	23	15	25
$d\sigma/d\theta_p$	58	35	34	26
$d\sigma/d\delta p_T$	102	95	31	24
$d\sigma/d\delta\phi_T$	87	32	18	23
$d\sigma/d\delta\alpha_T$	15	17	14	12
$d\sigma/d\delta p_{Tx}$	159	61	48	32
$d\sigma/d\delta p_{Ty}$	127	40	42	33
MicroBooNE CCNp0 π data				
$d\sigma/dp_\mu^{reco}$	71	35	32	10
$d\sigma/d\cos\theta_\mu^{reco}$	413	137	123	12
$d\sigma/dp_p^{reco}$	33	25	27	10
$d\sigma/d\cos\theta_p^{reco}$	100	49	42	9
$d\sigma/d\theta_{\mu p}^{reco}$	549	195	155	6

Table 5.9: Summary of χ^2 values associated the CCNp0 π datasets specified in each row. The χ^2 values are calculated using the NS method for three GENIE predictions. The GENIE predictions are calculated with the G18_10a_02_11b tune. Each prediction uses a different nuclear model: RFG, LFG or CFG.

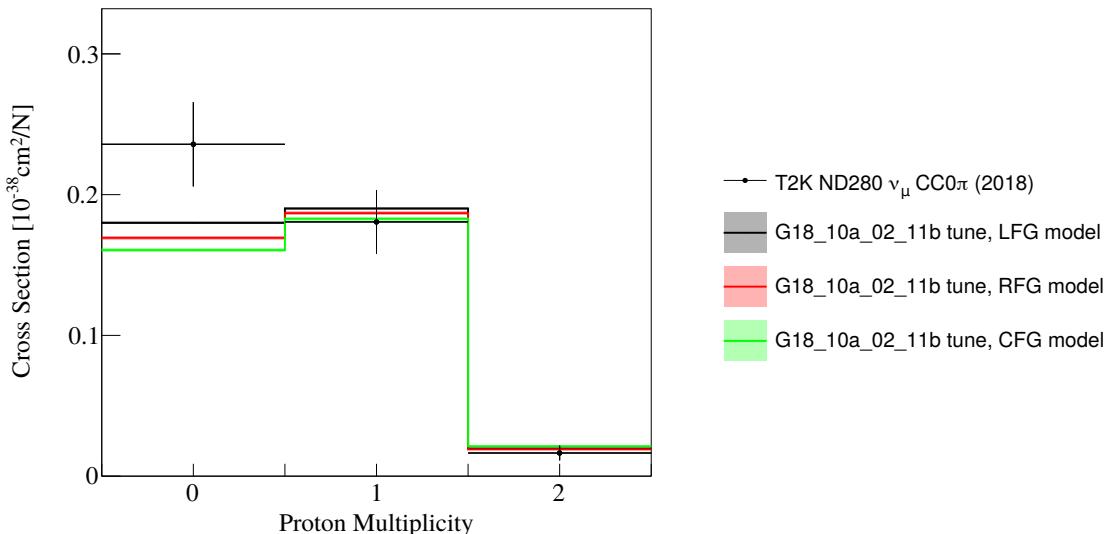


Figure 5.37: Comparison of G18_10a_02_11b against T2K ND280 ν_μ CCNp0 π total cross section data as a function of the proton multiplicity [119]. Three calculations are shown for different nuclear models: LFG (black), RFG (red) and CFG (green).

might improve the agreement with CCNp 0π data from T2K ND280 and MINERvA. This possibility is explored in this section.

A crucial test for FSI models is to be able to reproduce nuclear transparency data from electron-scattering experiments. Transparency is defined as the probability for the knocked-out nucleon to not suffer FSI in the nuclear environment and it can be measured using noninteracting probes, such as electrons. In transparency measurements, the final-state nucleon is produced inside the nucleus. This feature is common with neutrino experiments, making transparency data extremely valuable to characterize and test FSI modeling uncertainties. Unfortunately, nuclear transparency measurements are scarce. Few data points on proton transparency on carbon as a function of the proton momentum are available in Ref. [297–300].

Transparency can be easily calculated within MC event generators as a ratio between the distribution of final-state protons which did or did not rescatter while leaving the nuclear environment. Ref. [295] provided with the first direct comparison of transparency calculations using a neutrino event generator. In their analysis, they take into account in the transparency definition the experimental acceptances of the electron-scattering experiments. Such analysis could be replicated in GENIE; however, it is out of the scope of this work. To be able to compare GENIE’s transparency calculation with data, we scale the GENIE predictions by the ratio between the transparency prediction from Ref. [295] with and without acceptance cuts. This approach was used in Ref. [143].

Fig. 5.38 compares the corrected transparency predictions for G18_10a_02_11b with the available data. The red and blue bands show the effect on the predictions when scaling up and down the nucleon mean-free path by 10% and 30% respectively. The 10% variation describes the data points with proton kinetic energies above 600 MeV within the 1σ error bound. The 30% variation covers all the data available.

The impact of the mean-free path uncertainties is also shown in Fig. 5.39. It is observed that a higher nucleon mean-free path results in an increase of the proton multiplicity. An interesting observation is that when reducing the mean-free path, a higher cross-section is predicted for events with no protons above detection threshold, indicating a reduction of the mean-free path could be key to resolve the tension between CC 0π and CCN π datasets.

Ref. [143] compares the different FSI models implemented in MC event generators. These differ mainly in the implementation of nuclear effects, such as Pauli Blocking and nucleon-nucleon correlations, in FSI codes. They observed that at high kinetic energies, differences between the calculations do not play a major role. However, for $T_p < 200$ MeV, the spread between the different transparency calculations is signifi-

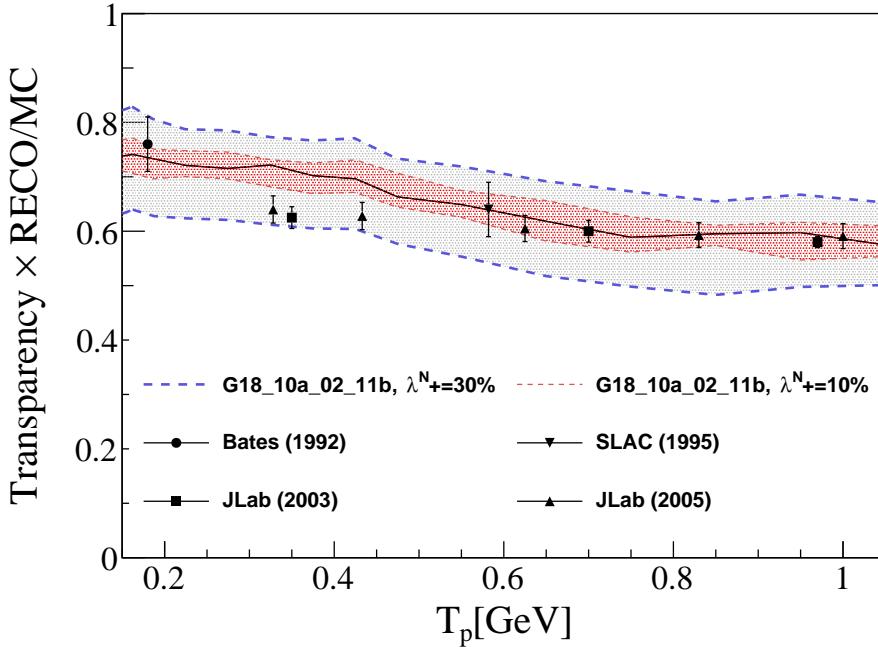


Figure 5.38: Comparison of proton transparency in carbon against the available electron-scattering data [297–300]. The GENIE prediction (black) is calculated using the G18_10a_02_11b. The error bands correspond to the expected uncertainty when varying the nucleon mean-free path by 10% (red) and 30% (grey). The predictions have been corrected according to experiments acceptance effects as determined in Ref. [295].

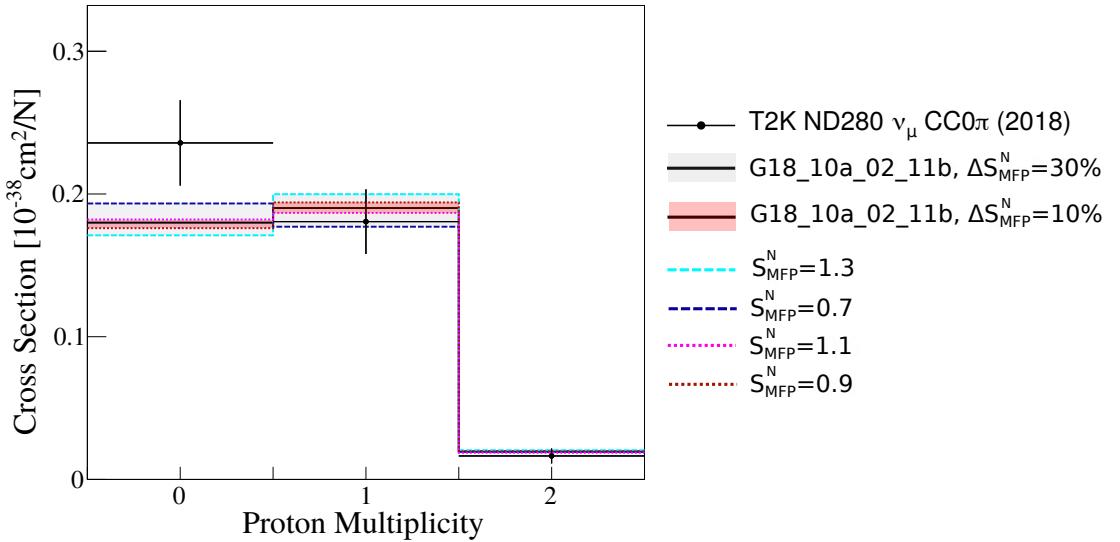


Figure 5.39: Comparison of the G18_10a_02_11b tune against T2K ND280 ν_μ CC0 π total cross-section data as a function of the proton multiplicity. The GENIE prediction (black) is calculated using the G18_10a_02_11b. The error bands correspond to the expected uncertainty when varying the nucleon mean-free path by 10% (red) and 30% (grey).

cant, indicating a stronger model dependency at low proton kinetic energies.

The T2K ND280 and MINERvA ν_μ CCNp 0π measurements require at least one proton above the detection threshold of $p_p \sim 450$ MeV/c. Therefore, FSI modeling uncertainties, which affect transparency calculations at low energies, are not expected to have a big impact on MINERvA and T2K ND280 CCNp 0π measurements. These could affect MINERvA $\bar{\nu}_\mu$ CC $0p0\pi$ measurements, which require no protons with momentum above 120 MeV/c. However, variations of the mean-free path have a small effect on MINERvA $\bar{\nu}_\mu$ CC $0p0\pi$ predictions.

The hA and hN FSI models are build on a simplistic view of the nuclear environment. More complex approaches which offer an improved description of CCNp 0π [119, 301, 302]. Another possible line of study would be to determine whether more elaborated can solve the tension. This study is out of the scope of this paper but will be explored in future iterations of this work.

5.7 Conclusions

This article describes the first neutrino-nucleus cross-section tuning effort within the GENIE Collaboration. The goal of this work is to tune GENIE against CC 0π data from either MiniBooNE, T2K ND280 or MINERvA using the same analysis procedure.

This tune incorporated new challenges which lead to important changes in the GENIE tuning software. To tune against CC 0π data, new parametrizations that encapsulate possible nuclear uncertainties are developed. These are now viable for wider use in GENIE. In addition, modern nuclear data provides the full correlation between the data release bins due to systematic uncertainties. This information, which was missing in previous tuning efforts, is incorporated in this analysis. This required a modification of the χ^2 definition to avoid the Peele's pertinent puzzle, which leads to nonphysical normalization factors. The *nominal* CMC used in the analysis is the G18_10a_02_11b, which was previously tuned against hydrogen and deuterium data. The correlation between parameters from the free-nucleon tune, provided in Ref. [217], is included in this analysis.

In total, five partial tunes using ν_μ or $\bar{\nu}_\mu$ CC 0π data on carbon are reported. Each partial tune is performed using data from either MiniBooNE, T2K ND280 or MINERvA. The focus is on tuning CC 0π datasets which are reported as a function of the muon kinematics. Whenever these datasets are not available, the tune is performed with CC $0p0\pi$ data.

All tunes increase the QEL and 2p2h cross section and there is a preference for the QEL cross section with RPA corrections. A clear energy dependence is observed in

5.7. CONCLUSIONS

the 2p2h parameters: the G10a, G11a, G20a tunes have a preference for a higher 2p2h cross section at the nucleon region, while suppressing it at the Δ region. The G30a and G31a tunes enhance the cross section in both regions, with a higher scaling factor at the Δ region. This empathises a dependency on the neutrino energy.

Tensions between the tunes exist. In particular, the G31a tune is in clear disagreement with the rest. This tune is performed using $\bar{\nu}_\mu$ CC0p0 π data, which requires no protons above 120 MeV. This analysis is the first hint of an underlying tension between CC0 π and CCNp0 π data in this work. This tension is further explored by comparing the partial tune results against CCNp0 π data from T2K ND280, MINERvA and MicroBooNE. In all cases, the G10a, G20a and G30a tunes over-predict CCNp0 π data. This tension is studied further with the G35a tune, which is performed using MINERvA CCNp0 π data. In this case, the result suggests decreasing the QEL and 2p2h cross section, contradicting the previous results. However, the G35a tune offers a much better description of all CCNp0 π data, including MicroBooNE’s data. This suggests that a possible A-dependency on the parameters is small.

The tension observed between CC0 π and CCNp0 π data cannot be ignored and it needs to be addressed before attempting to perform a global tune with all the available data. Two preliminary studies which aim to characterize possible modeling uncertainties for CCNp0 π data are also presented. The first one considers different nuclear models, suggesting that a more elaborated nuclear model is required to improve the agreement with CCNp0 π data. In particular, the agreement with the data improves significantly when the CFG model is used. The second study considers possible sources of uncertainties in the modeling of nucleon FSI and compares the results with transparency data. From this study it is concluded that a 30% decrease of the nucleon mean-free path, which effectively reduces the transparency of protons in the nuclear environment, would improve the agreement with both types of datasets. However, such variation is not supported by data. This discrepancy could also be caused by the simplistic nuclear model used in the *hA* implementation. This will be studied in future iterations of this work.

Chapter 6

Concluding remarks

Modern neutrino experiments strive to understand and control the systematic uncertainties associated with neutrino interactions. Unfortunately, current neutrino event generators lack a complete theory to model neutrino interactions at the energies of interest for neutrino experiments. In order to simulate neutrino interactions with all targets, event generators blend theory with empirical approaches.

GENIE has released several comprehensive model configurations which provide the most predictive, robust, and self-consistent model that are built by combining the different GENIE modeling elements with empirical models. Each of these configurations must be validated, characterized and tuned as a whole in order to address possible double-counting issues or disagreements with data due to limitations of the models.

Previous tuning efforts, mostly lead by neutrino experiment collaborations external to neutrino event generators, relied on reweighting procedures; these are limited to simple, reweightable, degrees of freedom that do not correspond to the underlying model parameters used in the event generation. In addition, this approach cannot be used to tune nonreweightable models, such as hadronization models.

The GENIE global analysis offers the most powerful tuning program in the neutrino community to date. One of the main benefits is that the GENIE approach is fully decoupled from event reweighting procedures. In addition, whilst previous efforts were external to event generators, this analysis effort is developed by the team of scientists maintaining the GENIE event generator itself. Hence, it provides a systematic analysis procedure that can be repeated as the models evolve and easily deploy the new tunes to the neutrino community through the GENIE platform. In addition, models can be developed further in response to findings from tuning. The ultimate goal of the GENIE global analysis effort is to tune all aspects relevant to the modeling of neutrino scattering on nuclei, such as bare-nucleon interaction, multi-nucleon emission, nuclear effects, final-state interactions or hadronization processes.

This thesis focuses on the development of the GENIE global analysis of scattering data for the tuning and uncertainty characterization of comprehensive model configurations. It describes the first three analyses performed within the GENIE tuning program. All the analyses presented in my thesis were possible through the continued development curation of data archives, the development of the GENIE Comparisons software for the evaluation of the corresponding predictions and their interface to the Professor tool, which enabled the efficient implementation of complex multi-parameter brute-force scans.

Chapter 3 describes the retune of the bare-nucleon cross-section model for all GENIE comprehensive models available. The bare-nucleon analysis focuses on the tuning of the shallow-inelastic scattering region using hydrogen and deuterium data from ANL 12 ft, BNL 7 ft, BEBC and FNAL 15 ft bubble chamber experiments. GENIE models the nonresonant background by extrapolating the deep-inelastic scattering contribution to the resonant region. The nonresonant background contribution needs to be tuned to data to avoid the double counting of the resonant cross-section. Previous GENIE tuning attempts failed to describe both inclusive and exclusive channels due to unresolved tensions with historical data from bubble chamber experiments. The tensions highlighted the need to improve the tuning procedure with a systematic treatment of correlations between datasets.

The bare-nucleon analysis focused on the tuning of the shallow-inelastic scattering region against ν_μ and $\bar{\nu}_\mu$ CC inclusive, quasielastic, one-pion and two-pion integrated cross sections as a function of E_ν . The global fit successfully describes both inclusive and exclusive cross sections simultaneously. The effect of the tune is to decrease the inclusive cross section at the 1–10 GeV energy region with respect to the historical *default* predictions. In terms of exclusive predictions, the single-pion production cross section on free nucleons is reduced as a consequence of a decrease of the nonresonant background component. On the contrary, two-pion production mechanisms are enhanced. This tune offers with an improved description of hydrogen and deuterium inclusive and exclusive data and it is used as starting point for all the other analyses presented in this thesis.

Hadronization uncertainties are also relevant for the modeling of pion production processes. Chapter 4 presents the first GENIE tune of the AGKY hadronization model. The goal of this analysis is to improve the agreement with neutrino charged averaged multiplicity data on hydrogen and deuterium from FNAL 15 ft and BEBC and provide the first data-driven uncertainties for hadronization to date. The quality of neutrino hadron shower data is the main limiting factor of this analysis. Tensions between hydrogen and deuterium data are observed. To further explore these tensions, two tunes

are performed. A global tune that aims to improve the agreement with both types of datasets, AGKY 2021 global tune, and a deuterium-only tune to investigate the tensions. The main effect of both tunes is to increase the averaged charged multiplicity for $W > 3$ GeV/c^2 , modeled with PYTHIA. The low- W region is also affected but constraints due to energy, momentum, charge, baryon number and strangeness conservation laws reduce the available phase space and the effect of the tuning procedure. The 2021 AGKY global tune under-predicts the deuterium data at the PYTHIA region whereas the deuterium-only tune over-predicts hydrogen data. Despite the tensions, the global tune has a better agreement with the charged averaged multiplicity data and provides the first data-driven analysis of this kind using neutrino interactions. The results from this analysis are an important input for systematic studies of modern neutrino experiments. This tune is an excellent example of the GENIE analysis capabilities, as it is the first analysis of nonreweightable models.

Ideally, the shallow-inelastic scattering region and neutrino hadron shower models should be tuned together as the multiplicity of nonresonant processes is determined by the hadronization model. This correlation was neglected in the bare-nucleon analysis aiming to reduce the computational power required to run the tune. Whilst variations on the shallow-inelastic scattering cross section are not relevant at energies of interest for the hadronization tune, the effect of the hadronization tune on inclusive and exclusive pion production processes is not negligible. In fact, the 2021 GENIE AGKY global tune increases the two-pion production cross section, losing the agreement with CC inclusive data. This tension can be resolved by tuning the shallow-inelastic scattering region and hadronization models altogether and it will be a subject for future work. Despite this fact, the information on the 2021 GENIE AGKY global tune systematic uncertainties is very valuable for neutrino experiments interested in the $W < 2 \text{ GeV}/c^2$ region.

Finally, Chapter 5 describes the first neutrino-nucleus cross-section tuning effort within the GENIE Collaboration. Whilst several attempts to tune GENIE against neutrino-nucleus data exist, this is the first analysis of neutrino-nucleus scattering data from different experiments using the same analysis method. The goal of this analysis is to explore avenues for improving the agreement with MiniBooNE, T2K ND280 or MINERvA CC0 π data, consolidate the main elements of the CC0 π tuning analysis and investigate possible tensions between the different experiments. In total, five partial tunes using ν_μ or $\bar{\nu}_\mu$ CC0 π data on carbon are reported. Whilst all the partial tunes have a preference to increase the 2p2h and QEL cross-section, differences between the best-fit predictions exist. The main observation is a clear preference from MiniBooNE and T2K data for a higher 2p2h cross section at $W = 938 \text{ GeV}/c^2$. Alternatively, MIN-

ERvA data prefers to scale it at higher W . This suggests a possible dependency on the neutrino energy for the 2p2h scaling. In addition, the tunes highlight a tension between datasets that require a different number of protons in the event. In particular, GENIE predictions under-predict CC0p0 π data, whilst CCNp0 π data is over-predicted. Preliminary studies suggest that a more elaborate nuclear model is required to describe CCNp0 π measurements. Decreasing nucleon transparency can also help resolve this tension, but this is not supported by transparency data. This tension will be the main area of study in future analysis, which will aim towards a global tune of CC0 π and CCNp0 π data.

Future GENIE tunes will be built on top of this successful paradigm. Some examples of possible lines of work are joint tunes with CC0 π and CC1 π data, a more systematic tuning of final-state interactions with hadron-nucleus data or the tuning of the vector part of neutrino interaction models with electron-scattering data. Many more possibilities will arise as new data and models become available.

List of acronyms

MC	Monte Carlo
SM	Standard model
CC	Charged-current
NC	Neutral-current
CKM	Cabibbo-Kobayashi-Maskawa
QEL	Quasielastic
RES	Resonant
NRB	Nonresonant background
DIS	Deep-inelastic scattering
SIS	Shallow-inelastic scattering
LS	Llewellyn-Smith model
RS	Rein-Seghal model
BS	Berner-Seghal model
MK	Monireh Kabirnezhad model
AGKY	Andreopoulos-Gallagher-Kehayias-Yang
QCD	Quantum-Chromodynamics
BY	Bodek and Yang model
PWIA	Plane wave impulse approximation
FSI	Final-state interactions
RFG	Relativistic Fermi gas
LFG	Local Fermi gas
CFG	Correlated Fermi gas
1p1h	One-particle-one-hole

RPA	Random-phase approximation
2p2h	Two-particles-two-holes
INC	Intra-nuclear cascade
MFP	Mean-free path
CMC	Comprehensive model configuration
KNO	Koba-Nielsen-Olesen scaling law
BEBC	Big European Bubble Chamber
EMI	External-muon identifier
LPS	Longitudinal phase-space model
STKI	Single-transverse kinematic invariance
BNB	Booster Neutrino Beam
NuMI	Neutrino at the Main Injector
PID	Particle identification
LArTPC	Liquid Argon Time Projection Chamber
PPP	Peele's pertinent puzzle
NS	Norm-shape

Appendices

Appendix A

Neutrino-nucleon cross-section model tuning in GENIE v3

A.1 Comprehensive model configuration naming convention

A comprehensive model configuration is identified by at least 7-character string in the form:

Gdd_MMv

where

- G is a capital letter string of arbitrary length that identifies the authors of the tune (GENIE).
- dd is a number describing the year during which the model configuration was first developed.
- MM is a number (00, 01, 02, ...) identifying a family of model configurations.
- v is a character (a, b, c, ...) enumerating different members of the given family of model configurations.

Once a comprehensive model configuration is defined, a number of different tunes may be produced. These may be produced, for example, by a) incorporating different combinations of experimental data, b) considering variations in different combinations of our modeling elements (e.g. bare-nucleon cross sections, nuclear model and nuclear cross sections, neutrino-induced hadronization etc), c) considering different degrees of freedom (different parametrizations) for the variation of each model, or d) incorporating different parameter priors and/or different strategies for eliminating nuisance

parameters. A tune is identified by the model configuration name, and additional information enumerating the parameters and datasets. This is at least 14-character string in the form:

Gdd_MMv_PP_xxx

where

- Gdd_MMv describes the model configuration (see above).
- PP is a number identifying the set of tuned parameters. This parameter set is defined uniquely only in the context of a particular model configuration.
- xxx is a number that identifies the dataset used for the model configuration tuning. This may include a unique set of weights associated with each component dataset.

A.2 Tunes for z-expansion CMCs

There is a group of CMCs that are derived from the G18_10[a-d] ones, by replacing the dipole axial form factor, used in the calculations of quasi-elastic cross sections, with the better-motivated z-expansion model [30], providing a richer set of degrees of freedom for parametrizing quasielastic model uncertainties. They are labelled G18_10[i-l]. As in all previous families of models, four comprehensive model variations (i-l) are constructed by using alternative intranuclear hadron transport models (i: INTRANUKE/hA, j: INTRANUKE/hN, k: GEANT4/Bertini, and l: INCL++).

A specific tune has not been performed by the GENIE collaboration for these CMCs, yet their tuned versions are available in the generator. The released tunes (G18_10[i-l]_02_11b) are constructed by applying all the values obtained in the fit, Tab. 3.7 (G18_02a column), with the exception of M_A^{QE} . Instead, the z-expansion parameters provided in [30] are used for the QEL axial form factor. This is considered good enough for the time being, since the fit targets specifically the SIS region and the uncertainty of the QEL axial form factor is only used to improve the fitting procedure. In addition, the M_A^{QEL} fit result is not far from its initial value.

In the future we expect to be able to release a dedicated tune for these CMCs that uses the full z-expansion uncertainty from [30] as a prior in a similar exercise as described in this paper.

Appendix B

Hadronization model tuning in GENIE v3

B.1 Dataset compatibility study

The tensions highlighted in the paper were investigated to understand if their source could be caused by a specific dataset or analysis procedure. In order to do that, we performed a series of tunes using all the data used in the AGKY 2021 global fit leaving one dataset out at a time. For each fit, we plotted the parameter best fit value, see Fig. B.1. Most of the partial fits results are compatible with the AGKY 2021 global tune predicted values. There are only two datasets in disagreement with the rest: BEBC,1 and 2 (hydrogen dataset) and FNAL,1 (deuterium dataset)¹. Those datasets use different targets, they come from different experiments, and they were analyzed in different years. Hence, the cause of the tension cannot be due to a specific dataset or experiment.

¹Please note that BEBC 1 is only one point, which is why we count it together with BEBC 2.

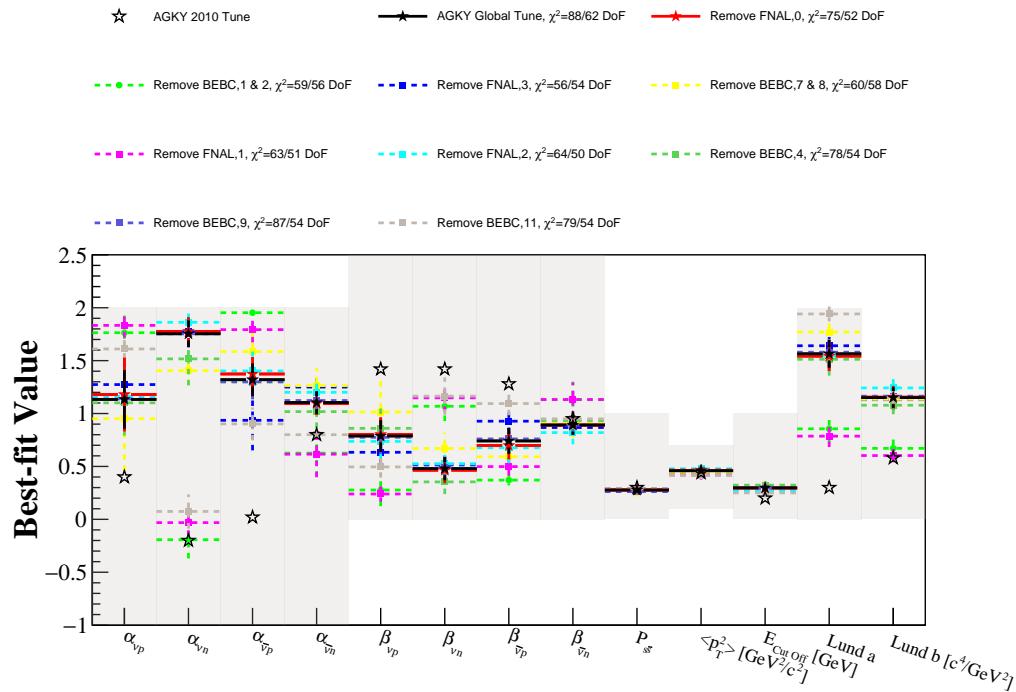


Figure B.1: Summary of best-fit values for a series of partial tunes obtained by using all the data from the AGKY 2021 global tune leaving one dataset out at a time. The result of the fit when removing a H dataset is shown in circles, and squares when removing a ${}^2\text{H}$ dataset. The χ^2 for each tune is listed in the legend. The grey area represents the available range for each parameter. The AGKY 2021 global tune best fit result is shown in filled black stars. The previous GENIE parameters are shown in empty stars (AGKY 2010 tune). The plotted errors are the square roots of the covariance diagonal elements.

Bibliography

- [1] M. Sajjad Athar and S. K. Singh. *The physics of neutrino interactions*. Cambridge University Press, 2020. DOI: 10.1017/9781108489065.
- [2] B. Abi et al. “Long-baseline neutrino oscillation physics potential of the DUNE experiment”. In: *The European Physical Journal C* 80.10 (Oct. 2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-08456-z.
- [3] B. Abi et al. “The DUNE far detector interim design report volume 1: physics, technology and strategies”. In: (July 2018).
- [4] Teppei Katori and Marco Martini. “Neutrino–nucleus cross sections for oscillation experiments”. In: *J. Phys. G* 45.1 (2018), p. 013001. DOI: 10.1088/1361-6471/aa8bf7.
- [5] L. Alvarez-Ruso et al. “NuSTEC White Paper: Status and challenges of neutrino–nucleus scattering”. In: *Prog. Part. Nucl. Phys.* 100 (2018), pp. 1–68. DOI: 10.1016/j.ppnp.2018.01.006.
- [6] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.
- [7] Abdus Salam. “Weak and Electromagnetic Interactions”. In: *Conf. Proc. C* 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.
- [8] S. L. Glashow, J. Iliopoulos, and L. Maiani. “Weak Interactions with Lepton–Hadron Symmetry”. In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285.
- [9] Y. Fukuda et al. “Evidence for oscillation of atmospheric neutrinos”. In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562.
- [10] Carlo Giunti and Chung W. Kim. *Fundamentals of neutrino physics and astrophysics*. Oxford Scholarship Online, 2010. DOI: DOI : 10.1093/acprof:oso/9780198508717.001.0001.

BIBLIOGRAPHY

- [11] Makoto Kobayashi and Toshihide Maskawa. “CP-violation in the renormalizable theory of weak interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652.
- [12] J. H. Christenson et al. “Evidence for the 2π decay of the K_2^0 Mmeson”. In: *Phys. Rev. Lett.* 13 (1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138.
- [13] M. Tanabashi et al. “Review of particle physics”. In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.
- [14] F. Reines et al. “Detection of the free antineutrino”. In: *Phys. Rev.* 117 (1960), pp. 159–173. DOI: 10.1103/PhysRev.117.159.
- [15] F.J. Hasert et al. “Search for elastic muon-neutrino electron scattering”. In: *Physics Letters B* 46.1 (1973), pp. 121–124. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(73\)90494-2](https://doi.org/10.1016/0370-2693(73)90494-2).
- [16] A. Benvenuti et al. “Observation of muonless neutrino-induced inelastic interactions”. In: *Phys. Rev. Lett.* 32 (14 Apr. 1974), pp. 800–803. DOI: 10.1103/PhysRevLett.32.800.
- [17] G. Arnison et al. “Experimental observation of isolated large transverse energy electrons with associated missing energy at $s=540$ GeV”. In: *Physics Letters B* 122.1 (1983), pp. 103–116. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2).
- [18] M. Banner et al. “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{p} p$ collider”. In: *Phys. Lett. B* 122 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2.
- [19] G. Arnison et al. “Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider”. In: *Phys. Lett. B* 126 (1983), pp. 398–410. DOI: 10.1016/0370-2693(83)90188-0.
- [20] P. Bagnaia et al. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p} p$ collider”. In: *Phys. Lett. B* 129 (1983), pp. 130–140. DOI: 10.1016/0370-2693(83)90744-X.
- [21] Georges Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.

- [22] S. Galster et al. “Elastic electron-deuteron scattering and the electric neutron form factor at four-momentum transfers $5 \text{ fm}^{-2} < q^2 < 14 \text{ fm}^{-2}$ ”. In: *Nuclear Physics B* 32.1 (1971), pp. 221–237. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(71\)90068-X](https://doi.org/10.1016/0550-3213(71)90068-X).
- [23] H. Budd, A. Bodek, and J. Arrington. “Vector and axial form factors applied to neutrino quasielastic Scattering”. In: *Nuclear Physics B - Proceedings Supplements* 139 (2005). Proceedings of the third international workshop on neutrino-nucleus interactions in the few-GeV region, pp. 90–95. ISSN: 0920-5632. DOI: <https://doi.org/10.1016/j.nuclphysbps.2004.11.181>.
- [24] A. Papadopoulou et al. “Inclusive electron scattering and the GENIE neutrino event generator”. In: *Phys. Rev. D* 103 (11 June 2021), p. 113003. DOI: 10.1103/PhysRevD.103.113003.
- [25] R. Bradford et al. “A new parameterization of the nucleon elastic form factors”. In: *Nuclear Physics B - Proceedings Supplements* 159 (2006). Proceedings of the 4th International Workshop on Neutrino-Nucleus Interactions in the Few-GeV Region, pp. 127–132. ISSN: 0920-5632. DOI: <https://doi.org/10.1016/j.nuclphysbps.2006.08.028>.
- [26] A. Bodek et al. “Vector and axial nucleon form factors: A duality constrained parameterization”. In: *Eur. Phys. J. C* 53 (2008), pp. 349–354. DOI: 10.1140/epjc/s10052-007-0491-4.
- [27] J. J. Kelly. “Simple parametrization of nucleon form factors”. In: *Phys. Rev. C* 70 (6 Dec. 2004), p. 068202. DOI: 10.1103/PhysRevC.70.068202.
- [28] V. Punjabi et al. “The structure of the nucleon: elastic electromagnetic form Factors”. In: *Eur. Phys. J. A* 51 (2015), p. 79. DOI: 10.1140/epja/i2015-15079-x.
- [29] P.A. Zyla and others (Particle Data Group). In: *Prog. Theor. Exp. Phys.* 083C01 (2020).
- [30] Aaron S. Meyer et al. “Deuterium target data for precision neutrino-nucleus cross sections”. In: *Phys. Rev. D* 93.11 (2016), p. 113015. DOI: 10.1103/PhysRevD.93.113015.
- [31] C.H. Llewellyn Smith. “Neutrino reactions at accelerator energies”. In: *Phys. Rept.* 3.5 (1972), pp. 261–379. ISSN: 0370-1573. DOI: [https://doi.org/10.1016/0370-1573\(72\)90010-5](https://doi.org/10.1016/0370-1573(72)90010-5).

- [32] L.A. Ahrens et al. “Measurement of neutrino-proton and antineutrino-proton elastic scattering”. In: *Phys. Rev. D* 35 (1987), pp. 785–809. DOI: 10.1103/PhysRevD.35.785.
- [33] A Pais. “Weak interactions at high energies”. In: *Annals of Physics* 63.2 (1971), pp. 361–392. ISSN: 0003-4916. DOI: [https://doi.org/10.1016/0003-4916\(71\)90018-2](https://doi.org/10.1016/0003-4916(71)90018-2).
- [34] Dieter Rein and Lalit M. Sehgal. “Neutrino-excitation of baryon resonances and single pion production”. In: *Ann. Phys.* 133 (1981), pp. 79–153. DOI: 10.1016/0003-4916(81)90242-6.
- [35] Konstantin S. Kuzmin, Vladimir V. Lyubushkin, and Vadim A. Naumov. “Axial masses in quasielastic neutrino scattering and single-pion neutrino production on nucleons and nuclei”. In: *Acta Phys. Polon.* B37 (2006), pp. 2337–2348.
- [36] R. E. Cutkosky et al. “Pion-nucleon partial-wave analysis”. In: *Phys. Rev. D* 20 (11 Dec. 1979), pp. 2804–2838. DOI: 10.1103/PhysRevD.20.2804.
- [37] C. Andreopoulos et al. *The GENIE neutrino Monte Carlo generator: physics and user manual*. 2015.
- [38] Ch. Berger and L. M. Sehgal. “Lepton mass effects in single pion production by neutrinos”. In: *Phys. Rev. D* 76 (2007). [Erratum-ibid. D 77, 059901(E) (2008)], p. 113004. DOI: 10.1103/PhysRevD.76.113004.
- [39] M. Kabirnezhad. “Single pion production in neutrino-nucleon interactions”. In: *Phys. Rev. D* 97 (1 Jan. 2018), p. 013002. DOI: 10.1103/PhysRevD.97.013002.
- [40] Luis Alvarez-Ruso et al. “Recent highlights from GENIE v3”. In: (June 2021). DOI: 10.1140/epjs/s11734-021-00295-7.
- [41] T. Yang et al. “A hadronization model for few-GeV neutrino interactions”. In: *Eur. Phys. J. C* 63 (2009), pp. 1–10. DOI: 10.1140/epjc/s10052-009-1094-z.
- [42] M. Sajjad Athar and Jorge G. Morfin. *Neutrino (antineutrino)-nucleus interactions in the shallow- and deep-Inelastic scattering regions*. 2020.
- [43] P. C. Bosetti et al. “Analysis of nucleon structure functions in CERN bubble chamber neutrino experiments”. In: *Nucl. Phys. B* 142 (1978), pp. 1–28. DOI: 10.1016/0550-3213(78)90399-1.
- [44] J. D. Bjorken. “Asymptotic sum rules at infinite momentum”. In: *Phys. Rev.* 179 (5 Mar. 1969), pp. 1547–1553. DOI: 10.1103/PhysRev.179.1547.

BIBLIOGRAPHY

- [45] A. C. Benvenuti et al. “Test of QCD and a measurement of Λ from scaling violations in the proton structure function $F(2)(X, Q^2)$ at high Q^2 ”. In: *Phys. Lett. B* 223 (1989), pp. 490–496. DOI: 10.1016/0370-2693(89)91638-9.
- [46] M. Arneodo et al. “The A dependence of the nuclear structure function ratios”. In: *Nucl. Phys. B* 481 (1996), pp. 3–22. DOI: 10.1016/S0550-3213(96)90117-0.
- [47] A. C. Benvenuti et al. “A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2 ”. In: *Phys. Lett. B* 223 (1989), pp. 485–489. DOI: 10.1016/0370-2693(89)91637-7.
- [48] Ingo Schienbein et al. “Target mass corrections”. In: *Journal of Physics G: Nuclear and Particle Physics* 35.5 (Mar. 2008), p. 053101. DOI: 10.1088/0954-3899/35/5/053101.
- [49] Howard Georgi and H. David Politzer. “Freedom at moderate energies: Masses in color dynamics”. In: *Phys. Rev. D* 14 (7 Oct. 1976), pp. 1829–1848. DOI: 10.1103/PhysRevD.14.1829.
- [50] R.K. Ellis, W. Furmanski, and R. Petronzio. “Unravelling higher twists”. In: *Nuclear Physics B* 212.1 (1983), pp. 29–98. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(83\)90597-7](https://doi.org/10.1016/0550-3213(83)90597-7).
- [51] A. Bodek and U.K. Yang. “Modeling deep inelastic cross sections in the few GeV region”. In: *Nucl. Phys. B (Proc. Suppl.)* 112.1 (2002), pp. 70–76. ISSN: 0920-5632. DOI: [https://doi.org/10.1016/S0920-5632\(02\)01755-3](https://doi.org/10.1016/S0920-5632(02)01755-3).
- [52] Krzysztof M. Graczyk, Cezary Juszczak, and Jan T. Sobczyk. “Quark–hadron duality in the Rein–Sehgal model”. In: *Nuclear Physics A* 781.1 (2007), pp. 227–246. ISSN: 0375-9474. DOI: <https://doi.org/10.1016/j.nuclphysa.2006.10.036>.
- [53] E. D. Bloom and F. J. Gilman. “Scaling, duality, and the behavior of resonances in inelastic electron-proton scattering”. In: *Phys. Rev. Lett.* 25 (16 Oct. 1970), pp. 1140–1143. DOI: 10.1103/PhysRevLett.25.1140.
- [54] S.A. Kulagin and R. Petti. “Global study of nuclear structure functions”. In: *Nuclear Physics A* 765.1 (2006), pp. 126–187. ISSN: 0375-9474. DOI: <https://doi.org/10.1016/j.nuclphysa.2005.10.011>.
- [55] I. Niculescu et al. “Experimental verification of quark-hadron duality”. In: *Phys. Rev. Lett.* 85 (6 Aug. 2000), pp. 1186–1189. DOI: 10.1103/PhysRevLett.85.1186.

BIBLIOGRAPHY

- [56] C. Andreopoulos et al. “Summary of the NuSTEC workshop on shallow- and deep-inelastic scattering”. In: *NuSTEC Workshop on Shallow- and Deep-Inelastic Scattering*. July 2019.
- [57] M. Leuschner et al. “Quasielastic proton knockout from ^{16}O ”. In: *Phys. Rev. C* 49 (2 Feb. 1994), pp. 955–967. DOI: 10.1103/PhysRevC.49.955.
- [58] T. De Forest. “Off-Shell electron nucleon cross-sections. The impulse approximation”. In: *Nucl. Phys. A* 392 (1983), pp. 232–248. DOI: 10.1016/0375-9474(83)90124-0.
- [59] Omar Benhar, Adelchi Fabrocini, and Stefano Fantoni. “The nucleon spectral function in nuclear matter”. In: *Nuclear Physics A* 505.2 (1989), pp. 267–299. ISSN: 0375-9474. DOI: [https://doi.org/10.1016/0375-9474\(89\)90374-6](https://doi.org/10.1016/0375-9474(89)90374-6).
- [60] Or Hen et al. “Nucleon-nucleon correlations, short-lived excitations, and the quarks within”. In: *Rev. Mod. Phys.* 89 (4 Nov. 2017), p. 045002. DOI: 10.1103/RevModPhys.89.045002.
- [61] A. Bodek and J. L. Ritchie. “Fermi-motion effects in deep-inelastic lepton scattering from nuclear targets”. In: *Phys. Rev. D* 23 (5 Mar. 1981), pp. 1070–1091. DOI: 10.1103/PhysRevD.23.1070.
- [62] K. S. Egiyan et al. “Measurement of two- and three-nucleon short-range correlation probabilities in nuclei”. In: *Phys. Rev. Lett.* 96 (8 Mar. 2006), p. 082501. DOI: 10.1103/PhysRevLett.96.082501.
- [63] Yoshinari Hayato and Luke Pickering. “The NEUT neutrino interaction simulation program library”. In: *The European Physical Journal Special Topics* 230 (24 Dec. 2021), pp. 4469–4481. DOI: 10.1140/epjs/s11734-021-00287-7.
- [64] O. Benhar et al. “Spectral function of finite nuclei and scattering of GeV electrons”. In: *Nuclear Physics A* 579.3 (1994), pp. 493–517. ISSN: 0375-9474. DOI: [https://doi.org/10.1016/0375-9474\(94\)90920-2](https://doi.org/10.1016/0375-9474(94)90920-2).
- [65] A. Gil, J. Nieves, and E. Oset. “Many-body approach to the inclusive (e, e') reaction from the quasielastic to the Δ excitation region”. In: *Nuclear Physics A* 627.4 (1997), pp. 543–598. ISSN: 0375-9474. DOI: [https://doi.org/10.1016/S0375-9474\(97\)00513-7](https://doi.org/10.1016/S0375-9474(97)00513-7).
- [66] J. Nieves, J. E. Amaro, and M. Valverde. “Inclusive quasielastic charged-current neutrino-nucleus reactions”. In: *Phys. Rev. C* 70 (5 Nov. 2004), p. 055503. DOI: 10.1103/PhysRevC.70.055503.

BIBLIOGRAPHY

- [67] W.M. Alberico, M. Ericson, and A. Molinari. “Quenching and hardening in the transverse quasi-elastic peak”. In: *Nuclear Physics A* 379.3 (1982), pp. 429–448. ISSN: 0375-9474. DOI: [https://doi.org/10.1016/0375-9474\(82\)90007-0](https://doi.org/10.1016/0375-9474(82)90007-0).
- [68] S. K. Singh and E. Oset. “Inclusive quasielastic neutrino reactions in ^{12}C and ^{16}O at intermediate energies”. In: *Phys. Rev. C* 48 (3 Sept. 1993), pp. 1246–1258. DOI: 10.1103/PhysRevC.48.1246.
- [69] Teppei Katori. “Meson Exchange Current (MEC) models in neutrino interaction generators”. In: *AIP Conf. Proc.* 1663.1 (2015). Ed. by H. Da Motta, Jorge G. Morfin, and M. Sakuda, p. 030001. DOI: 10.1063/1.4919465.
- [70] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas. “Inclusive charged-current neutrino-nucleus reactions”. In: *Phys. Rev. C* 83 (4 Apr. 2011), p. 045501. DOI: 10.1103/PhysRevC.83.045501.
- [71] Jackie Schwehr, Dan Cherdack, and Rik Gran. “GENIE implementation of IFIC Valencia model for QE-like 2p2h neutrino-nucleus cross section”. In: (Jan. 2016).
- [72] J. E. Amaro et al. “Implementation of the SuSAv2-meson exchange current 1p1h and 2p2h models in GENIE and analysis of nuclear effects in T2K measurements”. In: *The European Physical Journal Special Topics* (Oct. 2021). DOI: 10.1140/epjs/s11734-021-00289-5.
- [73] Maria B. Barbaro, Arturo De Pace, and Luisa Fiume. “The SuSA model for neutrino oscillation experiments: from quasielastic scattering to the resonance region”. In: *Universe* 7.5 (2021), p. 140. DOI: 10.3390/universe7050140.
- [74] S. Dolan, G. D. Megias, and S. Bolognesi. “Neutrino-nucleus scattering in the SuSA model”. In: *Phys. Rev. D* 101 (3 Feb. 2020), p. 033003. DOI: 10.1103/PhysRevD.101.033003.
- [75] J. E. Amaro et al. “Using electron scattering superscaling to predict charge-changing neutrino cross sections in nuclei”. In: *Phys. Rev. C* 71 (1 Jan. 2005), p. 015501. DOI: 10.1103/PhysRevC.71.015501.
- [76] T. W. Donnelly and Ingo Sick. “Superscaling of inclusive electron scattering from nuclei”. In: *Phys. Rev. C* 60 (6 Nov. 1999), p. 065502. DOI: 10.1103/PhysRevC.60.065502.
- [77] T. W. Donnelly and Ingo Sick. “Superscaling in inclusive electron-nucleus scattering”. In: *Phys. Rev. Lett.* 82 (16 Apr. 1999), pp. 3212–3215. DOI: 10.1103/PhysRevLett.82.3212.

BIBLIOGRAPHY

- [78] K. Abe, N. Abgrall, H. Aihara, et al. “T2K neutrino flux prediction”. In: *Phys. Rev. D* 87 (1 Jan. 2013), p. 012001. DOI: 10.1103/PhysRevD.87.012001.
- [79] T. Leitner, L. Alvarez-Ruso, and U. Mosel. “Charged current neutrino-nucleus interactions at intermediate energies”. In: *Phys. Rev. C* 73 (6 June 2006), p. 065502. DOI: 10.1103/PhysRevC.73.065502.
- [80] G. D. Harp. “Extension of the isobar model for intranuclear cascades to 1 GeV”. In: *Phys. Rev. C* 10 (6 Dec. 1974), pp. 2387–2396. DOI: 10.1103/PhysRevC.10.2387.
- [81] D. S. Baranov et al. “Estimation of the hadron formation length in neutrino interactions”. In: *Sov. J. Nucl. Phys. (Engl. Transl.), (United States)* (Dec. 1984).
- [82] Davide Mancusi et al. “Improving the description of proton-induced one-nucleon removal in intranuclear-cascade models”. In: *Phys. Rev. C* 91 (3 Mar. 2015), p. 034602. DOI: 10.1103/PhysRevC.91.034602.
- [83] A. Boudard et al. “New potentialities of the Liège intranuclear cascade model for reactions induced by nucleons and light charged particles”. In: *Phys. Rev. C* 87 (1 Jan. 2013), p. 014606. DOI: 10.1103/PhysRevC.87.014606.
- [84] D. H. Wright and M. H. Kelsey. “The Geant4 Bertini cascade”. In: *Nucl. Instrum. Meth. A* 804 (2015), pp. 175–188. DOI: 10.1016/j.nima.2015.09.058.
- [85] V. R. Pandharipande and Steven C. Pieper. “Nuclear transparency to intermediate-energy nucleons from (e,e'p) reactions”. In: *Phys. Rev. C* 45 (2 Feb. 1992), pp. 791–798. DOI: 10.1103/PhysRevC.45.791.
- [86] L Bettelli, Marilena Bianchi-Streit, and G Giacomelli. *Particle physics with bubble chamber photographs*. Tech. rep. Geneva: CERN, July 1993.
- [87] Giorgio Giacomelli. *Introduction to the workshop "30 years of bubble chamber physics"*. 2006.
- [88] Sophia Elizabeth Bennett. “Gargamelle bubble chamber pictures c. 1976-7. Originals in the CERN Archive (reference CERN-ARCH-GGM-111)”. In: (June 2017). General Photo.
- [89] Sacha E. Kopp. “Accelerator neutrino beams”. In: *Physics Reports* 439.3 (2007), pp. 101–159. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2006.11.004>.
- [90] S. J. Barish et al. “Inclusive νp and νn charged-current neutrino reactions below 6 GeV”. In: *Phys. Lett. B* 66 (1977), pp. 291–294. DOI: 10.1016/0370-2693(77)90883-8.

BIBLIOGRAPHY

- [91] S. J. Barish et al. “Study of neutrino interactions in hydrogen and deuterium. Description of the experiment and study of the reaction $\nu + d \rightarrow \mu^- + p + p_s$ ”. In: *Phys. Rev. D* 16 (1977), pp. 3103–3121. DOI: 10.1103/PhysRevD.16.3103.
- [92] Joseph Grange and Teppei Katori. “Charged current quasi-elastic cross section measurements in MiniBooNE”. In: *Mod. Phys. Lett. A* 29.12 (2014), p. 1430011. DOI: 10.1142/S0217732314300110.
- [93] Callum Wilkinson et al. “Reanalysis of bubble chamber measurements of muon-neutrino induced single pion production”. In: *Phys. Rev. D* 90.11 (2014), p. 112017. DOI: 10.1103/PhysRevD.90.112017.
- [94] R. Gran et al. “Measurement of the quasi-elastic axial vector mass in neutrino-oxygen interactions”. In: *Phys. Rev. D* 74 (2006), p. 052002. DOI: 10.1103/PhysRevD.74.052002.
- [95] K. Hiraide. “Recent results from SciBooNE”. In: *Acta Phys. Polon. B* 40 (2009). Ed. by Arthur Ankowski and Jan Sobczyk, pp. 2659–2664.
- [96] J. Altegoer et al. “The NOMAD experiment at the CERN SPS”. In: *Nucl. Instrum. Meth. A* 404 (1998), pp. 96–128. DOI: 10.1016/S0168-9002(97)01079-6.
- [97] Justin Evans. “The MINOS experiment: Results and prospects”. In: *Adv. High Energy Phys.* 2013 (2013), p. 182537. DOI: 10.1155/2013/182537.
- [98] A.A. Aguilar-Arevalo et al. “The MiniBooNE detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 599.1 (Feb. 2009), pp. 28–46. ISSN: 0168-9002. DOI: 10.1016/j.nima.2008.10.028.
- [99] Alexis Armando Aguilar-Arevalo et al. “First measurement of the muon neutrino charged current quasielastic double differential cross section”. In: *Phys. Rev. D* 81 (2010), p. 092005. DOI: 10.1103/PhysRevD.81.092005.
- [100] A. A. Aguilar-Arevalo et al. “Neutrino flux prediction at MiniBooNE”. In: *Physical Review D* 79.7 (Apr. 2009). ISSN: 1550-2368. DOI: 10.1103/physrevd.79.072002.
- [101] R. Gran et al. “Measurement of the quasielastic axial vector mass in neutrino interactions on oxygen”. In: *Phys. Rev. D* 74 (5 Sept. 2006), p. 052002. DOI: 10.1103/PhysRevD.74.052002.

BIBLIOGRAPHY

- [102] Jose Luis Alcaraz-Aunion and Joseph Walding. “Measurement of the nu(mu)-CCQE cross-section in the SciBooNE experiment”. In: *AIP Conf. Proc.* 1189.1 (2009). Ed. by Federico Sanchez, M. Sorel, and Luis Alvarez-Ruso, pp. 145–150. DOI: 10.1063/1.3274145.
- [103] P. Adamson et al. “Study of quasielastic scattering using charged-current ν_μ -iron interactions in the MINOS near detector”. In: *Phys. Rev. D* 91 (1 Jan. 2015), p. 012005. DOI: 10.1103/PhysRevD.91.012005.
- [104] V Lyubushkin et al. “A Study of quasi-elastic muon neutrino and antineutrino scattering in the NOMAD experiment”. In: *Eur. Phys. J. C* 63 (2009), pp. 355–381. DOI: 10.1140/epjc/s10052-009-1113-0.
- [105] Masashi Yokoyama. “The Hyper-Kamiokande experiment”. In: *Prospects in neutrino physics*. Apr. 2017.
- [106] M. Antonello et al. “A proposal for a three detector short-baseline neutrino oscillation program in the Fermilab booster neutrino beam”. In: (Mar. 2015).
- [107] Pedro A.N. Machado, Ornella Palamara, and David W. Schmitz. “The Short-Baseline neutrino program at Fermilab”. In: *Annual Review of Nuclear and Particle Science* 69.1 (Oct. 2019), pp. 363–387. ISSN: 1545-4134. DOI: 10.1146/annurev-nucl-101917-020949.
- [108] K. Abe, N. Abgrall, H. Aihara, et al. “The T2K experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 659.1 (2011), pp. 106–135. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2011.06.067>.
- [109] Kevin Scott McFarland. “MINERvA: A Dedicated neutrino scattering experiment at NuMI”. In: *Nucl. Phys. B Proc. Suppl.* 159 (2006). Ed. by F. Cavanna, J. G. Morfin, and T. Nakaya, pp. 107–112. DOI: 10.1016/j.nuclphysbps.2006.08.073.
- [110] R. Acciarri et al. “Design and construction of the MicroBooNE detector”. In: *Journal of Instrumentation* 12.02 (Feb. 2017), P02017–P02017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/02/p02017.
- [111] K. Abe et al. “Measurement of the charged-current electron (anti-)neutrino inclusive cross-sections at the T2K off-axis near detector ND280”. In: *JHEP* 10 (2020), p. 114. DOI: 10.1007/JHEP10(2020)114.
- [112] K. Abe et al. “Measurement of the inclusive electron neutrino charged current cross section on carbon with the T2K near detector”. In: *Phys. Rev. Lett.* 113 (24 Dec. 2014), p. 241803. DOI: 10.1103/PhysRevLett.113.241803.

BIBLIOGRAPHY

- [113] M. V. Ascencio et al. “Measurement of inclusive charged-current ν_μ scattering on hydrocarbon at $\langle E_\nu \rangle \sim 6$ GeV with low three-momentum transfer”. In: (Oct. 2021).
- [114] A. Filkins et al. “Double-differential inclusive charged-current ν_μ cross sections on hydrocarbon in MINERvA at $\langle E_\nu \rangle \sim 3.5$ GeV”. In: *Phys. Rev. D* 101.11 (2020), p. 112007. DOI: [10.1103/PhysRevD.101.112007](https://doi.org/10.1103/PhysRevD.101.112007).
- [115] P. Abratenko et al. “First measurement of energy-dependent inclusive muon neutrino charged-current cross sections on argon with the MicroBooNE detector”. In: (Oct. 2021).
- [116] P. Abratenko, C. Adams, M. Alrashed, et al. “First measurement of inclusive muon neutrino charged current differential cross sections on argon at $E_\nu \sim 0.8$ GeV with the MicroBooNE detector”. In: *Phys. Rev. Lett.* 123 (13 Sept. 2019), p. 131801. DOI: [10.1103/PhysRevLett.123.131801](https://doi.org/10.1103/PhysRevLett.123.131801).
- [117] D. Ruterbories et al. “Measurement of quasielastic-like neutrino scattering at $\langle E_\nu \rangle \sim 3.5$ GeV on a hydrocarbon target”. In: *Phys. Rev. D* 99 (1 Jan. 2019), p. 012004. DOI: [10.1103/PhysRevD.99.012004](https://doi.org/10.1103/PhysRevD.99.012004).
- [118] M. F. Carneiro et al. “High-statistics measurement of neutrino quasielastic-like scattering at 6 GeV on a hydrocarbon target”. In: *Phys. Rev. Lett.* 124.12 (2020), p. 121801. DOI: [10.1103/PhysRevLett.124.121801](https://doi.org/10.1103/PhysRevLett.124.121801).
- [119] K. Abe et al. “Characterization of nuclear effects in muon-neutrino scattering on hydrocarbon with a measurement of final-state kinematics and correlations in charged-current pionless interactions at T2K”. In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 032003. DOI: [10.1103/PhysRevD.98.032003](https://doi.org/10.1103/PhysRevD.98.032003).
- [120] D. Coplowe et al. “Probing nuclear effects with neutrino-induced charged-current neutral pion production”. In: *Phys. Rev. D* 102.7 (2020), p. 072007. DOI: [10.1103/PhysRevD.102.072007](https://doi.org/10.1103/PhysRevD.102.072007).
- [121] D. Ruterbories et al. “Simultaneous measurement of proton and lepton kinematics in quasielastic-like ν_μ -hydrocarbon interactions from 2 to 20 GeV”. In: (Mar. 2022).
- [122] P. Abratenko et al. “Measurement of differential cross sections for ν_μ -Ar charged-current interactions with protons and no pions in the final state with the MicroBooNE detector”. In: *Phys. Rev. D* 102.11 (2020), p. 112013. DOI: [10.1103/PhysRevD.102.112013](https://doi.org/10.1103/PhysRevD.102.112013).

BIBLIOGRAPHY

- [123] K. Abe et al. “First measurement of the muon neutrino charged current single pion production cross section on water with the T2K near detector”. In: *Phys. Rev. D* 95.1 (2017), p. 012010. DOI: 10.1103/PhysRevD.95.012010.
- [124] C. L. McGivern et al. “Cross sections for ν_μ and $\bar{\nu}_\mu$ induced pion production on hydrocarbon in the few-GeV region using MINERvA”. In: *Phys. Rev. D* 94.5 (2016), p. 052005. DOI: 10.1103/PhysRevD.94.052005.
- [125] O. Altinok et al. “Measurement of ν_μ charged-current single π^0 production on hydrocarbon in the few-GeV region using MINERvA”. In: *Phys. Rev. D* 96.7 (2017), p. 072003. DOI: 10.1103/PhysRevD.96.072003.
- [126] X.-G. Lu, M. Betancourt, T. Walton, et al. “Measurement of final-state correlations in neutrino muon-proton mesonless production on hydrocarbon at $\langle E_\nu \rangle = 3$ GeV”. In: *Phys. Rev. Lett.* 121 (2 July 2018), p. 022504. DOI: 10.1103/PhysRevLett.121.022504.
- [127] MicroBooNE Collaboration. *Event displays: first neutrino images*. 2015. URL: https://microboone-exp.fnal.gov/public/approved_plots/Event_Displays.html.
- [128] C. Andreopoulos et al. “The GENIE neutrino Monte Carlo generator”. In: *Nucl. Instrum. Meth. A* 614 (2010), pp. 87–104. DOI: 10.1016/j.nima.2009.12.009.
- [129] Andy Buckley et al. “Systematic event generator tuning for the LHC”. In: *Eur. Phys. J. C* 65.1-2 (2010), p. 331. DOI: 10.1140/epjc/s10052-009-1196-7.
- [130] Holger Schulz et al. *Professor web page*. <https://professor.hepforge.org>. 2020.
- [131] Philip Rodrigues, Callum Wilkinson, and Kevin McFarland. “Constraining the GENIE model of neutrino-induced single pion production using reanalyzed bubble chamber data”. In: (2016). DOI: 10.1140/epjc/s10052-016-4314-3.
- [132] Konstantin S. Kuzmin, Vladimir V. Lyubushkin, and Vadim A. Naumov. “Lepton polarization in neutrino–nucleon interactions”. In: *Mod. Phys. Lett. A* 19 (2004). *Phys. Part. Nucl.* 35, S133 (2004), pp. 2815–2829. DOI: 10.1142/S0217732304016172.
- [133] Konstantin S. Kuzmin, Vladimir V. Lyubushkin, and Vadim A. Naumov. “Extended Rein-Sehgal model for tau lepton production”. In: *Nucl. Phys. B (Proc. Suppl.)* 139 (2005), pp. 158–161. DOI: 10.1016/j.nuclphysbps.2004.11.213.

BIBLIOGRAPHY

- [134] J. Mandula et al. “Duality and the hadron spectrum”. In: *Phys. Rev. Lett.* 22 (1969), pp. 1147–1149. DOI: 10.1103/PhysRevLett.22.1147.
- [135] H. Grassler et al. “Multiplicities of secondary hadrons produced in νp and $\bar{\nu} p$ charged current interactions”. In: *Nucl. Phys. B* 223.2 (1983), pp. 269–295. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(83\)90057-3](https://doi.org/10.1016/0550-3213(83)90057-3).
- [136] Daria Ziemska et al. “Charged particle multiplicity distributions in νn and νp charged current interactions”. In: *Phys. Rev. D* 27 (1983), pp. 47–57. DOI: 10.1103/PhysRevD.27.47.
- [137] S. Barlag et al. “Charged hadron multiplicities in high energy $\bar{\nu}_\mu n$ and $\bar{\nu}_\mu p$ interactions”. In: *Z. Phys. C* 11 (1982). [Erratum: *ibid.* **14**, 281 (1982)], pp. 283–292. DOI: 10.1007/BF01578279.
- [138] Dieter Rein and Lalit M. Sehgal. “Coherent π^0 production in neutrino reactions”. In: *Nucl. Phys. B* 223 (1983), pp. 29–44. DOI: 10.1016/0550-3213(83)90090-1.
- [139] S.G. Kovalenko. “Quasielastic neutrino production of charmed baryons from the point of view of local duality”. In: *Sov. J. Nucl. Phys.* 52 (1990), pp. 934–936.
- [140] M. A. G. Aivazis, Wu-Ki Tung, and Fredrick I. Olness. “Next-to-leading order QCD formulation of deep inelastic scattering”. In: *Proceedings, The Vancouver Meeting, Particles and Fields 91: Vancouver, Canada, August 18–22, 1991*. 1991, pp. 663–665.
- [141] Dieter Rein. “Diffractive pion production in neutrino reactions”. In: *Nucl. Phys. B* 278.1 (1986), pp. 61–77. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(86\)90106-9](https://doi.org/10.1016/0550-3213(86)90106-9).
- [142] M. Rafi Alam et al. “Weak kaon production off the nucleon”. In: *Phys. Rev. D* 82 (2010), p. 033001. DOI: 10.1103/PhysRevD.82.033001.
- [143] S. Dytman et al. “Comparison of validation methods of simulations for final state interactions in hadron production experiments”. In: *Phys. Rev. D* 104 (5 Sept. 2021), p. 053006. DOI: 10.1103/PhysRevD.104.053006.
- [144] S. Agostinelli et al. “Geant4 - a simulation toolkit”. In: *Nucl. Instrum. and Meth. in Phys. Res. A* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).

BIBLIOGRAPHY

- [145] Aatos Heikkinen, Nikita Stepanov, and Johannes Peter Wellisch. “Bertini intranuclear cascade implementation in GEANT4”. In: *eConf* C0303241 (2003), MOMT008.
- [146] Davide Mancusi et al. “Improving the description of proton-induced one-nucleon removal in intranuclear-cascade models”. In: *Phys. Rev. C* 91.3 (2015), p. 034602. DOI: [10.1103/PhysRevC.91.034602](https://doi.org/10.1103/PhysRevC.91.034602).
- [147] P. C. Bosetti et al. “Total Cross Sections for Charged-Current Neutrino and anti-neutrino Interactions in BEBC in the Energy Range 20–200 GeV”. In: *Phys. Lett.* B70 (1977), pp. 273–277. DOI: [10.1016/0370-2693\(77\)90537-8](https://doi.org/10.1016/0370-2693(77)90537-8).
- [148] P. Bosetti et al. “Total cross-sections for ν_μ and $\bar{\nu}_\mu$ charged-current interactions between 20 and 200 GeV”. In: *Phys. Lett.* B110 (1982), pp. 167–172. DOI: [10.1016/0370-2693\(82\)91028-0](https://doi.org/10.1016/0370-2693(82)91028-0).
- [149] C. Baltay et al. “Cross-sectons and scaling variable distributions of neutral and charged current neutrino nucleon interactions from a low-energy narrow band beam”. In: *Phys. Rev. Lett.* 44 (1980), pp. 916–919. DOI: [10.1103/PhysRevLett.44.916](https://doi.org/10.1103/PhysRevLett.44.916).
- [150] William Glenn Seligman. “A next-to-leading order QCD analysis of neutrino–iron structure functions at the Tevatron”. PhD thesis. Nevis Labs, Columbia U., 1997. DOI: [10.2172/1421736](https://doi.org/10.2172/1421736).
- [151] M. Jonker et al. “Experimental study of neutral-current and charged-current neutrino cross sections”. In: *Phys. Lett.* B99 (1981). [Erratum-ibid. B 100, 520 (1981)], pp. 265–270. DOI: [10.1016/0370-2693\(81\)91123-0](https://doi.org/10.1016/0370-2693(81)91123-0).
- [152] T. Kitagaki et al. “Charged-current exclusive pion production in $\nu_\mu n \rightarrow \mu^- p$ interactions”. In: *Phys. Rev.* D34 (1986), pp. 2554–2565. DOI: [10.1103/PhysRevD.34.2554](https://doi.org/10.1103/PhysRevD.34.2554).
- [153] T. Eichten et al. “Measurement of the neutrino-nucleon antineutrino-nucleon total cross sections”. In: *Phys. Lett.* B46 (1973), pp. 274–280. DOI: [10.1016/0370-2693\(73\)90702-8](https://doi.org/10.1016/0370-2693(73)90702-8).
- [154] J. G. Morfin et al. “Total cross-sections and nucleon structure functions in the Gargamelle SPS neutrino/antineutrino experiment”. In: *Phys. Lett.* B104 (1981), pp. 235–238. DOI: [10.1016/0370-2693\(81\)90598-0](https://doi.org/10.1016/0370-2693(81)90598-0).

- [155] A. S. Vovenko. “Total cross section measurements for ν_μ , $\bar{\nu}_\mu$ interactions in 3–30 GeV energy range with IHEP-JINR detector and future plans”. In: *Nucl. Phys. B (Proc. Suppl.)* 112 (2002), pp. 116–123. DOI: 10.1016/S0920-5632(02)01770-X.
- [156] V Lyubushkin et al. “A study of quasi-elastic muon neutrino and antineutrino scattering in the NOMAD experiment”. In: *Eur. Phys. J. C63* (2009), pp. 355–381. DOI: 10.1140/epjc/s10052-009-1113-0.
- [157] P. Adamson et al. “Neutrino and antineutrino inclusive charged-current cross section measurements with the MINOS near detector”. In: *Phys. Rev. D81* (2010), p. 072002. DOI: 10.1103/PhysRevD.81.072002.
- [158] S. J. Barish et al. “Study of neutrino interactions in hydrogen and deuterium: inelastic charged current reactions”. In: *Phys. Rev. D19* (1979), pp. 2521–2542. DOI: 10.1103/PhysRevD.19.2521.
- [159] D. C. Colley et al. “Cross-sections for charged current ν and $\bar{\nu}$ interactions in the energy range 10 to 50 GeV”. In: *Z. Phys. C2* (1979), pp. 187–224. DOI: 10.1007/BF01474659.
- [160] Michael Andrew Parker et al. “A comparison of charged current cross sections and structure functions for neutrino and anti-neutrinos beams on hydrogen and neon”. In: *Nucl. Phys. B232* (1984), pp. 1–20. DOI: 10.1016/0550-3213(84)90358-4.
- [161] N. J. Baker et al. “Total cross-sections for $\nu_\mu n$ and $\nu_\mu p$ charged current interactions in the 7-foot bubble chamber”. In: *Phys. Rev. D25* (1982), pp. 617–623. DOI: 10.1103/PhysRevD.25.617.
- [162] D. MacFarlane et al. “Nucleon structure functions from high-energy neutrino interactions with iron and QCD results”. In: *Z. Phys. C26* (1984), pp. 1–12. DOI: 10.1007/BF01572534.
- [163] J. V. Allaby et al. “Total cross sections of charged-current neutrino and antineutrino interactions on isoscalar nuclei”. In: *Z. Phys. C38* (1988), pp. 403–410. DOI: 10.1007/BF01584388.
- [164] N. J. Baker et al. “Measurement of the ν_μ charged-current cross section”. In: *Phys. Rev. Lett.* 51 (1983), pp. 735–738. DOI: 10.1103/PhysRevLett.51.735.
- [165] S. Ciampolillo et al. “Total cross-section for neutrino charged current interactions at 3 and 9 GeV”. In: *Phys. Lett. B84* (1979), pp. 281–284. DOI: 10.1016/0370-2693(79)90303-4.

BIBLIOGRAPHY

- [166] A. E. Asratian et al. “Total antineutrino-nucleon charged current cross section in the energy range 10–50 GeV”. In: *Phys. Lett.* B137 (1984), pp. 122–124. DOI: 10.1016/0370-2693(84)91118-3.
- [167] V. B. Anikeev et al. “Total cross-section measurements for ν_μ , $\bar{\nu}_\mu$ interactions in 3–30 GeV energy range with IHEP-JINR neutrino detector”. In: *Z. Phys.* C70 (1996), pp. 39–46. DOI: 10.1007/s002880050078.
- [168] D. S. Baranov et al. “Measurements of the $\nu_\mu N$ total cross section at 2–30 GeV in SKAT neutrino experiment”. In: *Phys. Lett.* B81 (1979), pp. 255–257. DOI: 10.1016/0370-2693(79)90536-7.
- [169] Jose Luis Alcaraz-Aunion and Joseph Walding. “Measurement of the nu(mu)-CCQE cross-section in the SciBooNE experiment”. In: *AIP Conf. Proc.* 1189.1 (2009). Ed. by Federico Sanchez, M. Sorel, and Luis Alvarez-Ruso, pp. 145–150. DOI: 10.1063/1.3274145.
- [170] G. Fanourakis et al. “Study of low-energy antineutrino interactions on protons”. In: *Phys. Rev.* D21 (1980), pp. 562–568. DOI: 10.1103/PhysRevD.21.562.
- [171] G. N. Taylor et al. “ $\bar{\nu}_\mu$ -nucleon charged current total cross section for 5–250 GeV”. In: *Phys. Rev. Lett.* 51 (1983), pp. 739–742. DOI: 10.1103/PhysRevLett.51.739.
- [172] O. Erriquez et al. “Antineutrino-nucleon total cross section and ratio of antineutrino cross section on neutrons and protons”. In: *Phys. Lett.* B80 (1979), pp. 309–313. DOI: 10.1016/0370-2693(79)90224-7.
- [173] A. E. Asratian et al. “Charged current neutrino interactions below 30 GeV”. In: *Phys. Lett.* B76 (1978), pp. 239–242. DOI: 10.1016/0370-2693(78)90286-1.
- [174] W. A. Mann et al. “Study of the reaction $\nu + n \rightarrow \mu^- + p$ ”. In: *Phys. Rev. Lett.* 31 (1973), pp. 844–847. DOI: 10.1103/PhysRevLett.31.844.
- [175] D. Allasia et al. “Investigation of exclusive channels in $\nu/\bar{\nu}$ -deuteron charged current interactions”. In: *Nucl. Phys.* B343 (1990), pp. 285–309. DOI: 10.1016/0550-3213(90)90472-P.
- [176] T. Kitagaki et al. “High-energy quasielastic $\nu_\mu n \rightarrow \mu^- p$ scattering in deuterium”. In: *Phys. Rev.* D28 (1983), pp. 436–442. DOI: 10.1103/PhysRevD.28.436.
- [177] S. V. Belikov et al. “Quasielastic $\nu_\mu n$ scattering at 3–30 GeV energy”. In: *Yad. Fiz.* 35 (1982), pp. 59–63.

BIBLIOGRAPHY

- [178] J. Brunner et al. “Quasielastic nucleon and hyperon production by neutrons and antineutrinos with energies below 30 GeV”. In: *Z. Phys.* C45 (1990), pp. 551–555. DOI: 10.1007/BF01556267.
- [179] N. J. Baker et al. “Quasielastic neutrino scattering: a measurement of the weak nucleon axial vector form-factor”. In: *Phys. Rev.* D23 (1981), pp. 2499–2505. DOI: 10.1103/PhysRevD.23.2499.
- [180] S. Bonetti et al. “Study of Quasi-elastic Reactions of ν and $\bar{\nu}$ in Gargamelle”. In: *Nuovo Cim.* A38 (1977), pp. 260–270. DOI: 10.1007/BF02730023.
- [181] S. V. Belikov et al. “Restraints on parameters of oscillations of muon neutrinos from quasielastic scattering data.” In: *Yad. Fiz.* 41 (1985), pp. 919–924.
- [182] N. Armenise et al. “Charged current elastic antineutrino interactions in propane”. In: *Nucl. Phys.* B152 (1979), pp. 365–375. DOI: 10.1016/0550-3213(79)90087-7.
- [183] J. Campbell et al. “Study of the reaction $\nu p \rightarrow \mu^- \pi^+ p$ ”. In: *Phys. Rev. Lett.* 30 (1973), pp. 335–339. DOI: 10.1103/PhysRevLett.30.335.
- [184] G. M. Radecky et al. “Study of single-pion production by weak charged currents in low-energy μd interactions”. In: *Phys. Rev.* D25 (1982). [Erratum-*ibid.* D 26, 3297 (1982)], pp. 1161–1173. DOI: 10.1103/PhysRevD.25.1161.
- [185] W. Lerche et al. “Experimental Study of the Reaction $\nu p \rightarrow \mu^- \pi^+ p$ ”. In: *Phys. Lett.* 78B (1978), pp. 510–514. DOI: 10.1016/0370-2693(78)90499-9.
- [186] P. Allen et al. “Single π^+ production in charged current neutrino-hydrogen interactions”. In: *Nucl. Phys.* B176 (1980), pp. 269–284. DOI: 10.1016/0550-3213(80)90450-2.
- [187] J. Bell et al. “A study of the reaction $\nu p \rightarrow \mu^- \Delta^{++}$ at high-energies and comparisons with theory”. In: *Phys. Rev. Lett.* 41 (1978), pp. 1012–1015. DOI: 10.1103/PhysRevLett.41.1012.
- [188] P. Allen et al. “A study of single meson production in neutrino and antineutrinos charged-current interactions on protons”. In: *Nucl. Phys.* B264 (1986), pp. 221–242. DOI: 10.1016/0550-3213(86)90480-3.
- [189] S. J. Barish et al. “Study of the reaction $\bar{\nu}_\mu p \rightarrow \mu^+ p \pi^-$ ”. In: *Phys. Lett.* B91 (1980), pp. 161–164. DOI: 10.1016/0370-2693(80)90684-X.
- [190] T. Kitagaki et al. “Neutrino flux and total charged current cross sections in high-energy neutrino-deuterium interactions”. In: *Phys. Rev. Lett.* 49 (1982), pp. 98–101. DOI: 10.1103/PhysRevLett.49.98.

BIBLIOGRAPHY

- [191] D. Day et al. “Study of vd charged current two pion production in the threshold region”. In: *Phys. Rev.* D28 (1983), pp. 2714–2720. DOI: 10.1103/PhysRevD.28.2714.
- [192] J. P. Berge et al. “Inclusive negative-hadron production from high-energy \bar{v} -nucleus charged-current interactions”. In: *Phys. Rev. D* 18 (11 Dec. 1978), pp. 3905–3910. DOI: 10.1103/PhysRevD.18.3905.
- [193] J. Bell et al. “Cross section measurements for the reactions $\nu_\mu p \rightarrow \mu^- \pi^+ p$ and $\nu_\mu p \rightarrow \mu^- K^+ p$ at high energies”. In: *Phys. Rev. Lett.* 41 (1978), pp. 1008–1011. DOI: 10.1103/PhysRevLett.41.1008.
- [194] Konstantin S. Kuzmin, Vladimir V. Lyubushkin, and Vadim A. Naumov. “Quasielastic axial-vector mass from experiments on neutrino-nucleus scattering”. In: *Eur. Phys. J. C* 54 (2008), pp. 517–538. DOI: 10.1140/epjc/s10052-008-0582-x.
- [195] Andy Buckley et al. “Systematic event generator tuning for the LHC”. In: *Eur. Phys. J. C* 65 (2010), pp. 331–357. DOI: 10.1140/epjc/s10052-009-1196-7.
- [196] M. Betancourt et al. “Comparisons and challenges of modern neutrino scattering experiments (TENSIONS2016 report)”. In: *Physics Reports* 773-774 (2018). Comparisons and challenges of modern neutrino scattering experiments (TENSIONS2016 report), pp. 1–28. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2018.08.003>.
- [197] Ko Abe et al. “Observation of electron neutrino appearance in a muon neutrino beam”. In: *Phys. Rev. Lett.* 112 (2014), p. 061802. DOI: 10.1103/PhysRevLett.112.061802.
- [198] S. Childress and J. Strait. “Long baseline neutrino beams at Fermilab”. In: *Proceedings of the 13th International Workshop on Neutrino Factories, Superbeams and Betabeams (NUFACT 2011), Geneva, Switzerland, August 1–6, 2011*. Vol. 408. 2013, p. 012007. DOI: 10.1088/1742-6596/408/1/012007.
- [199] L. Aliaga et al. “Design, calibration, and performance of the MINERvA detector”. In: *Nucl. Instrum. Meth. A* 743 (2014), pp. 130–159. DOI: 10.1016/j.nima.2013.12.053.
- [200] Alexis Armando Aguilar-Arevalo et al. “The MiniBooNE detector”. In: *Nucl. Instrum. Meth. A* 599 (2009), pp. 28–46. DOI: 10.1016/j.nima.2008.10.028.

BIBLIOGRAPHY

- [201] Luis Alvarez-Ruso. “Neutrino interactions: Challenges in the current theoretical picture”. In: *Proceedings of the 24th International Conference on Neutrino Physics and Astrophysics (Neutrino 2010), Athens, Greece, June 14–19, 2010*. Vol. 229–232. 2012, pp. 167–173. DOI: [10.1016/j.nuclphysbps.2012.09.027](https://doi.org/10.1016/j.nuclphysbps.2012.09.027).
- [202] H. Gallagher, G. Garvey, and G. P. Zeller. “Neutrino–nucleus interactions”. In: *Ann. Rev. Nucl. Part. Sci.* 61 (2011), pp. 355–378. DOI: [10.1146/annurev-nucl-102010-130255](https://doi.org/10.1146/annurev-nucl-102010-130255).
- [203] Luis Alvarez-Ruso, Y. Hayato, and Juan Nieves. “Progress and open questions in the physics of neutrino cross sections at intermediate energies”. In: *New J. Phys.* 16 (2014), p. 075015. DOI: [10.1088/1367-2630/16/7/075015](https://doi.org/10.1088/1367-2630/16/7/075015).
- [204] Juan Nieves et al. “Neutrino–nucleus CCQE-like scattering”. In: *Proceedings of the 37th International Conference on High Energy Physics (ICHEP 2014), Valencia, Spain, July 2–9, 2014*. Vol. 273–275. 2016, pp. 1830–1835. DOI: [10.1016/j.nuclphysbps.2015.09.295](https://doi.org/10.1016/j.nuclphysbps.2015.09.295).
- [205] Alexis Armando Aguilar-Arevalo et al. “Measurement of muon neutrino quasi-elastic scattering on carbon”. In: *Phys. Rev. Lett.* 100 (2008), p. 032301. DOI: [10.1103/PhysRevLett.100.032301](https://doi.org/10.1103/PhysRevLett.100.032301).
- [206] Kunxian Huang. “Measurement of the neutrino-oxygen neutral current quasi-elastic interaction cross-section by observing nuclear de-excitation γ -rays in the T2K experiment”. PhD thesis. Kyoto University, 2015.
- [207] D. Ruterbories et al. “Measurement of quasielastic-like neutrino scattering at $\langle E_\nu \rangle \sim 3.5$ GeV on a hydrocarbon target”. In: *Phys. Rev. D* 99.1 (2019), p. 012004. DOI: [10.1103/PhysRevD.99.012004](https://doi.org/10.1103/PhysRevD.99.012004).
- [208] Ko Abe et al. “Measurement of the ν_μ charged current quasielastic cross section on carbon with the T2K on-axis neutrino beam”. In: *Phys. Rev. D* 91 (2015), p. 112002. DOI: [10.1103/PhysRevD.91.112002](https://doi.org/10.1103/PhysRevD.91.112002).
- [209] P. Abratenko et al. “First measurement of differential charged current quasielastic-like ν_μ –Argon scattering cross sections with the MicroBooNE Detector”. In: *Phys. Rev. Lett.* 125.20 (2020), p. 201803. DOI: [10.1103/PhysRevLett.125.201803](https://doi.org/10.1103/PhysRevLett.125.201803).
- [210] Olga Lalakulich and Ulrich Mosel. “Pion production in the MiniBooNE experiment”. In: *Phys. Rev. C* 87.1 (2013), p. 014602. DOI: [10.1103/PhysRevC.87.014602](https://doi.org/10.1103/PhysRevC.87.014602).

BIBLIOGRAPHY

- [211] K. Hiraide. “Measurement of charged current charged single pion production in SciBooNE”. In: *Proceedings of the 34th International Conference on High Energy Physics (ICHEP 2008), Philadelphia, Pennsylvania, USA, July 30 – August 5, 2008.*
- [212] Alexis Armando Aguilar-Arevalo et al. “Measurement of neutrino-induced charged-current charged pion production cross sections on mineral oil at $E_\nu \sim 1$ GeV”. In: *Phys. Rev. D* 83 (2011), p. 052007. DOI: 10.1103/PhysRevD.83.052007.
- [213] A. Rodriguez et al. “Measurement of single charged pion production in the charged-current interactions of neutrinos in a 1.3 GeV wide band beam”. In: *Phys. Rev. D* 78 (2008), p. 032003. DOI: 10.1103/PhysRevD.78.032003.
- [214] Ko Abe et al. “First measurement of the muon neutrino charged current single pion production cross section on water with the T2K near detector”. In: *Phys. Rev. D* 95 (2017), p. 012010. DOI: 10.1103/PhysRevD.95.012010.
- [215] R. Acciarri et al. “First measurement of the cross section for ν_μ and $\bar{\nu}_\mu$ induced single charged pion production on argon using ArgoNeuT”. In: *Phys. Rev. D* 98.5 (2018), p. 052002. DOI: 10.1103/PhysRevD.98.052002.
- [216] C. Adams et al. “First measurement of ν_μ charged-current π^0 production on argon with the MicroBooNE detector”. In: *Phys. Rev. D* 99.9 (2019), p. 091102. DOI: 10.1103/PhysRevD.99.091102.
- [217] Júlia Tena-Vidal et al. “Neutrino-nucleon cross-section model tuning in GENIE v3”. In: *Phys. Rev. D* 104 (7 Oct. 2021), p. 072009. DOI: 10.1103/PhysRevD.104.072009.
- [218] P. Adamson et al. “Measurement of neutrino oscillations with the MINOS detectors in the NuMI beam”. In: *Phys. Rev. Lett.* 101 (2008), p. 131802. DOI: 10.1103/PhysRevLett.101.131802.
- [219] T. Yang et al. “A hadronization model for few-GeV neutrino interactions”. In: *Eur. Phys. J. C* 63 (2009), pp. 1–10. DOI: 10.1140/epjc/s10052-009-1094-z.
- [220] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. “PYTHIA 6.4 physics and manual”. In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026.
- [221] Aya Ishihara. *The IceCube upgrade – Design and science goals*. 2019.

BIBLIOGRAPHY

- [222] Dawn Williams. “Status and prospects for the IceCube Neutrino Observatory”. In: *Nucl. Instrum. Meth. A* 952 (2020). 10th International Workshop on Ring Imaging Cherenkov Detectors (RICH 2018), p. 161650. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2018.11.109>.
- [223] Antoine Kouchner. “Next-generation atmospheric neutrino experiments”. In: *Proceedings of the 13th International Conference on Topics in Astroparticle and Underground Physics (TAUP 2013), Asilomar, California, USA, September 8–13, 2013*. Vol. 4. 2014, pp. 60–74. DOI: 10.1016/j.dark.2014.09.001.
- [224] M. Honda et al. “Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model”. In: *Phys. Rev. D* 92 (2 July 2015), p. 023004. DOI: 10.1103/PhysRevD.92.023004.
- [225] E. Valencia et al. “Constraint of the MINERvA medium energy neutrino flux using neutrino-electron elastic scattering”. In: *Phys. Rev. D* 100.9 (Nov. 2019), p. 092001. ISSN: 2470-0029. DOI: 10.1103/physrevd.100.092001.
- [226] M. A. G. Aivazis, Fredrick I. Olness, and Wu-Ki Tung. “Leptoproduction of heavy quarks. I. General formalism and kinematics of charged current and neutral current production processes”. In: *Phys. Rev. D* 50 (5 Sept. 1994), pp. 3085–3101. DOI: 10.1103/PhysRevD.50.3085.
- [227] Makoto Sakuda. “Results from low-energy neutrino–nucleus scattering experiments”. In: *Proceedings of the 1st International Workshop on Neutrino–Nucleus Interactions in the Few GeV Region (NuInt 2001), Tsukuba, Japan, December 13–16, 2001*. Vol. 112. 2002, pp. 109–115. DOI: 10.1016/S0920-5632(02)01769-3.
- [228] Teppei Katori and Shivesh Mandalia. “PYTHIA hadronization process tuning in the GENIE neutrino interaction generator”. In: *J. Phys. G* 42.11 (2015), p. 115004. DOI: 10.1088/0954-3899/42/11/115004.
- [229] Konstantin S. Kuzmin and Vadim A. Naumov. “Mean charged multiplicities in charged-current neutrino scattering on hydrogen and deuterium”. In: *Phys. Rev. C* 88 (2013), p. 065501. DOI: 10.1103/PhysRevC.88.065501.
- [230] N. Agafonova et al. “Study of charged hadron multiplicities in charged-current neutrino–lead interactions in the OPERA detector”. In: *Eur. Phys. J. C* 78.1 (2018). [Erratum: *ibid.* **78**, 747 (2018)], p. 62. DOI: 10.1140/epjc/s10052-018-6223-0, 10.1140/epjc/s10052-017-5509-y.

BIBLIOGRAPHY

- [231] A. Kayis-Topaksu et al. “Charged particle multiplicities in charged-current neutrino and anti-neutrino nucleus interactions”. In: *Eur. Phys. J. C* 51 (2007), pp. 775–785. DOI: [10.1140/epjc/s10052-007-0366-8](https://doi.org/10.1140/epjc/s10052-007-0366-8).
- [232] W. Wittek et al. “Production of π^0 mesons and charged hadrons in $\bar{\nu}$ neon and ν neon charged current interactions”. In: *Z. Phys. C* 40 (1988), pp. 231–251. DOI: [10.1007/BF01555886](https://doi.org/10.1007/BF01555886).
- [233] Z. Koba, Holger Bech Nielsen, and P. Olesen. “Scaling of multiplicity distributions in high-energy hadron collisions”. In: *Nucl. Phys. B* 40 (1972), pp. 317–334. DOI: [10.1016/0550-3213\(72\)90551-2](https://doi.org/10.1016/0550-3213(72)90551-2).
- [234] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. “A brief introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: [10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036).
- [235] Joshua George Rubin. “Polarization, motion, and fragmentation: Exploring the role of quarks in the nucleon through semi-inclusive longitudinal spin asymmetries at HERMES”. PhD thesis. University of Illinois, 2009. DOI: [10.3204/DESY-THESIS-2009-045](https://doi.org/10.3204/DESY-THESIS-2009-045).
- [236] Silvia Ferreres-Solé and Torbjörn Sjöstrand. “The space-time structure of hadronization in the Lund model”. In: *Eur. Phys. J. C* 78.11 (2018), p. 983. DOI: [10.1140/epjc/s10052-018-6459-8](https://doi.org/10.1140/epjc/s10052-018-6459-8).
- [237] Judith M. Katzy. “QCD Monte-Carlo model tunes for the LHC”. In: *Prog. Part. Nucl. Phys.* 73 (2013), pp. 141–187. DOI: [10.1016/j.ppnp.2013.08.002](https://doi.org/10.1016/j.ppnp.2013.08.002).
- [238] Peter Zeiler Skands. “Tuning Monte Carlo generators: The Perugia tunes”. In: *Phys. Rev. D* 82 (2010), p. 074018. DOI: [10.1103/PhysRevD.82.074018](https://doi.org/10.1103/PhysRevD.82.074018).
- [239] Richard Corke and Torbjörn Sjöstrand. “Interleaved parton showers and tuning prospects”. In: *JHEP* 03 (2011), p. 032. DOI: [10.1007/JHEP03\(2011\)032](https://doi.org/10.1007/JHEP03(2011)032).
- [240] G. Aad et al. “Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC”. In: *New J. Phys.* 13 (2011), p. 053033. DOI: [10.1088/1367-2630/13/5/053033](https://doi.org/10.1088/1367-2630/13/5/053033).
- [241] ATLAS Collaboration. *New ATLAS event generator tunes to 2010 data*. 2011.
- [242] ATLAS Collaboration. *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*. 2011.
- [243] ATLAS Collaboration. *Summary of ATLAS Pythia 8 tunes*. 2012.

- [244] Rick Field. “The underlying event in hadronic collisions”. In: *Ann. Rev. Nucl. Part. Sci.* 62 (2012), pp. 453–483. DOI: 10.1146/annurev-nucl-102711-095030.
- [245] J. Alcaraz Maestre et al. “The SM and NLO Multileg and SM MC Working Groups: Summary Report”. In: *Proceedings of the 7th Les Houches Workshop on Physics at TeV Colliders, Les Houches, France, May 30 – June 17, 2011*. 2012, pp. 1–220.
- [246] Nameeqa Firdous and Gerald Rudolph. “Tuning of PYTHIA 6 to minimum bias data”. In: *Proceedings of the 1st Large Hadron Collider Physics Conference (LHCP 2013), Barcelona, Spain, May 13–18, 2013*. Vol. 60. 2013, p. 20056. DOI: 10.1051/epjconf/20136020056.
- [247] Torbjörn Sjöstrand. *PYTHIA tunes*. <http://home.thep.lu.se/~torbjorn/pythia81html/Tunes.html>. 2014.
- [248] Peter Skands, Stefano Carrazza, and Juan Rojo. “Tuning PYTHIA 8.1: The Monash 2013 Tune”. In: *Eur. Phys. J. C* 74.8 (2014), p. 3024. DOI: 10.1140/epjc/s10052-014-3024-y.
- [249] A. Airapetian et al. “Multiplicities of charged pions and kaons from semi-inclusive deep-inelastic scattering by the proton and the deuteron”. In: *Phys. Rev. D* 87 (2013), p. 074029. DOI: 10.1103/PhysRevD.87.074029.
- [250] N. M. Agababyan et al. “Charged ρ -meson production in neutrino induced reactions at $\langle E_\nu \rangle \approx 10$ GeV”. In: *Phys. Atom. Nucl.* 74 (2011). [Yad. Fiz. **74**, 238 (2011)], pp. 221–228. DOI: 10.1134/S1063778811020025.
- [251] J. W. Chapman et al. “Multiplicity distributions in high-energy neutrino interactions”. In: *Phys. Rev. Lett.* 36 (1976), pp. 124–126. DOI: 10.1103/PhysRevLett.36.124.
- [252] G. T. Jones et al. “ W^2 and Q^2 dependence of charged hadron and pion multiplicities in νp and $\bar{\nu}p$ charged current interactions”. In: *Z. Phys. C* 46 (1990), pp. 25–34. DOI: 10.1007/BF02440830.
- [253] G. T. Jones et al. “Multiplicity distributions of charged hadrons in νp and $\bar{\nu}p$ charged current interactions”. In: *Z. Phys. C* 54 (1992), pp. 45–54. DOI: 10.1007/BF01881707.
- [254] M. Derrick et al. “Multiplicity distributions in $\bar{\nu}_\mu p$ interactions”. In: *Phys. Rev. D* 25 (1982), pp. 624–633. DOI: 10.1103/PhysRevD.25.624.

BIBLIOGRAPHY

- [255] J. Hanlon et al. “Comparison of νn and νp charged current cross-sections from high-energy neutrino interactions in deuterium”. In: *Phys. Rev. Lett.* 45 (1980), pp. 1817–1821. DOI: 10.1103/PhysRevLett.45.1817.
- [256] V. E. Barnes et al. “Measurement of quark momentum distributions in the proton using an antineutrino probe”. In: *Phys. Rev. D* 25 (1982), pp. 1–21. DOI: 10.1103/PhysRevD.25.1.
- [257] J. Blietschau et al. “Observation of scaling deviations in the energy distribution of secondary hadrons in inelastic neutrino–proton interactions”. In: *Phys. Lett.* 87 B (1979), pp. 281–286. DOI: 10.1016/0370-2693(79)90983-3.
- [258] B. Jongejans et al. “Multiplicity distributions of charged hadrons produced in (anti)neutrino–deuterium charged- and neutral-current interactions”. In: *Nuovo Cim. A* 101 (1989), pp. 435–453. DOI: 10.1007/BF02789427.
- [259] D. Allasia et al. “Fragmentation in neutrino and antineutrino charged current interactions on proton and neutron”. In: *Z. Phys. C* 24 (1984), pp. 119–131. DOI: 10.1007/BF01571716.
- [260] C. C. Chang et al. “Study of diquark fragmentation into Λ and Y^{*+} in νn and νp interactions”. In: *Phys. Rev. D* 27 (11 1983), pp. 2776–2779. DOI: 10.1103/PhysRevD.27.2776.
- [261] J. Bell et al. “Experimental study of hadrons produced in high-energy charged current neutrino–proton interactions”. In: *Phys. Rev. D* 19 (1979), pp. 1–16. DOI: 10.1103/PhysRevD.19.1.
- [262] A. Kayis-Topaksu et al. “Charged-particle multiplicities in charged-current neutrino- and anti-neutrino–nucleus interactions”. In: *Eur. Phys. J. C* 51 (2007), pp. 775–785. DOI: 10.1140/epjc/s10052-007-0366-8.
- [263] F. Vannucci. “The NOMAD experiment at CERN”. In: *Adv. High Energy Phys.* 2014 (2014), p. 129694. DOI: 10.1155/2014/129694.
- [264] Artem Chukanov and Roberto Petti. “Study of fragmentation parameters in deep inelastic scattering neutrino interactions”. In: *Proceedings of the 10th International Workshop on Neutrino–Nucleus Interactions in the Few GeV Region (NuInt 2015), Osaka, Japan, November 16–21, 2015*. Vol. 12. 2016, p. 010026. DOI: 10.7566/JPSCP.12.010026.
- [265] Emilio Radicioni. “Results from the CHORUS experiment at CERN”. In: *Proceedings of the 6th Topical Seminar on Neutrino and Astro-Particle Physics, San Miniato, Italy, May 17–21, 1999*. Vol. 85. 2000, pp. 95–100. DOI: 10.1016/S0920-5632(00)00489-8.

BIBLIOGRAPHY

- [266] G. Giacomelli and M. Giorgini. “The OPERA experiment”. In: *Proceedings of the Vulcano Workshop: Frontier Objects in Astrophysics and Particle Physics, Vulcano, Italy, May 22–27, 2006*.
- [267] C. Adams et al. “Comparison of ν_μ -Ar multiplicity distributions observed by MicroBooNE to GENIE model predictions”. In: *Eur. Phys. J. C* 79.3 (2019), p. 248. DOI: 10.1140/epjc/s10052-019-6742-3.
- [268] L. Aliaga et al. “Neutrino flux predictions for the NuMI beam”. In: *Phys. Rev. D* 94 (9 Nov. 2016), p. 092005. DOI: 10.1103/PhysRevD.94.092005.
- [269] Júlia Tena-Vidal et al. “Hadronization model tuning in GENIE v3”. In: *Phys. Rev. D* 105 (1 Jan. 2022), p. 012009. DOI: 10.1103/PhysRevD.105.012009.
- [270] The MicroBooNE Collaboration, P. Abratenko, R. An, et al. *New theory-driven GENIE tune for MicroBooNE*. 2021.
- [271] A. Filkins et al. “Double-differential inclusive charged-current ν_μ cross sections on hydrocarbon in MINERvA at $\langle E_\nu \rangle \sim 3.5$ GeV”. In: *Phys. Rev. D* 101 (11 June 2020), p. 112007. DOI: 10.1103/PhysRevD.101.112007.
- [272] K. Abe et al. “Improved constraints on neutrino mixing from the T2K experiment with 3.13×10^{21} protons on target”. In: *Phys. Rev. D* 103 (11 June 2021), p. 112008. DOI: 10.1103/PhysRevD.103.112008.
- [273] A. A. Aguilar-Arevalo et al. “First measurement of the muon anti-neutrino double-differential charged-current quasielastic cross section”. In: *Physical Review D* 88.3 (Aug. 2013). ISSN: 1550-2368. DOI: 10.1103/physrevd.88.032001.
- [274] T. Cai et al. “Nucleon binding energy and transverse momentum imbalance in neutrino-nucleus reactions”. In: *Phys. Rev. D* 101 (9 May 2020), p. 092001. DOI: 10.1103/PhysRevD.101.092001.
- [275] P. Abratenko, M. Alrashed, R. An, et al. “Measurement of differential cross sections for ν_μ -Ar charged-current interactions with protons and no pions in the final state with the MicroBooNE detector”. In: *Phys. Rev. D* 102 (11 Dec. 2020), p. 112013. DOI: 10.1103/PhysRevD.102.112013.
- [276] C. E. Patrick et al. “Measurement of the muon anti-neutrino double-differential cross section for quasielastic-like scattering on hydrocarbon at $E_\nu \sim 3.5$ GeV”. In: *Phys. Rev. D* 97 (5 Mar. 2018), p. 052002. DOI: 10.1103/PhysRevD.97.052002.

BIBLIOGRAPHY

- [277] X.-G. Lu et al. “Measurement of nuclear effects in neutrino interactions with minimal dependence on neutrino energy”. In: *Phys. Rev. C* 94 (1 July 2016), p. 015503. DOI: 10.1103/PhysRevC.94.015503.
- [278] T. Cai et al. “Nucleon binding energy and transverse momentum imbalance in neutrino-nucleus reactions”. In: *Phys. Rev. D* 101.9 (2020), p. 092001. DOI: 10.1103/PhysRevD.101.092001.
- [279] K. Abe, C. Andreopoulos, M. Antonova, et al. “Measurement of double-differential muon neutrino charged-current interactions on C₈H₈ without pions in the final state using the T2K off-axis beam”. In: *Phys. Rev. D* 93 (11 June 2016), p. 112012. DOI: 10.1103/PhysRevD.93.112012.
- [280] D.G. Michael et al. “The magnetized steel and scintillator calorimeters of the MINOS experiment”. In: *Nuclear instruments and methods in physics research section A: Accelerators, spectrometers, detectors and associated equipment* 596.2 (2008), pp. 190–228. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2008.08.003>.
- [281] R. Acciarri et al. “Noise characterization and filtering in the MicroBooNE Liquid Argon TPC”. In: *Journal of Instrumentation* 12.08 (Aug. 2017), P08003–P08003. DOI: 10.1088/1748-0221/12/08/p08003.
- [282] C. Adams et al. “Ionization electron signal processing in single phase LArTPCs. Part II. Data/simulation comparison and performance in MicroBooNE”. In: *Journal of Instrumentation* 13.07 (July 2018), P07007–P07007. DOI: 10.1088/1748-0221/13/07/p07007.
- [283] R. Acciarri et al. “The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector”. In: *Eur. Phys. J. C* 78.1 (2018), p. 82. DOI: 10.1140/epjc/s10052-017-5481-6.
- [284] P. Abratenko, M. Alrashed, R. An, et al. “First measurement of differential charged current quasielastic-like ν_μ-argon scattering cross sections with the MicroBooNE detector”. In: *Phys. Rev. Lett.* 125 (20 Nov. 2020), p. 201803. DOI: 10.1103/PhysRevLett.125.201803.
- [285] C. Adams, M. Alrashed, et al. “First measurement of ν_μ charged-current π⁰ production on argon with the MicroBooNE detector”. In: *Phys. Rev. D* 99 (9 May 2019), p. 091102. DOI: 10.1103/PhysRevD.99.091102.

BIBLIOGRAPHY

- [286] Giampaolo Co'. "Random phase approximation and neutrino-nucleus cross sections". In: *Acta Phys. Polon. B* 37 (2006). Ed. by K. M. Graczyk and J. T. Sobczyk, pp. 2235–2242.
- [287] Richard Gran. *Model uncertainties for Valencia RPA effect for MINERvA*. 2017. DOI: 10.48550/ARXIV.1705.02932.
- [288] P. A. Rodrigues et al. "Identification of nuclear effects in neutrino-carbon interactions at low three-momentum transfer". In: *Phys. Rev. Lett.* 116 (2016). [Addendum: *Phys.Rev.Lett.* 121, 209902 (2018)], p. 071802. DOI: 10.1103/PhysRevLett.116.071802.
- [289] Arie Bodek and Tejin Cai. "Removal energies and final state interaction in lepton nucleus scattering". In: *The European Physical Journal C* 79.4 (Apr. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6750-3.
- [290] E. S. Pinzon Guerra et al. "Using world π^\pm -nucleus scattering data to constrain an intranuclear cascade model". In: *Phys. Rev. D* 99 (5 Mar. 2019), p. 052007. DOI: 10.1103/PhysRevD.99.052007.
- [291] Tomasz Bonus et al. "Data-based two-body current contribution to the neutrino-nucleus cross section". In: *Physical Review C* 102.1 (July 2020). ISSN: 2469-9993. DOI: 10.1103/physrevc.102.015502.
- [292] Kenneth M. Hanson, Toshihiko Kawano, and Patrick Talou. "Probabilistic interpretation of Peelle's pertinent puzzle and its resolution". In: *AIP Conf. Proc.* 769.1 (2005). Ed. by Robert C. Haight et al., p. 304. DOI: 10.1063/1.1945011.
- [293] G. D'Agostini. "On the use of the covariance matrix to fit correlated data". In: *Nucl. Instrum. Meth. A* 346 (1994), pp. 306–311. DOI: 10.1016/0168-9002(94)90719-6.
- [294] Aaron S. Meyer, André Walker-Loud, and Callum Wilkinson. *Status of lattice QCD determination of nucleon form factors and their relevance for the few-GeV neutrino program*. 2022.
- [295] Kajetan Niewczas and Jan T. Sobczyk. "Nuclear transparency in Monte Carlo neutrino event generators". In: *Phys. Rev. C* 100 (1 July 2019), p. 015505. DOI: 10.1103/PhysRevC.100.015505.
- [296] S. Dolan. "Exploring nuclear effects in neutrino-nucleus interactions using measurements of transverse kinematic imbalance from T2K and MINERvA". In: (Oct. 2018).

BIBLIOGRAPHY

- [297] G. Garino et al. “Proton propagation in nuclei studied in the (e,e'p) reaction”. In: *Phys. Rev. C* 45 (2 Feb. 1992), pp. 780–790. DOI: 10.1103/PhysRevC.45.780.
- [298] T.G. O’Neill et al. “A-dependence of nuclear transparency in quasielastic A(e, e'p) at high Q2”. In: *Physics Letters B* 351.1 (1995), pp. 87–92. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(95\)00362-0](https://doi.org/10.1016/0370-2693(95)00362-0).
- [299] D. Dutta et al. “Quasielastic ($e, e' p$) reaction on ^{12}C , ^{56}Fe , and ^{197}Au ”. In: *Phys. Rev. C* 68 (6 Dec. 2003), p. 064603. DOI: 10.1103/PhysRevC.68.064603.
- [300] D. Rohe et al. “Nuclear transparency from quasielastic $^{12}\text{C}(e, e' p)$ ”. In: *Phys. Rev. C* 72 (5 Nov. 2005), p. 054602. DOI: 10.1103/PhysRevC.72.054602.
- [301] A. Ershova et al. “Study of final-state interactions of protons in neutrino-nucleus scattering with INCL and NuWro cascade models”. In: (Feb. 2022).
- [302] Alexis Nikolopoulos et al. “Benchmarking intra-nuclear cascade models for neutrino scattering with relativistic optical potentials”. In: (Feb. 2022).