

# First combined tuning on transverse kinematic imbalance data with and without pion production constraints

Weijun Li<sup>1,\*</sup> Marco Roda<sup>2,†</sup> Júlia Tena-Vidal<sup>3,‡</sup> Costas Andreopoulos<sup>2,§</sup> Xianguo Lu<sup>4,||</sup> Adi Ashkenazi,<sup>3</sup> Joshua Barrow,<sup>5</sup> Steven Dytman,<sup>6</sup> Hugh Gallagher,<sup>7</sup> Alfonso Andres Garcia Soto,<sup>8</sup> Steven Gardiner,<sup>9</sup> Matan Goldenberg,<sup>3</sup> Robert Hatcher,<sup>9</sup> Or Hen,<sup>10</sup> Igor D. Kakorin,<sup>11</sup> Konstantin S. Kuzmin,<sup>11,12</sup> Anselmo Meregalia,<sup>13</sup> Vadim A. Naumov,<sup>11</sup> Afroditi Papadopoulou,<sup>14</sup> Gabriel Perdue,<sup>9</sup> Komninos-John Plows,<sup>2</sup> Alon Sportes,<sup>3</sup> Noah Steinberg,<sup>9</sup> Vladyslav Syrotenko,<sup>7</sup> Jeremy Wolcott,<sup>7</sup> and Qiyu Yan<sup>15,4</sup>

(GENIE Collaboration)

<sup>1</sup>*University of Oxford, Dept. of Physics, Oxford OX1 3RH, United Kingdom*

<sup>2</sup>*University of Liverpool, Dept. of Physics, Liverpool L69 7ZE, United Kingdom*

<sup>3</sup>*Tel Aviv University, Tel Aviv 69978, Israel*

<sup>4</sup>*University of Warwick, Coventry CV4 7AL, United Kingdom*

<sup>5</sup>*University of Minnesota Twin Cities, Minneapolis, Minnesota 55455, USA*

<sup>6</sup>*University of Pittsburgh, Dept. of Physics and Astronomy, Pittsburgh, Pennsylvania 15260, USA*

<sup>7</sup>*Tufts University, Dept. of Physics and Astronomy, Medford, Massachusetts 02155, USA*

<sup>8</sup>*Instituto de Física Corpuscular, Paterna 46980, València, Spain*

<sup>9</sup>*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*

<sup>10</sup>*Massachusetts Institute of Technology, Dept. of Physics, Cambridge, Massachusetts 02139, USA*

<sup>11</sup>*Joint Institute for Nuclear Research (JINR), Dubna, Moscow Region 141980, Russia*

<sup>12</sup>*Alikhanov Institute for Theoretical and Experimental Physics (ITEP) of NRC “Kurchatov Institute”, Moscow 117218, Russia*

<sup>13</sup>*CENBG, Université de Bordeaux, CNRS/IN2P3, 33175 Gradignan, France*

<sup>14</sup>*Argonne National Laboratory, Argonne, Illinois 60439, USA*

<sup>15</sup>*University of the Chinese Academy of Sciences, Beijing 100864, China*



(Received 21 May 2024; accepted 6 September 2024; published 22 October 2024)

We present the first combined tuning, using GENIE, of four transverse kinematic imbalance measurements of neutrino-hydrocarbon scattering, both with and without pion final states, from the T2K and MINERvA experiments. As a proof of concept, we have simultaneously tuned the initial state and final-state interaction models (SF-CFG and hA, respectively), producing a new effective model that more accurately describes the data.

DOI: 10.1103/PhysRevD.110.072016

## I. INTRODUCTION

Neutrinos play a central role in advancing our understanding of physics and addressing fundamental inquiries in contemporary science. The Hyper-Kamiokande experiment

[1] aims to conduct precise measurements of charge parity violation within the neutrino sector, a phenomenon believed to be closely linked to the observed matter-anti-matter asymmetry in the Universe. Similarly, the Deep Underground Neutrino Experiment (DUNE) [2–6] promises the same, in addition to elucidating the neutrino mass ordering. Whether seeking to ascertain Standard Model parameters or probing for exotic phenomena, significant enhancements in both theoretical frameworks and computational simulations of neutrino-nucleus interactions are imperative. As neutrinos interact mainly with nucleons inside the nucleus, the interaction is subject to nuclear effects, namely the nucleon initial state and final-state interactions (FSIs). These effects are difficult to calculate and can alter the final event topology by changing the number of final state pions with respect to the interactions on

\*Contact author: weijun.li@physics.ox.ac.uk

†Contact author: mroda@liverpool.ac.uk

‡Contact author: jtenavidal@tauex.tau.ac.il

§Contact author: c.andreopoulos@cern.ch

||Contact author: xianguo.lu@warwick.ac.uk

Published by the American Physical Society under the terms of the *Creative Commons Attribution 4.0 International license*. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

free nucleons. Notably, the neutrino flux at DUNE yields a comparable proportion of events with and without pions in the final state, highlighting the pressing need for a generator capable of accurately describing both event topologies.

The large data samples and superb imaging capabilities of modern neutrino experiments offer us a detailed new look at neutrino interaction physics. Recently, the GENIE [7,8] Collaboration has made substantial progress toward a global tuning using neutrino, charged lepton, and hadron scattering data, in an attempt to integrate new experimental constraints with state-of-the-art theories and construct robust and comprehensive simulations of neutrino interactions with matter. Cross-experiment and cross-topology analyses are challenging tasks as each measurement features its unique selection criteria and various other aspects, such as the neutrino flux. GENIE has built an advanced tuning framework that enables the validation and tuning of comprehensive interaction models using an extensive curated database of measurements of neutrino, charged lepton, and hadron scattering off nucleus and nuclei. So far, the nonresonant backgrounds [9], hadronization [10], and the quasielastic (QE) and 2-particle-2-hole (2p2h) components [11] of the neutrino-nucleus interaction have been tuned with  $\nu_\mu$  and  $\bar{\nu}_\mu$  charged-current (CC) pionless ( $0\pi$ ) data from MiniBooNE, T2K, and MINERvA. A partial tune was performed for each experiment, highlighting the neutrino energy dependence on the QE and 2p2h tuned cross sections. Even though post-tune predictions enhanced the data description for each experiment, the added degrees of freedom were not sufficient to fully describe all CC $0\pi$  data and exhibited tensions with some proton observables [11]. More exclusive measurements result in additional model constraints. In addition, observables that are sensitive to targeted aspects of the complex dynamics of neutrino interactions are invaluable for model tuning. The transverse kinematic imbalance (TKI) [12,13], a final-state correlation between the CC lepton and the hadronic system, is a good example since it is sensitive to the initial-state nuclear environment and hadronic FSI. Our next step is to incorporate TKI data from experiments where various exclusive topologies at different energies are considered. This marks the first combined tuning on TKI data with and without pions in the final states and serves as the starting point of a more comprehensive tuning effort in the energy region most relevant for future accelerator-based neutrino experiments.

This paper is structured as follows: In Secs. II and III, we review the TKI measurements and GENIE models respectively. In Sec. IV, we detail the tuning considerations and procedures. Results are summarized in Sec. V, highlighting how the data-MC discrepancy in the MINERvA- $\pi^0$  TKI measurement [14] is resolved while maintaining good data-MC agreement elsewhere. We conclude in Sec. VI.

## II. TKI MEASUREMENTS

TKI is a methodology based on the conservation of momentum in neutrino interactions. In essence, it involves quantifying the imbalance between the observed transverse momentum of the final-state particles and the expected transverse momentum from neutrino interactions with free nucleons [12,13]. This “kinematic mismatch,” together with its longitudinal and three-dimensional variations [15,16] and the derived asymmetry [17], has been a crucial set of observables, establishing a pathway to extract valuable information about the participating particles and the underlying nuclear processes. Recent experimental results from neutrino experiments such as T2K [18,19], MINERvA [14,20–22], and MicroBooNE [23–27], as well as electron scattering experiments such as CLAS [28], highlight the efficacy of TKI.

In neutrino scattering off a free nucleon, the sum of the transverse components of the final products is expected to be zero, visualized through a back-to-back configuration between the final-state lepton and hadronic system in the plane transverse to the neutrino direction. Hence, in a neutrino interaction with a nucleus, the transverse momentum imbalance,  $\delta p_T$  [13], results from intranuclear dynamics, including Fermi motion and FSIs as shown in Fig. 1. The deviation from being back-to-back is quantified by the coplanarity angle  $\delta\phi_T$  [13], while the transverse boosting angle,  $\delta\alpha_T$  [13], represents the direction of  $\delta p_T$  within the transverse plane. Furthermore, analyzing the energy and longitudinal momentum budget [15,16] enables the conversion of  $\delta p_T$  to the emulated (initial) nucleon momentum,

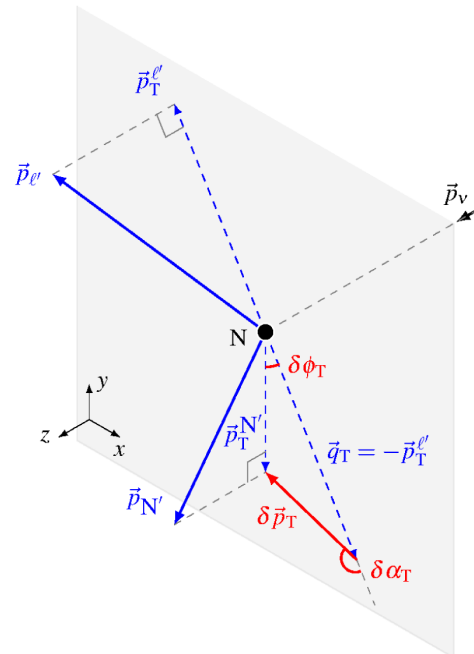


FIG. 1. Schematic illustration of the TKI variables. Diagram taken from Ref. [13].

$p_N$ , providing further insight into the Fermi motion; this conversion amounts to a correction on the order of  $\mathcal{O}(20\%)$  [29]. With one-body currents in the absence of FSIs,  $\delta\alpha_T$  remains uniform (except for second-order effects, such as variations in the center-of-mass energy), given the isotropic nature of the initial nucleon motion. However, as the final products propagate through the nuclear medium, they experience FSIs, thereby disturbing the isotropy and the Fermi motion peak of the  $\delta\alpha_T$  and  $p_N$  ( $\delta p_T$ ) distributions, respectively. Hence,  $\delta p_T$  and  $p_N$  elucidate the Fermi motion details, while  $\delta\alpha_T$  characterizes the FSI strength—crucial for understanding medium effects in neutrino interactions. A notable advantage of these observables is their minimal dependence on neutrino energy [13]. Moreover, the double TKI variable,  $\delta p_{TT}$  [12], is the projection of  $\delta\vec{p}_T$  along the axis perpendicular to the lepton scattering plane (hence “double”). In addition to its use for extracting neutrino-hydrogen interactions [12,30], it has also been applied to study nuclear effects in neutrino pion productions [14,19]. Its equivalent in pionless production,  $\delta p_{Tx}$ , has been proposed and studied together with its orthogonal companion,  $\delta p_{Ty}$ , in MINERvA [21].

This work surveyed four TKI data sets: T2K  $0\pi$  [18], T2K  $\pi^+$  [19], MINERvA  $0\pi$  [20,21], and MINERvA  $\pi^0$  [14] measurements. All four measurements require the presence of one CC muon and at least one proton in the final state. While the T2K- $\pi^+$  measurement requires exactly one  $\pi^+$ , and MINERvA- $\pi^0$  requires at least one  $\pi^0$ , the other two datasets require the absence of any pions. The definitions of the kinematic cuts for these samples are summarized in Table I.

Given the significant physics potential of TKI measurements, a wealth of literature has emerged on the combined

analysis of T2K and MINERvA CC $0\pi$  TKI data [11,31–37]. However, model studies incorporating TKI data with pion production have been notably absent. Since the TKI measurements are sensitive to the initial state and FSI, this work investigates a partial tune of these two processes, in an effort to derive an effective model describing neutrino-nuclear pion production, while maintaining the efficacy with the pionless production.

### III. GENIE MODEL SELECTION

GENIE uses formatted strings to uniquely define configurations, which are usually called tunes. Regardless of the name, they can refer to configurations either before or after an actual tuning procedure.

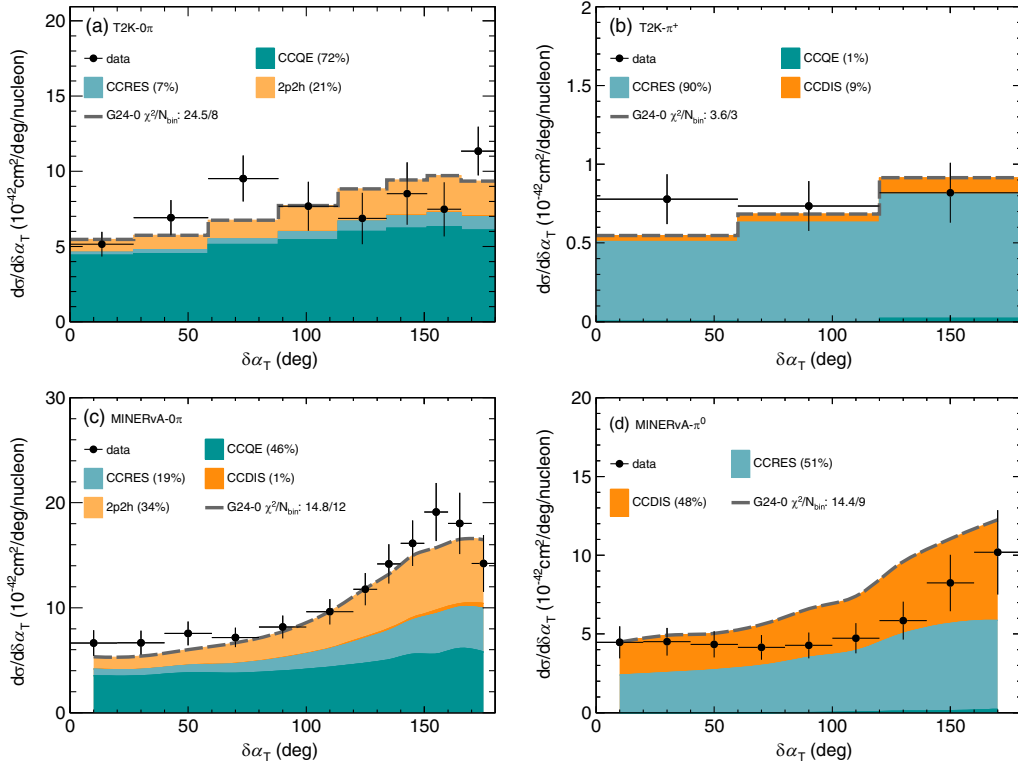
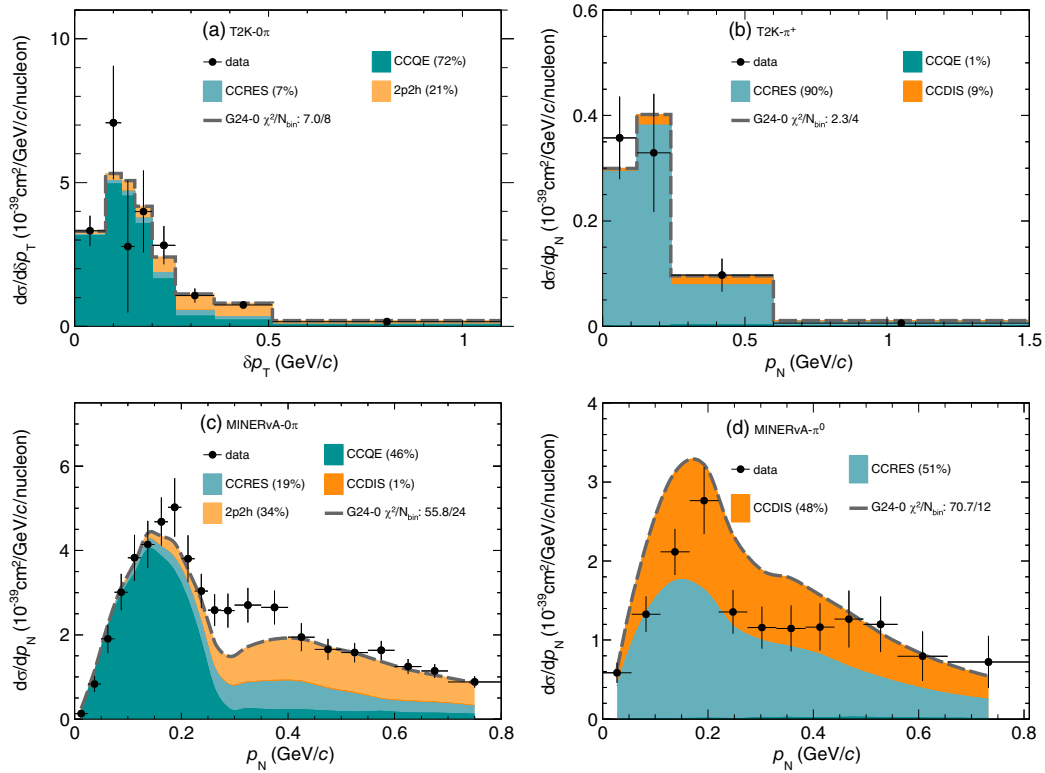
Free nucleon tuning effort in Ref. [9] has produced a tune, G18\_10a\_02\_11b, with better description of bubble chamber data. This is the starting point of the model we want to tune. The first effect to apply to this model is adding the proper initial state. Among the available GENIE initial state models, the best choice is the spectral-function-like correlated Fermi gas model (SF-CFG) [8,38,39]. SF-CFG is essentially based on the previously implemented local Fermi gas (LFG) model in GENIE with a larger number of adjustable degrees of freedom for tuning and improved physics. It differs from the previous implementation of LFG in two aspects. “Spectral-function-like” refers to the implementation of removal energy as a varying function rather than a fixed value, while “correlated” highlights the incorporation of the high-momentum tail above the Fermi momentum due to nucleon-nucleon short-range correlations (SRC), as evidenced by electron-scattering data [40]. These two improvements are considered and modeled in the Valencia model in Ref. [41]. Hence, SF-CFG resembles closer the initial state of the Valencia model than previous LFG implementations. More details on the SF-CFG can be found in Ref. [8]. Note that the original Valencia model in Ref. [41] has its own

TABLE I. Kinematic cuts for the samples of the TKI measurements.

Variables	Cuts ( $p$ in GeV/ $c$ )
	T2K- $0\pi$ [18]
$\vec{p}_\mu$	$0.25 < p_\mu, \cos\theta_\mu > -0.6$
$\vec{p}_p$	$0.45 < p_p < 1.0, \cos\theta_p > 0.4$
	T2K- $\pi^+$ [19]
$\vec{p}_\mu$	$0.25 < p_\mu < 7, \theta_\mu < 70^\circ$
$\vec{p}_p$	$0.45 < p_p < 1.2, \theta_p < 70^\circ$
$\vec{p}_\pi$	$0.15 < p_\pi < 1.2, \theta_\pi < 70^\circ$
	MINERvA- $0\pi$ [20,21]
$\vec{p}_\mu$	$1.5 < p_\mu < 10, \theta_\mu < 20^\circ$
$\vec{p}_p$	$0.45 < p_p < 1.2, \theta_p < 70^\circ$
	MINERvA- $\pi^0$ [14]
$\vec{p}_\mu$	$1.5 < p_\mu < 20, \theta_\mu < 25^\circ$
$\vec{p}_p$	$0.45 < p_p$

TABLE II. Model components of G24-0. Processes with nontrivial  $\Delta S$  and  $\Delta C$  are those with strangeness and charm production, respectively.

Simulation component	Model
Nuclear state	SF-CFG [8,38,39]
QE	Valencia [41]
2p2h	SuSAv2 [44]
QE $\Delta S = 1$	Pais [45]
QE $\Delta C = 1$	Kovalenko [46]
Resonance (RES)	Berger-Sehgal [47]
Shallow/Deep inelastic scattering (SIS/DIS)	Bodek-Yang [48]
DIS $\Delta C = 1$	Aivazis-Tung-Olness [49]
Coherent $\pi$ production	Berger-Sehgal [50]
Hadronization	AGKY [51]
FSI	INTRANUKE hA [52]

FIG. 2.  $\delta\alpha_T$  measurements decomposed in interaction types, compared to G24-0 prediction.FIG. 3. Similar to Fig. 2 but for the  $p_N$  (T2K- $\pi^+$ , MINERvA- $0\pi$  and MINERvA- $\pi^0$ ) and  $\delta p_T$  (T2K- $0\pi$ ) measurements.

method of modeling FSI. However, since FSI is factorized from other models in implementation in GENIE, this method is not used. Instead, GENIE develops the INTRANUKE hA FSI model (hA for short) for its simplicity and interpretability, which is the FSI candidate for the model to be tuned in this work.

Further improvements to the configuration to be tuned include: (1) using  $z$ -expansion axial-vector form factor [42] rather than the dipole form factor of the Valencia model in QE processes [41]; and (2) replacing the Valencia model [43] in 2p2h processes with the SuSAv2 model [44] since it covers larger region of the  $q^0$ ,  $q^3$  phase space.

The complete list of the model components is given in Table II. In the GENIE code, this configuration is identified as G24\_20i\_00\_000, and it is currently the default GENIE configuration for a number of LAr based experiments. As G24\_20i\_00\_000 is widely used in the paper, we will call it G24-0 for simplicity.

Comparing the GENIE predictions for G24-0 against TKI datasets, Figs. 2 and 3, shows that the model fails to accurately describe the MINERvA  $\pi^0$  [14] TKI measurement.

#### IV. TUNING INITIAL-STATE AND FSI MODELS ON TKI DATA

This study adopts the tuning procedure outlined in Ref. [11], utilizing  $N_{\text{par}}$  model parameters. The objective is to identify the best fit—that is, the optimal set of parameter values—within the parameter space that minimizes the  $\chi^2$  between model predictions and data. Given the high dimensionality of the parameter space, conducting a brute-force, point-by-point scan on a grid is infeasible. Therefore, a sufficient number of points is randomly sampled. Each point is used to run a full simulation. The simulation output for each data observable bin is then parametrized using Professor [53] with a polynomial of the model parameters of a chosen order,  $N_{\text{ord}}$ . The minimal number of points ( $N_s$ ) required for tuning  $N_{\text{par}}$  parameters with an order  $N_{\text{ord}}$  polynomial is determined by the combination formula,

$$N_s = \binom{N_{\text{par}} + N_{\text{ord}}}{N_{\text{ord}}}. \quad (1)$$

Given that  $N_s$  increases factorially, caution is essential when augmenting the number of parameters or the polynomial degree. Given that our tuning incorporates all 2 SF-CFG and 12 hA parameters (see Sec. IV B), a degree-4 parametrization necessitates over 6,000 generations. Hence, only up to order 4 polynomials are explored, and it has shown to reproduce the MC predictions well, as shown in Fig. 4.

Furthermore, to circumvent Peelle’s Pertinent Puzzle [36,56], the norm-shape (NS) transformation prescription [54,55] is applied. The extremal point is then found from a minimization of  $\chi^2$  between this NS-transformed

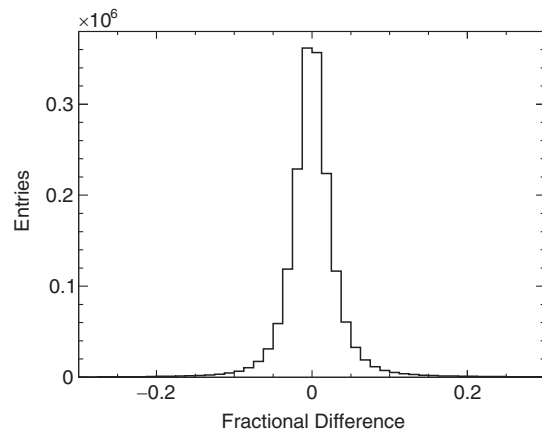


FIG. 4. Fractional difference for the bin-by-bin cross sections between MC truth and the parametrized approximation with order-4 polynomials, both in the norm-shape (NS) [54,55] space with AllPar. See Table IV for the definition of AllPar. The residual has a mean of 0.003 and a standard deviation of 0.073.

polynomial approximation and NS-transformed data. In minimization, priors, usually based on systematic uncertainty, can be imposed on each parameter to penalize it from getting too far from its default value. The following subsections elaborate on the specific choice of measurement observables and model parameters in this work.

##### A. Data points

The observables for the reported differential cross sections across the four TKI measurements—T2K- $0\pi$ , T2K- $\pi^+$ , MINERvA- $0\pi$ , and MINERvA- $\pi^0$ —are detailed in Table III. To identify the most sensitive variables, various combinations were evaluated for tuning purposes. A total of 26 combinations were assessed, with the superset comprising  $\delta\alpha_T$ ,  $\delta p_T$ ,  $\delta\phi_T$ ,  $p_N$ , and  $\delta p_{TT}$  (see Table VI in the Appendix for all combinations). Each combination was evaluated such that nonselected variables—including proton momentum and angle ( $p_p$  and  $\theta_p$ ) [20], as well as  $\delta p_{Tx}$  and  $\delta p_{Ty}$  [21] in MINERvA- $0\pi$ —were used solely for validation; the model is expected to accurately predict these variables post-tuning. When constructing the combinations, it was observed that  $\delta\phi_T$  is strongly dependent on neutrino beam energy [13], which suppresses its sensitivity to nuclear effects. Additionally, correlations exist between various observables, notably between  $\delta p_T$  and  $p_N$ . Therefore, determining the optimal combination is not straightforward; this study employs a criterion based on the  $\chi^2$  with the complete (tuned plus validation) observable set (Fig. 5).

##### B. Model parameters

As outlined in Sec. III, we focus exclusively on the currently tunable parameters within the SF-CFG and hA models. Therefore, 14 parameters (collectively

TABLE III. Observables of the TKI measurements and their binning. Those with “✓”s are used for tuning, while those without are for the respective validation. See Table IV for definitions of Combi-Best-RedPar and Combi-Best-AllPar.

Observables	No. of bins	Combi-Superset	Combi-Best-AllPar	Combi-Best-RedPar
T2K-0 $\pi$				
$\delta\alpha_T$	8	✓		✓
$\delta p_T$	8	✓	✓	✓
$\delta\phi_T$	8	✓		
T2K- $\pi^+$				
$\delta\alpha_T$	3	✓		✓
$p_N$	4	✓	✓	✓
$\delta p_{TT}$	5	✓		✓
MINERvA-0 $\pi$				
$\delta\alpha_T$	12	✓		✓
$p_N$	24	✓	✓	✓
$\delta p_T$	24	✓	✓	
$\delta\phi_T$	23	✓		
$p_p$	25			
$\theta_p$	26			
$\delta p_{Tx}$	32			
$\delta p_{Ty}$	33			
MINERvA- $\pi^0$				
$\delta\alpha_T$	9	✓		✓
$p_N$	12	✓	✓	✓
$\delta p_{TT}$	13	✓		✓

referred to as the AllPar set) are utilized for tuning, as detailed in Table IV. Nominal values and associated uncertainties for the hA model are taken from Table 17.3 in Ref. [52].

In our analysis of the SF-CFG model, the two most relevant parameters are  $R_{\text{SRC}}$ , representing the fraction of

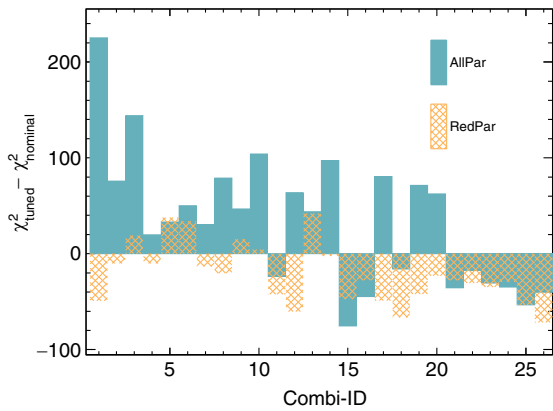


FIG. 5. Change of  $\chi^2$  calculated for the full (i.e., tuned plus validation) observable set as a function of the tuned combination (cf. Table VI in the Appendix). The two model parameter sets (AllPar and RedPar, see Table IV for definitions) are compared, and it can be seen that the respective minima happen at Combi-15 and 26.

the high-momentum tail from SRC, notably above the Fermi surface, and  $E_{\text{RM}}^{\text{C}}$ , marking the commencement of the nuclear removal energy distribution for carbon. A larger  $R_{\text{SRC}}$  indicates the presence of more energetic initial nucleons, while an increased  $E_{\text{RM}}^{\text{C}}$  implies that greater energy is necessary to liberate a nucleon from the carbon nucleus, so the product particles will possess lower energy. Given the novelty and limited constraints of the spectral-function-like implementation, this study employs relatively relaxed priors for  $R_{\text{SRC}}$  and  $E_{\text{RM}}^{\text{C}}$ , at  $0.12 \pm 0.12$  and  $0.01 \pm 0.005$  GeV, respectively.

GENIE users employ the hA model [52] for the majority of neutrino analysis, owing to its straightforward reweighting capability. Unlike cascade models such as the hN model [52], this model determines the FSI type for each hadron exactly once, without considering further propagation of rescattered hadrons within the nucleus. Its model parameters are explained as follows.

- (1) The model evaluates a mean free path (MFP),  $\lambda$ , which varies with the hadron’s energy,  $E$ , and its radial distance,  $r$ , from the nucleus center, as follows:

$$\lambda(E, r) = \frac{1}{\sigma_{\text{hN,tot}}(E)\rho(r)}, \quad (2)$$

TABLE IV. Tuneable parameters and their ranges in the SF-CFG (*uppermost* group) and hA (*lower* groups, uncertainties from Ref. [52]) models. Parameters to be tuned in the two sets are marked with “✓”. See later text for definitions of G24-t and G24-c.

Parameter	Nominal (G24-0)	Range In Tuning SF-CFG	AllPar (G24-t)	RedPar (G24-c)
$R_{\text{SRC}}$	0.12	(0.0, 0.5)	✓	✓
$E_{\text{RM}}^{\text{C}}$	0.01	(0.0, 0.2)	✓	
		hA		
$S_{\lambda}^{\pi^{\pm}}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	
$S_{\lambda}^{\pi^0}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	✓
$S_{\lambda}^{\text{N}}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	
$S_{\text{CEX}}^{\pi}$	$1.0 \pm 0.5$	(0.0, 3.0)	✓	✓
$S_{\text{CEX}}^{\text{N}}$	$1.0 \pm 0.5$	(0.0, 3.0)	✓	✓
$S_{\text{INEL}}^{\pi}$	$1.0 \pm 0.4$	(0.0, 3.0)	✓	
$S_{\text{INEL}}^{\text{N}}$	$1.0 \pm 0.4$	(0.0, 3.0)	✓	
$S_{\text{ABS}}^{\pi^{\pm}}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	
$S_{\text{ABS}}^{\pi^0}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	
$S_{\text{ABS}}^{\text{N}}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	✓
$S_{\text{PIPD}}^{\pi}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	
$S_{\text{PIPD}}^{\text{N}}$	$1.0 \pm 0.2$	(0.0, 3.0)	✓	✓

where  $\sigma_{\text{hN,tot}}$  is the total hadron-nucleon cross section, and  $\rho$  is the number density of the nucleons inside the nucleus. Once a hadron is generated inside the nucleus, it travels a distance  $\lambda$  before a probabilistic determination of rescattering occurs, inversely related to  $\lambda$ . If no rescattering happens, the hadron will be propagated for another  $\lambda$ , and the dice will be thrown again. The cycle repeats until a rescattering happens or the hadron is propagated far enough to escape the nucleus without any interactions. If a rescattering happens, the type of rescattering is evaluated according to their cross section. The relative probabilities of each type are extracted from stored hadron scattering data [57–63]. After a specific rescattering type is chosen, the rescattered product will be generated. These rescattered products will not undergo further rescattering and will be considered out of the nucleus as final state particles. The MFP values for pions and nucleons can be modified using scaling factors ( $S_{\lambda}^{\pi^{\pm}}$ ,  $S_{\lambda}^{\pi^0}$ , and  $S_{\lambda}^{\text{N}}$ ) detailed in Table IV. Both charged pions,  $\pi^+$  and  $\pi^-$ , are scaled by the same factor,  $S_{\lambda}^{\pi^{\pm}}$ , due to isospin symmetry. The  $\lambda$  for  $\pi^0$  is calculated from those of charged pions based on isospin symmetry as well and can be adjusted separately by  $S_{\lambda}^{\pi^0}$ . If the factor is larger than 1,  $\lambda$  will increase, and hence, the total number of steps required for the hadron to escape the nucleus reduces, thereby reducing the average probability of rescattering. Given the high precision with which total hadron-nucleus cross sections are determined [57–61], strict Gaussian priors (error— $\sigma$ , not to be confused with the various cross sections—of 0.2, same as the

systematic uncertainties shown in Table IV) are applied to the MFP scaling factors. Varying  $S_{\lambda}^{\pi^0}$  by  $\pm 2.5\sigma$  significantly affects the MINERvA- $\pi^0$   $d\sigma/dp_{\text{N}}$ , as depicted in Fig. 6. Decreasing MFP naturally leads to increased rescattering. Hence, fewer  $\pi^0$ s can escape the nucleus, and fewer events will have  $\pi^0$  in the final states, thereby reducing the cross section.

- (2) The relative probability of each rescattering type can be adjusted by a scale factor. There are four rescattering types for both nucleons and pions available for tuning. They are charge exchange (CEX), inelastic scattering (INEL), absorption (ABS), and pion production (PIPD). The default energy-dependent probability distributions come from hadron data tuning

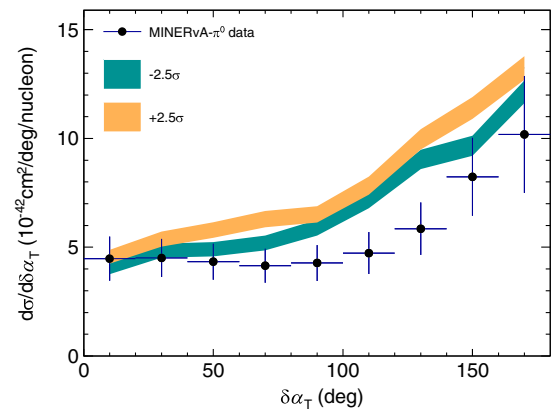


FIG. 6. Effect of varying  $S_{\lambda}^{\pi^0}$  by  $\pm 2.5\sigma$  compared to the MINERvA- $\pi^0$  measurement. Each band’s width indicates the GENIE prediction’s statistical uncertainty from  $10^5$  events.

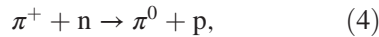
[57–61] and can be adjusted by respective scaling factors, like  $S_{\text{CEX}}^\pi$  for CEX of all pions, listed in Table IV. The relative probabilities of all four rescattering types will be normalized such that they sum up to one. Scaling factors directly multiply the probability of each type, with new relative probabilities subsequently renormalized. More specifically, suppose the initial fraction of absorption,  $f_{\text{ABS}}$ , is 0.5, which means that the sum of the other fractions is  $f_{\text{other}} = f_{\text{INEL}} + f_{\text{CEX}} + f_{\text{PIPD}} = 1 - 0.5 = 0.5$ . If  $f_{\text{ABS}}$  is scaled up with  $S_{\text{ABS}} = 2$ , the new absorption fraction is

$$f'_{\text{ABS}} = \frac{f_{\text{ABS}} * S_{\text{ABS}}}{f_{\text{ABS}} * S_{\text{ABS}} + f_{\text{oth}}} = 2/3, \quad (3)$$

which is not equal to simply multiplying  $f_{\text{ABS}}$  by  $S_{\text{ABS}}$  due to the presence of the other rescattering types. Hence, it is unavoidable that a change of the scaling factor for one type would affect others. Nonetheless, when the cross section is plotted with a breakdown according to rescattering type, the interpretation will be clear regarding the increase or decrease of a particular FSI type. This correlation implies that the scaling parameters' effective impact on FSI fate is often less pronounced than what the raw numbers suggest. As illustrated in the above example,  $f_{\text{ABS}}$  is only scaled up by  $f'_{\text{ABS}}/f_{\text{ABS}} = 4/3 = 1.3$  instead of  $S_{\text{ABS}} = 2$ . Hence, a relatively more relaxed Gaussian prior, with a  $\sigma$  of 0.5, is placed on the FSI fate scales, slightly larger than the systematic uncertainties shown in Table IV.

Understanding each rescattering type is crucial for interpreting tuning results. A detailed elaboration of the four tunable FSI types is given as follows.

(a) CEX involves changing the charge of the participating particles; for example,

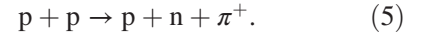


or vice versa. This rescattering type is crucial for event topologies requiring the presence of a pion; depending on the signal pion charge, CEX could migrate events between signal and background.

(b) INEL is the case where the nucleus is left in an excited state after the rescattering. This category only contains the situation where a single additional nucleon is emitted/knocked out after rescattering. Since it does not affect the number of pions produced, it will not convert an event from a pionless topology to a pion-production topology. The effects on nucleons are twofold. Firstly, it can alter the number of signal events within each event topology. If the inelastic rescattering leads to two low-momentum protons below the detection threshold as opposed to a

high-momentum proton, this signal event will be discarded as no protons are observed. Secondly, INEL invariably changes the kinematics of the rescattering particle. Be it the leading proton or the leading pion, based on which the TKI observables are calculated, the TKI distribution shape will be affected. Hence, while  $S_{\text{INEL}}^{\text{N}}$  would affect all four datasets,  $S_{\text{INEL}}^\pi$  would only affect T2K- $\pi^+$  and MINERvA- $\pi^0$ .

- (c) ABS refers to the case where the particle undergoes an interaction so that it does not emerge as a final particle. For example,  $\pi^+$  can interact with two or more nucleons, initially forming a baryon resonance that subsequently interacts with other nucleons, emitting multiple nucleons rather than pions. Hence, the  $\pi^+$  would not emerge from the nucleus anymore.
- (d) PIPD happens for energetic particles where an extra pion emerges as a result of the rescattering, such as



Such an interaction significantly alters the event topology.

## V. THE FIRST TKI-DRIVEN GENIE TUNES

Given the intricate correlations among model parameters, determining which ones are most sensitive to the data proves challenging. Considerable correlation and anticorrelation are observed between different FSI fates for the same particle, e.g.,  $S_{\text{ABS}}^{\pi^\pm}$  and  $S_{\text{CEX}}^\pi$ , as shown in Figs. 7(a) and 7(b). Our tuning began with the AllPar parameter set, achieving the best result with Combi-15, which is denoted as Combi-Best-AllPar in Table III (cf. Fig. 5 for the  $\chi^2$  distribution of all combinations). Observations showed that for most combinations, certain parameters remained close to their default values, within  $1\sigma$  of the imposed priors, such as  $S_{\text{PIPD}}^\pi$  since none of the datasets has a significant contribution from PIPD of pions—removing these parameters from the tuning process does not significantly affect the outcome. Therefore, we then constructed a reduced set comprising only  $R_{\text{SRC}}$ ,  $S_\lambda^{\pi^0}$ ,  $S_{\text{CEX}}^\pi$ ,  $S_{\text{CEX}}^{\text{N}}$ ,  $S_{\text{ABS}}^{\text{N}}$ , and  $S_{\text{PIPD}}^{\text{N}}$  (denoted as RedPar in Table IV) and ran the tuning on the 26 combinations with it. Note that the parameters most correlated with  $S_{\text{CEX}}^\pi$  and  $S_\lambda^{\pi^0}$  are not present in RedPar. The RedPar tuning proved more stable, with nearly all combinations showing a more negative  $\chi^2$  change than the AllPar tuning, as shown in Fig. 5. Moreover, the single-observable (Combi-1 to 5), single-measurement (Combi-6 to 9), single-experiment (Combi-10 and 11), and single-topology (Combi-12 and 13) combinations all systematically yielded no or limited overall (tuned plus validation) improvement post-tuning. The best tuning then occurred with Combi-26

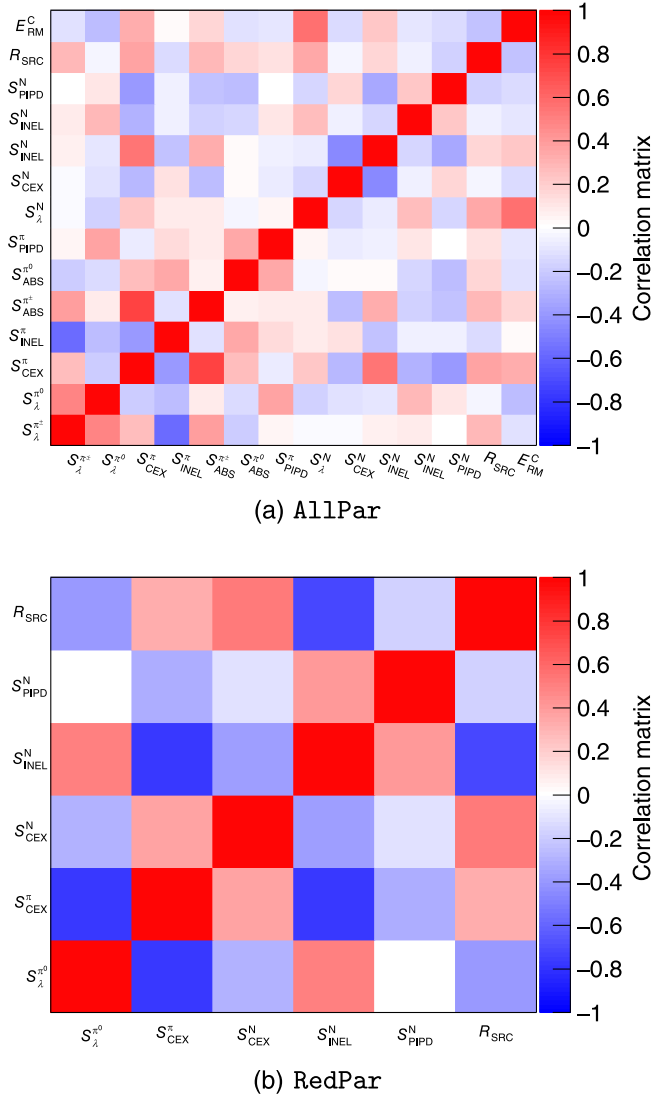


FIG. 7. Postfit Correlation coefficient on Combi-26.

(Combi-Best-RedPar in Table III), that is, all TKI variables excluding  $\delta\phi_T$  and  $\delta p_T$  (unless no  $p_N$  is available) — $\delta\alpha_T$ ,  $p_N$  (or  $\delta p_T$  if  $p_N$  is unavailable), and  $\delta p_{TT}$ —from pionless and pion-production measurements in both T2K and MINERvA.

Table V summarizes the primary results of our tuning process. The tuned RedPar, G24\_20i\_06\_22c (G24-c) is derived from the observable set Combi-Best-RedPar, and the tuned AllPar, G24\_20i\_fullpara\_alt (G24-t) is from Combi-Best-AllPar. The upper part of Table V contains the parameter values. For the G24-c tune, changes to the SF-CFG model were moderately sized ( $R_{\text{SRC}}$  increases from 0.12 to 0.15), while in the hA model,  $S_{\lambda}^{\pi^0}$  and  $S_{\text{CEX}}^{\pi}$  are highly suppressed and the nucleon FSI has larger CEX and PIPD and lower ABS. The tune G24-t differs most significantly from G24-c in  $R_{\text{SRC}}$  and  $S_{\text{CEX}}^{\pi}$ . The table's lower section illustrates  $\chi^2$  improvements for selected observables and their corresponding validation sets.

TABLE V. Parameters in G24-0, G24-c, and G24-t. Those without errors are not tuned. Lower section: `combi` indicates that the following  $\chi^2$  sums are calculated for the tuned measurements [i.e., Combi-Best-RedPar (98 bins) and Combi-Best-AllPar (72 bins) for G24-c and G24-t, respectively], while `vald` indicates that the respective validation sets are used; `untuned` means that the  $\chi^2$  calculation uses the nominal values of the parameters, while `tuned` denotes the tuned ones. `diff` displays the improvement achieved by the respective tuning.

Parameter	Nominal (G24-0)	RedPar (G24-c)	AllPar (G24-t)
SF-CFG			
$R_{\text{SRC}}$	0.12	$0.15 \pm 0.08$	$0.30 \pm 0.05$
$E_{\text{RM}}^{\text{C}}$	0.01	0.01	$0.011 \pm 0.003$
hA			
$S_{\lambda}^{\pi^{\pm}}$	$1.0 \pm 0.2$	1.0	$1.11 \pm 0.16$
$S_{\lambda}^{\pi^0}$	$1.0 \pm 0.2$	$0.22 \pm 0.07$	$0.17 \pm 0.06$
$S_{\lambda}^{\text{N}}$	$1.0 \pm 0.2$	1.0	$1.20 \pm 0.12$
$S_{\text{CEX}}^{\pi}$	$1.0 \pm 0.5$	$0.26 \pm 0.12$	$1.53 \pm 0.37$
$S_{\text{CEX}}^{\text{N}}$	$1.0 \pm 0.4$	$1.43 \pm 0.34$	$1.41 \pm 0.38$
$S_{\text{INEL}}^{\pi}$	$1.0 \pm 0.4$	1.0	$0.67 \pm 0.30$
$S_{\text{INEL}}^{\text{N}}$	$1.0 \pm 0.4$	1.0	$1.26 \pm 0.48$
$S_{\text{ABS}}^{\pi^{\pm}}$	$1.0 \pm 0.2$	1.0	$1.59 \pm 0.31$
$S_{\text{ABS}}^{\pi^0}$	$1.0 \pm 0.2$	1.0	$0.90 \pm 0.28$
$S_{\text{ABS}}^{\text{N}}$	$1.0 \pm 0.2$	$0.25 \pm 0.28$	$0.28 \pm 0.27$
$S_{\text{PIPD}}^{\pi}$	$1.0 \pm 0.2$	1.0	$1.12 \pm 0.30$
$S_{\text{PIPD}}^{\text{N}}$	$1.0 \pm 0.2$	$2.05 \pm 0.48$	$1.27 \pm 0.48$
$\chi^2$ for combi			
untuned		231.75	161.26
tuned		174.84	122.53
diff		-56.91	-38.73
$\chi^2$ for vald			
untuned		229.5	299.99
tuned		214.7	263.41
diff		-14.8	-36.58
$\chi^2$ for combi+vald			
untuned		461.25	461.25
tuned		389.54	385.94
diff		-71.71	-75.31

Crucially, tuning should also enhance the accuracy of validation set descriptions. Overall, both tunes have substantial reduction in the total  $\chi^2$ . In the following, we will discuss the two tunes.

### A. Reduced tune: G24-c

To assess the tune's quality, we plot predictions using G24-c for  $\delta\alpha_T$  and  $p_N$  in Figs. 8 and 9, respectively. For

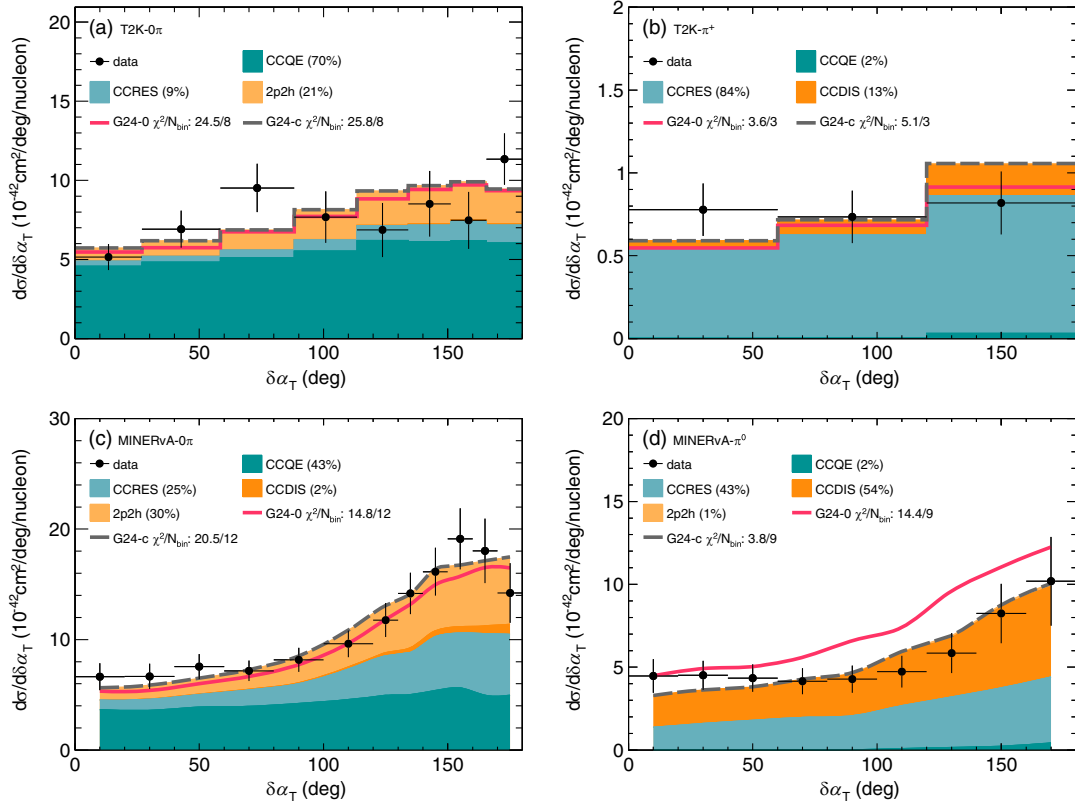
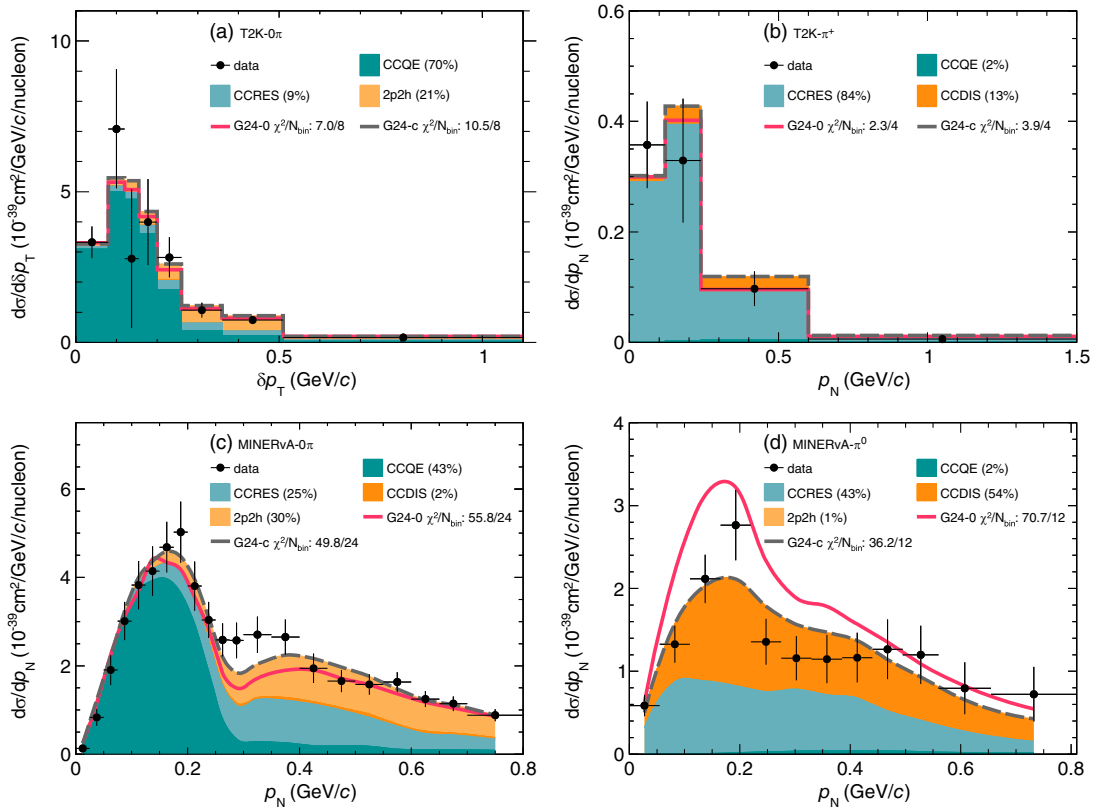


FIG. 8. Similar to Fig. 2 but with G24-c. The G24-0 prediction is also plotted for comparison.

FIG. 9. Similar to Fig. 8 but for the  $p_N$  (T2K- $\pi^+$ , MINERvA- $0\pi$  and MINERvA- $\pi^0$ ) and  $\delta p_T$  (T2K- $0\pi$ ) measurements.

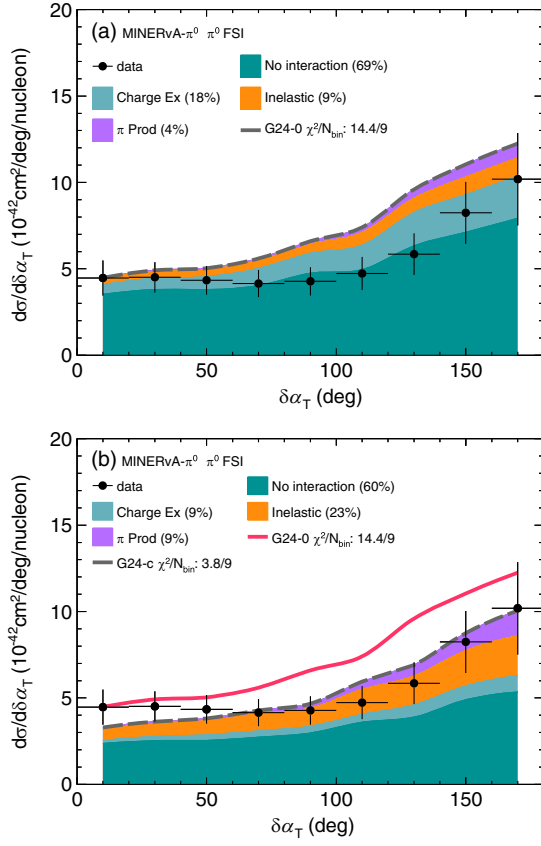


FIG. 10. MINERvA- $\pi^0$   $\delta\alpha_T$  measurement compared to GENIE predictions decomposed in  $\pi^0$  FSI fates with (a) G24-0 and (b) G24-c.

T2K- $0\pi$ , T2K- $\pi^+$  and MINERvA- $0\pi$ , the new  $\chi^2$  values are comparable to those obtained with G24-0, changing less than the number of bins. For MINERvA- $\pi^0$ , G24-c is distinctly better than G24-0, thereby demonstrating the possibility of simultaneous good descriptions of both pionless and pion production samples with constrained parameters from cross-topology TKI tuning. Note that in the region of  $p_N \sim 0.3$  GeV/c, the model deficit previously reported in Ref. [20] (and also shown in Fig. 3) persists even after the fit. The physical origin of this deficit might be attributed to the strength of the 2p2h contribution [20], but it is still a topic of significant discussion within the community.

Examining individual datasets closely helps to understand the origin of the improvement. The decomposition of the cross section for  $\delta\alpha_T$  according to the FSI fates of the  $\pi^0$  is presented in Fig. 10. The hA model rescatters primary interaction products only once and records this event with a rescattering code for these particles. The rescattered products, which are the final outputs, are stored as daughters of the primary interaction products. Hence, the FSI fate that produces the final products can be checked by the rescattering code of their first parent. More specifically, we loop through all final-state particles to look for the leading  $\pi^0$  and check the rescattering code of its first parent.

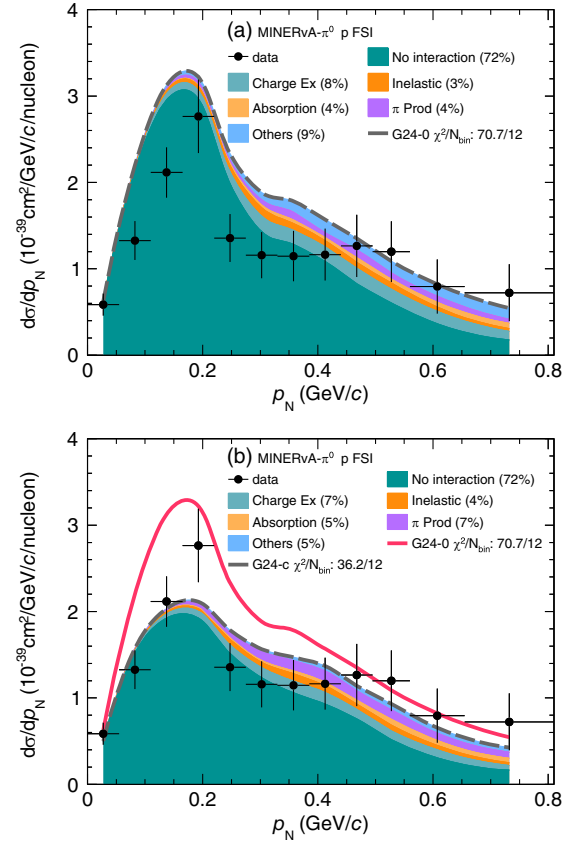


FIG. 11. Similar to Fig. 10 but for  $p_N$  with proton FSI fate decomposition. Note that the “absorption” of the proton is referring to the absorption of the  $\pi^+$  from the decay of  $\Delta^{++}$ , which could lead to emission of nucleons as discussed in Sec. IV B.

In MINERvA- $\pi^0$ , the number of  $\pi^0$ s undergoing no FSI (“no interaction,” as shown in the figures) is adjustable via the  $S_\lambda^{\pi^0}$  parameter. A decrease in  $S_\lambda^{\pi^0}$  reduces the size of the “no interaction” events, as is shown in Fig. 10(b) when compared to Fig. 10(a). Meanwhile, an increase in  $\pi^0$  rescattering only manifests in pionless measurements through ABS. There is indeed an increase of ABS for  $\pi^0$  in T2K- $0\pi$  and MINERvA- $0\pi$ , but the fraction is small to begin with, so the impact is small. The increase of  $\pi^0$  rescattering can increase T2K- $\pi^+$  via CEX as discussed below. However, due to the significant suppression of CEX, its impact on T2K- $\pi^+$  is also minimal (for detailed breakdown, see Figs. 14–17 in the Appendix).

As shown in Fig. 10(a), the MINERvA- $\pi^0$  prediction from the nominal tune has considerable contributions from CEX controlled by  $S_{\text{CEX}}^\pi$ . It does not impact T2K- $0\pi$  and MINERvA- $0\pi$  as CEX only changes the pion type without removing them. Hence, events with pions in the final state would be rejected in these two datasets regardless of the pion charge. In principle, CEX will also migrate events between the signal and background definitions for T2K- $\pi^+$  when a  $\pi^0$  is converted to a  $\pi^+$  and vice versa. However, T2K- $\pi^+$  does not have a considerable CEX fraction to

begin with, so changing CEX will have a minimal impact on its prediction. Hence, this is one of the most effective paths to decrease the MINERvA- $\pi^0$  cross section prediction without affecting other measurements, and indeed the  $S_{\text{CEX}}^\pi$  is heavily suppressed as shown in Table V and illustrated in Fig. 10(b) compared to Fig. 10(a). However, due to the intricate correlation between the FSI fates in the hA implementation discussed in Sec. IV B, although  $S_{\text{INEL}}^\pi$  and  $S_{\text{PIPD}}^\pi$  are not modified explicitly, a considerable increase is observed in both INEL and PIPD for MINERvA- $\pi^0$ .

Suppressing both “no interaction” and CEX for  $\pi^0$  adjusts the MINERvA- $\pi^0$  prediction to the appropriate magnitude with small impacts on other datasets. The effects of the other parameter changes are more transparent in the comparison of the  $p_N$  distribution of MINERvA- $\pi^0$  between Figs. 11(a) and 11(b). A relatively large increase in  $S_{\text{CEX}}^\pi$  and  $S_{\text{PIPD}}^\pi$  moves events away from the Fermi motion peak at  $p_N \leq 0.25$  GeV/c; the same effect can be achieved by an increased  $R_{\text{SRC}}$ .

Another impact of the larger  $S_{\text{PIPD}}^\pi$  manifests in the appearance of 2p2h (1%) and CCQE (2%) in the MINERvA- $\pi^0$  cross section, as shown in Fig. 9(d). These neutrino-nucleon interactions do not produce pions, but the product nucleons can produce pions via rescattering in the nucleus, leading to a contribution to topologies with pions in the final states. However, such contribution remains small.

Although the interaction models, e.g., 2p2h, are not directly tuned, the respective fractions have undergone small changes in Figs. 8 and 9, compared to Figs. 2 and 3. This is a result of the increased  $R_{\text{SRC}}$  which leads to more events with higher initial nucleon momenta such that interactions requiring higher initial state energy, such as RES and DIS, will be more frequent. This is reflected by the increase in the ratios of these interactions: For example, RES increases from 19% (Fig. 3) to 25% [Fig. 9(c)] for MINERvA- $0\pi$ . Hence, the other interactions, although not directly tuned, will experience a decrease in the fraction of total cross section: For example, 2p2h drops from 34% [Fig. 3(c)] to 30% [Fig. 9(c)] for MINERvA- $0\pi$ .

### B. Alternative tune: G24-t

A similar improvement is achieved by the intermediate tune, G24-t, illustrated in Fig. 12. Instead of significantly increasing proton PIPD to elevate the  $p_N$  tail, this tune notably enhances  $R_{\text{SRC}}$ —more-energetic nucleons are present, and therefore, the RES  $p_N$  peak is shifted to the right in Fig. 12(a) with respect to Fig. 9(d). Rather than suppressing  $S_{\text{CEX}}^\pi$  [cf. Fig. 17(d) in the Appendix], this tune significantly increases it [Fig. 12(b)] to offset the further reduction of “no interaction”  $\pi^0$  events (smaller  $S_\lambda^{\pi^0}$  in AllPar, cf. Table V). Overall, this leads to a reduction of 75.31 in  $\chi^2$  (Table V), similar to G24-c, as well as a good data-MC agreement across all datasets.

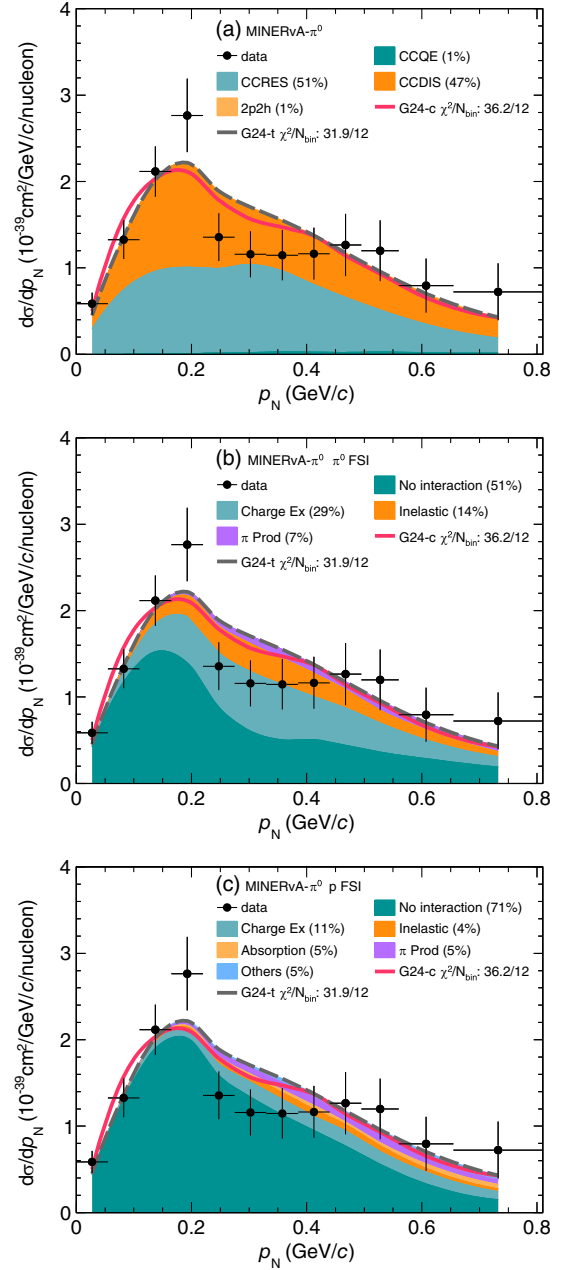


FIG. 12. MINERvA- $\pi^0$   $p_N$  measurement compared to GENIE predictions decomposed in (a)  $\nu$ -N interaction, (b)  $\pi^0$  FSI, and (c) proton FSI for the alternative tune, G24-t.

### C. Discussion

For both G24-c and G24-t, the fit results suggest extreme values, but the effect of the parameter change is less effective than the change in the parameter suggests. For the individual FSI fate cross section, due to the renormalization step after scaling, the effective change for one particular fate is smaller than the scaling parameter magnitude as explained in Sec. IV B. As for the total FSI cross sections, only that of  $\pi^0$  is changed. The total  $\pi^0$ -C scattering cross section for pion kinetic energy is plotted in Fig. 13 to assess the overall

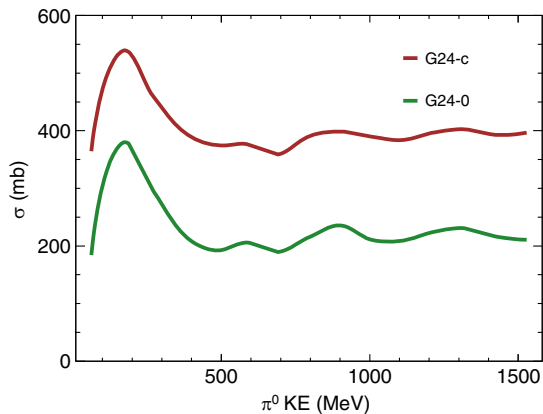


FIG. 13. Change in MC prediction for  $\pi^0$  cross section between G24-0 and G24-c.

impact. The shape remains similar and consistent with Fig. 2.23 ( $\pi^+$ -C reactions) in Ref. [52], while the peak magnitude increases from 380 to 538 mb, a 48% increase, considerably less than the scaling factor,  $S_\lambda^{\pi^0} = 0.22$ , seemingly suggests. Moreover, the default  $\pi^0$  parameter values are calculated from charged pion data assuming isospin symmetry rather than extracted directly from experimental data. Hence, this modification does not violate any existing agreement with hadron scattering data. In addition, agreement of this tune with most of the non-TKI neutrino datasets available is unchanged before and after the tune, thereby demonstrating the physicality of the tune.

Exclusive electron scattering data is better suited to constrain FSI models. Regrettably, most electron scattering data to date is inclusive [64], and it is not well suited for the tuning of FSI. The “electrons for neutrinos” (e4nu) collaboration exploits data from the CLAS6 spectrometer to measure exclusive final states, such as a  $1p0\pi$  cross-section measurement on carbon as a function of TKI variables [28], but for electron-scattering pion production, currently, the GENIE predictions overpredict data due to the lack of tuning of the GENIE models on free nucleon data. Such big uncertainties in the pion production model mask possible nuclear model or FSI dependence. In order to be sensitive to these effects, one needs to first tune electron-scattering to free nucleon data. The current tuning study could be extended in the future when data from dedicated electro-pion production measurements are available.

In conclusion, this tune is a valid and effective model that can be used as a starting point for an analysis. Further refinement will be available once more data will be included in the fit, including non-TKI observables and electron-scattering data.

## VI. SUMMARY AND OUTLOOK

This work represents the first global tuning effort on TKI data. Our partial tune of the SF-CFG and hA models, G24\_20i\_06\_22c (G24-c), provides an effective theory to better describe both the neutrino-hydrocarbon

pionless and pion production. The largest change in the model is demanded by the MINERvA- $\pi^0$  TKI measurement [14] that was significantly overestimated in GENIE. The improvement is crucial for future precision GeV neutrino experiments. This tuning configuration has been integrated into the master branch of GENIE and is slated for inclusion in the upcoming release.

To develop an effective model, we focused on the most sensitive parameters,  $R_{\text{SRC}}$ ,  $S_\lambda^{\pi^0}$ ,  $S_{\text{CEX}}^\pi$ ,  $S_{\text{CEX}}^N$ ,  $S_{\text{ABS}}^N$ , and  $S_{\text{PIPD}}^N$ , of the SF-CFG and hA models, reducing the total number of parameters from 14 to 6. The corresponding best combination of observables is  $\delta\alpha_T$ ,  $p_N$  (or  $\delta p_T$  if no  $p_N$  is available), and  $\delta p_{\text{TT}}$ , from pionless and pion-production measurements in both T2K and MINERvA. In doing so, the pion production model is held fixed such that the tuned values of  $S_\lambda^{\pi^0}$ ,  $S_{\text{CEX}}^\pi$ ,  $S_{\text{ABS}}^N$ , and  $S_{\text{PIPD}}^N$  depart from the established constraints from hadron measurements. A different pion production model might compensate part of these model tuning effects [65]. We also chose the hA FSI model for practical reasons. Its sequential treatment of MFP and rescattering should be considered as a numerical advantage (as was also pointed out in Ref. [11]). Other more sophisticated models could be considered in future tunes. With the existing four TKI measurements from T2K and MINERvA, we have also derived other effective tunes like G24\_20i\_fullpara\_alt (G24-t). The degeneracy can be resolved with additional data, particularly from argon-based measurements [23–27,66] and e4nu data [28].

## ACKNOWLEDGMENTS

We thank the CC-IN2P3 Computing Center providing computing resources and for their support. X. L. is supported by the STFC (UK) Grant No. ST/S003533/2. J. T.-V. acknowledges the support of the Raymond and Beverly Sackler scholarship.

## APPENDIX: TEST COMBINATIONS AND COMPLEMENTARY TUNES

### 1. Combinations of Observables

Table VI displays the tested combinations of observables:

- (i) *Combi-1 through 5* are cross-experiment selections of a single observable. For example, *Combi-3* uses only  $\delta\alpha_T$  from all four datasets. If a chosen observable is absent from a dataset, that dataset is excluded from the specific combination. For example, T2K- $0\pi$  is not used in *Combi-1* due to the absence of  $p_N$ .
- (ii) *Combi-6 through 9* incorporate all variables from a single measurement. For example, *Combi-9* uses MINERvA- $\pi^0$  only.
- (iii) *Combi-10 to 13* uses two out of four measurements according to the experiment or topology. For example, *Combi-10* uses only the two datasets from T2K, and *Combi-13* only with pion production.

TABLE VI. Specifications of observable combinations within the tuning superset in Table III. Combi-15 is Best-AllPar, Combi-24 is Superset, and Combi-26 is Best-RedPar.

Combi-	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15 (Best- AllPar)	16	17	18	19	20	21	22	23	24 (Super- set)	25	26 (Best- RedPar)	
T2K-0 $\pi$																											
$\delta\alpha_T$			✓			✓				✓	✓		✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
$\delta p_T$		✓				✓				✓	✓			✓			✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
$\delta\phi_T$				✓		✓				✓	✓									✓	✓		✓				
T2K- $\pi^+$																											
$\delta\alpha_T$			✓				✓			✓			✓	✓		✓	✓	✓	✓	✓	✓		✓		✓	✓	✓
$p_N$	✓						✓			✓			✓		✓			✓			✓	✓	✓	✓	✓	✓	✓
$\delta p_{TT}$					✓		✓			✓			✓	✓				✓	✓	✓		✓		✓			✓
MINERvA-0 $\pi$																											
$\delta\alpha_T$			✓				✓			✓	✓		✓		✓	✓	✓	✓	✓	✓	✓		✓		✓	✓	✓
$p_N$	✓						✓			✓	✓		✓		✓			✓			✓	✓	✓	✓	✓	✓	✓
$\delta p_T$		✓					✓			✓	✓		✓		✓		✓		✓		✓	✓	✓	✓	✓	✓	✓
$\delta\phi_T$				✓			✓			✓	✓		✓							✓	✓		✓				✓
MINERvA- $\pi^0$																											
$\delta\alpha_T$			✓				✓			✓	✓		✓	✓		✓	✓	✓	✓	✓	✓		✓		✓	✓	✓
$p_N$	✓						✓			✓	✓		✓		✓			✓			✓	✓	✓	✓	✓	✓	✓
$\delta p_{TT}$					✓		✓			✓	✓		✓	✓				✓	✓	✓		✓		✓			✓

(iv) Combi-14 to 17 combine two separate combinations:

$$\begin{aligned} \text{Combi} - 14 &= \text{Combi} - 3 \cup \text{Combi} - 5, \\ \text{Combi} - 15 &= \text{Combi} - 1 \cup \text{Combi} - 2, \\ \text{Combi} - 16 &= \text{Combi} - 1 \cup \text{Combi} - 3, \\ \text{Combi} - 17 &= \text{Combi} - 2 \cup \text{Combi} - 3. \end{aligned}$$

(v) Combi-18 through 22 result from merging three combinations:

$$\begin{aligned} \text{Combi} - 18 &= \text{Combi} - 1 \cup \text{Combi} - 3 \cup \text{Combi} - 5, \\ \text{Combi} - 19 &= \text{Combi} - 2 \cup \text{Combi} - 3 \cup \text{Combi} - 5, \\ \text{Combi} - 20 &= \text{Combi} - 3 \cup \text{Combi} - 4 \cup \text{Combi} - 5, \\ \text{Combi} - 21 &= \text{Combi} - 1 \cup \text{Combi} - 2 \cup \text{Combi} - 3, \\ \text{Combi} - 22 &= \text{Combi} - 1 \cup \text{Combi} - 2 \cup \text{Combi} - 4. \end{aligned}$$

(vi) Combi-23 merges four different combinations:

$$\begin{aligned} \text{Combi} - 23 &= \text{Combi} - 1 \cup \text{Combi} - 2 \cup \\ &\quad \text{Combi} - 3 \cup \text{Combi} - 5. \end{aligned}$$

(vii) Combi-24 encompasses all variables, acting as the superset.

(viii) Combi-25 and 26 are the same as Combi-21 and 23, respectively, except that  $\delta p_T$  in MINERvA-0 $\pi$  is removed to avoid correlation with  $p_N$  in the same dataset.



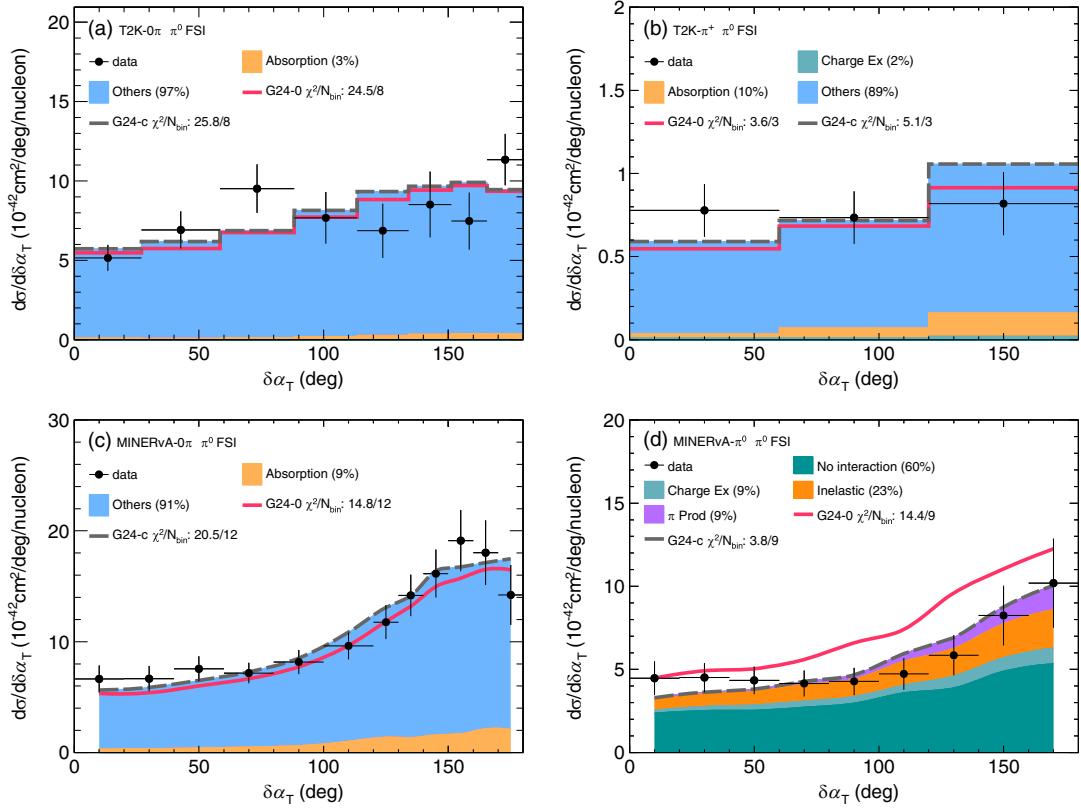
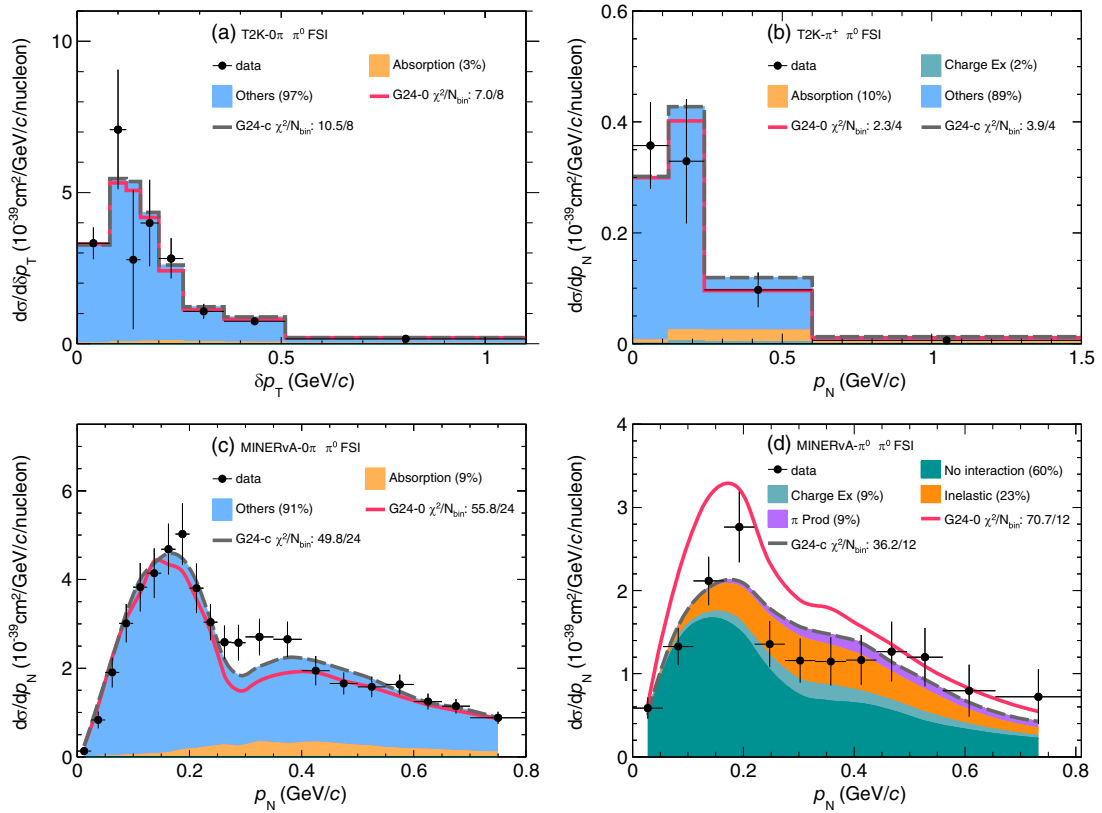
FIG. 16. Extension of Fig. 10(b) to all four  $\delta\alpha_T$  measurements.FIG. 17. Extension of Fig. 10(b) but to all four  $p_N$  measurements.

Figure 5 shows the change in  $\chi^2$  for the complete observable set (tuned plus validation) as a function of the tuning combination, where the two model parameter sets, AllPar and RedPar, are compared, and it can be seen that the respective minima happen at Combi-15 and 26.

## 2. $\pi^0$ FSI fate decomposition for G24-0 and G24-c

Figures 14–17 display comparisons of G24-0 and G24-c predictions to data, detailing the composition by  $\pi^0$  FSI fate.

- 
- [1] K. Abe *et al.* (Hyper-Kamiokande Collaboration), [arXiv:1805.04163](#).
- [2] R. Acciarri *et al.* (DUNE Collaboration), [arXiv:1601.05471](#).
- [3] R. Acciarri *et al.* (DUNE Collaboration), [arXiv:1512.06148](#).
- [4] J. Strait *et al.* (DUNE Collaboration), [arXiv:1601.05823](#).
- [5] R. Acciarri *et al.* (DUNE Collaboration), [arXiv:1601.02984](#).
- [6] V. Hewes *et al.* (DUNE Collaboration), *Instruments* **5**, 31 (2021).
- [7] C. Andreopoulos *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **614**, 87 (2010).
- [8] L. Alvarez-Ruso *et al.* (GENIE Collaboration), *Eur. Phys. J. Special Topics* **230**, 4449 (2021).
- [9] J. Tena-Vidal *et al.* (GENIE Collaboration), *Phys. Rev. D* **104**, 072009 (2021).
- [10] J. Tena-Vidal *et al.* (GENIE Collaboration), *Phys. Rev. D* **105**, 012009 (2022).
- [11] J. Tena-Vidal *et al.* (GENIE Collaboration), *Phys. Rev. D* **106**, 112001 (2022).
- [12] X.-G. Lu, D. Coplowe, R. Shah, G. Barr, D. Wark, and A. Weber, *Phys. Rev. D* **92**, 051302 (2015).
- [13] X.-G. Lu, L. Pickering, S. Dolan, G. Barr, D. Coplowe, Y. Uchida, D. Wark, M. O. Wascko, A. Weber, and T. Yuan, *Phys. Rev. C* **94**, 015503 (2016).
- [14] D. Coplowe *et al.* (MINERvA Collaboration), *Phys. Rev. D* **102**, 072007 (2020).
- [15] A. P. Furmanski and J. T. Sobczyk, *Phys. Rev. C* **95**, 065501 (2017).
- [16] X.-G. Lu and J. T. Sobczyk, *Phys. Rev. C* **99**, 055504 (2019).
- [17] T. Cai, X.-G. Lu, and D. Ruterbories, *Phys. Rev. D* **100**, 073010 (2019).
- [18] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **98**, 032003 (2018).
- [19] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **103**, 112009 (2021).
- [20] X.-G. Lu *et al.* (MINERvA Collaboration), *Phys. Rev. Lett.* **121**, 022504 (2018).
- [21] T. Cai *et al.* (MINERvA Collaboration), *Phys. Rev. D* **101**, 092001 (2020).
- [22] X.-G. Lu *et al.* (MINERvA Collaboration), *Eur. Phys. J. Special Topics* **230**, 4243 (2021).
- [23] P. Abratenko *et al.* (MicroBooNE Collaboration), [arXiv:2211.03734](#).
- [24] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **108**, 053002 (2023).
- [25] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. Lett.* **131**, 101802 (2023).
- [26] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **109**, 092007 (2024).
- [27] P. Abratenko *et al.* (MicroBooNE Collaboration), [arXiv:2403.19574](#).
- [28] M. Khachatryan *et al.* (CLAS, e4v Collaborations), *Nature (London)* **599**, 565 (2021).
- [29] K. Yang, Measurement of the pion charge exchange differential cross section on Argon with the ProtoDUNE-SP detector, Ph.D. thesis, Oxford University, FERMILAB-THESIS-2023-19, 2023.
- [30] P. Hamacher-Baumann, X.-G. Lu, and J. Martín-Albo, *Phys. Rev. D* **102**, 033005 (2020).
- [31] S. Dolan, [arXiv:1810.06043](#).
- [32] B. Bourguille, J. Nieves, and F. Sánchez, *J. High Energy Phys.* **04** (2021) 004.
- [33] J. M. Franco-Patino, M. B. Barbaro, J. A. Caballero, and G. D. Megias, *Phys. Rev. D* **104**, 073008 (2021).
- [34] A. Ershova *et al.*, *Phys. Rev. D* **106**, 032009 (2022).
- [35] J. M. Franco-Patino, R. González-Jiménez, S. Dolan, M. B. Barbaro, J. A. Caballero, G. D. Megias, and J. M. Udias, *Phys. Rev. D* **106**, 113005 (2022).
- [36] J. Chakrani *et al.*, *Phys. Rev. D* **109**, 072006 (2024).
- [37] A. Ershova *et al.*, *Phys. Rev. D* **108**, 112008 (2023).
- [38] L. Munteanu, Modeling neutrino-nucleus interaction uncertainties for dune long-baseline sensitivity studies (2022), nuInt, Seoul, KOREA [Conference presentation], <https://indico.cern.ch/event/881216/contributions/5048776/>.
- [39] L. Munteanu, Genie generator (2023), commit SHA: 64135ac, <https://github.com/GENIE-MC/Generator/pull/249>.
- [40] K. S. Egiyan *et al.* (CLAS Collaboration), *Phys. Rev. Lett.* **96**, 082501 (2006).
- [41] J. Nieves, J. E. Amaro, and M. Valverde, *Phys. Rev. C* **70**, 055503 (2004); **72**, 019902 (2005).
- [42] R. J. Hill and G. Paz, *Phys. Rev. D* **82**, 113005 (2010).
- [43] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas, *Phys. Rev. C* **83**, 045501 (2011).
- [44] R. González-Jiménez, G. D. Megias, M. B. Barbaro, J. A. Caballero, and T. W. Donnelly, *Phys. Rev. C* **90**, 035501 (2014).
- [45] A. Pais, *Ann. Phys. (N.Y.)* **63**, 361 (1971).
- [46] S. G. Kovalenko, *Sov. J. Nucl. Phys.* **52**, 934 (1990), [https://inis.iaea.org/search/search.aspx?orig\\_q=RN:22053782](https://inis.iaea.org/search/search.aspx?orig_q=RN:22053782).
- [47] C. Berger and L. M. Sehgal, *Phys. Rev. D* **76**, 113004 (2007).
- [48] A. Bodek and U. K. Yang, *Nucl. Phys. B, Proc. Suppl.* **112**, 70 (2002).

- [49] M. A. G. Aivazis, W.-K. Tung, and F. I. Olness, in *Particles & Fields 91: Meeting of the Division of Particles & Fields of the APS* (1991), pp. 0663–665, [arXiv:hep-ph/9302305](#).
- [50] C. Berger and L. M. Sehgal, *Phys. Rev. D* **79**, 053003 (2009).
- [51] T. Yang, C. Andreopoulos, H. Gallagher, K. Hoffmann, and P. Kehayias, *Eur. Phys. J. C* **63**, 1 (2009).
- [52] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, [arXiv:1510.05494](#).
- [53] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, *Eur. Phys. J. C* **65**, 331 (2010).
- [54] G. D’Agostini, *Nucl. Instrum. Methods Phys. Res., Sect. A* **346**, 306 (1994).
- [55] K. M. Hanson, T. Kawano, and P. Talou, *AIP Conf. Proc.* **769**, 304 (2005).
- [56] R. Frühwirth, D. Neudecker, and H. Leeb, *EPJ Web Conf.* **27**, 00008 (2012).
- [57] D. Rowntree *et al.* (LADS Collaboration), *Phys. Rev. C* **60**, 054610 (1999).
- [58] I. Navon, D. Ashery, J. Alster, G. Azuelos, B. M. Barnett, W. Gyles, R. R. Johnson, D. R. Gill, and T. G. Masterson, *Phys. Rev. C* **28**, 2548 (1983).
- [59] A. S. Carroll, I. H. Chiang, C. B. Dover, T. F. Kycia, K. K. Li, P. O. Mazur, D. N. Michael, P. M. Mockett, D. C. Rahm, and R. Rubinstein, *Phys. Rev. C* **14**, 635 (1976).
- [60] A. S. Clough *et al.*, *Nucl. Phys.* **B76**, 15 (1974).
- [61] W. Bauhoff, *At. Data Nucl. Data Tables* **35**, 429 (1986).
- [62] S. G. Mashnik, R. J. Peterson, A. J. Sierk, and M. R. Braunstein, *Phys. Rev. C* **61**, 034601 (2000).
- [63] K. Ishibashi *et al.*, *J. Nucl. Sci. Technol.* **34**, 529 (1997).
- [64] A. Papadopoulou *et al.* (electrons for neutrinos Collaboration), *Phys. Rev. D* **103**, 113003 (2021).
- [65] Q. Yan, K. Niewczas, A. Nikolakopoulos, R. González-Jiménez, N. Jachowicz, X.-G. Lu, J. Sobczyk, and Y. Zheng, [arXiv:2405.05212](#).
- [66] R. Acciarri *et al.* (MicroBooNE, LAr1-ND, ICARUS-WA104 Collaborations), [arXiv:1503.01520](#).