

Neutrino-nucleus cross section impacts on neutrino oscillation measurements

Nina M. Coyle^{1,*}, Shirley Weishi Li^{1,†} and Pedro A. N. Machado^{2,‡}

¹*Department of Physics and Astronomy, University of California, Irvine, California 92697, USA*

²*Theoretical Physics Department, Fermilab, P.O. Box 500, Batavia, Illinois 60510, USA*



(Received 6 March 2025; accepted 7 May 2025; published 30 May 2025)

The challenges in neutrino-nucleus cross section modeling and its impact on neutrino oscillation experiments are widely recognized. However, a comprehensive and theoretically robust estimation of cross section uncertainties has been lacking, and few studies have quantitatively examined their impact on oscillation measurements. In this work, we evaluate the effect of cross section uncertainties on oscillation parameters using setups inspired by NOvA and DUNE. To characterize these uncertainties, we adopt multiple neutrino-nucleus event generators and simulate a realistic experimental procedure that incorporates near-detector data and near-to-far-detector extrapolation. Our results confirm that cross section uncertainties cannot significantly bias oscillation results in current statistics-dominated experiments like NOvA. However, they could lead to substantial bias for future systematics-dominated experiments like DUNE, even when near-detector data are employed to mitigate uncertainties. These findings underscore the need for further studies on the quantitative impacts of cross section modeling, improved strategies to utilize near-detector data and the PRISM concept, and more robust cross section models to optimize the success of future experiments.

DOI: [10.1103/PhysRevD.111.093010](https://doi.org/10.1103/PhysRevD.111.093010)

I. INTRODUCTION

As neutrino physics enters a precision era with future experiments like DUNE, Hyper-Kamiokande, and JUNO [1–3], a rich physics program emerges for exploration, alongside an increasing need to understand neutrino-nucleus interaction physics [4–7]. Long-baseline experiments such as DUNE and Hyper-Kamiokande, which utilize broadband neutrino beams, require inferring the neutrino energy of individual events from observed final states. This process relies on cross section models that predict the amount of missing energy due to factors like escaping neutrons or charged particles below detection thresholds. However, in the GeV energy range of these neutrino beams, the nuclear degree of freedom involves partons, nucleons, and baryonic resonances, making first-principle calculations of cross sections currently unfeasible [6].

Given this complexity, experiments must adopt pragmatic approaches to model neutrino-nucleus interactions.

To this end, they employ neutrino event generators [8–18]. These generators compute cross sections across different channels, each corresponding to a specific type of hadronic degree of freedom: nucleons for quasielastic scattering, nucleon pairs for meson exchange current, baryonic resonances for resonance production, and partons for deep-inelastic scattering (which may also include shallow inelastic scattering). All these channels are subject to nuclear effects, such as final-state interactions [11,12,19–30]. While the models implemented in event generators provide reasonable estimates of cross sections, quantitative studies have revealed their limitations in accurately reproducing neutrino-nucleus and electron-nucleus scattering data [31–44].

To address cross section mismodeling and uncertainties, long-baseline experiments employ an additional tool by leveraging near detectors (ND). These detectors are placed sufficiently close to the neutrino beam source, where no measurable oscillation effects are expected. This proximity allows ND data to serve as a validation sample for refining both beam simulations and cross section models. When predictions from the default generator models deviate from ND data, experiments frequently perform ND tuning, adjusting the generator models to match the observed data. The resulting tuned generator is then used to predict oscillations in the far detector (FD).

Despite recent intensive efforts in neutrino cross section studies [4–6,45,46], including theoretical calculations and experimental measurements, few studies have investigated

*Contact author: ncoyle@uci.edu

†Contact author: shirley.li@uci.edu

‡Contact author: pmachado@fnal.gov

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 SCOAP³.

how cross section mismodeling and uncertainties impact oscillation measurements. While existing oscillation measurements published by experimental collaborations account for cross section uncertainties and assign associated systematic errors, the full analysis details are not publicly available. Moreover, these analyses often include experimental effects that ideally should be disentangled from cross section physics. Although future experiments like DUNE and Hyper-Kamiokande have produced sensitivity studies using early detector simulation and cross section models, neither has published an analysis with an uncertainty model sufficient for analyzing real data. While Refs. [47–51] were the first to explore cross section uncertainties and their propagation to oscillation parameters in neutrino experiments, they only consider the FD.

There is an urgent need to answer the following question: given current cross section uncertainties, can we achieve the designed *accuracy* on oscillation parameters in next-generation experiments? Our paper aims to address this question. We characterize current cross section uncertainties by using multiple event generators; GENIE [9,10], GiBUU [11], and NuWro [12]. For the utilization of ND data, we closely follow NOvA’s analysis strategy as an instructive example, including ND tuning and near-to-far-detector extrapolation.

The paper is organized as follows. In Sec. II, we describe our analysis procedure and methodology, as well as the results of performing this analysis for the NOvA oscillation setup. In Sec. III, we present results for the DUNE oscillation analysis. Finally, we discuss our conclusions in Sec. IV.

II. METHODS

Our aim is to select an analysis strategy that closely approximates DUNE’s future approach. DUNE is currently in the construction phase, and experiments like SBND [52] are actively collecting ν -Ar cross section data that may inform DUNE’s cross section model. Consequently, the full DUNE analysis procedure and cross section mitigation strategy have not yet been published. Given these circumstances, our analysis adopts the existing NOvA procedure [53–55] as a proxy, leveraging the similarities between the energy ranges of the two experiments. It is important to note that this approach is not tailored to DUNE’s specific needs and design, nor does it utilize DUNE-PRISM [56]. As such, we expect our findings to represent a conservative estimate of DUNE’s potential accuracy. The actual DUNE oscillation analysis, when finalized, may achieve significant improvements beyond our current projections. In what follows, we describe the details of our analysis for both the NOvA and DUNE experiments.

First, to characterize the current theoretical uncertainties in cross section modeling, we face the challenge of lacking robust uncertainty estimates. To address this, we employ three different event generators; GENIE (version 3.04.00

with G18_02a_00_000 tune), NuWro (version 21.09.2), and GiBUU (release 2023). We treat the differences between these generators as a proxy for cross section uncertainties. While this approach is not entirely accurate—it may underestimate or overestimate true uncertainties—it provides a workable framework for our analysis. Specifically, we simulate mock data using GENIE, assuming it represents the true cross section model. We then conduct the analysis using either NuWro or GiBUU, simulating a scenario where these are the available generators. This method allows us to assess the impact of potential discrepancies between the assumed cross section model and the one used in the analysis.

Second, because the impact of cross section calculations arises primarily in the predicted relation between true and reconstructed neutrino energies, as will be discussed in Sec. II B, we must consider detector effects and reconstruction capabilities. We employ calorimetric energy reconstruction, in which the incoming neutrino energy is reconstructed as the sum of the total energy of outgoing lepton and mesons, plus the kinetic energy of visible hadrons. For NOvA, we include a 30% smearing on hadronic final-state particles and 3% on leptons [53]. We include proton quenching effects with Birks’ constant and dE/dx values taken from Ref. [57] and NIST Standard Reference Database 124, respectively [58]; a proton kinetic energy threshold of 50 MeV and a pion kinetic energy threshold of 20 MeV. Neutrons are considered invisible for NOvA. These thresholds are choices that enable us to approximately reproduce the NOvA visible hadronic energy spectrum. For DUNE, we employ the resolutions and kinematic thresholds listed in Table 3.3 of Ref. [59] along with proton quenching effects. As described in the caption of that table, we consider 60% of postsmeared neutron energy to be visible, and include a 10% probability that a neutron with momentum less than 1 GeV will be missed by reconstruction techniques (see also Ref. [60]).

In the following subsections, we provide a detailed description of NOvA’s primary mitigation strategies—ND tuning and near-to-far-detector extrapolation—which we adopt in our analysis.

A. Near-detector tuning procedure

Experiments frequently observe significant discrepancies between measured neutrino spectra and baseline generator predictions in the ND [53,61,62]. To address this, experiments utilize ND data to adjust the models implemented in event generators. The resulting modified generator is then employed in the unfolding and smearing procedures. It is important to note that the precise tuning method varies widely between experiments, and the resulting cross section model may not be universally applicable across different experiments or analyses.

NOvA employs the GENIE event generator for their analysis. There are two primary pieces of the NOvA tune;

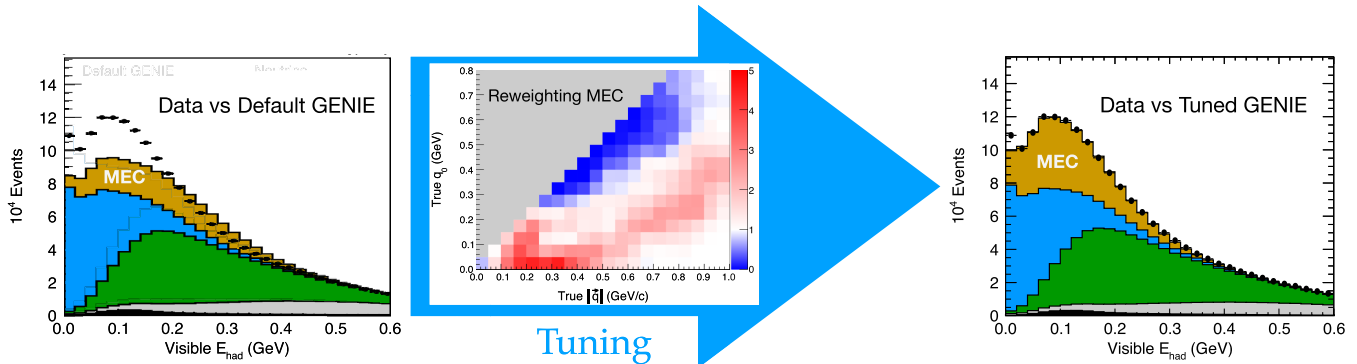


FIG. 1. Illustration of NOvA ND tuning procedure; figure taken from Ref. [7], modified from Ref. [53].

adjustments to the base generator based on data from previous experiments, and a subsequent tune of the updated generator to NOvA ND data [53]. Examples of the first step include adjusting the axial mass from GENIE default value 0.99 to 1.04 GeV, based on a reanalysis of the deuterium bubble chamber data [31]. We do not implement these initial modifications because we are using multiple different generators as a proxy for the possible impacts of cross section modeling. Regardless, significant discrepancies remain between the simulated and measured ND spectra even after these changes, which motivates the second step of tuning the generator to NOvA ND data.

The NOvA Collaboration reweights the meson exchange current (MEC) contribution in the energy and three-momentum transfer plane. This approach is based on the assumption that other channels—quasielastic, resonance production, and deep-inelastic scattering—either have significantly less cross section mismodeling, given the initial adjustments made before fitting to the ND data, or have small contributions to the NOvA event rate. Note that while this tuning procedure is specific to the NOvA experiment, it provides an example of tuning from which we can take away some lessons. This procedure, as illustrated in Fig. 1, involves the following:

- (i) The simulated events are binned in a two-dimensional plane of true energy transfer q^0 and true three-momentum transfer $|\vec{q}|$ to the hadronic system.
- (ii) Weights are applied independently to the MEC cross section in each bin of the two-dimensional true plane.
- (iii) The reweighted distribution in the true plane is migrated into the reconstructed $E_{\text{had}}^{\text{vis}}$ and $|\vec{q}|^{\text{reco}}$ plane.
- (iv) The simulation is compared to data in the reconstructed plane, with the tune weights in the true plane iteratively updated to fit the data in the reconstructed plane.

The fit in this procedure employs a χ^2 fit with statistical uncertainties for each bin in the $(|\vec{q}|^{\text{reco}}, E_{\text{had}}^{\text{vis}})$ plane; the tune weights themselves are included as Gaussian penalty terms with an uncertainty of 100%. In the experimental oscillation analyses, systematic uncertainties arising from

cross section modeling are identified by adjusting the relevant GENIE inputs within their uncertainties. The uncertainties arising specifically from the ND MEC tuning procedure are determined by performing two different versions of the tune: one in which the QE contribution is enhanced by 1σ before the ND tune, and another in which the RES contribution is instead enhanced. The spread of the resulting tuned spectra provides a systematic uncertainty for the MEC part of the ND tune. It is clear from examining the resulting spectra in Fig. 1 that an initial simulation with a significant deficit in visible hadronic energy has been tuned to a simulation that matches the ND data well in $E_{\text{had}}^{\text{vis}}$.

In our analysis of NOvA, we follow their procedure closely, but we have checked that including other channels in the tuning does not significantly impact oscillation sensitivities. In the case of DUNE, the choice of tuning only the MEC contribution needs to be modified. We will discuss this further in our mock DUNE analysis.

B. Near-to-far-detector extrapolation

To further reduce cross section dependence in the oscillation analyses, NOvA employs a data-driven approach to predict FD spectra [54,55]. In particular, the FD flux prediction incorporates the unfolded ND flux, rather than relying only on beam simulation. Figure 2 illustrates NOvA's near-to-far-detector extrapolation procedure. It involves the following steps:

- (i) Unfold the observed ND data in reconstructed neutrino energy (E_{ν}^{reco}) to an inferred spectrum in true neutrino energy (E_{ν}^{true}). This unfolding employs an ND migration matrix simulated using the tuned generator.
- (ii) Propagate the resulting spectrum in E_{ν}^{true} to the FD. This involves the oscillation probabilities as well as factors that account for detector acceptances, solid angles, and other experimental details.
- (iii) Migrate the predicted FD E_{ν}^{true} spectrum to a predicted E_{ν}^{reco} spectrum using an FD migration matrix

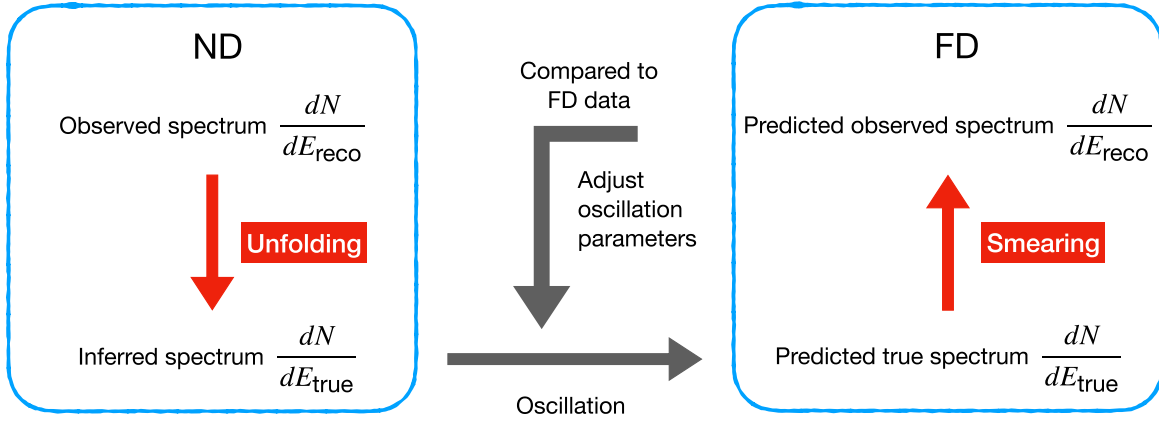


FIG. 2. Illustration of NOvA near-to-far-detector extrapolation procedure. The steps utilizing cross section predictions are in red.

simulated using the tuned generator. The predicted E_{ν}^{reco} spectrum may then be compared with data.

This approach utilizes cross section calculations in the unfolding and migration steps, which require simulations to model the migration matrices between true and reconstructed neutrino energies. Consequently, the predicted distributions of final-state particles and their kinematics become crucial, driving the cross section dependence of the analysis. To illustrate, consider two events with the same true incoming neutrino energy but different interaction types. A deep-inelastic scattering event with multiple knocked-out neutrons and pions may yield a significantly different E_{ν}^{reco} compared to a quasielastic scattering event with a single outgoing proton. Therefore, the relative modeling of different interaction channels and final-state interactions plays a pivotal role in shaping the analysis outcomes.

The unfolding step in this analysis can be conceptualized as an inversion problem. While various methods exist for unfolding, we employ iterative Bayesian unfolding [63] through the RooUnfold package [64]. This procedure involves an iterative fit process: A test spectrum in true neutrino energy is migrated to reconstructed energy using a generator-based migration matrix. The resulting spectrum is compared to data using a χ^2 statistic, and the test spectrum is iteratively updated to minimize the difference between the resulting reconstructed spectrum and the observed data. The process is terminated after a specified number of iterations. We have tested this unfolding procedure and found that our results are robust against the choice of the unfolding methods.

To propagate the unfolded ND spectrum to the FD, we employ an FD-to-ND ratio that accounts for the different expected fluxes and normalizations in the FD and ND for the respective experiments. For NOvA, the ratio is taken from Fig. 4.16 in Ref. [55]. The oscillations are calculated including matter effects, with a density of $\rho = 2.828 \text{ g/cm}^3$.

C. NOvA oscillation results

As a check of the above mock analysis procedure, we examine the results of employing this procedure for the NOvA experiment. Since the NOvA experiment is primarily statistics-limited, we expect that the cross section dependence will not strongly influence the results of the analysis; this is in contrast to DUNE, which will be systematics-dominated. The statistics we use for this mock NOvA analysis match the event rate for 13.6×10^{20} protons on target (POT) in the forward horn current (FHC) mode and 12.5×10^{20} POT in the reverse horn current (RHC) mode [54], leading to $\mathcal{O}(100)$ disappearance signal events and $\mathcal{O}(10)$ appearance signal events in the FD for each mode. We scale the untuned simulation generator by an overall factor of 0.9 to approximately mimic the scenario in the NOvA ND tune without overly specifying to their particular spectra; this choice of normalization does not meaningfully impact the resulting best-fit regions.

We perform an example analysis for simultaneous fits of Δm_{32}^2 , $\sin^2 \theta_{23}$ and of δ_{CP} , $\sin^2 \theta_{23}$. We assume normal hierarchy and fix the other oscillation parameters to those used in Ref. [54]: $\Delta m_{21}^2 = 7.53 \times 10^{-5} \text{ eV}^2$, $\sin^2 \theta_{12} = 0.307$, and $\sin^2 \theta_{13} = 0.021$. Implementing a marginalization over these parameters slightly enlarges the size of the best-fit regions but does not meaningfully affect the conclusions of our analysis. For flux uncertainties, we employ values of 20% overall normalization, 4% relative ND-to-FD normalization, 2% bin-to-bin uncorrelated, and 2% correlated bin-to-bin spectral uncertainty.

Figure 3 shows the results. The black region corresponds to performing the full analysis procedure using GENIE; because the mock data is also generated using GENIE, this region represents the case in which one employs a cross section model that accurately models nature. The results using NuWro are plotted in blue, while the ones for GiBUU are shown in red. Because the regions obtained using the tuned and untuned generators differ only slightly, we only show the tuned ones.

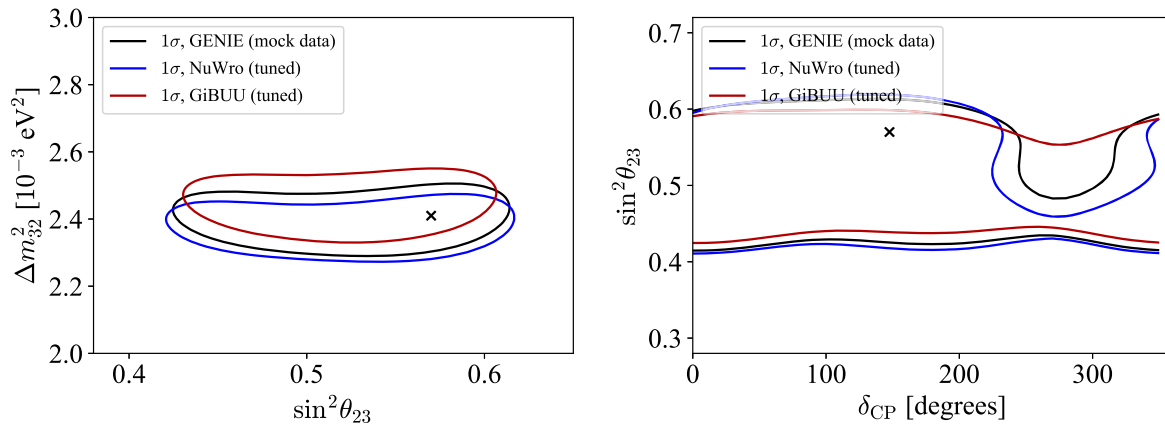


FIG. 3. Oscillation analyses for NOvA setup. For all regions, the mock data is generated using GENIE, whereas the analysis is performed with either GENIE, NuWro, or GiBUU. The input truth oscillation parameters are indicated by the black crosses.

Our analysis yields region sizes and shapes comparable to those reported in the NOvA analyses [54], validating our procedure. The regions using different simulators vary due to their distinct underlying physics models. Notably, for both analyses, the NuWro region closely resembles the “Truth” GENIE region, while the GiBUU result shows more divergence. This difference likely stems from GiBUU’s significantly different nuclear physics modeling, which we discuss in detail in Sec. III. Despite these differences, even the GiBUU region has substantial overlap with the true region in both panels of Fig. 3. We conclude that the choice of simulation generator does not significantly impact the results of either analysis. This finding aligns with expectations, given that the NOvA analysis is predominantly statistics-driven.

III. DUNE OSCILLATION ANALYSIS

We now move to employing this oscillation analysis procedure for DUNE to examine the cross section impact.

A. Tuning and FD prediction setup

We make a few changes to the details of the analysis in order to model the DUNE scenario rather than NOvA. The event files for DUNE are generated for an argon target rather than carbon, and are weighted by the expected DUNE ND flux available at Ref. [65]. The ND-to-FD ratio, which accounts for detector sizes and solid angles, is derived from a ratio of the predicted FD and ND fluxes, and is similar to that shown in Fig. 3 of Ref. [66]. Meanwhile, the event rate is scaled to have statistics equivalent to an exposure of 624 kt-MW-years, or 10 years of running with the staging assumptions outlined in Ref. [66]. Following Ref. [67], we employ a simple overall normalization uncertainty of 8% on the $\nu_{\mu}(\bar{\nu}_{\mu})$ events and 5% on $\nu_e(\bar{\nu}_e)$ events; we have taken slightly larger values than those in Ref. [67] to approximately account for background uncertainties, which we do not explicitly include. We find

similar region sizes and shapes to those generated by the GLOBES code in Ref. [67].

The ND tuning procedure requires a slightly more meaningful modification. While the MEC event distribution covers most of the parameter space in the two-dimensional ND tune plane occupied by the NOvA events, the same is not true for the expected DUNE flux; in particular, a greater contribution from deep-inelastic scattering events leads to a large portion of the parameter space that is not covered by MEC. As such, a choice must be made as to which interaction process or set of events to tune. For our analysis, we choose to tune the full cross section rather than a specific interaction process.

B. DUNE oscillation results

As before, we generate mock ND and FD data using GENIE, while employing NuWro and GiBUU as simulation generators. For each generator, we perform the full analysis procedure, including ND tuning and near-to-far-detector extrapolation using both its tuned and untuned versions. The results for NuWro and GiBUU are presented in Figs. 4 and 5, respectively. The left panel of each figure shows the 1-sigma region for each of the three generators used in the analysis. The right panel shows a comparison of the ND mock data with ND spectra simulated using both the untuned and tuned generators; the upper subplot shows histograms of the spectra, while the lower subplot presents the ratio of simulation to data for both generator versions.

From Fig. 4, we see that employing the NuWro generator to fit GENIE mock data results in best-fit regions that are similar to the truth region. While the untuned region shows a slight bias, the tuned region has a significant overlap with the truth region. The ND spectra, meanwhile, are in good agreement: although the untuned NuWro generator underpredicts the ND spectrum, the tuned NuWro generator matches the ND data well, and the ratio between the simulation and data varies only slightly around 1.

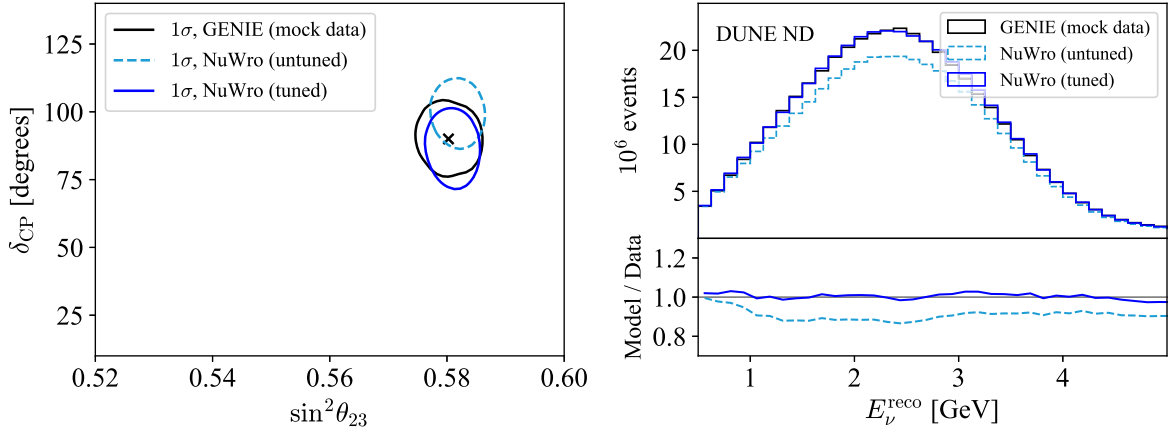


FIG. 4. Oscillation analyses for DUNE setup. For all cases, the mock data is generated using GENIE, whereas the analysis is performed with GENIE, both untuned and tuned NuWro. The input parameters are indicated by the black cross.

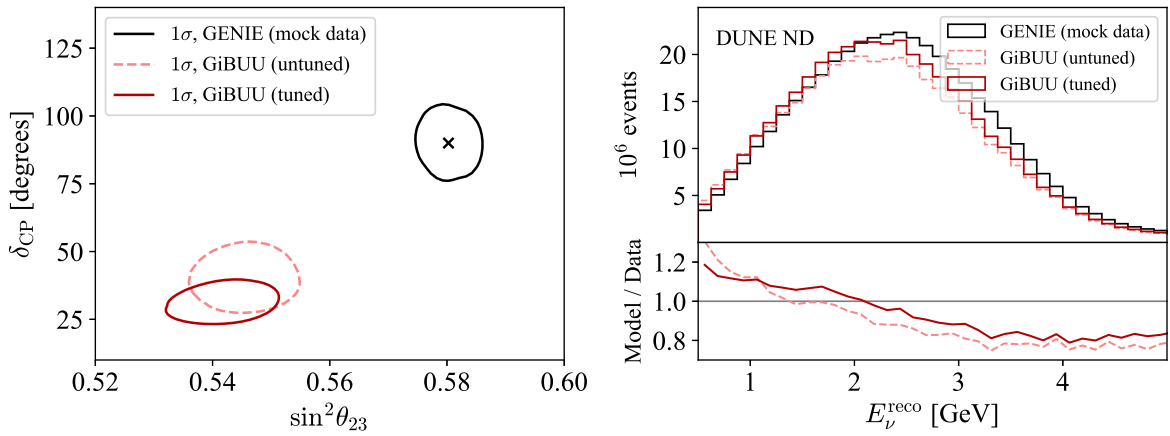


FIG. 5. Oscillation analyses for DUNE setup. For all cases, the mock data is generated using GENIE, whereas the analysis is performed with GENIE, both untuned and tuned GiBUU. The input parameters are indicated by the black cross.

From these results, one may infer that this oscillation analysis procedure does not strongly depend on the cross section model, and that the tuning procedure can successfully limit what impacts remain. However, Fig. 5 indicates that this is not the case for all possible implementations of current cross section modeling. The untuned GiBUU region shows significant bias relative to the true region; the tuning, meanwhile, does not mitigate this bias, despite a successful fit to the tune parameters (see the Appendix for comparison spectra). This illustrates that a naive implementation of the oscillation analysis procedure could lead to biased results for the oscillation parameters. It is then important to ask if there is a way to detect this potential bias by leveraging specific observables. In this case, the right panel makes it clear that the simulation—even after tuning—does not successfully model the observed ND data. One may therefore examine the ND spectra to conclude that the tuning procedure fails and that an oscillation analysis employing the tuned cross section model could present strong biases.

There are many differences in how individual event generators implement the primary cross sections and the nuclear effects. However, given that cross section models affect the analysis through migration matrices, a potential reason that GiBUU results are so different from GENIE and NuWro is how final-state interactions (FSI) are implemented. With the impulse approximation, after the primary neutrino-nucleon interaction, the secondary hadrons still need to propagate inside of the nucleus from their production point to where they exit the nucleus. During propagation, the secondary hadrons could scatter with other nucleons in the nucleus, producing new particles or transferring energy. GENIE 3.04.00 (with G18_02a_00_000 tune) and NuWro 21.09.2 treatments of FSI are similar in that both employ hadronic transport models. In these models, nucleons propagate as “billiard balls” throughout the nucleus (or some simplified version of this), possibly scattering off other nucleons according to nucleon-nucleon cross sections (see e.g., Refs. [9,12,24]). In-medium effects and the nucleon configuration may differ between these generators, but the

overall physics is similar. GiBUU, on the other hand, solves the quantum kinetic Boltzmann equations, also known as Kadanoff-Baym equations [68,69]. Since obtaining exact solutions is currently not possible, GiBUU adopts the gradient approximation, which assumes the nuclear density distribution to be fairly flat.

To be quantitative, we show in Fig. 6 the distribution of the difference between true neutrino energy and total available energy in outgoing particles, that is $\Delta \equiv E_\nu^{\text{true}} - E_{\text{out}}$. The total available energy E_{out} includes the total energy of leptons and pions (and other mesons) and the kinetic energy of nucleons. As we see in Fig. 6 the distribution of Δ is similar for GENIE and NuWro, while GiBUU predicts a wider distribution due to the different modeling of FSI. Therefore, when inferring the neutrino energy from a given calorimetric energy, GiBUU would predict higher neutrino energies compared to GENIE and NuWro. Note that this difference Δ for monochromatic neutrinos from kaon decay-at-rest has been studied

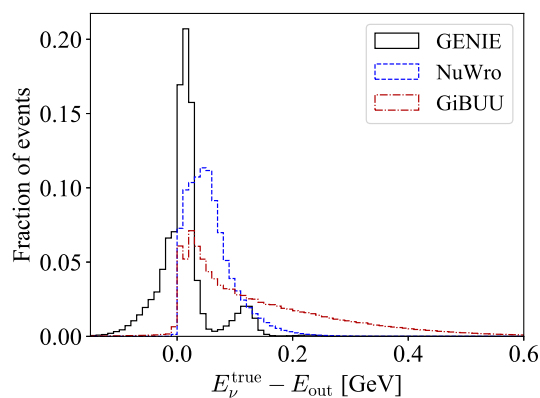


FIG. 6. Distribution of difference between true neutrino energy and total available energy in an event, $E_\nu^{\text{true}} - E_{\text{out}}$. The total available energy E_{out} includes the total energy of leptons and pions (and other mesons) and the kinetic energy of nucleons.

experimentally by the JSNS² experiment, and it was found that GiBUU predicts a broader Δ distribution than data, while NuWro presents a better fit [70]. Motivated by Fig. 6, Gallmeister and Mosel have further investigated this wide distribution in Ref. [71].

To test our hypothesis and further evaluate the range of biases generators can give, we examine a phenomenologically modified version of GiBUU, denoted by GiBUU-mod, which enforces Δ closer to zero values. We add the difference $E_\nu^{\text{true}} - E_{\text{out}}$ to the hadronic energy, with a 5% smearing included on this contribution when calculating the visible hadronic energy. This gives us an example generator more similar to GENIE and NuWro in terms of this energy difference, while still using a completely different FSI framework.

Figure 7 shows the results of this analysis. As we can see in the right panel, this modified version of GiBUU provides a modeling of the ND mock data after tuning that is closer to GENIE. However, the relationship between true and reconstructed energy still does not match the one used in generating mock data, as there remains a bias in the preferred regions in δ_{CP} . This seems to indicate that the ND data is not sufficient to constrain the cross section models accurately, and the oscillation analysis could be biased outside of the error budget.

Although this may be worrisome at first, we note that more experimental handles can be used to flesh out the discrepancies even in this case. DUNE near detector will be movable to probe different off-axis angles, a technique coined DUNE-PRISM [56]. By probing different angles with respect to the neutrino beam axis, DUNE-PRISM can probe different neutrino spectra correlated by the pion energy spectrum.

We show an example off-axis spectrum in Fig. 8, where the flux has been pulled from a simple Gaussian distribution centered at $E_\nu^{\text{true}} = 1.5$ GeV with a width of 0.5 GeV. We do not employ a specific prediction for the number of

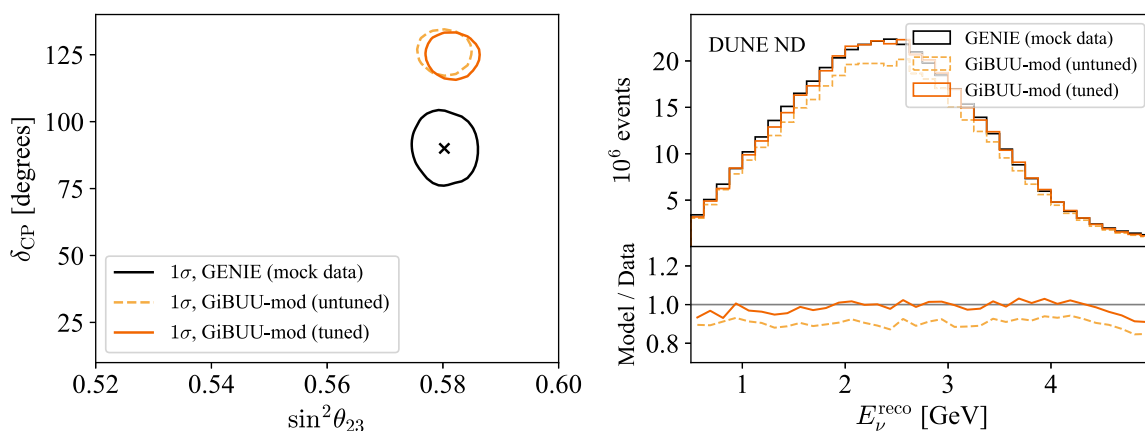


FIG. 7. Oscillation analyses for DUNE setup. For all cases, the mock data is generated using GENIE, whereas the analysis is performed with GENIE, both untuned and tuned GiBUU-mod (our phenomenologically modified GiBUU). The input parameters are indicated by the black cross.

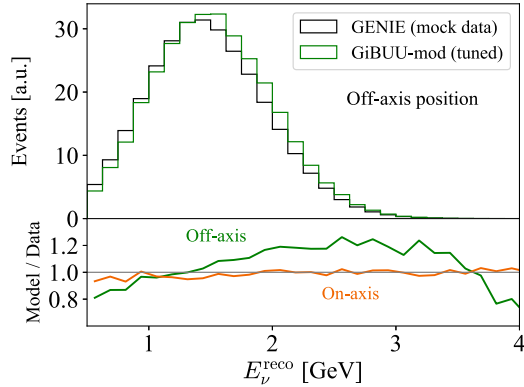


FIG. 8. An illustration of the power of DUNE-PRISM. In the top panel, we show the ND spectra for an example off-axis flux for the data and the tuned generator. In the bottom panel, we show the ratio of the spectra in green. The orange line shows the ratio for on-axis spectra, taken from Fig. 7.

events relative to the nominal on-axis flux, instead focusing on the relative shapes of the predicted spectra. Although the on-axis post-tune E_ν^{reco} spectrum agrees well with the mock GENIE data, the off-axis flux presents a large discrepancy. This is due to the different relationships between E_ν^{true} and E_ν^{reco} in GENIE and GiBUU-mod. Our findings are similar to those found by DUNE in the TDR [3].

This result suggests that more off-axis ND data, using the DUNE-PRISM technique, is crucial for breaking the degeneracies in specific experimental observables and providing potentially faithful constrained cross section models. Investigating how well this works quantitatively requires detailed DUNE-PRISM information and analysis strategies that are not publicly available. We, therefore, leave it to future work for quantitative studies.

IV. CONCLUSION

In this work, we studied in detail the impact of neutrino-nucleus cross section modeling on the robustness of long-baseline oscillation experiments. We have considered three neutrino-nucleus event generators to capture the current cross section uncertainties. We then performed mock oscillation analyses with strategies inspired by NOvA’s procedure, which leverages near detector data to mitigate uncertainties, and considered setups similar to both NOvA and DUNE.

We find that the NOvA analyses are largely unaffected by biases arising from cross section modeling due to limited statistics, which is consistent with official results [54]. However, DUNE sensitivity could be significantly biased due to mis-modeling. Extreme biases arise when we simulate data using GENIE and perform the analysis using GiBUU; this is due to the drastic differences in the

neutrino-nucleus interaction models. Nevertheless, one can identify these issues by examining E_ν^{reco} spectra after tuning and observing whether discrepancies remain. However, in other cases such as when we perform the analysis using a phenomenologically modified version of GiBUU, the ND spectra match well for on-axis fluxes, although biases in the oscillation analysis remain. To combat this scenario, utilizing off-axis ND data with DUNE-PRISM is crucial. We leave quantitative studies along this line for future work.

Our results illustrate a need for improvements in the theoretical cross section modeling, as well as new experimental approaches to address cross section uncertainties. One way of testing the tuning procedure would be to perform experimental analyses with several event generators. Measurements of ν -argon cross sections from experiments such as SBND may contribute to the input cross section models employed in future DUNE analyses. At the same time, DUNE-PRISM off-axis ND data can provide further constraints on cross section modeling. A procedure that employs these and other strategies will be necessary to ensure the success of future long-baseline neutrino experiments.

ACKNOWLEDGMENTS

We thank Jianming Bian, Joshua Isaacson, Kevin Kelly, Joachim Kopp, Ulrich Mosel, Afroditi Papadopoulou, Xin Qian, Linyan Wan, and Daniel Whiteson for helpful discussions. The work of N. C. is supported by the U.S. National Science Foundation under Grant No. PHY-2210283. This manuscript has been authored by Fermi Forward Discovery Group, LLC under Contract No. 89243024CSC000002 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The work and conclusions presented in this publication are not to be considered as results from the DUNE Collaboration.

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX: TUNED SPECTRA

In this appendix, we present the results of the tuning procedure for two illustrative cases in Figs. 9 and 10; tuning NuWro to data generated with GENIE, and tuning GiBUU to data generated by GENIE. We use event files simulated for neutrino interactions on argon. In both cases, we find that while there exists significant disagreement between the mock data and the untuned simulation, the tuned generator fits the data well in both observables.

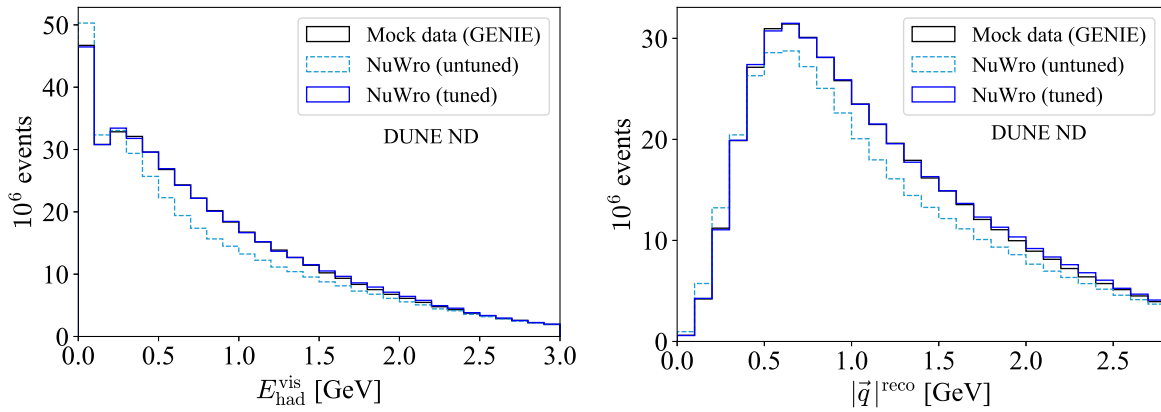


FIG. 9. Comparison of spectra in $E_{\text{had}}^{\text{vis}}$ (left) and $|\vec{q}|^{\text{reco}}$ (right) for GENIE mock data, untuned, and tuned NuWro.

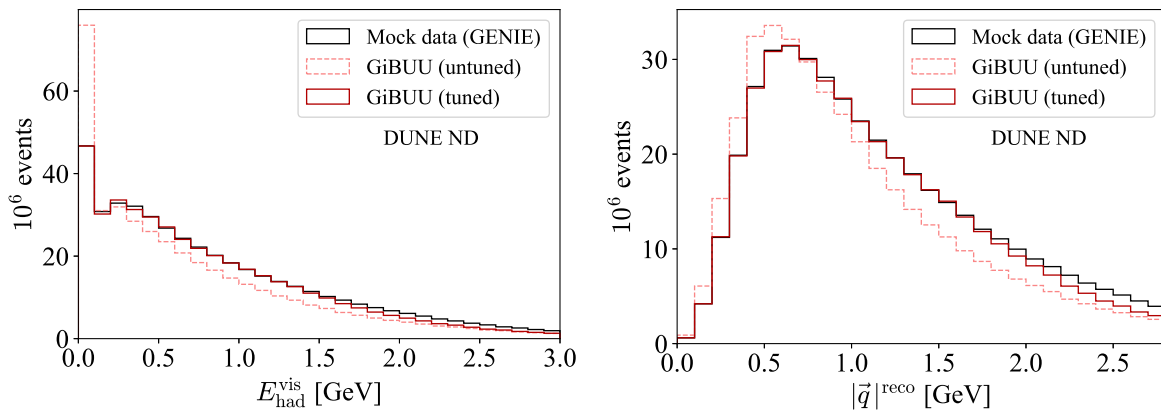


FIG. 10. Comparison of spectra in $E_{\text{had}}^{\text{vis}}$ (left) and $|\vec{q}|^{\text{reco}}$ (right) for GENIE mock data, untuned, and tuned GiBUU.

-
- [1] F. An *et al.* (JUNO Collaboration), *J. Phys. G* **43**, 030401 (2016).
- [2] K. Abe *et al.* (Hyper-Kamiokande Collaboration), [arXiv:1805.04163](https://arxiv.org/abs/1805.04163).
- [3] B. Abi *et al.* (DUNE Collaboration), [arXiv:2002.03005](https://arxiv.org/abs/2002.03005).
- [4] U. Mosel, *Annu. Rev. Nucl. Part. Sci.* **66**, 171 (2016).
- [5] L. Alvarez-Ruso *et al.* (NuSTEC Collaboration), *Prog. Part. Nucl. Phys.* **100**, 1 (2018).
- [6] L. A. Ruso *et al.*, *J. Phys. G* **52**, 043001 (2025).
- [7] N.M. Coyle, S.W. Li, and P.A.N. Machado, *J. High Energy Phys.* **12** (2022) 166.
- [8] D. Casper, *Nucl. Phys. B, Proc. Suppl.* **112**, 161 (2002).
- [9] C. Andreopoulos *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **614**, 87 (2010).
- [10] L. Alvarez-Ruso *et al.* (GENIE Collaboration), *Eur. Phys. J. Special Topics* **230**, 4449 (2021).
- [11] O. Buss, T. Gaitanos, K. Gallmeister, H. van Hees, M. Kaskulov, O. Lalakulich, A. B. Larionov, T. Leitner, J. Weil, and U. Mosel, *Phys. Rep.* **512**, 1 (2012).
- [12] T. Golan, C. Juszczak, and J. T. Sobczyk, *Phys. Rev. C* **86**, 015505 (2012).
- [13] T. T. Böhlen, F. Cerutti, M. P. W. Chin, A. Fassò, A. Ferrari, P. G. Ortega, A. Mairani, P. R. Sala, G. Smirnov, and V. Vlachoudis, *Nucl. Data Sheets* **120**, 211 (2014).
- [14] P. Stowell *et al.*, *J. Instrum.* **12**, P01016 (2017).
- [15] S. Gardiner, *Comput. Phys. Commun.* **269**, 108123 (2021).
- [16] Y. Hayato and L. Pickering, *Eur. Phys. J. Special Topics* **230**, 4469 (2021).
- [17] M. Khachatryan *et al.* (CLAS and e4v Collaborations), *Nature (London)* **599**, 565 (2021).
- [18] J. Isaacson, W. I. Jay, A. Lovato, P. A. N. Machado, and N. Rocco, *Phys. Rev. D* **107**, 033007 (2023).
- [19] H. W. Bertini, *Phys. Rev.* **131**, 1801 (1963).
- [20] J. Cugnon, *Phys. Rev. C* **22**, 1885 (1980).
- [21] G. F. Bertsch, H. Kruse, and S. D. Gupta, *Phys. Rev. C* **29**, 673(R) (1984); **33**, 1107(E) (1986).
- [22] G. F. Bertsch and S. Das Gupta, *Phys. Rep.* **160**, 189 (1988).

- [23] J. Cugnon, C. Volant, and S. Vuillier, *Nucl. Phys.* **A620**, 475 (1997).
- [24] A. Boudard, J. Cugnon, S. Leray, and C. Volant, *Phys. Rev. C* **66**, 044615 (2002).
- [25] Y. Sawada, Y. Uozumi, S. Nogamine, T. Yamada, Y. Iwamoto, T. Sato, and K. Niita, *Nucl. Instrum. Methods Phys. Res., Sect. B* **291**, 38 (2012).
- [26] Y. Uozumi, T. Yamada, S. Nogamine, and M. Nakano, *Phys. Rev. C* **86**, 034610 (2012).
- [27] N. Rocco, *Front. Phys.* **8**, 116 (2020).
- [28] J. Isaacson, W. I. Jay, A. Lovato, P. A. N. Machado, and N. Rocco, *Phys. Rev. C* **103**, 015502 (2021).
- [29] S. Dytman, Y. Hayato, R. Raboanary, J. T. Sobczyk, J. Tena Vidal, and N. Vololoniaina, *Phys. Rev. D* **104**, 053006 (2021).
- [30] J. Gonzalez-Rosa, G. D. Megias, J. A. Caballero, M. B. Barbaro, and J. M. Franco-Patino, *Phys. Rev. D* **108**, 113008 (2023).
- [31] A. S. Meyer, M. Betancourt, R. Gran, and R. J. Hill, *Phys. Rev. D* **93**, 113015 (2016).
- [32] T. Le *et al.* (MINERvA Collaboration), *Phys. Rev. D* **100**, 052008 (2019).
- [33] A. M. Ankowski, A. Friedland, S. W. Li, O. Moreno, P. Schuster, N. Toro, and N. Tran, *Phys. Rev. D* **101**, 053004 (2020).
- [34] A. M. Ankowski and A. Friedland, *Phys. Rev. D* **102**, 053001 (2020).
- [35] D. Coplowe *et al.* (MINERvA Collaboration), *Phys. Rev. D* **102**, 072007 (2020); **110**, 059903(E) (2024).
- [36] A. Filkins *et al.* (MINERvA Collaboration), *Phys. Rev. D* **101**, 112007 (2020).
- [37] K. Borah, R. J. Hill, G. Lee, and O. Tomalak, *Phys. Rev. D* **102**, 074012 (2020).
- [38] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **101**, 112004 (2020).
- [39] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **103**, 112009 (2021).
- [40] O. Tomalak, Q. Chen, R. J. Hill, and K. S. McFarland, *Nat. Commun.* **13**, 5286 (2022).
- [41] D. Ruterbories *et al.* (MINERvA Collaboration), *Phys. Rev. D* **104**, 092007 (2021).
- [42] M. A. Acero *et al.* (NOvA Collaboration), *Phys. Rev. D* **107**, 052011 (2023).
- [43] U. Mosel and K. Gallmeister, *Phys. Rev. D* **109**, 033008 (2024).
- [44] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **110**, 013006 (2024).
- [45] A. S. Kronfeld, D. G. Richards, W. Detmold, R. Gupta, H.-W. Lin, K.-F. Liu, A. S. Meyer, R. Sufian, and S. Syritsyn (USQCD Collaboration), *Eur. Phys. J. A* **55**, 196 (2019).
- [46] A. S. Meyer, A. Walker-Loud, and C. Wilkinson, *Annu. Rev. Nucl. Part. Sci.* **72**, 205 (2022).
- [47] A. M. Ankowski, P. Coloma, P. Huber, C. Mariani, and E. Vagnoni, *Phys. Rev. D* **92**, 091301 (2015).
- [48] A. M. Ankowski and C. Mariani, *J. Phys. G* **44**, 054001 (2017).
- [49] A. M. Ankowski, arXiv:1704.07835.
- [50] S. Nagu, J. Singh, J. Singh, and R. B. Singh, *Nucl. Phys.* **B951**, 114888 (2020).
- [51] R. Devi, J. Singh, and B. Potukuchi, *Int. J. Theor. Phys.* **61**, 243 (2022).
- [52] R. Acciarri *et al.* (MicroBooNE, LAr1-ND, and ICARUS-WA104 Collaborations), arXiv:1503.01520.
- [53] M. A. Acero *et al.* (NOvA and R. Group Collaborations), *Eur. Phys. J. C* **80**, 1119 (2020).
- [54] M. A. Acero *et al.* (NOvA Collaboration), *Phys. Rev. D* **106**, 032004 (2022).
- [55] A. C. Booth, Electron neutrino appearance at the NOvA experiment, Ph.D. thesis, University of Sussex, 2021.
- [56] V. Hewes *et al.* (DUNE Collaboration), *Instruments* **5**, 31 (2021).
- [57] C. Awe *et al.*, *J. Instrum.* **16**, P02035 (2021).
- [58] M. Berger, J. Coursey, M. Zucker, and J. Chang, ESTAR, PSTAR, and ASTAR: Computer programs for calculating stopping-power and range tables for electrons, protons, and helium ions (version 1.2.3) (National Institute of Standards and Technology, Gaithersburg, 2005), <http://physics.nist.gov/Star>.
- [59] R. Acciarri *et al.* (DUNE Collaboration), arXiv:1512.06148.
- [60] A. Friedland and S. W. Li, *Phys. Rev. D* **99**, 036009 (2019).
- [61] P. Stowell *et al.* (MINERvA Collaboration), *Phys. Rev. D* **100**, 072005 (2019).
- [62] K. Abe *et al.* (T2K Collaboration), *Eur. Phys. J. C* **83**, 782 (2023).
- [63] G. D'Agostini, *Nucl. Instrum. Methods Phys. Res., Sect. A* **362**, 487 (1995).
- [64] T. Adye, in *PHYSTAT 2011* (CERN, Geneva, 2011), pp. 313–318.
- [65] L. Fields, DUNE Fluxes, <https://glaucus.crc.nd.edu/DUNEFluxes/>.
- [66] B. Abi *et al.* (DUNE Collaboration), *Eur. Phys. J. C* **80**, 978 (2020).
- [67] B. Abi *et al.* (DUNE Collaboration), arXiv:2103.04797.
- [68] L. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962).
- [69] W. Botermans and R. Malfliet, *Phys. Rep.* **198**, 115 (1990).
- [70] E. Marzec *et al.*, *Phys. Rev. Lett.* **134**, 081801 (2025).
- [71] K. Gallmeister and U. Mosel, *Phys. Rev. D* **111**, L091301 (2025).