

# Testing charged current quasi-elastic and multinucleon interaction models in the NEUT neutrino interaction generator with published datasets from the MiniBooNE and MINERvA experiments

C. Wilkinson,<sup>1,2,\*</sup> R. Terri,<sup>3,†</sup> C. Andreopoulos,<sup>4,5</sup> A. Bercellie,<sup>6</sup> C. Bronner,<sup>7</sup> S. Cartwright,<sup>2</sup> P. de Perio,<sup>8,‡</sup> J. Dobson,<sup>9,§</sup> K. Duffy,<sup>10</sup> A. P. Furmanski,<sup>11,||</sup> L. Haegel,<sup>12</sup> Y. Hayato,<sup>13,14</sup> A. Kaboth,<sup>15,4</sup> K. Mahn,<sup>16</sup> K. S. McFarland,<sup>6</sup> J. Nowak,<sup>17</sup> A. Redij,<sup>1</sup> P. Rodrigues,<sup>6</sup> F. Sánchez,<sup>18</sup> J. D. Schwehr,<sup>19</sup> P. Sinclair,<sup>9</sup> J. T. Sobczyk,<sup>20</sup> P. Stamoulis,<sup>21</sup> P. Stowell,<sup>2</sup> R. Tacik,<sup>22,23</sup> L. Thompson,<sup>2</sup> S. Tobayama,<sup>24</sup> M. O. Wascko,<sup>9</sup> and J. Żmuda<sup>20</sup>

<sup>1</sup>University of Bern, Albert Einstein Center for Fundamental Physics,  
Laboratory for High Energy Physics (LHEP), Bern 3012, Switzerland

<sup>2</sup>University of Sheffield, Department of Physics and Astronomy, Sheffield S10 2TN, United Kingdom

<sup>3</sup>Queen Mary University of London, School of Physics and Astronomy, London E1 4NS, United Kingdom

<sup>4</sup>STFC, Rutherford Appleton Laboratory, Harwell Oxford OX11 0QS, and Daresbury Laboratory,  
Warrington WA4 4AD, United Kingdom

<sup>5</sup>University of Liverpool, Department of Physics, Liverpool L69 7ZE, United Kingdom

<sup>6</sup>University of Rochester, Department of Physics and Astronomy, Rochester 14627, New York, USA

<sup>7</sup>Kavli Institute for the Physics and Mathematics of the Universe (WPI),  
Todai Institutes for Advanced Study, University of Tokyo,  
Kashiwa 277-8583, Chiba, Japan

<sup>8</sup>University of Toronto, Department of Physics, Toronto M5S 1A7, Ontario, Canada

<sup>9</sup>Imperial College London, Department of Physics, London W7 2BB, United Kingdom

<sup>10</sup>Oxford University, Department of Physics, Oxford OX1 3RH, United Kingdom

<sup>11</sup>University of Warwick, Department of Physics, Coventry CV4 7AL, United Kingdom

<sup>12</sup>University of Geneva, Section de Physique, DPNC, Geneva 1211, Switzerland

<sup>13</sup>Kavli Institute for the Physics and Mathematics of the Universe (WPI),

Todai Institutes for Advanced Study, University of Tokyo, Kashiwa 277-8583, Chiba, Japan

<sup>14</sup>University of Tokyo, Institute for Cosmic Ray Research, Kamioka Observatory, Kamioka 277-8582,  
Japan

<sup>15</sup>Royal Holloway University of London, Department of Physics, Surrey TW20 0EX, United Kingdom

<sup>16</sup>Michigan State University, Department of Physics and Astronomy, East Lansing 48824, Michigan, USA

<sup>17</sup>Lancaster University, Physics Department, Lancaster LA1 4YB, United Kingdom

<sup>18</sup>Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology,  
Campus UAB, Bellaterra (Barcelona) 08193, Spain

<sup>19</sup>Colorado State University, Department of Physics, Fort Collins 80523, Colorado, USA

<sup>20</sup>Wrocław University, Faculty of Physics and Astronomy, Wrocław 50-204, Poland

<sup>21</sup>IFIC (CSIC & University of Valencia), Valencia 46980, Spain

<sup>22</sup>University of Regina, Department of Physics, Regina S4S 0A2, Saskatchewan, Canada

<sup>23</sup>TRIUMF, Vancouver, British Columbia V6T 2A3, Canada

<sup>24</sup>University of British Columbia, Department of Physics and Astronomy, Vancouver,  
British Columbia V6T 1Z1, Canada

(Received 23 January 2016; published 21 April 2016)

There has been a great deal of theoretical work on sophisticated charged current quasi-elastic (CCQE) neutrino interaction models in recent years, prompted by a number of experimental results that measured unexpectedly large CCQE cross sections on nuclear targets. As the dominant interaction mode at T2K energies, and the signal process in oscillation analyses, it is important for the T2K experiment to include realistic CCQE cross section uncertainties in T2K analyses. To this end, T2K's Neutrino Interaction Working Group has implemented a number of recent models in NEUT, T2K's primary neutrino interaction event generator. In this paper, we give an overview of the models implemented and present fits to published  $\nu_\mu$  and  $\bar{\nu}_\mu$  CCQE cross section measurements from the MiniBooNE and MINERvA experiments.

\*Corresponding author.  
callum.wilkinson@lhep.unibe.ch

†Corresponding author.  
r.terri@qmul.ac.uk

‡Present address: Columbia University, Physics Department, New York, NY 10027, USA.

§Present address: University College London, Department of Physics and Astronomy, London WC1E 6BT, United Kingdom.

||Present address: University of Manchester, School of Physics and Astronomy, Manchester M13 9PL, United Kingdom.

The results of the fits are used to select a default cross section model for future T2K analyses and to constrain the cross section uncertainties of the model. We find strong tension between datasets for all models investigated. Among the evaluated models, the combination of a modified relativistic Fermi gas with multinucleon CCQE-like interactions gives the most consistent description of the available data.

DOI: [10.1103/PhysRevD.93.072010](https://doi.org/10.1103/PhysRevD.93.072010)

## I. INTRODUCTION

Charged current quasi-elastic (CCQE) scattering ( $\nu_\mu + n \rightarrow p + \mu^-$ ) is the dominant neutrino interaction process for muon (anti)neutrinos impinging on a nuclear target at neutrino energies on the order of 1 GeV. Because CCQE is a two-body process and the incoming neutrino direction is known for an accelerator experiment, a reasonable approximation of the neutrino energy can be calculated using only the outgoing lepton kinematics. Because of this, CCQE is the preferred signal process for neutrino oscillation experiments, which generally require some handle on the incoming neutrino energy to extract neutrino oscillation parameters due to  $\bar{\nu}_\mu$  disappearance or  $\bar{\nu}_e$  appearance in this energy region. However, nuclear effects and interactions that are not distinguishable from CCQE in the final state bias or smear the reconstructed neutrino energy, so a good understanding of these effects is important.

Neutrino interaction generators typically use the relativistic Fermi gas (RFG) model of the nucleus for all neutrino-nucleus interactions because of its simplicity. In the RFG model, all possible nucleon momentum states are filled up to the Fermi momentum, there is a constant binding energy required to separate the nucleon from the nucleus, and the neutrino interacts with a single bound nucleon. Neutrino-nucleon CCQE scattering for free nucleons is described by the Llewellyn-Smith formalism [1], which has been extended to cover neutrino-nucleus CCQE scattering in the Smith-Moniz RFG model [2], where nucleons bound within the nucleus are described by the RFG nuclear model. The only parameter of the weak current or in the RFG model that is not well constrained by electron scattering data [3,4] is the axial mass  $M_A$ . Results from a global analysis of neutrino-deuteron scattering experiments and pion electroproduction data find  $M_A = 1.00 \pm 0.02$  GeV/ $c^2$  [5], which is consistent with other analyses [6–8]. These results are also consistent with high energy neutrino beam experiments on heavy nuclear targets [9].

Recent differential CCQE cross section results from the MiniBooNE Collaboration [10,11] are significantly higher than the expectation, which can only be accounted for in the framework of the Smith-Moniz RFG model by inflating the axial mass, giving rise to the term “MiniBooNE large axial mass anomaly.” A separate issue observed by both K2K

[12] and MiniBooNE was a deficit of events at low  $Q^2$ . K2K obtained an axial mass value of  $M_A = 1.20 \pm 0.12$  GeV/ $c^2$  by fitting the  $d\sigma/dQ^2$  shape. The low- $Q^2$  region can also be modeled by an inflation of the axial mass, as increasing  $M_A$  leads to suppression at low  $Q^2$  as well as an enhancement at high  $Q^2$ . However, the inflation of the axial mass required to model the low- $Q^2$  region is significantly larger than the value required to model the increased event rate, so these two discrepancies in the data are not consistent. Typically, the  $M_A$  values quoted from both K2K and MiniBooNE do not include the low- $Q^2$  region in the fit or add additional parameters to model this region. Other experiments using heavy nuclear targets with beam energies in the few-GeV region have also measured cross sections that are consistent with an inflation of the axial mass [13–16], although these results do not paint a coherent picture. More recently, the MINER $\nu$ A experiment [17,18], which is at a somewhat higher neutrino energy than MiniBooNE, has shown good agreement with the Smith-Moniz RFG model with  $M_A = 1.00$  GeV/ $c^2$ , but it requires an enhancement to the transverse component of the cross section, an effect also seen in electron-nucleus scattering [19]. These inconsistent results present a considerable challenge to neutrino oscillation experiments, which need to be able to model their signal processes well.

Recent theoretical efforts that have attempted to resolve the “MiniBooNE large axial mass anomaly” have focused on two main areas: more sophisticated descriptions of the initial ground state of the nucleus and additional nuclear effects, such as multinucleon interaction models, which go beyond interactions with a single nucleon within the nucleus. The combination of these models would allow for a consistent picture of an axial mass close to 1.00 GeV/ $c^2$ , with a suppressed cross section at low  $Q^2$  and larger cross section at higher  $Q^2$  relative to a simple RFG model. Comprehensive reviews of available CCQE cross section models can be found in Refs. [3,20–22].

Spectral function (SF) models are more sophisticated descriptions of the initial state of the nucleus than the RFG model and are available from a number of authors [23–26]. These models have a more realistic nucleon momentum distribution taking into account the shell structure of the nucleus and correlated pairs of nucleons within the nucleus, and they have nonconstant binding energies. Note that although these models include correlations between nucleons in the initial state, they still use the impulse

approximation and only consider interactions with a single nucleon. More complex models that go beyond the simple picture of noninteracting fermions are available [27–32]. However, with the exception of the GiBUU interaction model [30,33], these are not currently implemented in neutrino interaction generators. In these models, a mean-field potential due to the presence of other nucleons within the nucleus is calculated, which will generally depend on the position and momentum of the struck nucleon. These models are not discussed further as they cannot be easily implemented in the NEUT neutrino interaction generator [34] due to the models being computationally intensive, and therefore slow, and they may not be valid for the entire energy range required by the generator.

Although alternative nuclear models modify the cross section as a function of the outgoing lepton kinematics significantly, they do not change the total CCQE cross section significantly as a function of neutrino energy [21]. Additional nuclear effects are also likely to be required to explain the current global dataset. Multinucleon interaction (2p2h) models such as those from Martini *et al.* [35] and Nieves *et al.* [36,37] go beyond the impulse approximation and include diagrams where two nucleons are involved in the interaction. This kind of interaction adds significant strength to the CCQE-like cross section and explains the difference in normalization observed in the MiniBooNE data, which was previously modeled with a large axial mass [35,36,38–41]. Because these 2p2h models are not two-body processes, they are expected to lead to significant biases in the neutrino energy reconstruction from the outgoing lepton, which assumes CCQE kinematics [42–46]. Additionally, the random phase approximation (RPA) is a nuclear screening effect that modifies the propagator for interactions in nuclear matter [35,36,47–49] and has a significant effect on the differential cross section as a function of  $Q^2$ , suppressing the cross section in the low- $Q^2$  region and enhancing the cross section for  $Q^2 \gtrsim 0.5 \text{ GeV}^2$ . RPA needs to be included, both in interactions with a single nucleon (1p1h) and those from 2p2h calculations, to find good agreement with data. Note that both Nieves and Martini calculations are performed in the context of a local Fermi gas (LFG) model, where the Fermi momentum depends on the local nuclear density. Because these nuclear effects are nuclear model dependent, it is not necessarily easy to combine models of nuclear effects with new CCQE interaction models. While there have been rapid experimental and theoretical developments relating to CCQE cross sections, a limited number of new nuclear models and nuclear effects (such as 2p2h and RPA) have only recently been implemented into neutrino interaction generators or confronted with neutrino-nucleus scattering data, and no consistent picture has yet emerged. It is not clear which models fit the global data best, and where the deficiencies now lie, which should be a serious concern for neutrino oscillation experiments. This paper shows the

effect of fitting current CCQE and multinucleon models to the MiniBooNE [10,11] and MINER $\nu$ A [17,18] datasets to a variety of models implemented in NEUT by members of T2K’s Neutrino Interactions Working Group (NIWG). Previous constraints on the CCQE model produced by the NIWG and used in T2K oscillation analyses only considered a RFG model and recommended the NEUT default central value for the axial mass  $M_A = 1.21 \text{ GeV}/c^2$  based on the value found by the K2K experiment [12], with an error large enough to cover fits to the MiniBooNE neutrino mode CCQE dataset [10], as is fully described in Ref. [50]. This work improves on the previous situation by including more sophisticated effects proposed to explain the large axial mass anomaly and by using all of the newly available CCQE data to constrain all model parameters without reference to the default NEUT model.

The models that have been implemented in the NEUT generator are discussed in Sec. II, and Sec. III discusses cross-generator validation. Section IV gives a brief overview of the MiniBooNE and MINER $\nu$ A data used in the fit. The nominal NEUT predictions for these datasets are shown in Sec. V for a variety of models. Section VI discusses the fit procedure. The results of fake data studies and the fit to external data are given in Sec. VII. In Sec. VIII, we interpret the results and discuss possible implications in cross section and neutrino oscillation analyses, and Sec. IX summarizes the results.

## II. INTERACTION MODELS IN NEUT

The motivation for, and an overview of, new CCQE models has already been discussed. This section will briefly outline the important technical details of the models as implemented in NEUT and highlight any caveats that should be kept in mind when fitting with them. The models used in the fits include the SF model, multinucleon-neutrino interactions, and RPA.

The NEUT implementation of the SF model from Benhar and Fabrocini [23] is described fully in Ref. [51]. Although SF is a generic term, in this work it will specifically refer to the Benhar SF. The model information is all encoded in the initial state nucleon distribution shown in Fig. 1. Pauli blocking is implemented as a hard cutoff: final state nucleons with momenta less than the Fermi momentum  $p_F^{\text{SF}}$  are forbidden. There are two terms in the SF model: a short-range correlation term, which extends to higher initial state nucleon momenta, and a mean-field term, which contributes the main peak at lower momenta. These terms can be seen in Fig. 2, where the two-dimensional SF in terms of the removal energy and initial state nucleon momentum has been projected onto the momentum axis. There are three parameters in NEUT that modify the SF as illustrated in Fig. 2. The default values for these parameters are given in Table I. The mean-field width and normalization of the correlation term are well constrained by electron-scattering data [51] and have little effect on the

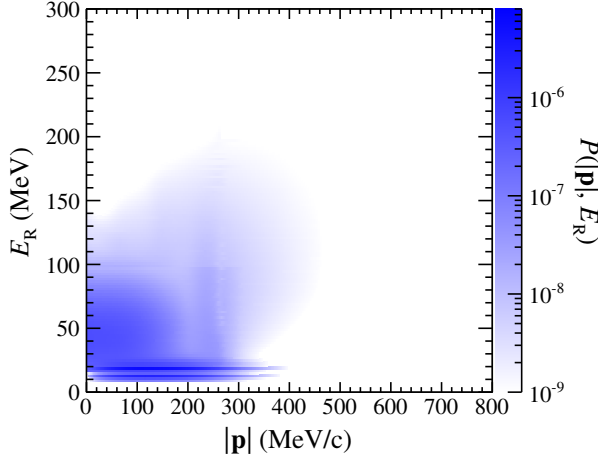


FIG. 1. The probability distribution for initial state protons within an oxygen nucleus for Benhar's SF model [23] as a function of the removal energy ( $E_R$ ) and the magnitude of the nucleon momentum ( $|\mathbf{p}|$ ). The SF is normalized such that the integral of this distribution is 1.

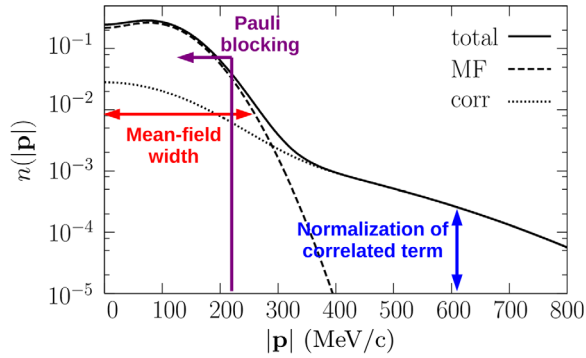


FIG. 2. SF parameters in NEUT that may be modified on the SF initial state momentum distribution. This figure has been adapted from Ref. [52].

shape or normalization of the cross section. Thus, they are not considered further in this work. Pauli blocking is modified by changing the Fermi momentum in the fits. It should be noted that in the RFG model, the Fermi momentum defines the Pauli blocking but *also* modifies the width of the initial state nucleon distribution. As a result, changing  $p_F^{\text{RFG}}$  affects a wide range of  $Q^2$ , whereas changing  $p_F^{\text{SF}}$  only affects very low  $Q^2$  events.

The 2p2h model from Nieves *et al.* [36,37] has been implemented in NEUT as described in Ref. [54]. The cross section as a function of neutrino energy and the outgoing lepton kinematics was made available by the authors of the model and is implemented as a series of lookup tables for various nuclear targets and neutrino species. The tables provided have hadronic variables integrated out, so a generic model based on Ref. [55] for simulating the initial and final hadronic states is used for generating NEUT

TABLE I. Nominal model parameters for the SF + 2p2h model.

Model parameter	NEUT default value
$M_A$	1.01 GeV/ $c^2$
Fermi momentum, $p_F^{\text{SF}}$	209 MeV/ $c$
Mean-field width	200 MeV/ $c$ (Benhar nominal [23])
Normalization of the correlation term	Benhar nominal [23] (correlated tail $\sim 20\%$ of total)
2p2h normalization	100% Nieves model [36,37]
Axial form factor	Dipole
Vector form factors	BBBA05 [53]

events.<sup>1</sup> The discrepancy between the leptonic and hadronic simulations makes the current NEUT implementation of the Nieves model inadequate for comparisons with experimental measurements of the final state hadrons from CCQE events (such as can be found in Ref. [56]). For this reason, only leptonic measurements are used in this analysis. As the Nieves model is very complex, the current NEUT implementation does not allow fundamental model parameters to be changed. For simplicity, only a simple scaling parameter that changes the normalization of 2p2h events has been considered in this analysis. Note that the Nieves 2p2h model included  $\pi$ -less  $\Delta$  decay contributions, where a nucleon excited into a  $\Delta(1232)$  resonance decays without producing a pion [57,58]. Contributions from  $\pi$ -less  $\Delta$  decay were previously implemented in NEUT and other generators, and have been treated as an intrinsic background in CCQE selections and corrected for. This leads to complications when comparing the full Nieves model to CCQE cross section measurements.

RPA [36] is implemented into NEUT as a modification to the 1p1h cross section as a function of  $E_\nu$  and  $Q^2$ . Figure 3 shows the ratio of the Nieves 1p1h cross section with RPA included over the bare 1p1h cross section; these two-dimensional tables of the ratio were supplied by the authors of Ref. [36] and are used to apply the RPA correction in NEUT. The Nieves RPA calculation uses the *local* Fermi gas nuclear model, and NEUT only has a *global* Fermi gas model implemented for 1p1h interactions. However, the authors of the RPA calculation have noted [37] that the same ratio can be applied, with reasonable precision, to a global Fermi gas. Because of the model dependence, the same ratios cannot be applied to modify the 1p1h interactions calculated with a SF model, and no RPA calculation performed in the context of the SF nuclear model is available. Two different RPA calculations are available from the same authors, termed relativistic and nonrelativistic, which affect the quenching of the RPA at high  $Q^2$

<sup>1</sup>This model simply enforces energy and momentum conservation, treats initial nucleons as uncorrelated and drawn from a local Fermi gas model, and shares momentum equally between final state nucleons [55].



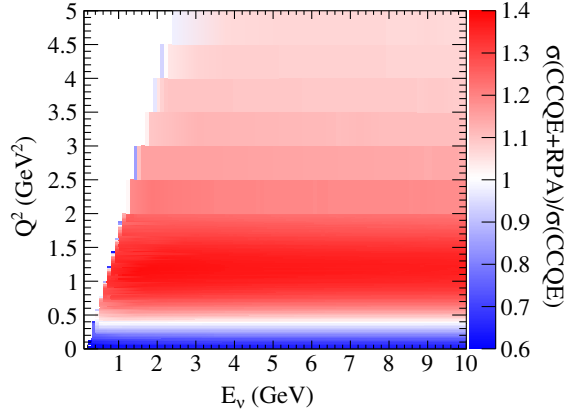
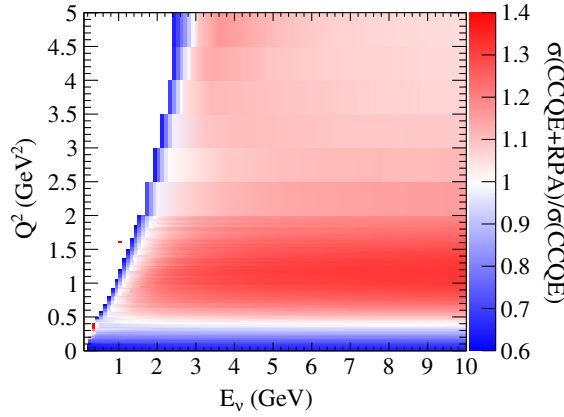
(a)  $\nu_\mu - {}^{12}\text{C}$ (b)  $\bar{\nu}_\mu - {}^{12}\text{C}$ 

FIG. 3. The ratio of the CCQE cross section including the nonrelativistic RPA model to the CCQE cross section without RPA, shown for both muon neutrino and muon antineutrino interactions on carbon. An enhancement of the ratio can be seen at high  $Q^2$ , and a suppression can be seen at low  $Q^2$  (and close to the kinematic boundary for antineutrinos). These  $E_\nu$ —and  $Q^2$ —dependent tables are used in NEUT to apply the RPA model. For these plots, an axial mass value of  $M_A = 1.01 \text{ GeV}/c^2$  was used.

( $\gtrsim 0.5 \text{ GeV}^2$ ). The ratio of nonrelativistic to relativistic RPA is shown in Fig. 4. Both RPA models are investigated in this analysis as it is not clear which model is expected to best fit the global dataset. In Ref. [38], nonrelativistic RPA corrections are used to compare with MiniBooNE neutrino mode data, whereas in Ref. [37], the relativistic RPA correction is introduced to compare with the MINERvA datasets. The “stray” points in Figs. 3 and 4 are artifacts from the authors of the RPA model, who provided the data used to produce these figures. The cause of these artifacts is unknown, but as these points lie outside the kinematically allowed region of ( $E_\nu$ ,  $Q^2$ ) space, they do not affect the RPA implementation in NEUT as no event outside this region can be generated.

With these different ingredients, three distinct candidate CCQE models are available in NEUT, which are all considered in this work:

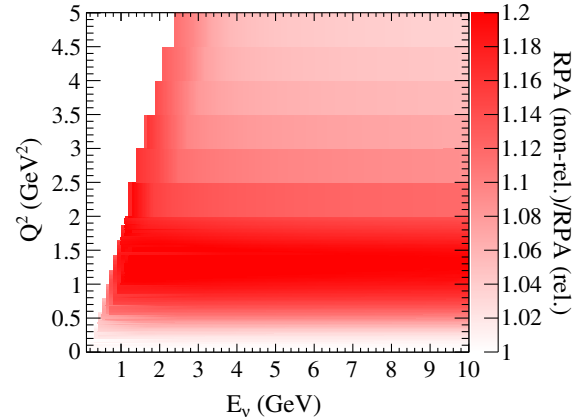
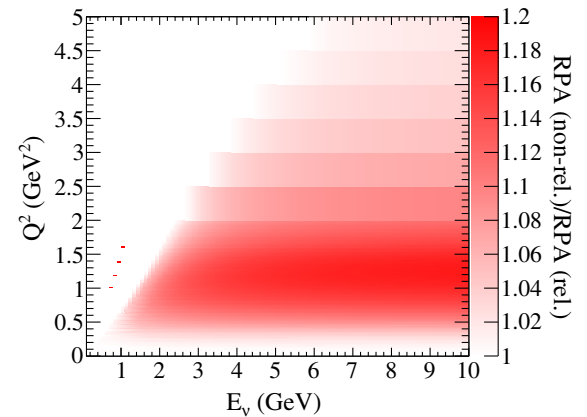
(a)  $\nu_\mu - {}^{12}\text{C}$ (b)  $\bar{\nu}_\mu - {}^{12}\text{C}$ 

FIG. 4. The ratio of the nonrelativistic RPA correction to the relativistic RPA correction, shown for both muon neutrino and muon antineutrino interactions on carbon. These  $E_\nu$ —and  $Q^2$ —dependent tables are used to reweight NEUT events from one RPA model to the other. By default, NEUT events are generated with the nonrelativistic RPA model.

- (1) RFG + relativistic RPA + 2p2h
- (2) RFG + nonrelativistic RPA + 2p2h
- (3) SF + 2p2h

The default values for all variable model parameters are listed in Tables I and II for both SF + 2p2h and RFG + RPA + 2p2h models, respectively.

TABLE II. Nominal model parameters for the relativistic and nonrelativistic RFG + RPA + 2p2h models.

Model parameter	NEUT default value
$M_A$	$1.01 \text{ GeV}/c^2$
Fermi momentum, $p_F^{\text{RFG}}$	$217 \text{ MeV}/c$
RPA	Nieves relativistic or Nonrelativistic model [36]
2p2h normalization	100% Nieves model [36,37]
Axial form factor	Dipole
Vector form factors	BBBA05 [53]

It should be noted that there are deficiencies for both models as currently implemented in NEUT. The RFG + RPA + 2p2h model is very like the full Nieves model as both the RPA and 2p2h calculations are taken from it. However, the Nieves model consistently uses a *local* Fermi gas, whereas NEUT uses a *global* Fermi gas model for the 1p1h calculation. Currently, we do not have the ability to vary the value of  $M_A$  used in the Nieves model prediction as implemented in NEUT, making the fits slightly inconsistent in this regard. We note that the value of the axial mass used for the 2p2h contribution to the cross section is fixed to  $M_A = 1.01 \text{ GeV}/c^2$  in NEUT, which is slightly different from the nominal  $M_A = 1.049 \text{ GeV}/c^2$  used in Ref. [36]. The SF + 2p2h model has no RPA correction applied, which is physically inconsistent as the 2p2h enhancement is used (both corrections are due to complications in heavy nuclear targets). As previously noted, no RPA calculation appropriate for a SF model is currently available, so this inconsistency is unavoidable. The nuclear models used for the 1p1h calculation (SF) and the 2p2h calculation (LFG) are also inconsistent, and it has been remarked that the short-range correlations included in the SF nuclear model may be the same as some contributions to the Nieves 2p2h interaction model, so some contributions may be included twice.

Additionally, the effective spectral function (ESF) [26,59] has been implemented in NEUT as described in Ref. [60] and is included for comparison with the other nominal models in Sec. V. The ESF enforces agreement with the longitudinal response function extracted from electron scattering data by modifying the initial state nucleon momentum distribution (using a simple parametrization of the Benhar SF model) and should be used with the transverse enhancement model (TEM), which parametrizes the observed discrepancy between the longitudinal and transverse response functions extracted from electron scattering data as an enhancement to the magnetic form factor [19]. By construction, the ESF + TEM agrees with elastic electron scattering data and is extended to neutrino scattering data by modifying the Llewellyn-Smith interaction formalism for nucleons bound in a nucleus described by the ESF (and with the modified magnetic form factor from the TEM). This model was implemented too late to be a candidate model for the T2K oscillation analysis and is not considered further in the fitting work described in this paper.

We note that neither of our candidate models is expected to describe experimental data at low momentum transfer because they do not include nuclear effects such as nuclear excitations and collective resonances which will affect the cross section. In other analyses that fit models to CCQE data, bins that are dominated by low momentum transfer events are excluded [61]. In this analysis, we have not followed any such bin masking procedure. Arguably, to obtain a realistic value of the model parameters, one should

only fit the model in its stated region of validity. However, the main focus of this analysis is to obtain central values and errors for the T2K oscillation analysis, where the cross section model is used for all regions of phase space, so some pragmatism is required.

### III. NUWRO AS A VALIDATION TOOL FOR NEW INTERACTION MODELS

The NuWro Monte Carlo (MC) generator for neutrino interactions has been developed over the past  $\sim 10$  years at the University of Wrocław [62]. It was the first MC generator to have an implementation of the Benhar SF [23] and the Nieves 2p2h model included [36,37], and it served as the benchmark for the NEUT development of both models. The implementation of the SF model in NuWro was based on the code written for Ref. [63] and subsequently optimized for NuWro. The Nieves model implementation in NuWro used a series of lookup tables for the 2p2h cross section as a function of leptonic variables for various nuclear targets and neutrino species, so it is very similar to NEUT; however, it has since been improved to use a more general formalism that depends on a number of nuclear response functions that can be extracted from the Nieves code and therefore reduces the number of lookup tables required. The same generic model [55] was used to simulate the initial and final hadronic states in NuWro as was used in NEUT. For both the SF and Nieves 2p2h models, NuWro and NEUT are in good agreement, which provides a useful validation of the NEUT implementations of these models.

### IV. EXTERNAL DATASETS

Four datasets are used in the CCQE fits presented in this work: the MiniBooNE neutrino [10] (2010) and antineutrino [11] (2013) results, and the MINERvA neutrino [17] (2013) and antineutrino [18] (2013) results. All experimental details and information about these results, which are reproduced here, are taken from the references cited above unless otherwise stated.

The single-differential cross section results are given in terms of  $Q_{QE}^2$ , the four-momentum transfer calculated from lepton kinematics under the quasi-elastic hypothesis, which is calculated using the equations

$$E_\nu^{QE} = \frac{2M'_n E_\mu - (M_n'^2 + m_\mu^2 - M_p^2)}{2(M'_n - E_\mu + \sqrt{E_\mu^2 - m_\mu^2} \cos \theta_\mu)}, \quad (1)$$

$$Q_{QE}^2 = -m_\mu^2 + 2E_\nu^{QE} \left( E_\mu - \sqrt{E_\mu^2 - m_\mu^2} \cos \theta_\mu \right), \quad (2)$$

where  $E_\mu$  is the muon energy;  $M_n$ ,  $M_p$  and  $m_\mu$  are the neutron, proton and muon masses, respectively; and  $M'_n = M_n - V$ , where  $V$  is the binding energy of carbon

assumed in the analysis.<sup>2</sup> For both MiniBooNE datasets and the MINER $\nu$ A neutrino dataset,  $V = 34$  MeV; for the MINER $\nu$ A antineutrino dataset,  $V = 30$  MeV.

In the MiniBooNE analysis,  $Q_{QE}^2$  is calculated from the unfolded  $T_\mu$  and  $\cos\theta_\mu$  distributions. The MINER $\nu$ A analysis unfolds the  $Q_{QE}^2$  distribution calculated with the reconstructed  $p_\mu$  and  $\cos\theta_\mu$  values. The errors on the  $Q_{QE}^2$  distributions for both experiments include the uncertainties relating to the muon reconstruction, so they should cover the difference in the method used to produce the  $Q_{QE}^2$  cross section results. We note that the main results of our analysis use the MiniBooNE double-differential results only, so there is no possible tension from differences between the methods used to produce  $Q_{QE}^2$  distributions.

### A. MiniBooNE neutrino

The MiniBooNE CCQE dataset has been released as a double-differential cross section as a function of  $(T_\mu, \cos\theta_\mu)$ , where  $T_\mu$  is the kinetic energy of the outgoing muon and  $\theta_\mu$  is the angle between the incoming neutrino and outgoing muon. Differential cross sections were also released as a function of  $Q_{QE}^2$  or  $E_\nu^{QE\text{ RFG}}$ , but the double-differential result was preferred as it contains the most information and has minimal model dependence. The MiniBooNE data release included central values for each bin and the diagonal elements of the shape-only covariance matrix; correlations between bins were not released. Additionally, the overall normalization uncertainty was given as 10.7% for neutrino running.

The MiniBooNE CCQE cross sections are released as both CCQE-corrected and CCQE-like measurements. The CCQE-like sample is obtained by selecting events in which a muon was detected with no pions, but no requirement was made on the proton. The CCQE-corrected measurement is produced by subtracting background events (where the primary interaction is not CCQE) based on the NUANCE [64] generator prediction. The dominant background is  $CC1\pi^+$ , and a dedicated sample was used to tune the NUANCE prediction, which was used in the background subtraction. It should be noted that the NUANCE  $CC1\pi^+$  simulation included  $\pi$ -less  $\Delta$  decay. The published signal purity for the neutrino dataset is 77%.

CCQE-like results are less model dependent than CCQE-corrected results (as they do not rely on the experiment's own MC correction strategy) but make the analysis dependent on the simulation of the background in the MiniBooNE detector, which cannot be tuned to the MiniBooNE data in the same way MiniBooNE's

background model could be. CCQE-corrected results are used in this analysis. A downside of using the CCQE-corrected data is the explicit subtraction of  $\pi$ -less  $\Delta$  decay events in the MiniBooNE analysis, which forms part of the Nieves multinucleon-neutrino prediction which we treat as a signal in the analysis. Unfortunately, there is no obvious way to account for this effect, so we ignore it for the analysis presented. We note that Nieves *et al.* also used the CCQE-corrected dataset to compare to their full models [38,40].

### B. MiniBooNE antineutrino

The MiniBooNE antineutrino data have been released in the same format as the neutrino mode data. Again, the double-differential CCQE-corrected results are used. The overall normalization uncertainty was given as 13.0% for antineutrino running. This is likely to be strongly correlated with the normalization uncertainty for the neutrino mode data, as the uncertainty comes mostly from the flux normalization uncertainty. However, as this information was not released, no correlation is assumed in this analysis.

The correction strategy for the antineutrino dataset is more complicated than for the neutrino mode sample because of the relatively high  $\nu_\mu$  contamination in the  $\bar{\nu}_\mu$  beam, which is the largest background in the antineutrino CCQE sample (MiniBooNE is an unmagnetized detector). There is also a large  $CC1\pi^-$  background, the analogue of the  $CC1\pi^+$  contamination in the neutrino dataset. Two properties are used to measure the  $\nu_\mu$  background [65]: 8% of  $\nu_\mu$ -induced CC interactions produce no decay electron due to muon-nucleus capture, and the  $\nu_\mu$ -induced  $CC1\pi^+$  events can be identified independently of  $\bar{\nu}_\mu$ -induced  $CC1\pi^-$  as most  $\pi^-$  mesons are absorbed. Unfortunately, this property makes  $CC1\pi^-$  a bigger background to the CCQE analysis in the antineutrino mode; this means that there is no sample with which to directly tune the  $CC1\pi^-$  production from the NUANCE resonance model, so the neutrino mode  $CC1\pi^+$  has to be used (as was done for the neutrino mode sample). Other backgrounds are subtracted using the NUANCE interaction model after some tuning and corrections. As a result of the two large backgrounds in the antineutrino sample, the purity of the CCQE-like sample is 61%, making the correction larger than for the neutrino mode sample.

### C. MINER $\nu$ A

The CCQE datasets from MINER $\nu$ A are released as CCQE-corrected single-differential flux-averaged cross sections as a function of  $Q_{QE}^2$ , where the flux has been averaged over the region  $1.5 \leq E_\nu \leq 10$  GeV. There is an additional requirement that  $1.5 \leq E_\nu^{QE} \leq 10$  GeV, with  $E_\nu^{QE}$  as defined in Eq. (1). Covariance matrices and central values have been released to perform fits to both shape-only and absolutely normalized neutrino and antineutrino

<sup>2</sup>Note that the binding energy  $V$  is just the value assumed when calculating  $Q_{QE}^2$ . Therefore, we must use the same value as the experiments when producing comparable  $Q_{QE}^2$  distributions, but it need not be consistent with the binding energy used in our simulation.

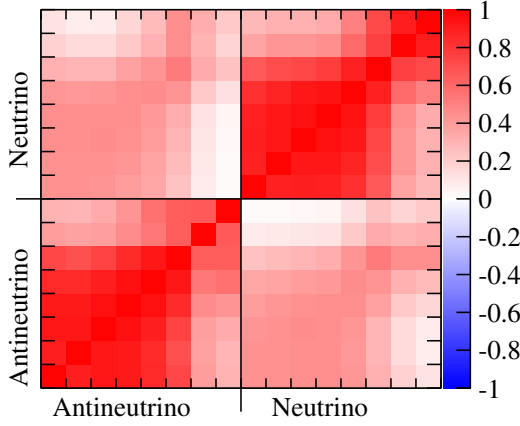


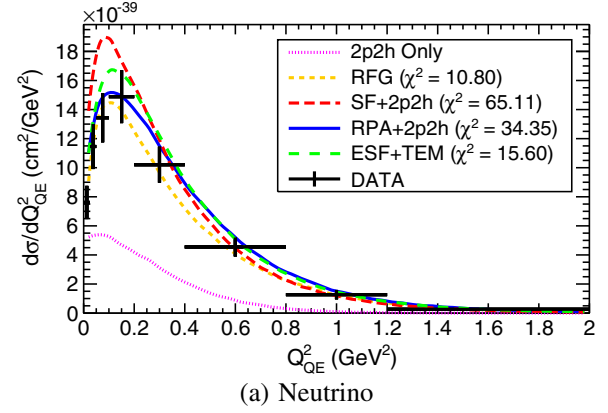
FIG. 5. Cross-correlation matrix including both shape and normalization uncertainties for the MINERνA neutrino and antineutrino samples. The eight neutrino and eight antineutrino bins shown here correspond to the eight  $Q_{QE}^2$  bins from the MINERνA datasets.

datasets. In this work, the absolutely normalized distributions have been used in the fit.

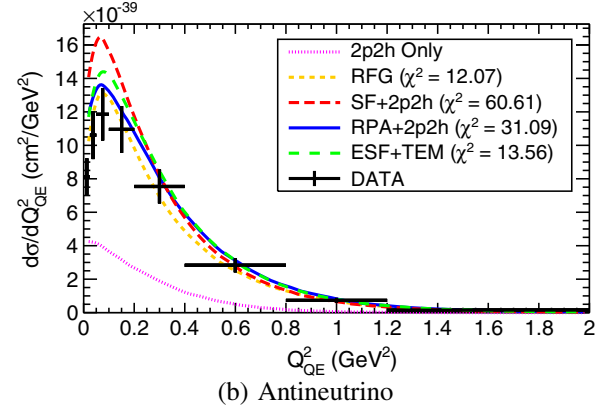
The correction strategy for the MINERνA data is to fit the relative normalizations of simulated background distributions to the data in terms of the recoil energy, energy deposited outside a vertex region (the recoil region), and then subtract the predicted background from the CCQE-like sample. The published purity for the neutrino dataset ranges from 65% at low  $Q_{QE}^2$  to 40% at high  $Q_{QE}^2$  (with an overall purity of 49%). The purity for the antineutrino dataset is given as 77%. The purity is lower for the neutrino analysis because events with a proton from the initial interaction are more complicated to reconstruct than those with a neutron.<sup>3</sup>

In the MINERνA CCQE analyses, the efficiency for selecting events with  $\theta_\mu > 20^\circ$  is very low because the MINOS near detector, downstream of MINERνA, is used to tag muons. This introduces a small model dependence on the results because a RFG model was used to correct for the unsampled region of phase space. The MINERνA Collaboration subsequently released a distribution where the cross section is measured for CCQE events with  $\theta_\mu \leq 20^\circ$ . As this dataset is less model dependent, it has been used in the fits and will be consistently used in this analysis. MINERνA also made cross-correlations between the neutrino and antineutrino datasets available in a data release after the publication of their CCQE papers. The correlation matrices released include both shape and normalization errors, but it is possible to extract shape-only correlation matrices using the method given in

<sup>3</sup>The antineutrino analysis has an additional cut requiring no additional (other than the muon) tracks from the vertex, and it allows only one isolated energy shower, whereas the neutrino mode analysis allows two [17,18].



(a) Neutrino



(b) Antineutrino

FIG. 6. Nominal model predictions for the MINERνA datasets with  $M_A = 1.01$  GeV/ $c^2$  and all other model parameters at their default values. The relativistic RPA model is shown.

Ref. [66]. The full matrix including both shape and normalization errors included is shown in Fig. 5.

## V. MONTE CARLO PREDICTION

For each of the four experimental results included in the fit, one million CCQE and 2p2h events were generated with NEUT for each model using the default parameters given in Tables I and II and the published flux for each dataset. The flux averaged cross section predictions were produced using the following method:

- (1) For each event, apply experiment-specific cuts and, if the event passes, calculate the relevant reconstructed quantity and fill the 1D or 2D event rate histogram.
- (2) Calculate the event rate by integrating the MC event rate histogram (flux  $\times$  cross section).
- (3) Integrate the published flux histogram to get the average flux.
- (4) Scale the filled histogram by the event rate divided by the average flux to get the flux averaged cross section per target nucleon.
- (5) Divide the content of each bin by the bin width.

The default predictions for a variety of models available in NEUT, as well as the data, are shown in Figs. 6, 7 and 8



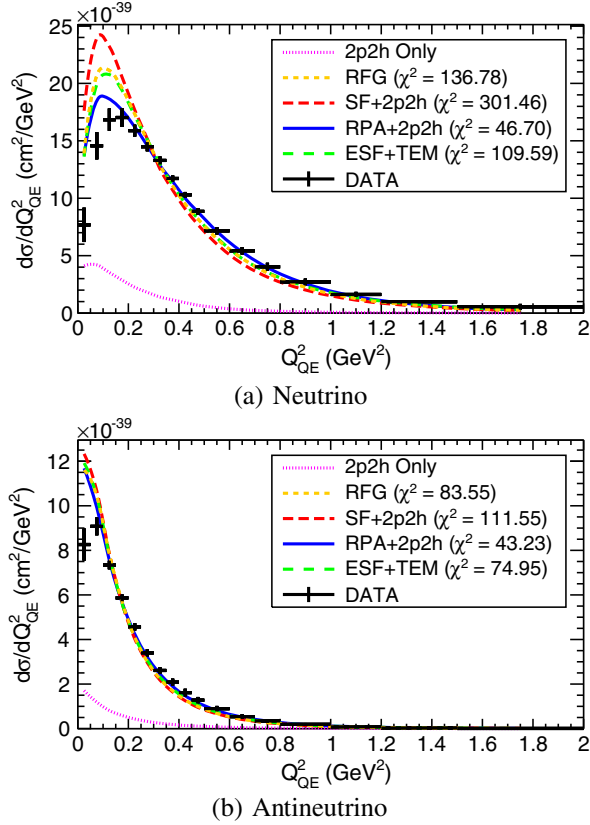


FIG. 7. Nominal model predictions for the MiniBooNE single-differential datasets with  $M_A = 1.01$  GeV/ $c^2$  and all other model parameters at their default values. The relativistic RPA calculation is shown. Normalization parameters are applied as given in Table III.

$$\begin{aligned}
 \chi^2(\vec{\mathbf{x}}) = & \left[ \sum_{k=0}^N \left( \frac{\nu_k^{\text{DATA}} - \lambda_\nu^{-1} \nu_k^{MC}(\vec{\mathbf{x}})}{\sigma_k} \right)^2 + \left( \frac{\lambda_\nu - 1}{\varepsilon_\nu} \right)^2 \right] \rightarrow \text{MiniBooNE } \nu \\
 & + \left[ \sum_{l=0}^M \left( \frac{\nu_l^{\text{DATA}} - \lambda_{\bar{\nu}}^{-1} \nu_l^{MC}(\vec{\mathbf{x}})}{\sigma_l} \right)^2 + \left( \frac{\lambda_{\bar{\nu}} - 1}{\varepsilon_{\bar{\nu}}} \right)^2 \right] \rightarrow \text{MiniBooNE } \bar{\nu} \\
 & + \left[ \sum_{i=0}^{16} \sum_{j=0}^{16} (\nu_i^{\text{DATA}} - \nu_i^{MC}(\vec{\mathbf{x}})) V_{ij}^{-1} (\nu_j^{\text{DATA}} - \nu_j^{MC}(\vec{\mathbf{x}})) \right] \rightarrow \text{MINER}\nu\text{A}
 \end{aligned} \quad (3)$$

where  $\vec{\mathbf{x}}$  are the model parameters varied in the fit,  $\sigma_k$  and  $\sigma_l$  are the diagonals of the MiniBooNE shape-only covariance matrices for the neutrino and antineutrino results,  $V_{ij}$  is the cross-covariance matrix provided by MINER $\nu$ A, and  $\lambda_\nu$  and  $\lambda_{\bar{\nu}}$  are the normalization parameters for MiniBooNE neutrino and antineutrino datasets, with published normalization uncertainties of  $\varepsilon_\nu$  (10.7%) and  $\varepsilon_{\bar{\nu}}$  (13.0%).<sup>4</sup>

<sup>4</sup>Note that the MINER $\nu$ A normalization uncertainty is included in the covariance matrix, so it also contributes a penalty term to the fit.

for the MINER $\nu$ A, MiniBooNE single-differential and MiniBooNE double-differential samples, respectively. The Nieves 2p2h contribution is also shown on these plots for reference.

Because of the way the data are presented, it is necessary to fit the MiniBooNE normalization parameters [see Eq. (3)] to produce a meaningful  $\chi^2$  value for the MiniBooNE datasets. It is not necessary to explicitly fit any parameters for MINER $\nu$ A as the normalization uncertainty is included in the covariance matrix. The fitted values of the normalization parameters  $\lambda_\nu^{\text{MB}}$  and  $\lambda_{\bar{\nu}}^{\text{MB}}$  are given in Table III. The single- and double-differential plots shown in Figs. 7 and 8 are scaled according to the normalization factors in Table III. For reference, the nominal predictions for the MiniBooNE double-differential datasets, without the scaling factor applied, are shown in Fig. 9 and are easier to interpret by eye.

Note that the double-differential cross section plots shown in Fig. 9 have been rebinned. In the distributions released by MiniBooNE, and used in the fits, there are 20  $\cos \theta_\mu$  bins uniformly distributed between  $-1$  and  $1$ . For ease of presentation, these have been rebinned, and the results are shown in eight  $\cos \theta_\mu$  slices of varying sizes, where merged bins have been averaged and their errors combined in quadrature.

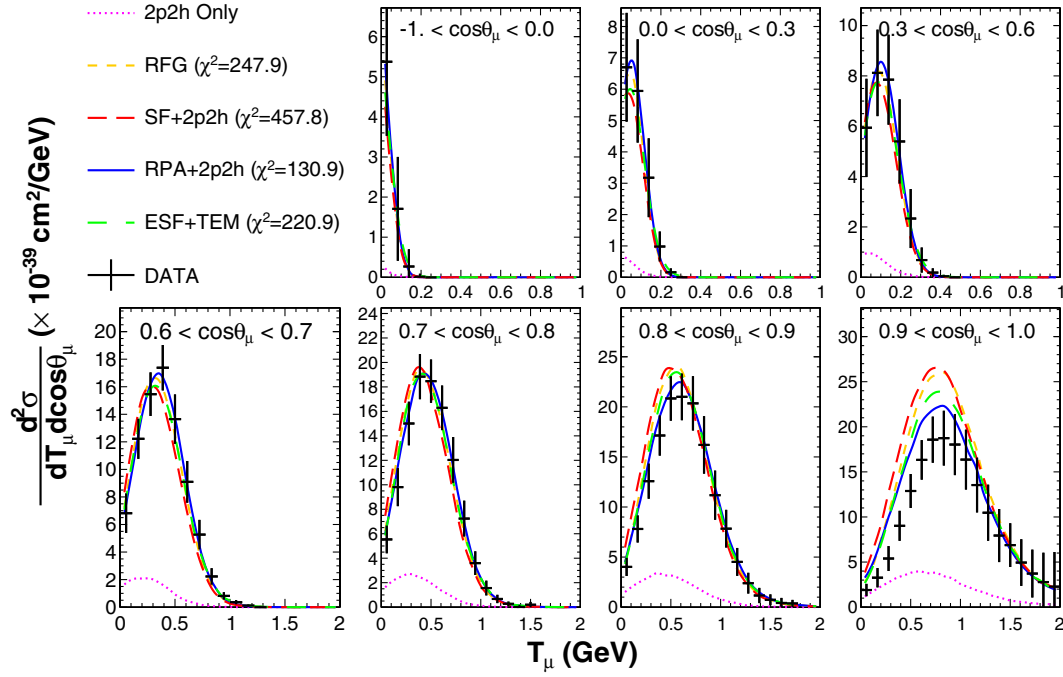
## VI. FIT PROCEDURE

All minimizations are performed using the MIGRAD algorithm of the MINUIT package [67], using the  $\chi^2$  statistic:

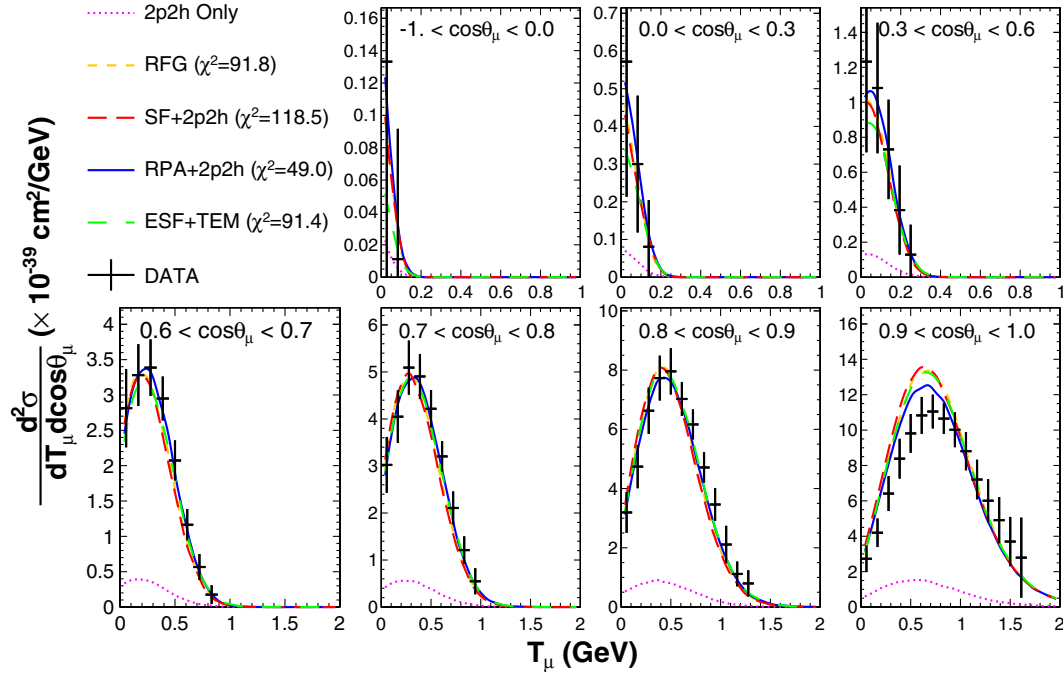
Fits to individual datasets only include the relevant terms from the  $\chi^2$  definition in Eq. (3), and fits to single MINER $\nu$ A datasets neglect cross-correlations (the summation is only over the relevant eight bins).

### A. Parameter goodness-of-fit (PGoF) test

Standard goodness-of-fit tests, such as the Pearson  $\chi^2_{\text{min}}$  test used as an example here, test the agreement between prediction and data; however, some issues can arise with their use in global fits, as discussed in Ref. [68]. The basic problem is that much of the data will have limited power to



(a) Neutrino



(b) Antineutrino

FIG. 8. Nominal model predictions for the MiniBooNE double-differential datasets with  $M_A = 1.01 \text{ GeV}/c^2$  and all other model parameters at their default values. The relativistic RPA calculation is shown. Normalization parameters are applied as given in Table III.

constrain any one parameter, but they agree well with the prediction regardless of the parameter values. These data will add little to the  $\chi^2$  but contribute another degree of freedom. Thus, the  $\chi^2_{\min}$  found may be deceptively good despite not agreeing well with those parts of the dataset that

actually have power to constrain key parameters. It is also possible that a dataset with a large number of data points (such as MiniBooNE) that does agree well with a model may hide disagreements with other datasets included in a global fit for which fewer data points are available

TABLE III. Table of best-fit MiniBooNE normalization parameter values for the nominal model comparisons shown in Figs. 7 and 8. The relativistic RPA calculation is shown.

Fit type		$\lambda_{\nu}^{\text{MB}}$	$\lambda_{\bar{\nu}}^{\text{MB}}$
Neutrino 1D	RFG	$0.732 \pm 0.007$	...
	SF + 2p2h	$0.741 \pm 0.007$	...
	RPA + 2p2h	$0.760 \pm 0.007$	...
	ESF + TEM	$0.804 \pm 0.008$	...
Antineutrino 1D	RFG	...	$0.805 \pm 0.011$
	SF + 2p2h	...	$0.826 \pm 0.011$
	RPA + 2p2h	...	$0.774 \pm 0.010$
	ESF + TEM	...	$0.803 \pm 0.011$
Neutrino 2D	RFG	$0.725 \pm 0.011$	...
	SF + 2p2h	$0.756 \pm 0.011$	...
	RPA + 2p2h	$0.760 \pm 0.011$	...
	ESF + TEM	$0.827 \pm 0.012$	...
Antineutrino 2D	RFG	...	$0.808 \pm 0.015$
	SF + 2p2h	...	$0.838 \pm 0.015$
	RPA + 2p2h	...	$0.802 \pm 0.015$
	ESF + TEM	...	$0.833 \pm 0.015$

(such as MINER $\nu$ A); again, the key problem is a dilution of the  $\chi^2$ .

This problem is worsened in the case of datasets for which correlations between data points have not been included, where the  $\chi^2/\text{DOF}$  can be much less than 1, such as is the case for MiniBooNE. Looking at the Pearson  $\chi^2_{\text{min}}$  test statistic is not very illuminating when fitting to both MiniBooNE and MINER $\nu$ A datasets.

The PGoF is a more rigorous test proposed in Ref. [68] for fitting to global datasets and has been used extensively in sterile neutrino literature [69,70], where there are often contradictory results coming from different experiments, and the fitters are fitting to many different experiments which are sensitive to different parameters. It is also referred to as the likelihood ratio test in both statistics and other high energy physics literature. The PGoF test statistic is given by

$$\chi^2_{\text{PGoF}}(\vec{\mathbf{x}}) = \chi^2_{\text{tot}}(\vec{\mathbf{x}}) - \sum_{r=1}^D \chi^2_{r,\text{min}}(\vec{\mathbf{x}}), \quad (4)$$

where  $\vec{\mathbf{x}}$  are the parameters floated in the fits,  $D$  is the number of datasets,  $\chi^2_{\text{tot}}$  is the minimum  $\chi^2$  in a fit to all  $D$  subsets of the data, and  $\chi^2_{r,\text{min}}$  is the minimum  $\chi^2$  obtained in a fit to the  $r$ th subset of the data. The PGoF test statistic forms a  $\chi^2$  distribution with the number of degrees of freedom,

$$P_{\text{PGoF}} = \sum_{r=1}^D P_r - P_{\text{tot}},$$

where  $P_r$  and  $P_{\text{tot}}$  are the number of degrees of freedom for each fit.

The aim of the PGoF is to test the compatibility of the different datasets in the framework of the model. Put simply, it tests whether the best-fit parameter values to subsets of the data pull the fit parameters far from the best-fit values found when fitting to all of the data. If different subsets favor very different values, then those subsets are not compatible in the framework of the model (though individually each may be able to find parameter combinations that produce a good fit).

A further advantage of the PGoF test in the situation where some of the data lack correlations is that the number of degrees of freedom come from the number parameters varied in the fits, not from the number of bins that the dataset contributes, which mitigates against the  $\chi^2_{\text{min}}/N_{\text{DOF}} \ll 1$  issue.

The PGoF test still assumes that the datasets follow a  $\chi^2$  distribution, but it allows for a lower effective number of degrees of freedom. This assumption is not quantitatively correct due to the aforementioned lack of correlations in the MiniBooNE CCQE data. The  $p$  values returned should be taken with the caution that they highlight tensions between datasets but are not to be interpreted in the same manner as they would if all correlations were reported.

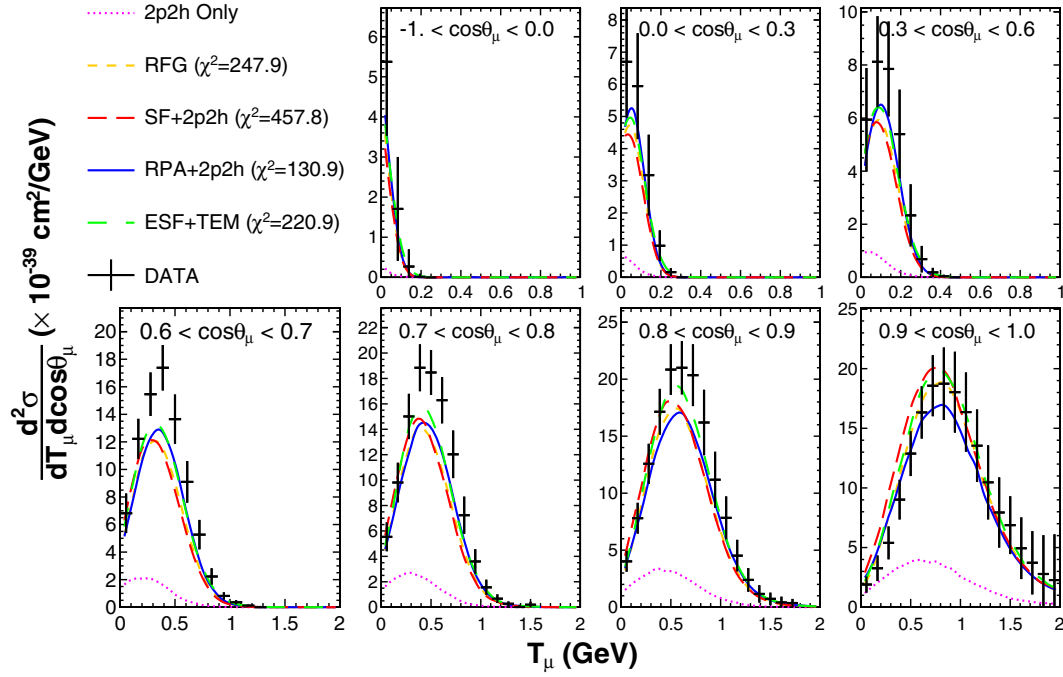
## VII. FIT RESULTS

### A. Fake data studies

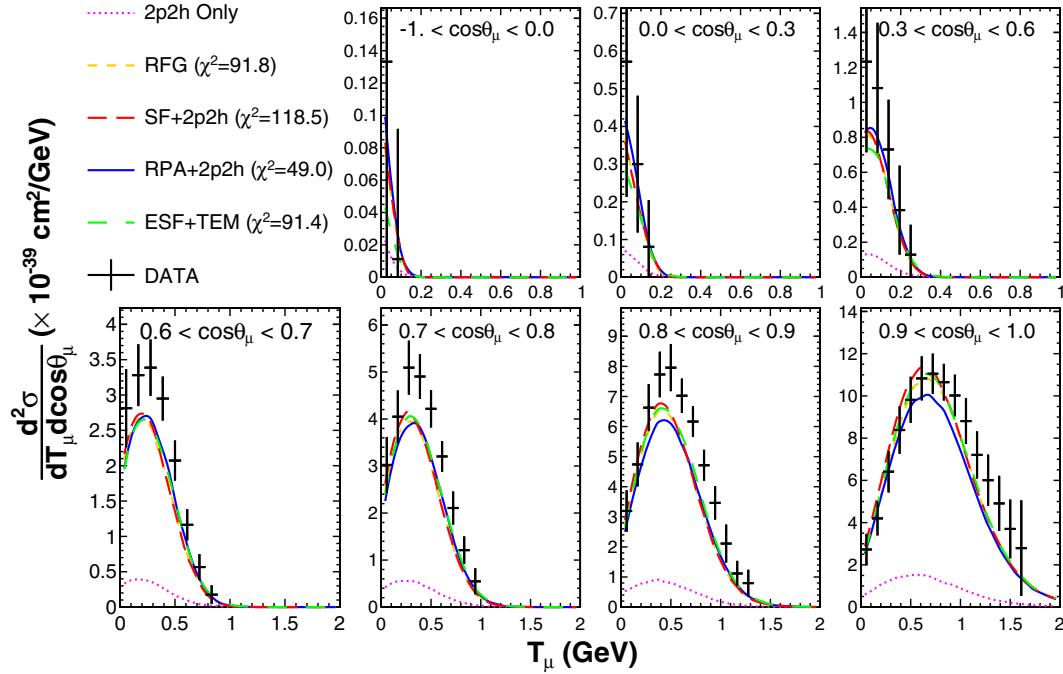
The fitter was validated in two ways. First, Asimov fake datasets [71] were produced to estimate the size of the errors that would be produced from the fit and used as a sanity check of the real fit results. The Asimov tests also provide a very basic test of the fitting framework developed for this analysis. Second, pull studies were performed to check that the  $\chi^2$  definition given in Eq. (3) is an unbiased estimator of the parameter central values and errors. For all parameters, the biases were less than 10% across the entire parameter range allowed in the fit, so we conclude that the fitter behaves well.

### B. Combined fit

The results for the combined fits to all four datasets are given for both relativistic and nonrelativistic RFG + RPA + 2p2h models and the SF + 2p2h model in Table IV. The best-fit distributions are compared with data for MINER $\nu$ A in Fig. 10, and for MiniBooNE in Fig. 11. Relativistic RPA is used in the figures, as this was the best fit of the two RPA models available. In the legends of these figures, each line is given two  $\chi^2$  values, the contribution from that dataset to the  $\chi^2_{\text{min}}$  in the combined fit and the total  $\chi^2_{\text{min}}$  in parentheses. Note that in Fig. 10, the contributions from MINER $\nu$ A are calculated for the individual datasets, which necessarily ignores cross-correlations and makes these numbers slightly misleading. Explicitly,  $\chi^2_{\text{MLA total}} \neq \chi^2_{\text{MLA}\nu} + \chi^2_{\text{MLA}\bar{\nu}}$  due to cross-correlations, so the values shown in the figure should be treated with caution.



(a) Neutrino



(b) Antineutrino

FIG. 9. Nominal model predictions for the MiniBooNE double-differential datasets with  $M_A = 1.01 \text{ GeV}/c^2$  and all other model parameters at their default values. Note that for each model, the relevant MiniBooNE normalization parameter has been allowed to vary to minimize the  $\chi^2$  value; however, the scaling factors (given in Table III) have not been applied in this figure. The relativistic RPA calculation is shown.

It is clear from Figs. 10 and 11 that MiniBooNE is not completely dominating the fits, as might be expected given the large number of bins in each of the MiniBooNE datasets. Indeed, these fits exploit the fact that, without

correlations,  $\chi^2_{\text{MB}} \approx \chi^2_{\text{M}\nu\text{A}}$ . It is also clear that neither model fits all of the datasets perfectly at the best-fit point, which is not reflected by the reduced  $\chi^2$  values of 97.5/228 and 97.8/228 for the SF + 2p2h and RFG + rel RPA + 2p2h



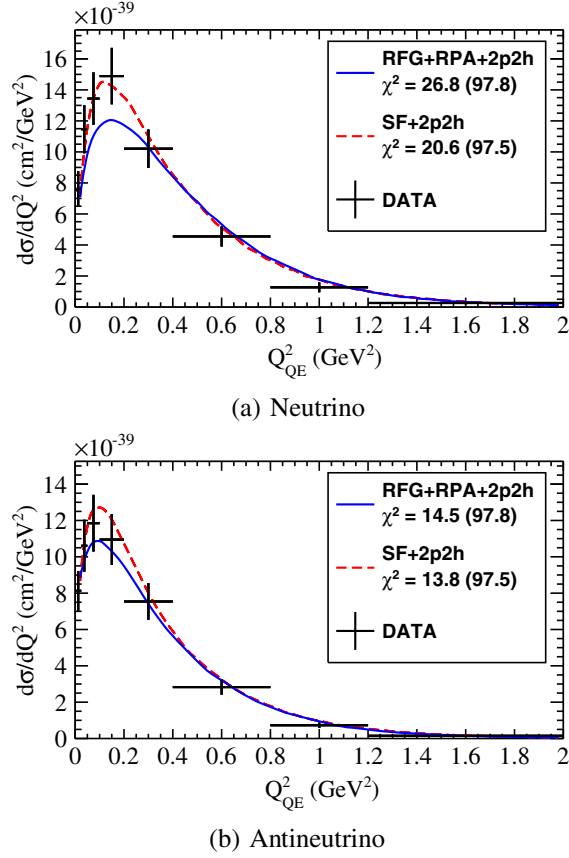


FIG. 10. Comparison of the best fit from the combined fits detailed in Table IV with the MINERνA datasets used in the fit. The  $\chi^2$  values in the legend are the contribution from each dataset at the best-fit point (and the total  $\chi^2_{\min}$  for the combined fit).

models, respectively. As the MiniBooNE public data release lacks bin-to-bin correlations, the  $\chi^2_{\text{MB}}$  contributions are not as large as would be expected for the number of bins contributed.

In all fits, it was observed that the MiniBooNE normalization values tended to be suppressed for both neutrino and antineutrino datasets indicating that the MC underestimated the published data by 20%–30% (10%–20%) for the RFG + RPA + 2p2h (SF + 2p2h) models. It is not possible to accurately determine the favored MINERνA normalization as the normalization uncertainty is included in the published covariance matrix, but the output distributions show that the MC normalization is approximately equal to the data normalization.

Because of the large pulls on the MiniBooNE normalization parameters, shape-only fits were also performed (see Ref. [60] for further details). It was found that the best-fit parameters were not significantly changed, indicating that there is not a significant bias to the other parameters caused by the large pulls on the MiniBooNE normalization parameters. A recent reanalysis of the MINERνA flux [72] results in an increase in the normalization of previous

MINERνA cross section results including the CCQE samples used in this analysis. Although these updated datasets are not included in this work, we note that as the results were found to be largely unchanged in a shape-only fit, the main results will not be significantly affected. Additionally, results from fits to individual datasets, and to various combinations of datasets, can be found in Ref. [60].

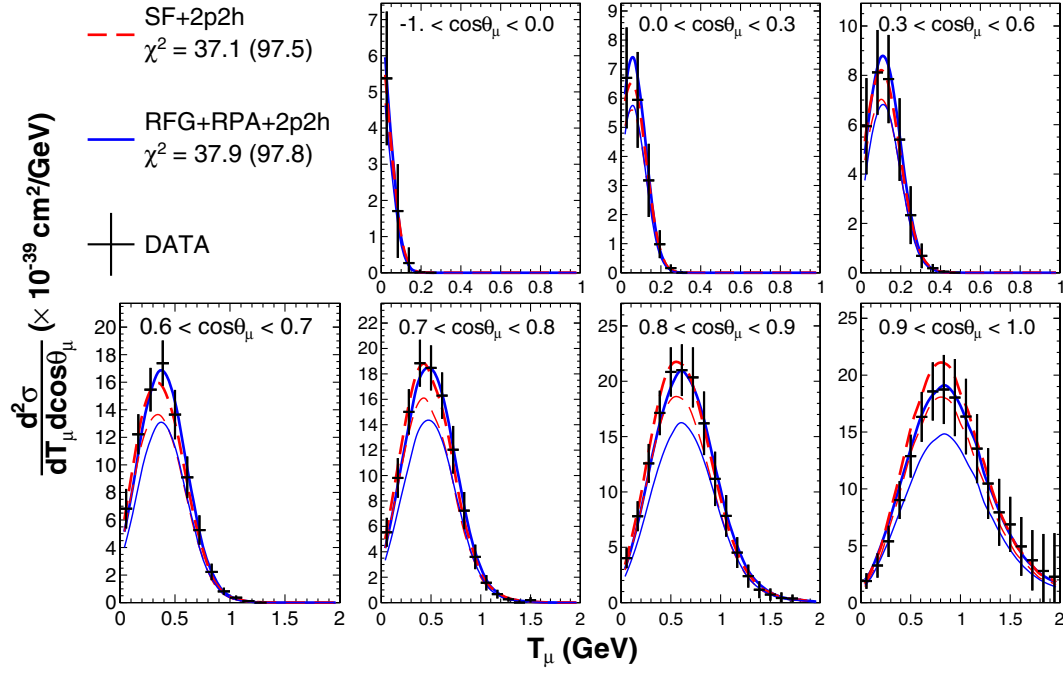
### C. PGoF results

Using the PGoF test defined in Sec. VI A, it is possible to test the compatibility between different subsets of the data. Tables VI–VIII show a breakdown of the four datasets used in the combined fits for each initial CCQE model assumption. The standard goodness of fit (SGoF) for each row is determined using Pearson's  $\chi^2_{\min}$  test, where  $\chi^2_{\min}$  is found by minimizing the function given in Eq. (3), including only the terms for the relevant datasets. The PGoF test is found by subtracting  $\chi^2_{\min}$  for each of the constituent datasets from the minimum of the combined dataset. The formulas for calculating the PGoF test statistic  $\chi^2_{\text{PGoF}}$  are given explicitly in Table V. The  $\chi^2_{\min}$  for each dataset is again determined by minimizing the function given in Eq. (3) with only the relevant terms included.

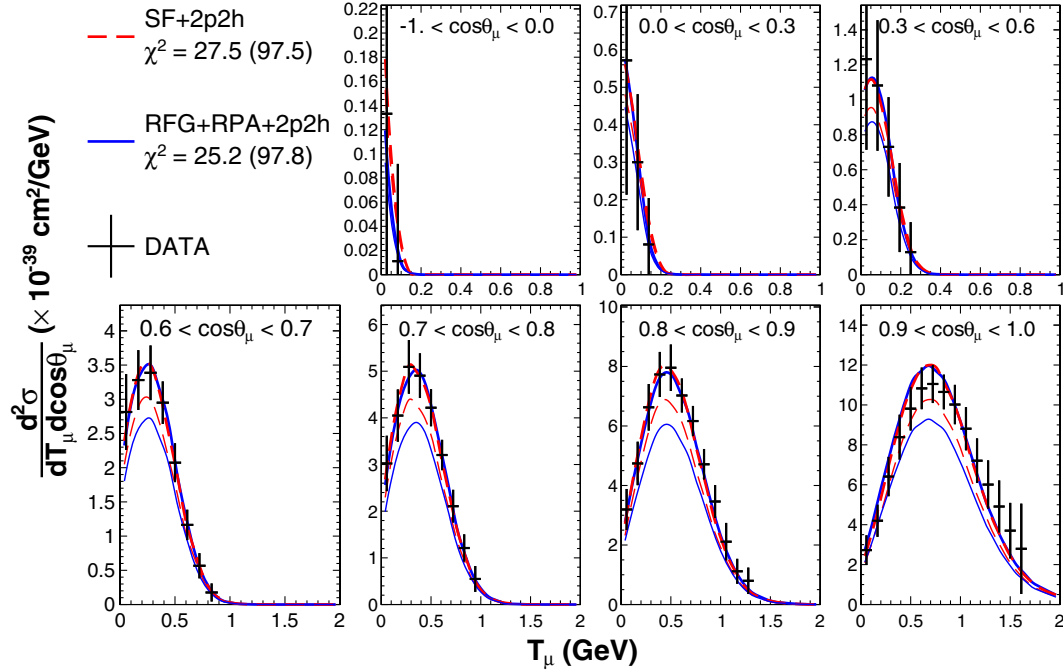
In each fit, the  $M_A$ , 2p2h normalization,  $p_F$ , and any MiniBooNE normalization terms are allowed to float.

One subtlety must be kept in mind when analyzing the results in Tables VI–VIII: the PGoF test is only appropriate for statistically independent datasets. This makes the interpretation difficult for MINERνA, where cross-correlations are provided and used in the fits. Whenever a subset of data includes both MINERνA  $\nu$  and  $\bar{\nu}$  datasets, the fits include cross-correlations, but if only one dataset is included, they do not. This means that two of the rows in each table give slightly unreliable results: “MINERνA” and “ $\nu$  vs  $\bar{\nu}$ .” In each case, the  $\chi^2$  function for the combined dataset includes cross-correlations, and the  $\chi^2$  functions for the subdivided dataset do not. The issue is most obvious in Table VII, where the “ $\nu$  vs  $\bar{\nu}$ ” row gives a negative PGoF  $\chi^2$ . These values are still useful as a comparison between models and to give a rough idea of compatibility between datasets, but the exact values must be treated with caution.

The PGoF test highlights the incompatibility of the various datasets within the framework of the SF + 2p2h and both RFG + RPA + 2p2h models, despite the apparent goodness of fit when only considering  $\chi^2_{\min}/N_{\text{DOF}}$ . The level of agreement given in the final column of Tables VI–VIII should be interpreted as the level of agreement between the datasets included in that row. For example, it is clear that for all models considered, the agreement found between the MINERνA and MiniBooNE datasets (which include both neutrino and antineutrino samples) have the lowest level of agreement as shown by



(a) Neutrino



(b) Antineutrino

FIG. 11. Comparison of the best fit from the combined fits detailed in Table IV with the MiniBooNE double-differential datasets used in the fit. The  $\chi^2$  values in the legend are the contribution from each dataset at the best-fit point (and the total  $\chi^2_{\min}$  for the combined fit). The thick lines have the MiniBooNE normalization factors applied (given in Table IV), while the thin lines do not, to indicate the large pulls on these parameters.

the “M $\nu$ A vs MB” row. In contrast, the level of agreement between the neutrino and antineutrino datasets (which include the MINER $\nu$ A and MiniBooNE samples) show relatively good agreement, indicating that fitting to the

neutrino and antineutrino datasets separately produces similar best-fit parameter values.

It is clear from Table VIII that the SF + 2p2h model does not fit the various datasets well; the poor PGoF statistics

TABLE IV. Best-fit parameter values for the fits to all datasets simultaneously.

Fit type	$\chi^2/N_{\text{DOF}}$	$M_A$ (GeV/ $c^2$ )	2p2h norm (%)	$p_F$ (MeV/ $c$ )	$\lambda_\nu^{\text{MB}}$	$\lambda_{\bar{\nu}}^{\text{MB}}$
RFG + rel RPA + 2p2h	97.8/228	$1.15 \pm 0.03$	$27 \pm 12$	$223 \pm 5$	$0.79 \pm 0.03$	$0.78 \pm 0.03$
RFG + nonrel RPA + 2p2h	117.9/228	$1.07 \pm 0.03$	$34 \pm 12$	$225 \pm 5$	$0.80 \pm 0.04$	$0.75 \pm 0.03$
SF + 2p2h	97.5/228	$1.33 \pm 0.02$	0 (at limit)	$234 \pm 4$	$0.81 \pm 0.02$	$0.86 \pm 0.02$

indicate that the datasets favor very different parameter values when fit separately. This is particularly true for any fits involving the MiniBooNE neutrino dataset, though there is no *a priori* reason to exclude this dataset and improve the fit results. The PGoF tests for RFG + RPA + 2p2h using both relativistic and nonrelativistic RPA, shown in Tables VI and VII, show much better compatibility between experiments than SF + 2p2h. There is still a considerable amount of tension, which is largely due to differences between MINER $\nu$ A and MiniBooNE. Because of the relatively poor consistency between datasets for the SF + 2p2h model compared with RFG + rel RPA + 2p2h, the latter model is a better choice as the default model for T2K oscillation analyses.

#### D. Rescaling parameter errors

Assuming Gaussian statistics,  $1\sigma$  errors on a single fit parameter are defined by the parameter value for which

TABLE V. Explicit formulas for calculating the  $\chi^2_{\text{PGoF}}$  test statistics for each of the subsets of the data investigated. Each  $\chi^2$  value listed in this table denotes the  $\chi^2$  at the minimum.

	$\chi^2_{\text{PGoF}}$
All	$\chi^2_{\text{ALL}} - \chi^2_{\text{MB}\nu} - \chi^2_{\text{MB}\bar{\nu}} - \chi^2_{[\text{M}\nu\text{A}\nu+\bar{\nu}]}$
MINER $\nu$ A	$\chi^2_{[\text{M}\nu\text{A}\nu+\bar{\nu}]} - \chi^2_{\text{M}\nu\text{A}\nu} - \chi^2_{\text{M}\nu\text{A}\bar{\nu}}$
MiniBooNE	$\chi^2_{[\text{MB}\nu+\bar{\nu}]} - \chi^2_{\text{MB}\nu} - \chi^2_{\text{MB}\bar{\nu}}$
$\nu$	$\chi^2_{[\text{MB}\nu+\text{M}\nu\text{A}\nu]} - \chi^2_{\text{MB}\nu} - \chi^2_{\text{M}\nu\text{A}\nu}$
$\bar{\nu}$	$\chi^2_{[\text{MB}\bar{\nu}+\text{M}\nu\text{A}\bar{\nu}]} - \chi^2_{\text{MB}\bar{\nu}} - \chi^2_{\text{M}\nu\text{A}\bar{\nu}}$
M $\nu$ A vs MB	$\chi^2_{\text{ALL}} - \chi^2_{[\text{MB}\nu+\text{MB}\bar{\nu}]} - \chi^2_{[\text{M}\nu\text{A}\nu+\bar{\nu}]}$
$\nu$ vs $\bar{\nu}$	$\chi^2_{\text{ALL}} - \chi^2_{[\text{MB}\nu+\text{M}\nu\text{A}\nu]} - \chi^2_{[\text{M}\nu\text{A}\bar{\nu}+\text{MB}\bar{\nu}]}$

TABLE VI. PGoF results for various subsets of the data for the RFG + nonrel RPA + 2p2h model.

	$\chi^2_{\text{min}}/N_{\text{DOF}}$	SGoF (%)	$\chi^2_{\text{PGoF}}/N_{\text{DOF}}$	PGoF (%)
All	117.9/228	100.00	25.3/6	0.03
MINER $\nu$ A	30.3/13	0.42	0.4/3	93.09
MiniBooNE	65.7/212	100.00	3.4/3	33.09
$\nu$	69.1/142	100.00	12.7/3	0.53
$\bar{\nu}$	46.1/83	99.97	10.4/3	1.55
M $\nu$ A vs MB	117.9/228	100.00	21.9/3	0.01
$\nu$ vs $\bar{\nu}$	117.9/228	100.00	2.6/3	45.12

$\chi^2 = \chi^2_{\text{min}} + 1$  [73]. MINUIT uses this assumption when calculating the errors at the minimum, which were included with the best-fit values for the combined fit in Table IV. However, as well as motivating the use of the PGoF test, the lack of bin correlations from MiniBooNE also means that Gaussian statistics no longer work as expected when estimating parameter errors.

There is a large body of literature looking at how this problem affects fits to parton density distributions, where global fits include a large number of datasets, many of which did not provide bin correlations [74–76]. A summary of the work of one PDF fitting group is given in Ref. [74] and was used as a guide here. Their solution for producing reasonable parameter error estimates is to inflate the value of the  $\Delta\chi^2$  used to define the  $1\sigma$  parameter errors, although no generic solution is offered for defining that value. In the case of the PDF fits in Ref. [74], the  $\Delta\chi^2$  used was very large,  $\sim 100$ , although it should be kept in mind that many more datasets are used in that fit than in the current work.

The PGoF gives a value for the incompatibility between the datasets: how much the  $\chi^2$  increases between the best-fit

TABLE VII. PGoF results for various subsets of the data for the RFG + rel RPA + 2p2h model.

	$\chi^2_{\text{min}}/N_{\text{DOF}}$	SGoF (%)	$\chi^2_{\text{PGoF}}/N_{\text{DOF}}$	PGoF (%)
All	97.8/228	100.00	17.9/6	0.66
MINER $\nu$ A	23.4/13	3.74	1.0/3	79.03
MiniBooNE	58.6/212	100.00	2.0/3	57.69
$\nu$	62.6/142	100.00	16.1/3	0.11
$\bar{\nu}$	38.5/83	100.00	6.1/3	10.75
M $\nu$ A vs MB	97.8/228	100.00	15.9/3	0.12
$\nu$ vs $\bar{\nu}$	97.8/228	100.00	−3.3/3	100.00

TABLE VIII. PGoF results for various subsets of the data for the SF + 2p2h model.

	$\chi^2_{\text{min}}/N_{\text{DOF}}$	SGoF (%)	$\chi^2_{\text{PGoF}}/N_{\text{DOF}}$	PGoF (%)
All	97.5/228	100.00	41.1/6	0.00
MINER $\nu$ A	12.6/13	47.75	1.0/3	79.49
MiniBooNE	50.2/212	100.00	6.5/3	8.92
$\nu$	54.8/142	100.00	25.1/3	0.00
$\bar{\nu}$	34.1/83	100.00	8.5/3	3.61
M $\nu$ A vs MB	97.5/228	100.00	34.6/3	0.00
$\nu$ vs $\bar{\nu}$	97.5/228	100.00	8.5/3	3.59

TABLE IX. The final errors for the RFG + relRPA + 2p2h parameters. Note that the scaled errors should be used by any analyses which use these results.

Fit type	$\chi^2/N_{\text{DOF}}$	$M_A$ (GeV/ $c^2$ )	2p2h (%)	$p_F$ (MeV/ $c$ )
Unscaled	97.8/228	$1.15 \pm 0.03$	$27 \pm 12$	$223 \pm 5$
PGoF scaling		$1.15 \pm 0.06$	$27 \pm 27$	$223 \pm 11$

points of each experiment and the best-fit point for the combined dataset. The PGoF value can therefore be used as a measure of how much the errors have to be inflated to cover the difference between the best-fit parameter values from the combined fit and the best-fit values of individual datasets; this is shown explicitly in Eq. (5), where the value used to define the  $1\sigma$  error is given by  $\Delta\chi^2$ , and the rescaling parameter is given by  $r$ .

$$\Delta\chi^2 = \chi_{\text{PGoF}}^2/N_{\text{DOF}},$$

$$r = \sqrt{\chi_{\text{PGoF}}^2/N_{\text{DOF}}}. \quad (5)$$

Note that this PGoF rescaling procedure does not modify the correlations between parameters; it simply rescales the error on each parameter.

There is some ambiguity over which PGoF statistic to use, the “All” or “ $M_A$  vs MB” row of Table VII, with  $\chi_{\text{PGoF}}^2/N_{\text{DOF}}$  values of 17.9/6 and 15.9/3, respectively. The more conservative value is from the “ $M_A$  vs MB” (because the greatest differences are between experiments, not between neutrino and antineutrino running), so this is used.<sup>5</sup> To be explicit, we multiply the parameter errors from MINUIT by  $r = \sqrt{15.9/3} \approx 2.3$  based on this statistic, as shown in Table IX. It can be seen from Table IX that the 2p2h normalization is strongly suppressed and, even with the rescaled error, is nearly  $3\sigma$  away from the Nieves nominal model prediction. It is also clear that although the axial mass value preferred in the fit is not as strongly inflated as in fits to MiniBooNE data alone [10,11,61], it is still significantly higher than the value of  $M_A \approx 1$  GeV/ $c^2$  found by fitting to light target data and pion electroproduction data [53], and the inflated  $1\sigma$  error does not cover this difference.

## VIII. DISCUSSION OF THE FIT RESULTS

The results from the fits presented in Sec. VII B show that none of the models that are currently available in NEUT describe all of the CCQE data adequately. In particular, there is a significant difference between MiniBooNE and MINER $\nu$ A data that forces the model parameters to compromise between the two, as well as a

<sup>5</sup>It should also be noted that this rescaling procedure more than covers the difference between neutrino and antineutrino datasets.

large change in the normalization for the MiniBooNE datasets. Although the  $M_A$  value obtained from the fit to the RFG + relRPA + 2p2h model is lower than that obtained from past fits of the RFG model to MiniBooNE data alone (see Ref. [50] as an example), it is still inconsistent with that obtained in global fits to light target bubble chamber data or high energy heavy target data [53]. Additionally, the data require a large suppression of the nominal 2p2h model. The SF + 2p2h model, in which the 2p2h component is completely suppressed, requires an inflated  $M_A$  value to fit the data. This is unsurprising as the SF model alone does not significantly change the total cross section. Including an RPA calculation appropriate for the SF model is likely to reduce the tension with the 2p2h model and is likely to change this conclusion significantly; this work will be revisited when such a calculation is available. Both fits also initially imply that there may need to be additional interactions used that may mimic CCQE interactions or change the shape of the distributions through additional, but currently unmodeled, effects in the nucleus.

The expectation is for the additional interaction strength from 2p2h to remove the need for an inflated axial mass, implying that these parameters should be anticorrelated. However, at the best-fit point 2p2h is suppressed to 27% of the Nieves nominal value as shown in Table IX. This suppression is driven by MINER $\nu$ A, which would completely suppress the 2p2h component of the model if MiniBooNE were not included in the fit [60]. As discussed, the cross section is smaller in the low- $Q^2$  region than the model prediction. Therefore, the model parameters act to suppress the low- $Q^2$  region in the fit. The Nieves 2p2h model is concentrated in the low- $Q^2$  region as can be seen in Fig. 6. As a result, the 2p2h contribution is largely suppressed, and an inflation of the axial mass for the 1p1h contribution is favored. This leads to a positive correlation between the 2p2h normalization and  $M_A$  parameters as shown in Fig. 12. We see very similar results in a

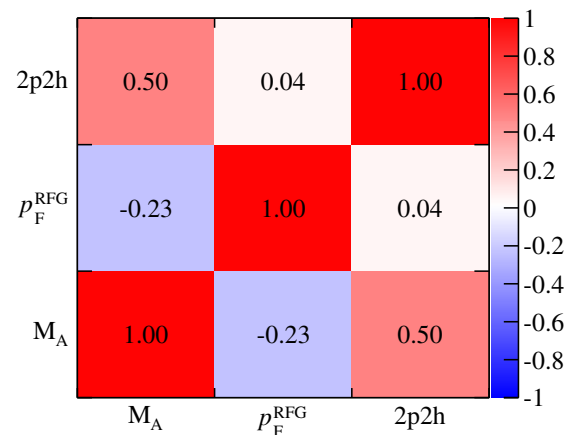


FIG. 12. Correlation matrix for the best-fit RFG + relRPA + 2p2h model parameters.



shape-only fit, indicating that the normalization disagreement that has previously led to an inflated axial mass is not responsible for the suppression of 2p2h normalization. For this reason, we are confident that the change in the normalization of the MINER $\nu$ A cross section results expected from the flux reanalysis [72] will not significantly modify the conclusions of the fit.

A further issue is that the CCQE-corrected cross section results from both MINER $\nu$ A and MiniBooNE have part of the 2p2h signal,  $\pi$ -less  $\Delta$  decay, removed as a background.<sup>6</sup> Because the current implementation of 2p2h in NEUT does not have the capability to separate out the contribution from  $\pi$ -less  $\Delta$  decay, we have simply ignored this potential issue, as has also been done previously by the 2p2h model builders [35,36]. Better agreement between the model and data might be found if it were possible to separate physics processes and to modify the shape of the 2p2h cross section in the fit. Future cross section measurements should be encouraged to focus on exclusive final states (CC0 $\pi$ ) rather than initial state processes (CCQE), which will avoid such an issue in the future. Finally, as previously remarked, the RFG + RPA + 2p2h model implemented in NEUT is not equivalent to the full Nieves model because the 1p1h component in NEUT uses a global, rather than local, Fermi gas nuclear model. The difference between the 1p1h models will be most pronounced in the low- $Q^2$  region, so it is possible that the 2p2h shape issue is due to a conflict with the 1p1h model and that a more consistent LFG + RPA + 2p2h model might alleviate this issue.

Although both the RFG + rel RPA + 2p2h model and SF + 2p2h model give reasonable agreement with data at the best-fit point, it is difficult to trust standard goodness-of-fit tests as the lack of MiniBooNE correlations means that Gaussian statistics no longer work correctly. An alternative measure of the goodness of fit, the PGoF, was used to try to improve the situation. Although the PGoF procedure still assumes Gaussian statistics, it highlights disagreements within the combined dataset by dividing the dataset into subsamples. These disagreements are completely hidden by the standard goodness-of-fit tests because the MiniBooNE  $\chi^2$  contribution is so low relative to the number of degrees of freedom it contributes to the fit. The PGoF shows that there is considerably better agreement between the best-fit parameter values obtained in fits to subsamples of the data for the RFG + rel RPA + 2p2h fit, which gives some confidence to the fit result. For the SF + 2p2h model, the fits to subsamples of the data pull to drastically different parameter values at the best-fit points, which is highly undesirable behavior if the fit results are to be used as prior uncertainties in oscillation analyses, and

this indicates that the model is a bad fit to the global dataset. But the SF + 2p2h model can fit individual datasets well (as is clear in Table VIII) so should not be discounted completely.

The lack of reported MiniBooNE correlations and non-Gaussian behavior of the test statistic also means that standard parameter error estimation does not work, and it returns smaller parameter errors than are reasonable given the level of disagreement between the datasets used in the fit. An unrealistically tight constraint on cross section parameters would lead to biases in the near detector fit for T2K. To circumvent this problem a PGoF error inflation procedure was defined to ensure that the  $1\sigma$  parameter errors cover the disagreement between the MINER $\nu$ A and MiniBooNE datasets. This is a conservative approach, but as no model seems able to describe all of the available data, such an approach was necessary. Such *ad hoc* procedures are necessary when incomplete information is available from some of the datasets included in the fit. The lack of information about bin-to-bin correlations for the MiniBooNE datasets significantly complicates this analysis and may significantly change the results. We note that in the literature, many statements about how well various models agree with the MiniBooNE datasets are made that assume Gaussian statistics. It is clear that an appreciation of this issue is important for future model comparisons and that the availability of complete information for new cross section results will be critical for building consistent CCQE models.

For T2K, the results of this fit are part of a larger set of cross section model systematic uncertainties recommended by the NIWG that can be used as prior inputs for the oscillation analyses and various cross section analyses. In this case, the model used for these analyses is the RFG + rel RPA + 2p2h model since the SF + 2p2h model is disfavored in the fits and relativistic RPA is preferred over nonrelativistic RPA. The best-fit parameters and uncertainties of the model are given in the second row of Table IX and are correlated according to the matrix shown in Fig. 12.

## IX. SUMMARY

In this paper, we have shown how T2K's NIWG uses previously published CCQE datasets from the MiniBooNE and MINER $\nu$ A experiments to test CCQE+2p2h models in the NEUT neutrino interaction generator. For each model, the parameters that describe the data are fit, with both the SGoF and PGoF used to select the model that best describes the data. In this case, the RFG + rel RPA + 2p2h model is considered the best candidate, with  $M_A = 1.15 \pm 0.03$  GeV/ $c^2$ , the normalization on the 2p2h model  $27 \pm 12\%$ , and  $p_F = 223 \pm 5$  MeV/ $c$ . Tensions between the two experiments require an error scaling procedure outlined by the PGoF test, with the final result providing prior inputs into various future T2K analyses. This is the first time a comprehensive analysis has been performed and published with these models using

<sup>6</sup>MINER $\nu$ A floats the normalizations of various backgrounds including pion absorption to fit a sideband sample, which may implicitly cut out  $\pi$ -less  $\Delta$  decay events. MiniBooNE explicitly subtracts a  $\pi$ -less  $\Delta$  decay as described in Sec. IV.

a neutrino interaction generator and the first time that such models have been used in an oscillation analysis with full detector simulations [77].

Moving away from the RFG model for CCQE interactions is an ambitious step for a neutrino experiment as it is a departure from the standard that has been used for decades [3]. The new models on the market are not perfect, and their implementation into NEUT and other neutrino interaction generators will always have technical foibles. However, further theoretical development of these models requires the engagement of the experimental community, and so using them in our simulations is essential to move the field onwards. It is also clear that the current approach of inflating  $M_A$  is inadequate, and something better must be done in order to make precision measurements of neutrino oscillation parameters.

The fitting framework developed by the NIWG for this analysis is extensible, and the general method for producing cross section errors developed in this work will be used with new CCQE models and datasets in the future and with new cross section channels entirely, to continue to constrain systematic errors for T2K oscillation and cross section analyses. The results from the CCQE fits presented here will also help inform the future model development required to fit the data. It is clear that alternative 2p2h models and fundamental parameters in the 2p2h model should be investigated to see whether the disagreement with the 2p2h shape is telling us something meaningful about the Nieves model. It is also probable that the current RPA model is too inflexible, and this is partially responsible for the disagreement between MiniBooNE and MINER $\nu$ A data. Both of these problems may relate to the fact that for several years, the only neutrino data available for theorists to confront their models with was the MiniBooNE neutrino dataset, which is difficult to use

due to the lack of correlations and the explicit subtraction of  $\pi$ -less  $\Delta$  decay from the CCQE-corrected result. Converging on a new CCQE model that adequately describes all current and future data is likely to require several iterations between experimentalists and model builders. Confronting all the available models with a variety of data, as has been done in this analysis, and including these models in full Monte Carlo simulations, as will be done in T2K with the output from this analysis, is an important step in this cycle from the experimental side.

## ACKNOWLEDGMENTS

The authors would like to thank the members of the T2K Collaboration and the authors of the NuWro generator for their help and support in this analysis. We thank the MINER $\nu$ A and MiniBooNE collaborations for assistance in understanding their results and studies beyond the published work. We are grateful to J. Nieves and his collaborators for providing the code required to implement their 2p2h and RPA models in NEUT. We are grateful to T. Katori and P. Litchfield for helpful comments on the manuscript. We thank the J-PARC staff for superb accelerator performance and the CERN NA61/SHINE Collaboration for providing valuable particle production data. We acknowledge the support of MEXT, Japan; NSERC, NRC and CFI, Canada; National Science Centre (NCN), Poland; MINECO and ERDF funds, Spain; SNSF and SER, Switzerland; STFC, UK; and DOE, USA. We also thank CERN for the UA1/NOMAD magnet, DESY for the HERA-B magnet mover system, NII for SINET4, the WestGrid and SciNet consortia in Compute Canada, GridPP, UK, and the Emerald High Performance Computing facility in the Centre for Innovation, UK.

- 
- [1] C. H. Llewellyn Smith, *Phys. Rep.* **3**, 261 (1972).
  - [2] E. J. Moniz and R. A. Smith, *Nucl. Phys.* **B101**, 605 (1972).
  - [3] H. Gallagher, G. Garvey, and G. P. Zeller, *Annu. Rev. Nucl. Part. Sci.* **61**, 355 (2011).
  - [4] E. J. Moniz, I. Sick, R. R. Whitney, J. R. Ficenec, R. D. Kephart, and W. P. Trower, *Phys. Rev. Lett.* **26**, 445 (1971).
  - [5] H. S. Budd, A. Bodek, and J. Arrington, *arXiv:hep-ex/0308005*.
  - [6] V. Bernard, L. Elouadrhiri, and U.-G. Meißner, *J. Phys. G* **28**, R1 (2002).
  - [7] K. S. Kuzmin, V. V. Lyubushkin, and V. A. Naumov, *Acta Phys. Polon. B* **37**, 2337 (2006).
  - [8] K. S. Kuzmin, V. V. Lyubushkin, and V. A. Naumov, *Eur. Phys. J. C* **54**, 517 (2008).
  - [9] V. Lyubushkin *et al.*, *Eur. Phys. J. C* **63**, 355 (2009).
  - [10] A. Aguilar-Arevalo *et al.* (MiniBooNE Collaboration), *Phys. Rev. D* **81**, 092005 (2010).
  - [11] A. Aguilar-Arevalo *et al.* (MiniBooNE Collaboration), *Phys. Rev. D* **88**, 032001 (2013).
  - [12] R. Gran *et al.* (K2K Collaboration), *Phys. Rev. D* **74**, 052002 (2006).
  - [13] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **92**, 112003 (2015).
  - [14] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **91**, 112002 (2015).
  - [15] M. Dorman (MINOS Collaboration), *AIP Conf. Proc.* **1189**, 133 (2009).
  - [16] M. O. Wascko (SciBooNE Collaboration), *J. Phys. Conf. Ser.* **136**, 042028 (2008).

- [17] G. Fiorentini *et al.* (MINER $\nu$ A Collaboration), *Phys. Rev. Lett.* **111**, 022502 (2013).
- [18] L. Fields *et al.* (MINER $\nu$ A Collaboration), *Phys. Rev. Lett.* **111**, 022501 (2013).
- [19] A. Bodek, H. S. Budd, and E. Christy, *Eur. Phys. J. C* **71**, 1726 (2011).
- [20] J. G. Morfin, J. Nieves, and J. T. Sobczyk, *Adv. High Energy Phys.* **2012**, 934597 (2012).
- [21] G. T. Garvey, D. A. Harris, H. A. Tanaka, R. Tayloe, and G. P. Zeller, *Phys. Rep.* **580**, 1 (2015).
- [22] L. Alvarez-Ruso, Y. Hayato, and J. Nieves, *New J. Phys.* **16**, 075015 (2014).
- [23] O. Benhar and A. Fabrocini, *Phys. Rev. C* **62**, 034304 (2000).
- [24] A. M. Ankowski and J. T. Sobczyk, *Phys. Rev. C* **74**, 054316 (2006).
- [25] A. Butkevich, *Phys. Rev. C* **80**, 014610 (2009).
- [26] A. Bodek, M. Christy, and B. Coopersmith, *Eur. Phys. J. C* **74**, 3091 (2014).
- [27] C. Maieron, M. Martinez, J. Caballero, and J. Udias, *Phys. Rev. C* **68**, 048501 (2003).
- [28] A. Meucci, C. Giusti, and F. D. Pacati, *Nucl. Phys. A* **739**, 277 (2004).
- [29] J. E. Amaro, M. B. Barbaro, J. A. Caballero, T. W. Donnelly, A. Molinari, and I. Sick, *Phys. Rev. C* **71**, 015501 (2005).
- [30] T. Leitner, O. Buss, L. Alvarez-Ruso, and U. Mosel, *Phys. Rev. C* **79**, 034601 (2009).
- [31] A. Lovato, S. Gandolfi, R. Butler, J. Carlson, E. Lusk, S. C. Pieper, and R. Schiavilla, *Phys. Rev. Lett.* **111**, 092501 (2013).
- [32] V. Pandey, N. Jachowicz, T. Van Cuyck, J. Ryckebusch, and M. Martini, *Phys. Rev. C* **92**, 024606 (2015).
- [33] 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).
- [34] Y. Hayato, *Acta Phys. Pol. B* **40**, 2477 (2009).
- [35] M. Martini, M. Ericson, G. Chanfray, and J. Marteau, *Phys. Rev. C* **80**, 065501 (2009).
- [36] J. Nieves, I. R. Simo, and M. J. V. Vacas, *Phys. Rev. C* **83**, 045501 (2011).
- [37] R. Gran, J. Nieves, F. Sanchez, and M. J. V. Vacas, *Phys. Rev. D* **88**, 113007 (2013).
- [38] J. Nieves, I. R. Simo, and M. V. Vacas, *Phys. Lett. B* **707**, 72 (2012).
- [39] M. Martini, M. Ericson, and G. Chanfray, *Phys. Rev. C* **84**, 055502 (2011).
- [40] J. Nieves, I. R. Simo, and M. Vicente Vacas, *Phys. Lett. B* **721**, 90 (2013).
- [41] M. Martini and M. Ericson, *Phys. Rev. C* **87**, 065501 (2013).
- [42] M. Martini, M. Ericson, and G. Chanfray, *Phys. Rev. D* **85**, 093012 (2012).
- [43] J. Nieves, F. Sanchez, I. R. Simo, and M. J. V. Vacas, *Phys. Rev. D* **85**, 113008 (2012).
- [44] O. Lalakulich, U. Mosel, and K. Gallmeister, *Phys. Rev. C* **86**, 054606 (2012).
- [45] P. Coloma and P. Huber, *Phys. Rev. Lett.* **111**, 221802 (2013).
- [46] A. M. Ankowski, O. Benhar, P. Coloma, P. Huber, C.-M. Jen, C. Mariani, D. Meloni, and E. Vagnoni, *Phys. Rev. D* **92**, 073014 (2015).
- [47] E. Kolbe, K. Langanke, and S. Krewald, *Phys. Rev. C* **49**, 1122 (1994).
- [48] N. Jachowicz, S. Rombouts, K. Heyde, and J. Ryckebusch (ISOLDE Collaboration), *Phys. Rev. C* **59**, 3246 (1999).
- [49] N. Jachowicz, K. Heyde, J. Ryckebusch, and S. Rombouts, *Phys. Rev. C* **65**, 025501 (2002).
- [50] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **91**, 072010 (2015).
- [51] A. Furmanski, Ph.D. thesis, University of Warwick, 2015.
- [52] C. C. d. Atti and S. Simula, *Phys. Rev. C* **53**, 1689 (1996).
- [53] R. Bradford, A. Bodek, H. Budd, and J. Arrington, *Nucl. Phys. B, Proc. Suppl.* **159**, 127 (2006).
- [54] P. Sinclair, Ph.D. thesis, Imperial College London, 2014.
- [55] J. T. Sobczyk, *Phys. Rev. C* **86**, 015504 (2012).
- [56] T. Walton *et al.* (MINER $\nu$ A Collaboration), *Phys. Rev. D* **91**, 071301 (2015).
- [57] E. Oset and L. Salcedo, *Nucl. Phys. A* **468**, 631 (1987).
- [58] S. K. Singh, M. J. Vicente-Vacas, and E. Oset, *Phys. Lett. B* **416**, 23 (1998); **423**, 428 (1998).
- [59] A. Bodek, M. E. Christy, and B. Coopersmith, *AIP Conf. Proc.* **1680**, 020003 (2015).
- [60] C. Wilkinson, Ph.D. thesis, University of Sheffield, 2015.
- [61] C. Juszczak, J. T. Sobczyk, and J. Żmuda, *Phys. Rev. C* **82**, 045502 (2010).
- [62] T. Golan, C. Juszczak, and J. T. Sobczyk, *Phys. Rev. C* **86**, 015505 (2012).
- [63] A. M. Ankowski and J. T. Sobczyk, *Phys. Rev. C* **77**, 044311 (2008).
- [64] D. Casper, *Nucl. Phys. B, Proc. Suppl.* **112**, 161 (2002).
- [65] A. Aguilar-Arevalo *et al.* (MiniBooNE Collaboration), *Phys. Rev. D* **84**, 072005 (2011).
- [66] T. Katori, Ph.D. thesis, Indiana University (2008).
- [67] F. James and M. Roos, *Comput. Phys. Commun.* **10**, 343 (1975).
- [68] M. Maltoni and T. Schwetz, *Phys. Rev. D* **68**, 033020 (2003).
- [69] J. Conrad, C. Ignarra, G. Karagiorgi, M. Shaevitz, and J. Spitz, *Adv. High Energy Phys.* **2013**, 163897 (2013).
- [70] J. Kopp, P. A. N. Machado, M. Maltoni, and T. Schwetz, *J. High Energy Phys.* **05** (2013) 050.
- [71] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* **71**, 1554 (2011).
- [72] T. Golan, *Proceedings of the 10th International Workshop on Neutrino-Nucleus Interactions in the Few-GeV Region* (2015).
- [73] K. A. Olive *et al.* (Particle Data Group), *Chin. Phys. C* **38**, 090001 (2014).
- [74] J. Pumplin, D. Stump, and W. Tung, *Phys. Rev. D* **65**, 014011 (2001).
- [75] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, *Phys. Rev. D* **65**, 014012 (2001).
- [76] J. C. Collins and J. Pumplin, *arXiv:hep-ph/0105207*.
- [77] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. Lett.* (to be published).